

---

Konrad-Zuse-Zentrum  
für Informationstechnik Berlin

ZIB

Takustraße 7  
D-14195 Berlin-Dahlem  
Germany

MARTIN GRÖTSCHEL UND JOACHIM LÜGGER

**Die Wissenschaftsgalaxis  
Wie soll die zukünftige wissenschaftliche  
Informationsversorgung aussehen?**

# Die Wissenschaftsgalaxis

## Wie soll die zukünftige wissenschaftliche Informationsversorgung aussehen?

Martin Grötschel und Joachim Lügger  
Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB)  
Takustrasse 7, D 14195 Berlin-Dahlem, Deutschland  
groetschel@zib.de, Tel.: +49/30/84185-210

### Abstract

Wissenschaftler sind an qualitativ hochwertiger Information interessiert, selten an Informationsversorgung selbst. Es ist ihnen in Fragen der Organisation meistens egal, von welchem Versorger sie das Material, das sie benötigen, bekommen. Sie wollen einfach nur alles, was sie brauchen, jederzeit, überall, sofort und möglichst kostengünstig oder gar kostenlos haben. Im Prinzip ist das möglich. Seit mindestens zehn Jahren werden so gut wie alle wissenschaftlichen Dokumente (Artikel, Messdaten,...) elektronisch erstellt. Eigentlich bräuchte man diese nur ins Internet zu stellen, und damit wäre alles verfügbar. Irgendwie funktioniert diese einfache Idee nicht so richtig, obwohl viele Anstrengungen von den verschiedensten Einrichtungen und Individuen unternommen werden, z. B. durch Open-Access-Bemühungen. Was sind die Gründe dafür? Die Fortschritte beim schnellen Zugriff auf Information im Internet sind atemberaubend (siehe *Google*), dennoch ist der zielgerichtete Zugriff auf wissenschaftlich relevante Information weiterhin nur oberflächlich vorhanden. Die Hindernisse sind vielfältiger Natur: Bequemlichkeit und Unwissen der Autoren, Gewohnheiten der Nutzer und Informationsvermittler, Barrieren durch Copyrights und Verlagsstrategien, Eitelkeiten. In diesem Artikel werden einige der Probleme erläutert, die dem Traum des Aufbaus der wissenschaftlichen Universalbibliothek entgegenstehen. Gleichzeitig wird skizziert, wie man auf dem Weg zur Verwirklichung des Traums ein Stück vorankommen und wie die zukünftige wissenschaftliche Informationsversorgung gestaltet werden kann.

## 1. Ein Blick zurück

Vor zehn Jahren war die Welt noch in Ordnung. Bücher und wissenschaftliche Journale standen in der Bibliothek und waren (meistens) sorgfältig im Katalog erfasst. Bibliothekare kannten noch keine E-Mail.<sup>1</sup> Verleger druckten Publikationen noch auf Papier, in der Regel jedoch nur nach erfolgreichem *peer-review*-Prozess. Fachinformationszentren und Datenbankhersteller erfassten und bewerteten wissenschaftliche Publikationen abermals und dokumentierten die Früchte der wissenschaftlichen Arbeit in gedruckten Review-Organen. In einigen wenigen „glücklichen Fächern“, wie z.B. der Mathematik, gab es *Online*-Datenbanken, die den Stand des Faches weltweit dokumentierten und lokal recherchierbar machten.<sup>2</sup> In fast allen Disziplinen beschicken und unterwerfen sich Wissenschaftler diesem Publikationsregime mit ihrem gesamten publizierbaren Output.<sup>3</sup> Sie ernten dafür – abhängig vom Publikationsorgan – lokale oder weltweite Verbreitung ihrer Schriften bei Lesern, Anerkennung bei Fachkollegen, Anstellung oder Beförderung in Institutionen und – in seltenen Fällen – Ruhm für die Ewigkeit. Der Preis, die Langsamkeit der Publikationskette und ihre Kosten, schienen hinnehmbar, manchem sogar vorteilhaft zu sein.

Gegen Ende des zwanzigsten Jahrhunderts vollzog sich mit dem rapiden Wachstum des Internet und der globalen Verbreitung des World Wide Web ein gesellschaftlicher Wandel, der dieses System in Frage stellt. Das Jahr 2000 ist mit dem Platzen der Internet-Blase der erste Höhepunkt und zugleich die Ursache der nun verstärkt einsetzenden Globalisierung, siehe „Die Welt ist flach“ (Friedmann, 2006). Die Leistungsfähigkeit von Computern, Speichern und Netzen schien ins Grenzenlose zu wachsen bei zugleich rapide fallenden Kosten. Die neuen Technologien des Internet und des Web wurden überall verfügbar. Jedermann hatte nun die Mittel, an der Informationsgesellschaft aktiv teilzunehmen, gerade in den Wissenschaften, die global agieren und explosionsartiges Wachstum bereits im Publikationswesen kennen. Die Zahl der weltweit publizierten Artikel verdoppelt sich derzeit rund alle zehn Jahre, doch die Leistungsfähigkeit der Computer, Speicher und Netze wächst schneller (Lyman & Varian et al. 2001). Im Bereich des Publizierens stiegen jedoch die Kosten, ein Phänomen, das Bibliotheken in die Krise trieb. Doch dazu später mehr; zunächst zu einem Phänomen der Informationsgesellschaft, das vermutlich erst ihr Beginn ist.

Jedermann konnte nun selbst „publizieren“, weltweit, blitzschnell und zu sehr günstigen Kosten. Preprints gab es in einigen Fächern schon lange vor dem Internet, doch jetzt gab es Preprint-Server, die von ganzen Forschungs-Communities zur Beschleunigung ihrer wissenschaftlichen Kommunikation genutzt werden konnten (Bachrach et al., 1998). Fachbereiche stellten neuartige Formen wissenschaftlicher Journale ins Netz ([www.combinatorics.org](http://www.combinatorics.org)), Verlage und wissenschaftliche Fachgesellschaften zogen mit digitalen Versionen traditioneller Print-Publikationen nach ([www.elsevier.com](http://www.elsevier.com) und [portal.acm.org](http://portal.acm.org)), Bibliotheken mischten sich mit verlagsähnlichen Produkten ins Verlagsgeschäft ein ([highwire.stanford.edu](http://highwire.stanford.edu)), Rechenzentren an Universitäten mit Servern für Dissertationen und Hochschulschriften ([edoc.hu-berlin.de](http://edoc.hu-berlin.de)). Forschungs- und Supercomputerzentren vertrieben Software und Testdatensammlungen mittels Netzbibliotheken ([www.netlib.org](http://www.netlib.org) und [elib.zib.de](http://elib.zib.de)) die von der zugehörigen Print-Publikation

---

<sup>1</sup> Falls das unglaublich erscheint: Bei der ersten Vollversammlung des „Kooperativen Bibliotheksverbundes Berlin-Brandenburg“ im Konrad-Zuse-Zentrum rief die Frage nach dem Einsatz von E-Mail in den Bibliotheken Gelächter hervor.

<sup>2</sup> Die Mathematik verfügt über zwei global operierende umfassende Fachdatenbanken, MATH aus Deutschland und MathSci aus den USA kommend.

<sup>3</sup> Ausnahmen finden sich z. B. in der Industrieforschung und bei geheimen Militärprojekten.

der Algorithmen unabhängig geführt wurden. Die Biologen entwickelten neue Formen der Kollaboration und trieben die Entwicklung internationaler Informationsnetze voran (Human Genome Projekt). Hochenergiephysiker, die sich von Beginn an in der technischen Entwicklung der Netze engagiert hatte, distribuierten Forschungsdaten aus Experimenten beim CERN im Terabyte-Bereich an über die ganze Welt verstreute Forschungsgruppen. Förderorganisationen unterstützten die Entwicklung von Virtuellen Fachbibliotheken und Suchmaschinen (www.dfg.de). Innerhalb kurzer Zeit entstanden im Internet neue Ebenen wissenschaftlicher Information, die es in der traditionellen Welt des Publizierens in dieser Form nicht gab. Ein alter Traum der Wissenschaft, die *Universelle Bibliothek*, schien Wirklichkeit zu werden: „Alles sofort, jederzeit, überall und kostenlos zur Verfügung zu haben“ (Grötschel, 2001), jetzt zeitgemäß elektronisch, „an den Fingerspitzen“, wie es Bill Gates formulierte.

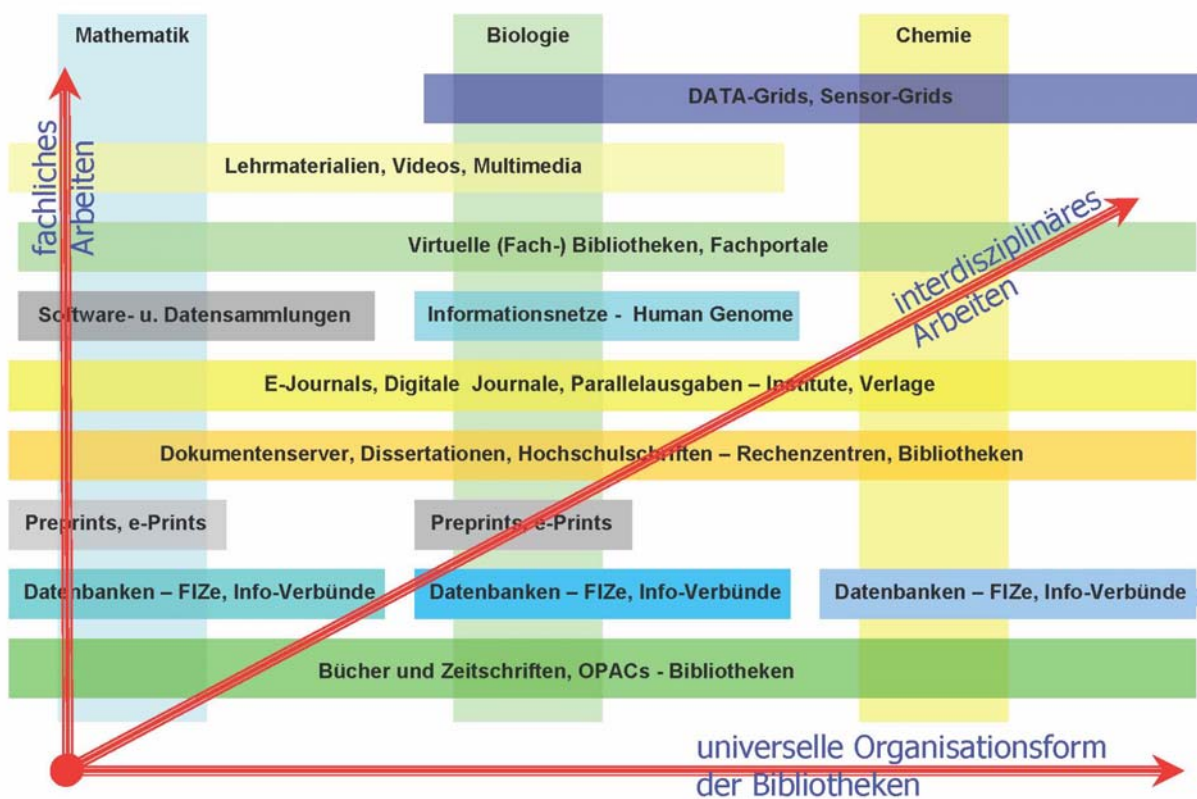


Abbildung: Ebenen wissenschaftlicher Information

Dies ist ein uralter Traum und zugleich ein moderner, ein „digitaler Traum“, der u. a. in der Mathematik seit mehreren Jahren verfolgt wird. In diesem Fachgebiet wird intensiv die Einrichtung der *World Digital Mathematics Library* im Zusammenhang mit der Digitalisierung mathematischer Publikationen diskutiert (IMU & CEIC, 2006): „*The grand vision of the World Digital Mathematics Library is to have all of the mathematical literature online and available through a central source to anyone who has a computer and an Internet connection.*“ (Jackson, 2003). Die *World Digital Mathematics Library* ist ein erstrebenswertes, aber noch fernes Ziel. „*This project will revolutionize the way in which mathematics conducts research – it is hard to imagine a single change that will have a greater influence. It remains a dream, of course, but an ideal dream on which to build foundations.*“ (Ewing, 2003).

Hier geht es jedoch um mehr als einen Traum. Der Wissenschaftsrat ordnet die Information im Internet den Hochschulbibliotheken zu und empfiehlt ihnen, zur Informationsversorgung der Wissenschaften unter anderem ihre vielfältigen Informationsangebote

- den Online-Katalog der eigenen Bibliothek und die Online-Kataloge anderer Bibliotheken des In- und Auslandes,
- den eigenen Verbundkatalog und die anderer Regionen,
- den nationalen Katalog und die anderer Länder,
- die Abstract- und Indexdienste von Verlagen, Zeitschriftenagenturen, Datenbankproduzenten und Bibliotheken,
- die digitalen Zeitschriften von Verlagen, wissenschaftlichen Gesellschaften und Institutionen,
- die digitalen Texte von Verlagen, von den Autoren direkt, von Bibliotheken und von Preprint-Archiven,
- die multimedialen Materialien von Verlagen, von Lehrenden direkt und von Bibliotheken,
- die Bestände in Spezialsammlungen (Karten, Audio-, Video-, Bildarchive),
- die retrodigitalisierten Materialien der eigenen Bibliothek oder die von anderen Bibliotheken und von Verlagen,
- die sonstigen Informationsressourcen im WWW (zu finden über Suchmaschinen),
- die Web-Verzeichnisse und fachlichen Information-Gateways

unter der einfachen und nutzerfreundlichen graphischen Oberfläche eines *Virtuellen OPAC* (Online Public Access Catalog) zu integrieren. „Durch die Verzahnung der Funktionalität und der Nutzungsmöglichkeiten der verschiedenen Datenbanken darf sich für den Anwender nicht mehr die Frage stellen, in welchem System er aktuell arbeitet.“ (Wissenschaftsrat, 2001).

Dieser Anspruch ist jedoch nicht einfach zu erfüllen. Hier kommen die unterschiedlichsten Ordnungsmechanismen ins Spiel. Die Welt der wissenschaftlichen Bibliotheken arbeitet mit universellen Schemata (alles muss hineinpassen), gewissermaßen „horizontal und quer“ durch alle Wissensgebiete. Die jeweiligen Wissenschaften und ihre Spezialfächer arbeiten „vertikal“ (vom Allgemeinen zum Speziellen und umgekehrt) mit ihren eigenen Begriffen und Taxonomien. Dort, wo Wissenschaft an ihre Grenzen gerät, gewissermaßen „zwischen“ den Fächern, sind jedoch andere Schemata erforderlich: notwendig sind „interdisziplinäre“ Begriffe und Schemata oder andere Darstellungsformen. Für diese gibt es jedoch selten gesicherte Taxonomien und Ontologien. Dies ist eine der Hauptschwierigkeiten interdisziplinären Arbeitens. Oft fehlen sogar die Worte gänzlich, z. B. bei Bildern, Fotografien und Experimentaldaten.

## **2. Informationsaustausch, autonomes Publizieren und *Open Access***

Die Naturwissenschaften und die Mathematik erkannten die neuen Möglichkeiten der Vernetzung schon in einem sehr frühen Stadium des Internet und des Web. Der erste und noch heute beispielhafte Preprint-Server von Paul Ginsparg stand im Los Alamos National Laboratory (Odlyzko, 1997). Dieses im August 1991 aufgelegte *e-print archive* (genannt *hep-th*, kurz für High Energy Physics – Theory) war noch an eine kleine Gruppe von Physikern in der String- und Gravitationstheorie gerichtet. Sehr schnell erweiterte sich das Fächerspektrum. Inzwischen ist daraus ein an der Cornell University angesiedelter *e-print service* für die Gebiete Physik, Mathematik, nichtlineare Wissenschaften, Computer Science, and quantitative Biologie geworden. Eine Kultur der frühzeitigen Information durch Preprints im Interesse der Beschleunigung der Forschung gab es auch in der Informatik und in der Mathematik. Zusammen mit der Chemie gründeten diese Fächer 1994 die IuK-Initiative in

Deutschland ([www.iuk-initiative.org](http://www.iuk-initiative.org)), die sich die Einrichtung und Verbesserung einer Infrastruktur zur wechselseitigen Information und Kommunikation zum Ziel setzte. Mit dem *Math-Net*, einem der ersten *Community-Networks* in den Fachwissenschaften ([www.math-net.de](http://www.math-net.de)), offerierte die Mathematik zunächst in Deutschland und später weltweit ein sehr breites Spektrum an Produkten. Man „nahm die Bibliotheken mit“, beschränkte sich aber nicht auf Bücher und Journale. *Math-Net* bezog auch mathematische Software und Datensammlungen, graphische und Multimedia-Informationen, Darstellungen von Projekten und Personen der Forschung bis hin zu Materialien der Lehre ein, alle Produkte also, die Mathematikern und Anwendern der Mathematik bei ihrer Forschung nutzen konnten. Alles das vernetzte das *Math-Net*-Projekt für jedermann nutzbar und kostenfrei im Internet und im World Wide Web.

Es war eine aufregende und euphorische Zeit. Ein gesellschaftlicher Umbruch lag in der Luft. Das globale Internet brachte das Versprechen des freien Informationsaustausches mit sich: Information ist jetzt blitzschnell, an jedem Ort der Welt erreichbar und kostenfrei. Andrew Odlyzko, ein international anerkannter Mathematiker, brachte es auf den Punkt, und er sah das nahende Hinsiechen der traditionellen wissenschaftlichen Journale voraus (Odlyzko, 1994). Er hatte diese Prognose weltweit mit Fachkollegen in einer persönlichen elektronischen Mailingliste diskutiert, einem der ersten „Blogs“. Die Mathematik müsse die Gelegenheit beim Schopfe greifen und das Publizieren umgehend selbst in die Hand nehmen, war die implizite Schlussfolgerung. Bibliotheken hatten überdies einen neuen Akzent in diese Diskussion gebracht: Wissenschaftliche Journale sind zu teuer. Schuld daran sind vor allem die großen europäischen Verlage. Würde es nicht genügen, so Stimmen in den USA, für Amerika immer nur ein einziges Exemplar einer Zeitschrift zu kaufen und diese per E-Mail – kostenfrei – an alle Bibliotheken in den Vereinigten Staaten zu verteilen (Okerson, 1997)?

Damit ist der Kern der inzwischen weltweiten *Open Access*-Bewegung kurz charakterisiert: eine von großen Verlagen in unverantwortlicher Art und Weise verschuldete Preis-Kosten-Spirale, die wissenschaftliche Bibliotheken in die Krise treibt und die – zeitgleich gegebene – Möglichkeit, dagegen mit den neuen Technologien des Internet und des Web anzugehen. Wissenschaftler in aller Welt sahen sich in Not. Bibliotheken gehören zu ihrer Infrastruktur.

Die *International Mathematical Union (IMU)*, der Weltverband der Mathematik, nahm sich 1998 dieses Problems durch Gründung des *Committee on Electronic Information and Communication (CEIC)* in besonderer Weise an und gab eine Reihe von Empfehlungen ([www.ceic.math.ca/Publications/index.shtml](http://www.ceic.math.ca/Publications/index.shtml)) für wissenschaftliche Autoren, Verlage und Bibliotheken heraus, die eine Neustrukturierung des wissenschaftlichen Publizierens der Mathematik in den Zeiten des Internet zum Ziel hatten:

- *Math-Net-Charter, Terms of References*
- *Best Current Practices for Mathematicians, Librarians and Publishers*
- *Call to all Mathematical Institutions to install Math-Net-Pages*
- *Call to all Mathematicians to make Publications Electronically Available*
- *Copy-Right-Recommendations for Mathematical Authors*
- *Journal Prices*

Eine wichtige und zugleich typische Internet-Idee des IMU/CEIC war das – in den Worten von Friedmann – „Hochladen“ der *Personal Collected Work* ins Netz und damit der Übergang von der passiven zur aktiven Teilnahme am Prozess des Publizierens. In den Worten der IMU: „*Mathematics ages slowly. Access to older literature is important for most mathematicians, and yet much of the older literature is likely to remain unavailable in*

*electronic form in the immediate future. Mathematicians can change that by taking collective action. Whenever legally and technically possible, mathematicians are encouraged to scan their old papers and post them on their homepages, making their "collected work" readily available to all. This relatively small effort on the part of every mathematician will provide enormous benefit to the entire community."*

Rückblickend betrachtet mag dieser Einsatz *Pro Open Access* als hoch und für manchen im Vergleich zu den Erwartungen als wenig erfolgreich angesehen werden. Das Verlagswesen ist durchaus beeindruckt, lernt aber schnell und reagiert. Man versteht sein Geschäft und experimentiert mit neuen Vertriebsmodellen. Heute machen große Verlage noch größere Gewinne als vor zehn Jahren. Mit dem Hochladen allein ist es offenbar nicht getan. Wertschöpfungsketten sind nicht einfach zu imitieren und noch schwieriger zu kopieren (Friedman, 2006).

Beim wissenschaftlichen Publizieren geht es in erster Linie um Renommee und Anerkennung (gelegentlich auch um Eitelkeit), technisch um „gefunden werden“, rechtlich um das internationale Copyright, aber nur ganz selten um Gewinnstreben von Autoren. Im Internet geht es nicht zuletzt um den Einsatz von Hochtechnologie zur Bereitstellung und Verbreitung von digitalem „Papier“. Aber digitale Publikationen sind kein Papier mehr, sondern Software, mit allen Konsequenzen und Komplexitäten, insbesondere technischer und rechtlicher Art. Software und ihre Verfügbarkeit muss permanent gepflegt werden. Digitaler Content wird deshalb für die Nutzung – befristet! – lizenziert und nicht mehr einfach nur verkauft. Wer Lizenzen kündigt, verliert den Zugriff auf zurückliegende Versionen (*Back-Files*) und muss dafür noch einmal bezahlen. Die digitalen Regale der wissenschaftlichen Journale stehen heute bei den Verlagen und nicht in den Bibliotheken. Im Bereich des Publizierens öffnet sich eine „*Digital Divide*“.

Deshalb ist der Übergang des mathematischen Publikationswesens (und analoge Bemühungen vieler anderer Fachgebiete) in Richtung auf *Open Access*, auf das „Hochladen“ der eigenen Information und auf aktive Teilnahme am Publizieren in einem ganz grundsätzlichen Sinne richtig, notwendig und wichtig. Preprint- und Dokumentenserver, Repositories und Informationsnetze, Informationsvermittlung und Austausch über E-Mail und Suchmaschinen sind nun fester Bestandteil der wissenschaftlichen Kultur und nicht mehr wegzudenken. Die *Bibliothek der Zukunft* muss diese Komponenten mit einbeziehen. Die Wissenschaftsgalaxis reicht jedoch weit darüber hinaus, und sie dehnt sich weiter aus. Wer hier navigiert, braucht Hochtechnologie.

Die *Bibliothek der Zukunft*, die die Wissenschaft selbst natürlich nicht treibt die aber ein äußerst wichtiges „Planetensystem der Wissenschaftsgalaxis“ ist, muss heute den unterschiedlichsten Anforderungen gerecht werden. Der Wissenschaftsrat hat diese in seinen Empfehlungen bereits genannt. Die *Bibliothek der Zukunft* wird *Digitalen Content* aus den unterschiedlichsten Quellen und Servern, mit Dokumenten in den unterschiedlichsten Formaten im eigenen Portal unter einer einheitlichen Oberfläche integrieren müssen. Schafft sie das nicht, dann zwingt sie den Nutzer zum Wechsel des Portals bzw. zu einem Bruch in der Navigation.

Die *Bibliothek der Zukunft* wird den freien und den lizenzierten Content in harmonischer Art und Weise integrieren müssen. Das ist keine einfache Aufgabe, denn dabei muss sie u. a. auf Verlagsserver zurückgreifen. Die Landschaft der Portale der Verlage sieht wenig einheitlich aus. Alle Verlagsportale sind unterschiedlich gebaut. Journale eines Faches sind in den Servern schwierig zu finden. Der Zugriff auf lizenzierte Dokumente ist von Verlag zu Verlag

unterschiedlich geregelt. Die Navigationsmethoden und Suchsysteme sind allesamt verschieden strukturiert, so dass manchmal auch der willigste Nutzer verzweifelt; mit anderen Worten, die Bibliothekare müssen nun auch Ordnung in dieses Chaos bringen, wenn sie den Bedürfnissen der Nutzer gerecht und einen guten Service anbieten wollen.

### 3. Informationsflut, Position der Bibliotheken, Selbstinformation

Um das Jahr 2000 erreichte die Expansion des Internet und des World Wide Web ihren ersten Höhepunkt. Die Welt des wissenschaftlichen Publizierens geriet in Bewegung. Die am Prozess des Publizierens beteiligten ehemaligen Partner wurden nun zu Konkurrenten, sahen neue Chancen oder fürchteten um ihre bisherige Position; und sie traten auch im offenen Wettbewerb gegeneinander an. Einzelne Wissenschaftler wandten sich dem Online-Verlegen von wissenschaftlichen Zeitschriften zu und legten auch Proceedingsbände elektronisch auf, wie z.B. die umfangreichen dreibändigen Proceedings des *International Congress of Mathematicians* 1998 ([www.mathematik.uni-bielefeld.de/documenta](http://www.mathematik.uni-bielefeld.de/documenta)). Die Bibliothek der Stanford Universität entwickelte Verfahren für Online-Produktion und -Vertrieb ganzer Serien von Journalen (HighWire Press). Verlage übernahmen mit dem Aufbau von Digitalen Bibliotheken Teile des Buchhandels- und des Bibliotheksgeschäfts (Elsevier Science). Förderorganisationen der Wissenschaften (DFG) trieben die Entwicklung von Verlagsinfrastrukturen bei Bibliotheken voran ([www.gap-portal.de](http://www.gap-portal.de)). Die großen Informationsbibliotheken, Fachinformationszentren und die DFG-geförderten Virtuellen Fachbibliotheken unternahmen den Versuch, mit *Google* gleichzuziehen ([www.vascoda.de](http://www.vascoda.de), ein Internet-Portal für alle, die wissenschaftliche Information suchen). Die Max-Planck-Gesellschaft baut gemeinsam mit dem Fachinformationszentrum Karlsruhe und mit Unterstützung des Bundesministeriums für Bildung und Forschung einen Publikationsserver auf ([www.escidoc-project.de](http://www.escidoc-project.de)), der eine richtungweisende Plattform für Kommunikation und Publikation in wissenschaftlichen Forschungsorganisationen sein soll.

Auslöser dieser dynamischen Entwicklung waren die hunderte von Protokollen und Standards des Internet und des Web, durch die eine globale Plattform entstand, deren Ressourcen von Internet-Suchmaschinen abgesehen, indiziert und in Kataloge integriert wurden, allen voran von *Google*. Das Netz begann, zum Gedächtnis der Welt zu werden. Die Treffergenauigkeit von *Google* erzeugte die erste wirklich nutzbare globale „Bibliothek der Welt“, genauer, der digitalen Ressourcen der Welt, die für *Google* offen und zugänglich waren. (Das sind bei weitem nicht alle Ressourcen, obwohl es dem Nutzer so erscheinen mag.) Heute gehört mit *Google Scholar* und der *Google Book Search* bereits ein großer Teil der Publikationen und Bücher, die früher allein den Bibliotheken vorbehalten waren, zur Domäne des Netzes. Doch wo stehen die Bibliotheken in diesem „Rennen“?

Die Bibliotheken werden hierzulande noch als die „Hüter des Wissens“ angesehen. Die Informationsmenge, die von großen Bibliotheken, wie zum Beispiel der Staatsbibliothek Preußischer Kulturbesitz zu Berlin (SBB) vorgehalten wird, ist durchaus vergleichbar mit der Menge an Information, die das Netz zu bieten hat, an Qualität ohnehin. Aber, macht es die Bibliothek ihren Nutzern leicht oder schwer? Die entscheidende Frage ist jedoch, wo geht der Nutzer hin, wenn er Information sucht? Geht er in die Bibliothek oder wendet er sich woanders hin? Dieser Frage ist das OCLC, der wohl größte Bibliotheksverbund, mit einer Umfrage nachgegangen, über die hier in ausgewählten Auszügen berichtet werden soll (OCLC, 2005). Die Umfrage richtete sich an Personen in Australien, Großbritannien, Kanada, Indien, Singapur und den Vereinigten Staaten, von denen insgesamt 3348 geantwortet haben. Alle Altersgruppen sind unter den „Rückläufern“ vertreten.



Auf die Frage „*Where do you typically begin your search for information on a specific topic*“ antworten 84% der befragten Personen mit „*Search Engine*“ und nur 1% mit „*Library Web site*“. Als weitere Informationsquellen werden u.a. genannt: E-Mail mit 6%, Topic-specific Web sites mit 2%, Online-news mit 2%, Online bookstore mit 1%, Online databases mit 1%. Was könnte den Wandel im Informationsverhalten deutlicher charakterisieren, als diese Ergebnisse zu der Frage, wo die Informationssuche beginnt. Die befragten Personen sind in der Nutzung elektronischer Mittel nicht unerfahren. Auf die Bitte „*Please indicate if you have used the following information sources, even if you have used them only once*“ antworten 74% der befragten Personen mit E-Mail, 72% mit Search Engine sowie, u.a., 51% mit Instant Messaging, 50% mit Online news, 47% mit Online bookstore, 30% mit Library Web site und 28% mit Electronic Magazines/Journals. Der größte Teil der befragten Personen benutzt mit 62% die Suchmaschinen von *Google.com*, zu 18% *Yahoo.com*, sowie zu 7% *MSN.com*. Alle anderen Suchmaschinen sind weit abgeschlagen. Weitere Ergebnisse der OCLC-Studie sind:

- „*Quality and quantity of information are top determinants of a satisfactory information search. Search engines are rated higher than librarians.*“
- „*The criterion selected by most information consumers to evaluate electronic resources is that the information is worthwhile. Free is a close second. Speed has less impact.*“
- „*Respondents do not trust purchased information more than free information. The verbatim comments suggest a high expectation of free information.*“

Dabei stehen Nutzer den Bibliotheken nicht unfreundlich gegenüber: „*Information consumers use the library*“, jedoch: „*They use the library less and read less since they began using the Internet. The majority of respondents anticipate their usage of libraries will be flat in the future.*“. Das sind keine besonders guten Aussichten für die *Bibliothek in der Zukunft*, zumal die Nutzer in der Bibliothek keine elektronischen Ressourcen erwarten: „*Most information consumers are not aware of, nor do they use, most libraries' Electronic information resources.*“, sondern Bücher: „*'Books' is the library brand. There is no runner-up.*“.

Bei Studenten schneiden Bibliotheken zwar insgesamt etwas besser ab: „*Most respondents agree the library is a place to learn.*“, jedoch: „*Only 10 percent of college students indicated that their library's collection fulfilled their information needs after accessing the library Web site from a search engine.*“.

Der Wandel im Verhältnis der Informationsverbraucher tritt deutlich in den Empfehlungen der Nutzer an ihre Bibliotheken zu Tage: „*Find ways to get material to people, rather than making them come to the library.*“. Dabei legen sie Wert auf Autonomie: „*Information consumers like to self-serve. They use personal knowledge and common sense to judge if electronic information is trustworthy. They cross-reference other sites to validate their findings.*“. Und „*Ninety percent of respondents are satisfied with their most recent search for information using a search engine.*“ (OCLC, 2005).

Diese Funde decken sich mit einem Phänomen, das Friedman „Selbstinformation“ nennt (Friedmann, 2006, S. 221 ff). „*Selbstinformation ist das, was für ein Unternehmen Open Source, Outsourcing, Insourcing, Offshoring und Wertschöpfungsketten sind.*“ Und: „*So paradox es klingen mag: Sie ist eine Form der Zusammenarbeit mit sich selbst – man wird zu seinem eigenen, selbstbestimmten Forscher, Redakteur, und Unterhaltungschef, ohne in eine Bibliothek oder in ein Kino zu gehen, ohne fernzusehen. Selbstinformation ist die Suche nach Wissen und die Suche nach gleichgesinnten Menschen und Gemeinschaften.*“. Das Internet bietet dem Wissensdurstigen heute das, was er früher in der Bibliothek fand, einen Treffpunkt

und Gemeinschaft im Wissen. Auf diese Art des Wissens sind die Leute heute begierig. Nicht überall haben das die älteren, in Führungspositionen befindlichen Personen bemerkt.

#### 4. Vom Suchen und Finden

Bibliotheken geraten in Deutschland heute auf mehrfache Art und Weise unter Druck. Im Zuge der fortschreitenden Globalisierung der Haushalte der Hochschulen stehen sie nun mit den Wissenschaftlern innerhalb der Universitäten in Konkurrenz um die Haushaltsmittel. Sie erleben stagnierende oder gar sinkende Bibliotheksetats und simultan steigende Beschaffungskosten, hier in erster Linie für wissenschaftliche Journale. Dieses führt zu starken Einschränkungen bei der Beschaffung von Büchern und Medien. Inzwischen wird auch in bibliothekarischen Kreisen realisiert, dass Internet-Firmen und Verlage ihnen Konkurrenz bereiten, jedoch halten Bibliotheken bis heute beharrlich an den hergebrachten Traditionen der Katalogisierung fest. „Back to the Basics“ ist das Leitmotiv einer aktuellen internationalen Konferenz, in der es um die Fragen der Verwendung der *Anglo American Cataloguing Rules (AACR2)* geht. Im Zentrum der bibliothekarischen Arbeit steht nach wie vor die Katalogkarte, auch wenn diese in digitaler Form im Bibliothekssystem verwaltet, dem Nutzer im OPAC zur Verfügung gestellt, nun als *Metadata* bezeichnet wird.

Die Katalogkarte dient dem Zweck, ein Buch in einer Bibliothek auf eindeutige Weise zu bezeichnen, damit es gefunden werden kann.<sup>4</sup> Ist die Katalogkarte zu mehr geeignet, etwa zur Orientierung in einem Fachgebiet? Dazu müsste die Katalogkarte qualitative Aussagen über ein Buch und dessen Stellung in seinem Fachgebiet enthalten. Dazu ist die Katalogkarte aber weder gedacht noch geeignet, auch nicht in ihrer modernen Form als Metadatensatz. Und wenn diesem Metadatensatz nun ein Abstrakt oder<sup>5</sup> ein Referat zugeordnet wird? Dann wird sie dem Kundigen, der ein Fach bereits gut kennt und den Überblick hat, vielleicht etwas sagen, den Unkundigen aber dennoch im Stich lassen, denn aus einer Katalogkarte erschließt sich weder der Inhalt eines Buches, noch ergibt sich aus ihr die thematische Übersicht eines Faches. Würde vielleicht ein Inhaltsverzeichnis weiterhelfen?<sup>6</sup> Das kommt darauf an, ob es im Buch um Allgemeinwissen oder um wissenschaftliches Fachwissen geht. Um das entscheiden zu können, sollte der Leser mehr von einem Buch sehen können, etwa kurze Inhaltsangaben vom Umschlag oder von den ersten Innenseiten. Auch die Copyrightseite wäre hilfreich, auf der sich wesentliche Daten über Ursprung und Aktualität finden, wie Informationen über den Autor oder den Herausgeber, vielleicht sogar eine international gebräuchliche Klassifikation.

Damit ist der „Kompetenzbereich“ der Katalogkarte bereits weit überschritten. Wir befinden uns bereits im klassischen *Information Retrieval*. Trotzdem ist es hilfreich, über ein Modell der erweiterten Repräsentation von wissenschaftlichen Büchern und Journalen für Bibliotheken „experimentell“ nachzudenken. Im nächsten Schritt fügen wir gedanklich Index-Seiten hinzu und befinden uns damit schon auf dem Weg zum Volltextretrieval, wie wir es heute von den großen Internet-Suchmaschinen kennen. Der Index eines Buches würde den o.g. erweiterten Metadaten einen fast vollständigen Satz von Schlüsselworten hinzufügen, sofern der Autor gut gearbeitet hat. Diese Erweiterung wird aber den Rücklauf (*Recall*) bei Nutzeranfragen drastisch erhöhen (kein Mensch kann Tausende „Hits“ anschauen), so dass dies nur im Zusammenhang mit einem leistungsfähigen Ranking zu empfehlen ist. Mit der Idee, nun die erweiterten Metadaten auch noch um Literaturzitate und andere Referenzen zu

---

<sup>4</sup> Zu diesem ist sie mit einem Standortverweis versehen und es kann der Ausleihzustand angehängt werden.

<sup>5</sup> Hier befinden wir uns bereits auf dem Gebiet der klassischen traditionellen Fachdatenbanken.

<sup>6</sup> Hier befinden wir uns im Bereich des *Catolog Enrichments*, das in den Bibliotheken diskutiert wird.

ergänzen, befinden wir uns fast schon im zu Google vergleichbaren Internet-orientierten Volltextretrieval. Bibliotheken sind mit ihrem OPAC, einer Ansammlung von digitalen Katalogkarten, von dieser Art der Informationssuche weit entfernt.

Die technischen Standards der Informationssuche werden heute von *Google.com.* gesetzt, jetzt auch im Bereich der wissenschaftlichen Literatur. Dazu arbeitet diese Firma mit einer Reihe von Verlagen und mehreren der weltweit bedeutendsten Bibliotheken zusammen, um digitale Journale zu indexieren und Bücher zu digitalisieren. Nutzer erhalten von *Google* wesentlich mehr als ein einfaches Information Retrieval. Die *Google Web*-Suche, *Google Scholar* und jetzt auch die *Google Book Search* bieten dem Nutzer ein hochklassiges *Information Mining* in sehr großen Datenbeständen. Ihr ursprüngliches und auch derzeitiges Erfolgsrezept sind der *Page Rank* im Web (Brin & Page 1998) und das dem *Page Rank* verwandte *Citation Ranking* bei *Google Scholar* und möglicherweise auch bei der *Google Book Search* (Vise, 2006). Google Inc. verknüpft Fundstellen durch intelligent eingesetzte *Open Linking*-Dienste (Van de Sompel & Beit-Arie 2001) mit Dokumenten im Internet, in den Bibliotheken und im Buchhandel.

Wer immer im Internet eine eigene Suchmaschine entwickelt oder eigene Content-Angebote erwägt, wird sich mit *Google* messen lassen müssen, ob er das will oder nicht. *Google* allein reicht auch nicht aus, denn *Google* und andere Internet-Suchmaschinen tauchen noch nicht tief genug in das so genannte *Deep Web* hinein. Zwar steht dem Wissenschaftler jetzt der digitale Zugriff auf „alles, jederzeit und überall“ zur Verfügung, jedoch nicht immer kostenlos und die zentrale Frage „gewusst wo“ ist nach wie vor nicht gelöst. Internet-Suchmaschinen indexieren nur die offene Oberfläche des Web. *Google* ist nur ein Anfang. Und wer von einer nahtlosen Navigation im World Wide Web träumt, wird bald enttäuscht. Die Wirklichkeit sieht anders aus. Die Welt der Publikationen im Netz ist in vielfältiger Weise fragmentiert und durch Zäune zertrennt, die nur mit hohem technischem und/oder finanziellem Aufwand überwunden werden können. Das liegt nicht im Interesse der Wissenschaft, aber durchaus im Interesse einiger, die mit Information handeln.

Dies soll am Beispiel der Mathematik kurz erläutert werden. Leser, die sich im Netz bewegen, werden mit einer Vielzahl unterschiedlich strukturierter Homepages konfrontiert. Oft ist die Mathematik dort nur schwer zu finden. Einheitliche Navigationsstrukturen existieren nicht. Nutzer, die nach einem Fund auf einen Artikel zugreifen wollen, werden auch immer wieder mit anderen Bezahl- und Zugangskontroll-Systemen konfrontiert<sup>7</sup>. Ist der gefundene Artikel vielleicht bereits von der eigenen Bibliothek oder durch die DFG bezahlt worden? Um das herauszufinden, muss erst der Server der Bibliothek aufgesucht werden – und dieser ist wie oben ausgeführt nicht einfach zu bedienen.

Die zentrale Frage aber ist, ob Internet-Suchmaschinen das „richtige“ Ergebnis liefern. Findet z.B. *Google Scholar* den für Mathematiker wichtigen „Content“ also das Material, nachdem inhaltlich gesucht wurde? Das ist unklar. *Google* gibt nicht an, was indexiert wird und was nicht. Wird *Google Scholar* vielleicht in Zukunft alle wesentlichen Journale der Mathematik in die Indexierung einbeziehen können? Auch das ist unklar, denn erhebliche Bestände wesentlicher mathematischer Literatur befinden sich in den Händen der kommerziellen Konkurrenz. Wird diese mit *Google Inc.* kooperieren?

Bei *Google Scholar* findet man wichtige neue Funktionen, wie z.B. das Ranking nach und die Navigation zu den Zitationen. Aber wie kann der Mathematiker denn sicher sein, bei seinen

---

<sup>7</sup> Dabei sind dies nicht einmal die neuen Verhältnisse, die Verlage mit DRM-Technologien (Digital Rights Management) einführen werden.

„Sprüngen in die Nachfolge-Publikationen“ einer Arbeit die wesentlichen Anwendungen zu finden? Hier ist die Vollständigkeit des mathematischen Basismaterials gefragt und die ist nach wie vor eine Domäne der Fachgesellschaften und der Fachdatenbanken.

Im *Open Access*-Bereich sind die Verhältnisse nicht sehr viel anders. Leider verbergen sich bei manchen *Open Access*-Servern die als „offen zugänglich“ gedachten Publikationen für Internet-Robots hinter CGI-Interfaces, speziellen oder komprimierten Dokumentenformaten oder anderen technischen Hindernissen. Nicht wenige Dokumente sind in (derzeit) nicht absuchbaren Formaten gespeichert. Diese bleiben damit, was Suchmaschinen betrifft, im *Deep Web* verborgen. *Open Access*-Server, die es bis ins World Wide Web geschafft haben, verwenden mangels Standards unterschiedliche Klassifikationen und Navigationsstrukturen. Solche Sammlungen sind nicht selten fachlich durchmixt und damit für Nutzer nur schwer durchschaubar. Relevante Artikel sind in ihnen aber schon allein deshalb schwierig zu finden, weil es in weiten Teilen des *Open Access*-Bereiches noch kein *peer-reviewing* oder eine vergleichbare Qualitätskontrolle gibt. Es bleibt dem Leser nichts anderes übrig, als auf die Seriosität der Institutionen zu vertrauen, die solche Server aufbauen. Diese werden Schriften ihrer eigenen Mitglieder aber nur selten ablehnen, auch nicht aus Qualitätsgründen.

## 5. Open Access – Ein Gegenentwurf

*Open Access* ist im Wissenschaftsbereich entstanden. Es ist eine Grundhaltung, die auch die Entwicklung des Internet als Netz der Netze und des Web erst möglich gemacht hat. Offene Software (*Open Source*), offen zur Erprobung gestellte Protokollentwicklungen, die für alle Beteiligten offene Diskussionen von Kommunikationsstandards und Schnittstellen sowie der offene Austausch von Anwendungen, wie Mail, Browser und andere Clients auf einer großen Vielfalt von unterschiedlichen Computern und Systemen stehen im Zentrum. Offenheit wurde zum Grundprinzip des Internet und zur Offenheit gehörte die Skalierbarkeit, d.h., das Prinzip weltweiter Einsetzbarkeit, mit anderen Worten, der Verbreitung, ohne primär auf die eigenen Verdienstmöglichkeiten zu achten. So entstand durch kostenfreien Tausch von Programmen, Protokollen und Prinzipien eine weltweite Plattform, die sich dann auch kommerziell nutzen ließ. Im Vordergrund stand aber immer die Kommunikation als solche. Die Protokolle des Internet und des World Wide Web stehen für jedermann nutzbar und einsehbar im Netz.

„*Open Access*“ ist in diesem Sinne auch ein Grundprinzip der Wissenschaften, insbesondere des wissenschaftlichen Publizierens, das in seiner ursprünglichen Wortbedeutung heißt, neue Erkenntnisse publik zu machen und sie der Öffentlichkeit zu übergeben. Wer zuerst etwas publiziert (nicht, wer zuerst etwas weiß), mit dessen Namen wird die Erkenntnis verbunden oder gar bezeichnet. Die Verbreitung einer Erkenntnis in der Öffentlichkeit und die Anerkennung dieses Verdienstes ist der Sinn des Publizierens. Dieses Prinzip scheint im Zuge der „Ökonomie des Publizierens“ zu einem großen Teil untergegangen zu sein. Der offene Zugang zum Wissen der Welt ist aber ein Gut, das für die Weltgemeinschaft ganz grundsätzlich notwendig ist. Ohne sie wäre der ärmere Teil der Welt „abgetrennt“ und könnte sich nicht an der Weiterentwicklung des Wissens, des gemeinsamen Gutes, beteiligen. Das ist für die *Have Nots* schlecht, aber auch für das Wissen selbst. Diese Grundhaltung formulieren zwei *Open Access*-Initiativen in folgenden Thesen.

Budapest Open Access Initiative ([www.soros.org/openaccess](http://www.soros.org/openaccess))

- *The public good they [scientists and scholars] make possible is the world-wide electronic distribution of the peer-reviewed journal literature and completely free*

*and unrestricted access to it by all scientists, scholars, teachers, students, and other curious minds.*

- *Removing access barriers to this [research] literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge.*

#### Berlin Declaration on Open Access to Knowledge (MPG 2003)

- *Promote the Internet as a functional instrument for a global scientific knowledge base and human reflection and to specify measures which research policy makers, research institutions, funding agencies, libraries, archives and museums need to consider.*
- *We define open access as a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community.*

*Open Access* bedeutet nicht, sich auf Preprint-Server und Autonomes Publizieren zu beschränken, es bedeutet auch nicht, allein auf (Forschungs-) Literatur zu fokussieren. *Open Access* umfasst die Ressourcen des Wissens ganz allgemein. Es bezieht zum Beispiel auch Daten der Forschung und Materialien der Lehre mit ein. *Open Access* darf auch nicht bedeuten, dass die Wissenschaften, aus welchen Gründen auch immer (z.B. wegen des Copyrights), nur noch mit lokal begrenzten Tauschzirkeln arbeiten dürfen. Und es darf auch nicht sein, sich allein an den Bedürfnissen der *Haves – excluding Have Nots* zu orientieren.

Die *Berlin 2 Open Access-Conference* ([www.zim.mpg.de/openaccess-cern/](http://www.zim.mpg.de/openaccess-cern/)) führt die Ziele der Implementierung von *Open Access* wie folgt weiter aus

*Our mission of disseminating knowledge is only half complete if the information is not made widely and readily available to society. New possibilities of knowledge dissemination not only through the classical form but also and increasingly through the open access paradigm via the Internet have to be supported. We define open access as a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community.*

*In order to realize the vision of a global and accessible representation of knowledge, the future Web has to be sustainable, interactive, and transparent. Content and software tools must be openly accessible and compatible.*

Die Erklärung definiert ferner das breite Spektrum der *Open Access*-Beiträge “*Open access contributions include original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material*“ und nennt zwei Bedingungen, die der Autor solcher Beiträge erfüllen muss. Die erste betrifft die Gewährung freizügiger Nutzungsrechte, die zweite ist eher technischer Art und betrifft das „Hochladen“ (*Posting*) von Artikeln durch Autoren.

1. *The author(s) and right holder(s) of such contributions grant(s) to all users a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship (community standards, will continue to provide the mechanism for enforcement of proper attribution and responsible use of the published work, as they*

*do now), as well as the right to make small numbers of printed copies for their personal use.*

2. *A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in an appropriate standard electronic format is deposited (and thus published) in at least one online repository using suitable technical standards (such as the Open Archive definitions) that is supported and maintained by an academic institution, scholarly society, government agency, or other well-established organization that seeks to enable open access, unrestricted distribution, interoperability, and long-term archiving.*

Die grundsätzliche Bedeutung der *Open Access*-Bewegung wird weithin anerkannt. Alle großen deutschen Forschungsorganisationen haben die *Open Access*-Erklärung (auch Berlin 1 genannt) unterschrieben. Trotzdem kommt *Open Access* nur langsam voran, nicht nur in Deutschland. Das lässt zumindest eine aktuelle, breit angelegte Studie des europäischen Publikationsmarkts (EU-Commission, 2006) ahnen, die in folgenden Empfehlungen gipfelt:

- A1 GUARANTEE PUBLIC ACCESS TO PUBLICLY-FUNDED RESEARCH RESULTS SHORTLY AFTER PUBLICATION (moving wall)
- A2 AIM AT A 'LEVEL-PLAYING FIELD' IN TERMS OF BUSINESS MODELS IN PUBLISHING
- A3 'EXTENDED QUALITY' RANKING OF SCIENTIFIC JOURNALS
- A4 GUARANTEE PERENNIAL ACCESS TO SCHOLARLY JOURNAL DIGITAL ARCHIVES
- A5 FOSTER INTEROPERABLE TOOLS TO IMPROVE KNOWLEDGE VISIBILITY, ACCESSIBILITY AND DISSEMINATION
- B1 PROMOTE PRO-COMPETITIVE PRICING STRATEGIES
- B2 SCRUTINIZE FUTURE SIGNIFICANT MERGERS
- B3 PROMOTE THE DEVELOPMENT OF ELECTRONIC PUBLICATIONS
- C1 SETTING-UP AN ADVISORY COMMITTEE
- C2 FURTHER INVESTIGATION

Die mit der angestrebten Änderung des Publikationsverhaltens der Wissenschaften einhergehenden strukturellen, rechtlichen, organisatorischen, kommerziellen und nicht zuletzt auch sozialen Modifikationen sind vielfältig und bisweilen komplexer Natur. Gesellschaften folgen einer neuen Technologie nur langsam, nicht nur im Feld des Publizierens, und die mit dem Publizieren verknüpften notwendigen Modifikationen sind weniger technischer Art. Wie kompliziert das gesamte Umfeld tatsächlich ist, tritt erst heute nach einer intensiven Phase des Experimentierens deutlich ans Licht. Trotzdem sind substantielle Fortschritte zu vermerken; vgl. das *Open Access*-Spezial 1/2006 des Forschungsjournals „Wissenschaftsmanagement“.

Globale Communities setzen heute auf der Grundlage der globalen Plattform des Internet und des Web wichtige *Open Access*-Akzente. *Open Access*-Literatur wird schneller zitiert und erste *Open Access*-Journale erreichen bereits hohe Impact-Faktoren. In einigen Science-Communities ist die Literatur mit Daten eng vernetzt. Es ist inzwischen Standard, dass auch die traditionelle wissenschaftliche Literatur in digitaler Form ins Netz gestellt wird. Diese neue Möglichkeit nutzen nicht nur traditionelle Verlage, sondern vor allem auch *Have Nots*, z.B. in Indien. Sie gründen eigene Wissenschaftsjournale oder führen sie im Web fort, wie z.B. die in Bangalore erscheinende „*Current Science*“ der Indian Academy of Science.

Im Netz bilden sich ferner erste Vorformen von elektronisch unterstützten Auswahl- und Bewertungsprozessen. Mit *Google Scholar* kommt *Citation Ranking* in das Web der digitalen Literatur, zu dem eine große Zahl von *Open Access*-gestellten Pre- und Post-Publikationen gehört. Mit *Google Scholar* und *Google Book Search* erhält das Web heute außerdem ein kaum zu spammendes „Rückgrat“. Neue Formen der Metasuchmaschinen, die selbst Internet-Robots und –Archive in ihre Suche einbeziehen, bieten *Semantic Clustering*, eine neue Art der fachspezifischen Orientierung und Navigation ([www.clusty.com](http://www.clusty.com)). *Wikipedia*- und *Blogger*-Communities ([www.technorati.com](http://www.technorati.com)) beginnen, Fachliteratur zu beschreiben, zu kommentieren, zu bewerten und gemeinsam miteinander zu vernetzen ([www.delicious.com](http://www.delicious.com)). Das Netz beginnt in seiner zweiten Phase (*Web2.0*) – seit der Jahrhundertwende – sich selbst zu organisieren und überlässt dieses „kostenfreie Geschäft“ nicht allein den Suchmaschinen.

## 6. Die Bibliothek der Zukunft der Wissenschaften

Wissenschaft entsteht und lebt in den Köpfen derjenigen, die Wissenschaft betreiben. Was nicht aufgeschrieben oder anderweitig dokumentiert und erreichbar gespeichert wird, geht verloren. Verfügbare Literatur ist Lebensquell und Nährboden für Fortschritt. Bibliotheken, sehr umfassend und weit verstanden, sind daher ein bedeutendes Betriebsmittel wissenschaftlicher Arbeit. Wir skizzieren die Bibliothek der Zukunft in der Mathematik, unserem eigenen Fachgebiet.

Die Mathematik ist in gewisser (zugegeben, etwas überzogen technokratischer Sicht-) Weise die vollständige Sammlung der publizierten mathematischen Ideen, in der jede Veröffentlichung auf mindestens einer neuen Idee beruht, die mit den ihr zu Grunde liegenden Bausteinen (Definitionen, Sätze, Lemmata) und Publikationen durch Literaturverweise und Referenzen vernetzt ist. Der Traum der *Bibliothek der Mathematik* ist, die gesamte Literatur zur Verfügung zu haben. Dies ist ein sehr alter Traum. Er ist bezogen auf die ganze Wissenschaft der Traum der *wissenschaftlichen Universalbibliothek*.

Seit Beginn des Internet erscheint der Traum der Mathematik in einer neuen Form: Die ganze Mathematik an den Fingerspitzen, *alles, jederzeit, überall und kostenlos* zur Verfügung zu haben. Es das Versprechen der Netze, des Internet und auch des World Wide Web. Die Wirklichkeit der Netze sieht heute noch anders aus, trotz der in den letzten Jahren erreichten großen Fortschritte. Doch es ist gut, seine Träume nicht zu vergessen. Deshalb soll der Traum der Mathematik, der auch der Traum anderer Wissenschaften ist, in einer neuen Form im Web und auf eine realistischere Art formuliert werden. Im Web geht es um Vernetzung.

Die „Vernetzung“ von Literatur ist für die Mathematik von hoher Bedeutung, insbesondere für die Gültigkeit der mathematischen Arbeit. Eine Literaturreferenz auf ihre mathematischen Grundlagen ordnet eine mathematische Idee bzw. eine Veröffentlichung in den historischen Kontext der mathematischen Ideen ein. Die Gesamtheit der Verweise und Referenzen bildet gewissermaßen die historische Struktur der Mathematik. Jeder Verweis und jede Referenz deutet nicht nur auf einen Ursprung einer Idee und eine frühere Publikation, sondern sie erzeugt dadurch implizite Rückverweise von den Ursprüngen auf nachfolgende Ideen und Publikationen. Verweise und Referenzen sind in diesem Sinne elementare und zugleich historische Bedeutungsträger der Mathematik. Mathematiker „navigieren“ bisweilen entlang dieser durch Verweise und Referenzen gegebenen historischen Linien.

Stichworte aus Titeln und Abstracts, Schlüsselworte und andere Elemente der Metadaten eines Artikels erschließen die mathematische Literatur aus einer anderen Perspektive,

gewissermaßen „von oben“, der „Vogelperspektive“. Sie führen den Leser mit Hilfe eines Index, eines Katalogs oder einer Suchmaschine direkt zu einer entsprechenden Titelauswahl. Sie bringen, für sich genommen, noch keine eigenständige mathematische Wertung mit sich. Es sind in erster Linie technische Hilfs- und Ordnungsmittel. Von einem solchen Wort „zitiert“ zu werden, stellt noch keinen Rang und keine Bewertung dar.

Die Ausnahme sind Namen von Autoren, die als Bezeichnungen von mathematischen Sätzen (Riemannscher Abbildungssatz) oder Definitionen (Hilbert-Raum) fungieren. Das ist in den Naturwissenschaften nicht sehr viel anders, nur werden hier die Objekte der jeweiligen Wissenschaft mit Namen in Verbindung gebracht. In der Mathematik hebt der Name eines Autors einen (seinen) Satz aus der Menge der mathematischen Sätze hervor und zeichnet diesen mit einer besonderen Bedeutung aus. Solche Namen gehören oft nicht zu den Metadaten von Artikeln, jedenfalls nicht in diesem Sinne. Leser können Publikationen, die solche mathematischen Sätze enthalten oder zitieren, oft nur mit Hilfe einer Volltextsuchmaschine finden. Derselbe Umstand trifft auch bei durch Namen im allgemeinen Sinne („freie Texte“) gekennzeichneten Sätzen zu.

Ein moderner Traum der Mathematik im World Wide Web ist es, jegliche Indexinformation und alle Namen und Literaturreferenzen öffentlich zugänglich zu haben. Literaturreferenzen unterliegen ja nicht dem Copyright. Mathematiker möchten entlang dieser Verweise und Links „nahtlos“ navigieren können, d.h. ohne Bruch der Navigation durch irgendeine Sperre technischer oder rechtlicher Art. Zu den Referenzen zählen auch Namen und Bezeichnungen von Sätzen, wichtigen Formeln und Algorithmen, Verweise auf Daten, Modelle, Animationen und mehr. Zu dem Traum gehört auch, dass die Navigation und der Zugriff einfach sind und die gesamte Mathematikliteratur einbeziehen, unabhängig davon, ob diese Literatur *Open Access*-gestellt oder durch Lizenzen geschützt ist. Wesentlicher ist, ob sie relevant und aktuell ist oder nicht.

Die großen Internet-Suchmaschinen und besonders *Google* mit *Google Scholar* und *Google Book Search* kommen diesem Traum schon beträchtlich nahe. Dennoch sind wir noch weit von dem Traumziel einer einfach zu nutzenden, digitalen Bibliothek der Mathematik entfernt. Eine solche „Bibliothek“ liegt sicher auch im Interesse anderer Wissenschaften. Diese „Bibliothek“ wäre fest im globalen Netz integriert, denn die Jugend der Welt trifft sich heute im Netz. Und zu ihr gehört auch der wissenschaftliche Nachwuchs. Es wäre eine Bibliothek, in der viele Sprachen gesprochen werden und in der es viele Such- und Ordnungsmechanismen gibt. Dieses deutlich zu machen, ist der Sinn dieses Artikels. Der Traum aber sollte plausibel machen, was uns für die Zukunft noch fehlt.

Bibliotheken jedoch reichen alleine nicht aus. Wir sollten uns davor hüten zu glauben, dass alleine die Bereitstellung von exzellent vernetzter Information (wie oben skizziert) zu Fortschritt oder besserer Wissenschaft führt. Man muss nicht nur lernen mit der Vernetzung umzugehen, das erschlossene Wissen muss auch verstanden werden. Dies erfordert qualitativ hochwertige Ausbildung, ein Thema, das wir am Ende nur erwähnen wollen aber natürlich hier nicht ausarbeiten können.

## Literaturverweise

Bachrach, S., Berry, S. Blume, M. Foerster, Th. von, Fowler, A., Ginsparg, P., Heller, S., Kestner, N., Odlyzko, A., Okerson, A. Wigington, R. Moffat, A. (1998). *Intellectual Property: Who Should Own Scientific Papers?* Science, 1998, Vol. 281(5382), 1459-1460



- MPG (2003). *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*. Conference on Open Access, Berlin 2001, Max-Planck-Gesellschaft, 2003, <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>
- Brin, S. & Page, L. (1998). *The anatomy of a large-scale hypertextual Web search engine*. Proc. of the 7th int. Conf. On the World Wide Web, Brisbane, Australia, 1998, 107-117
- EU Commission (2006). *Study on the economic and technical evolution of the scientific publication markets in Europe*. European Commission Directorate-General for Research, 2006, 108 S., [ec.europa.eu/research/science-society/pdf/scientific-publication-study\\_en.pdf](http://ec.europa.eu/research/science-society/pdf/scientific-publication-study_en.pdf)
- Ewing, J. (2003). *Twenty Centuries of Mathematics: Digitizing and Disseminating the Past Mathematical Literature*. AMS Notices, 49(7), 2002, 771-777
- Friedmann, Th.L. (2006). *Die Welt ist flach. Eine kurze Geschichte des 21. Jahrhunderts*. Suhrkamp, 1. Auflage, Gebundene Ausgabe, 800 S.
- Grötschel, M. (2001). *Mein digitaler Traum*. Gegenworte der BBAW, 8 (2001), 10-16
- IMU & CEIC (2006). *WDML: World Digital Mathematics Library*. Home Page der WDML [www.ceic.math.ca/WDML](http://www.ceic.math.ca/WDML) zuletzt besucht am 27. September 2006
- Jackson, A. (2003). *The Digital Mathematics Library*. AMS Notices, 50 (8), 2003, 918-923
- Lyman, P. & Varian, H.R., et al. (2003). *How much Information? 2003*. UC Berkeley, School of Information Management, 2001- 2003, Living Report, PDF at: [http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/printable\\_report.pdf](http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/printable_report.pdf)
- OCLC (2005). *Perception of Libraries and Information Resources. A Report to the OCLC Membership*. OCLC Online Computer Library Center, Inc., Dublin, Ohio, USA. PDF at: [www.oclc.org/reports/pdfs/Percept\\_all.pdf](http://www.oclc.org/reports/pdfs/Percept_all.pdf)
- Odlyzko, A. (1994). *Tragic loss or good riddance? The impending demise of traditional scholarly journals*. Journal of Universal Computer Science, 1994, [www.jucs.org](http://www.jucs.org)
- Odlyzko, A. (1997). *The Economics of Electronic Scholarly Journals*. First Monday 2(8), 1997. [/www.firstmonday.dk](http://www.firstmonday.dk) zuletzt besucht am 20. Dezember 2006.
- Okerson, A. (1995). *The electronic journal: what, whence, and when? Source Computerization and controversy (2nd ed.): value conflicts and social choices*. Academic Press, Orlando, 1995, 580-593
- Van de Sompel, H, & Beit-Arie, O. (2001). *Open Linking in the Scholarly Information Environment Using the OpenURL Framework*. D-Lib Magazin, 7(3), 2001, [www.dlib.org](http://www.dlib.org)
- Vise, D.A. (2006). *The Google Story*. Random House, 2006, 329 P.
- Wissenschaftsrat (2001). *Empfehlungen des Wissenschaftsrates zur Digitalen Informationsversorgung durch Hochschulbibliotheken*. Greifswald, 13. Juli 2001, 55 S.



# Abbilder

Abbildung: Ebenen Wissenschaftlicher Information

