

## Open Access – nur ein erster Schritt?

Freier Zugang nicht nur zu wissenschaftlichen Veröffentlichungen, sondern auch zu den eigentlichen Forschungsdaten

Von Donat Agosti\*

Bei der seit einigen Jahren unter Wissenschaftlern intensiv geführten Debatte um Open Access geht es vor allem um den freien Zugang zu Veröffentlichungen und damit um eine schnellere und effizientere Verbreitung neuer Erkenntnisse. Doch Open Access hat noch weit mehr Potenzial: Ein freier Zugang auch zu den eigentlichen Forschungsdaten könnte die wissenschaftliche Arbeitsweise grundlegend verändern.

Die Open-Access-Bewegung, die seit einigen Jahren für einen unentgeltlichen und freien Zugang zu Publikationen aller Art kämpft, ist für die einen ein Fluch, für die anderen ein Segen: Verleger wie auch Musik- und Filmproduzenten fürchten, dass sie mit der Kontrolle über ihre Produkte auch ihre Einnahmequellen verlieren. Auf der anderen Seite zeigt beispielsweise die Internet-Suchmaschine Google, dass durch den freien Zugang zu umfangreichen Datenquellen auch neue Erwerbsmöglichkeiten entstehen können. Die Nutzer oder Konsumenten profitieren ebenfalls von diesem Wandel – schliesslich kann man mit Hilfe raffinierter Internet-Suchmaschinen viel mehr Informationen finden, als dies früher anhand einzelner Bücher und Bibliothekskataloge möglich war.

### Auf der eigenen Internetseite

In den Naturwissenschaften hat Open Access bereits eine lange Tradition. Die Online-Datenbank arXiv.org etwa, über die inzwischen mehr als 390 000 Arbeiten von der Physik bis zur Strukturbiochemie in Form von «Preprints» – der letzten Version des Manuskripts vor dem Druck – frei zugänglich sind, wurde schon im August 1991 lanciert. Seit einigen Jahren gibt es zudem eigentliche Open-Access-Zeitschriften, bei denen die Produktionskosten aus den Forschungsbudgets bezahlt werden und daher nicht durch den Verkauf von Abonnements oder einzelnen Artikeln wieder wettgemacht werden müssen. Die bekanntesten Vertreter dieser Verlags-Gattung sind die «Public Library of Science» mit 7 und die Online-Datenbank «BioMedCentral» mit 120 biomedizinischen Zeitschriften. Bei «BioMedCentral» können Universitäten sogar eine Art Publikationsabonnements für bestimmte Zeitschriften kaufen; für einen Pauschalbetrag dürfen dann alle Mitarbeiter dieser Universität dort publizieren.

Dank dem Engagement des ungarischen Kognitionswissenschaftlers Stevan Harnad und des amerikanischen Philosophen Peter Suber, zweier Mitbegründer der Open-Access-Bewegung, erlauben heute zudem mehr als 70 Prozent der wissenschaftlichen Zeitschriften, dass Autoren ihre Arbeiten auf der eigenen oder der entsprechenden Instituts-Homepage zugänglich machen; weitere 23 Prozent, darunter auch die renommierten Zeitschriften «Nature» und «Science», gestatten diese sogenannte Selbstarchivierung zumindest für die letzte Manuskriptfassung, eben die «Preprints».

Die meisten selbstarchivierten Arbeiten lassen sich mit einer Kombination der Suchmaschine Google Scholar und der normalen Google-Suchfunktion finden: Erst durchsucht man mit Google Scholar wissenschaftliche Datenbanken, etwa von Verlagen und Bibliotheken, nach einem oder mehreren Stichworten. Die Originalartikel, die hier als Treffer geliefert werden, sind zwar oft nicht frei zugänglich. Sie sind jedoch immer mit einer Verknüpfung zur normalen Google-Suchmaschine versehen, die dann das gesamte Internet nach selbstarchivierten Fassungen der entsprechenden Arbeit durchsucht. Dank der Effizienz dieser Suchfunktionen können die kostenpflichtigen Angebote der Verlage also immer besser umgangen werden – und dies schnell und unkompliziert.

### Aktiver und passiver Widerstand

Die Umsetzung von Open Access wird von den traditionellen Verlagshäusern denn auch bekämpft, ebenso von einigen wissenschaftlichen Gesellschaften, die ihre Einnahmen zu einem wesentlichen Teil aus dem Verkauf von Zeitschriften beziehen. Wie die beiden englischen Forscher Ian Rowland und David Nicholas vom University College in London zeigen konnten, kommt zudem passiver Widerstand von etablierten angelsächsischen Forschern. Im Gegensatz zu fast allen anderen von Rowland und Nicholas untersuchten Nutzergruppen ist für sie der Zugang zu wissenschaftlichen Informationen nämlich kein Thema – an ihren wohlhabenden Institutionen ist er einfach vorhanden.<sup>1</sup>

Hinderlich ist schliesslich auch die Arbeitsweise der Forscher selbst. So ist die Vermittlung von Forschungsergebnissen meist mit der Publikation einer Arbeit abgeschlossen. Netzwerke für



Seit der Lancierung der Internet-Datenbank arXiv.org ist die Ameisen-spezifische Bibliothek der Harvard University für Forscher auf der ganzen Welt einfach und frei zugänglich. Die Grösse der Punkte auf der Karte spiegelt die Anzahl der Dateien, die in einem bestimmten Umkreis heruntergeladen wurden.

● 1-9 ● 10-99 ● 100-999 ● >1000  
Quelle: arXiv.org/ClustrMaps

den Datenaustausch sind ausser bei grossen Forschungsprojekten, etwa in der Elementarteilchenphysik oder Astronomie, wo wegen der teuren Anlagen zwangsläufig zusammengearbeitet wird, nur in einem sehr begrenzten Rahmen vorhanden und basieren oft auf persönlichen Kontakten. Ein allgemein zugängliches elektronisches Archiv wäre zwar zweifellos ein zusätzliches Hilfsmittel für die Forschung. Ein solches Archiv aufzubauen und zu betreiben, wäre aber mit zusätzlicher, noch kaum honorierter Arbeit und möglicherweise auch mit urheberrechtlichen Problemen verbunden.

Doch selbst wenn die entsprechende Infrastruktur zur Verfügung gestellt wird, wird sie von den Forschern meist nur zögerlich benutzt. So zeigt eine Studie des Tasmaniers Arthur King, dass von dem Zeitpunkt an, zu dem eine Institution ein eigenes Publikationsarchiv beschlossen hat und die entsprechende Infrastruktur eingerichtet hat, etwa drei Jahre vergehen, bis in diesem Archiv zumindest 80 Prozent der von den Institutsmitarbeitern veröffentlichten Arbeiten erfasst sind und neue Veröffentlichungen innerhalb eines Monats nach ihrem Erscheinen hinzugefügt werden. Weltweit werden laut Stevan Harnad derzeit nur gerade 15 Prozent aller neuen Veröffentlichungen über Institutsarchive und weitere 5 Prozent via Open-Access-Zeitschriften zugänglich gemacht.

### Mit Steuergeldern finanziert

Die Open-Access-Bewegung ist also immer noch stark vom Engagement einzelner Personen und Institutionen geprägt. Dies ist umso erstaunlicher, als die Abonnement-Preise für wissenschaftliche Zeitschriften seit dem Aufkommen der digitalen Versionen um 1990 herum stark angestiegen sind und viele Bibliotheken daher immer weniger Zeitschriften abonnieren können. Dabei, so ein Argument der Open-Access-Bewegung, würde ein möglichst umfassender, schneller und freier Zugang zu neuen Forschungsergebnissen den wissenschaftlichen Fortschritt beschleunigen – was etwa im Bereich der Medizin sogar Leben retten könne.

Die Forderung nach Open Access wird zudem damit begründet, dass ein Grossteil der Forschung mit Steuergeldern finanziert wird. Und nicht nur das – neben dem Inhalt der Publikationen wird in gewisser Weise sogar der «Peer Review»-Prozess, die durch Forscherkollegen vorgenommene Qualitätssicherung vor der Veröffentlichung einer Arbeit, von der öffentlichen Hand bestritten – die «Peers» werden von den Verlagen nämlich nicht für diese Dienstleistung bezahlt, sondern erledigen sie im Rahmen ihrer meist von öffentlichen Geldern finanzierten Forschung.

Die Rechte an den Veröffentlichungen werden von den Wissenschaftlern jedoch – besonders seit dem Aufkommen von digitalen Ausgaben – diskussionslos an die Verlage abgetreten. Diese steuern einen Redaktionsstab, den Druck, den oft weltweiten Vertrieb der gedruckten Ausgaben sowie den elektronischen Vertrieb bei, zudem Such- und andere Servicefunktionen auf ihren Internetseiten und in manchen Fällen auch noch einen wichtigen Markennamen wie «Nature» oder «Science». Dafür erhalten sie sämtliche Einnahmen aus dem Verkauf, womit die öffentliche Hand zweimal zur Kasse gebeten wird: zuerst für die Forschung und dann für den Zugang zu den von ihr finanzierten Arbeiten.

Zusammen mit der Einsicht, dass die Verbreitung von Forschungsdaten nicht optimal ist, solange nicht die Möglichkeiten des Internets ausgeschöpft werden, führte dies im Jahr 2003 zur «Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen»; bis heute wurde diese

Erklärung von 181 Organisationen unterzeichnet, darunter auch der Schweizerische Nationalfonds und die Universität Zürich.<sup>2</sup> Sie verlangt von den Unterzeichnenden unter anderem, dass sie die von ihnen finanzierten wissenschaftlichen Arbeiten frei zugänglich machen und jedermann das Recht gewähren, diese Arbeiten und die darin enthaltenen Resultate unter Angabe der Quelle weiterzuverwenden. Die «Berliner Erklärung» entspricht damit einer Lizenz der «Creative Commons», die in der Kunst- und Musikbranche als Alternative zum herkömmlichen Copyright bei über 130 Millionen Dateien und Songs angewandt und nun durch ihre Tochterorganisation, die «Science Commons», den Bedürfnissen der Forschung angepasst wird.<sup>3</sup>

Allein im Oktober haben elf weitere wissenschaftliche Institutionen, darunter etwa der britische Medical Research Council und die Universität Zürich, Open Access eingeführt. Die Universität Zürich hat dafür eigens das «Zürich Open Repository and Archive» eingerichtet.<sup>4</sup> Die Mitarbeiter der Universität sind verpflichtet, bereits erschienene Arbeiten dort zu archivieren, sofern keine rechtlichen Gründe – etwa Copyright-Ansprüche der Verlage – dagegen sprechen. Zudem sind sie gehalten, in Zukunft möglichst nur noch in Zeitschriften zu publizieren, die eine solche Selbstarchivierung gestatten. An einer Konferenz, die Anfang November am Cern in Genf stattgefunden hat, haben zudem die Teilchenphysiker beschlossen, dass Open Access als Standard für Veröffentlichungen des gesamten Forschungsgebietes verankert werden soll; wie, muss allerdings erst noch geklärt werden.

### Überdurchschnittlich häufig zitiert

Dass Open Access auch für die Autoren einer Veröffentlichung von Vorteil ist, konnte der Kanadier Gunther Eysenbach im Sommer mit einer Studie belegen.<sup>5</sup> Eysenbach hatte rund 1500 Veröffentlichungen verglichen, die in der Zeitschrift «Proceedings of the National Academy of

Sciences» erschienen waren, einer sogenannten Hybrid-Zeitschrift, in der seit nunmehr fünf Jahren kostenlose Artikel und solche, die nur für Abonnenten und zahlende Kunden zugänglich sind, nebeneinander publiziert werden. (Nach einer Frist von sechs Monaten sind jedoch alle Arbeiten frei zugänglich.) Dabei zeigte sich, dass die Open-Access-Artikel in den ersten vier bis zehn Monaten nach ihrem Erscheinen bis zu zweimal häufiger zitiert wurden als die nur begrenzt zugänglichen.

### Weltweit genutztes Angebot

Der enorme Einfluss von Open Access auf die Forschung wird aber auch anhand fachspezifischer Bibliotheken deutlich. So gab es früher nur an der Harvard University in Boston eine komplette Sammlung aller 2800 wissenschaftlichen Arbeiten, die die Beschreibungen der 12 000 bekannten Ameisenarten enthalten; diese Arbeiten waren in mehr als 500 Büchern und 850 Zeitschriften erschienen. Seit dem Jahr 2004 bietet nun auch die vom American Museum of Natural History und von der Ohio State University betriebene Internet-Datenbank arXiv.org mit Ausnahme einiger kopiergeschützter Artikel praktisch alle Arbeiten über Ameisenarten online an. Dieses Angebot wird weltweit genutzt: Allein innerhalb der vergangenen drei Monate wurde die Seite rund 23 000 Mal aus insgesamt 126 Ländern aufgerufen.

Eine ähnliche Erfahrung macht man derzeit am American Museum of Natural History in New York, wo die dort archivierten 6000 wissenschaftlichen Arbeiten in den ersten neun Monaten nach der Einführung von Open Access rund 500 000 Mal heruntergeladen wurden. Solche Angebote helfen gerade kleinen Forschungsgruppen mit wenigen Mitarbeitern oder solchen in Entwicklungsländern, die bis anhin ihre eigene fachspezifische Bibliothek in mühseliger Kleinarbeit aufbauen mussten oder dies wegen fehlender finanzieller Mittel gar nicht konnten.

### Freier Zugang auch zu den Daten

Da das Internet eine breite Vernetzung und Analyse von Daten zulässt, stellt sich allerdings die Frage, ob der Fokus der Open-Access-Debatte auf den Zugang zu klassischen Publikationen nicht zu eng gefasst ist. Einer der Vorzüge des Internets und digitaler Veröffentlichung besteht ja gerade darin, dass einzelne Arbeiten miteinander, mit Originaldaten und mit weiteren Medien wie Filmsequenzen verlinkt werden können. Eine solche Vernetzung würde beispielsweise auch die Überprüfung von Forschungsergebnissen durch Kollegen erleichtern.

Allerdings müssten dafür in vielen Fachgebieten erst noch die nötigen technischen Voraussetzungen geschaffen werden – rund 80 Prozent der Forschung finden laut einer Schätzung nämlich in kleinen Gruppen statt, die ihre Daten nur auf persönlichen Computern gespeichert haben, die maximal über ein internes Netzwerk zugänglich sind, nicht jedoch von aussen. In einigen Disziplinen wird dieser Datenaustausch jedoch bereits mit Erfolg praktiziert, etwa in den Sozialwissenschaften, wo es Archive für Umfragen gibt, in der Genomforschung, wo Daten auf zentralen Gen- und Protein-Datenbanken gespeichert werden, oder in der Ökologie, wo man über die «Global Biodiversity Information Facility» die Datenbanken verschiedener Museen abrufen kann.

### Der Text als dynamisches Produkt

Die von der Zeitschrift «Nature» lancierte «Open Text Mining Initiative» hat sogar ein noch höher gestecktes Ziel: dass irgendwann alle jemals ver-

öffentlichten wissenschaftlichen Arbeiten mit speziellen Suchmaschinen auch auf komplexe Zusammenhänge hin untersucht werden können.<sup>6</sup> Dafür müssten die Artikel derart aufgearbeitet werden, dass sie von Computern gefunden und automatisch eingesehen werden können. Das Potenzial dieser Initiative wird deutlich, wenn man bei Google Scholar das Stichwort «headline», also «Kopfschmerz», eingibt. Wären nämlich alle 364 000 Artikel, die die Suchmaschine zu diesem Thema findet, entsprechend markiert, zudem mit Verknüpfungen zu den Originaldaten versehen und kostenlos zugänglich, so könnte man tatsächlich auf das gesamte publizierte Wissen zu diesem Thema zurückgreifen.

Die Ideen von Mackenzie Smith vom Massachusetts Institute of Technology in Cambridge, USA, gehen sogar noch einen Schritt weiter: Die von ihr entwickelten «Enhanced Publications» («erweiterte Veröffentlichungen») sind nicht mehr statische, sondern vielmehr dynamische Produkte, die bei neuen Erkenntnissen immer weiterentwickelt werden können, etwa indem Kommentare angefügt werden. Damit sind allerdings nicht nur technische, sondern auch soziologische Herausforderungen verbunden – etwa, wie bei einem solchen Verfahren Beiträge einem einzelnen Autor zugeschrieben und entsprechend honoriert werden können.

<sup>1</sup> www.publishing.ucl.ac.uk/papers/2006Rowlands\_Nicholas.pdf;

<sup>2</sup> oa.mpg.de/openaccess-berlin/berlindeclaration.html;

<sup>3</sup> sciencecommons.org/; <sup>4</sup> www.zora.uni-zh.ch/zora/; <sup>5</sup> PLoS Biology 4, e157 (2006); <sup>6</sup> Nature 440, 1090 (2006).

\* Der Autor ist Biologe und als Mitarbeiter des American Museum of Natural History an Aufbau und Pflege der Online-Datenbank arXiv.org beteiligt. Darüber hinaus ist er freier wissenschaftlicher Mitarbeiter am Naturhistorischen Museum der Burgergemeinde Bern und beschäftigt sich mit allgemeinen Fragen zum Zugang zu wissenschaftlicher Information.