

Eprints in der wissenschaftlichen Kommunikation

Frank Havemann*

26. Oktober 2004

1 Die Zeitschriftenkrise und der Übergang zu Open Access

Wissenschaftliche Zeitschriften entstanden, als vor über 300 Jahren die Kommunikation europäischer Gelehrter über Briefe nicht mehr effektiv funktionierte, weil zuviele an ihr teilnahmen. Die Zahl der Zeitschriften wuchs seit damals exponentiell. Im 19. Jahrhundert war die Zeitschriftenliteratur so umfangreich geworden, dass Referate-Zeitschriften (die Vorläufer der bibliographischen Datenbanken) notwendig wurden, damit Forscher den Überblick über die aktuelle Entwicklung ihrer Fachgebiete behalten konnten (de Sol-la Price 1963). Heute stößt das System wissenschaftlicher Kommunikation erneut an Grenzen: die Etats der Bibliotheken wachsen nicht mehr in dem Maße, das notwendig wäre, um die wachsende Zahl immer umfangreicher und teurer werdender Journale erwerben zu können. Die steigenden Preise sind jedoch nur z.T. durch mehr Artikel pro Jahrgang gerechtfertigt. Bibliothekare weisen vor allem den großen Wissenschaftsverlagen eine Mitschuld an der Zeitschriftenkrise zu: "Wesentliche Ursachen dieser Entwicklung liegen im Konzentrationsprozess des Verlagswesens und in der Preispolitik von Marktführern wie Elsevier, Wiley, Kluwer oder Springer."¹

*Institut für Bibliothekswissenschaft der Humboldt-Universität zu Berlin. Der Text war Vorlage für einen Vortrag am 1. Juni 2004 im Rahmen der Ringvorlesung "Die Zukunft der Bibliotheken".

¹Webseite der Bibliothek der Universität Konstanz, am 28. April 2004.
<http://www.ub.uni-konstanz.de:8099/openaccess/index.htm>

Die Verlage wiederum können darauf verweisen, dass wegen der steigenden Zahl eingesandter Artikel die Organisation ihrer Begutachtung und Auswahl durch kompetente Fachkollegen teurer wird. Die Zeitschriftenkrise könnte überwunden werden, so die Verfechter von *Open Access* (OA), wenn Aufsätze in allen wissenschaftlichen Zeitschriften unentgeltlich über das Netz zugänglich wären. Die Qualitätskontrolle durch *peer review* muss dann von den Autoren finanziert werden.² Es wird erwartet, dass dadurch das System für die Wissenschaft insgesamt weniger kostspielig wird. Zur Zeit haben viele Verlage mit ihren spezialisierten Journalen gegenüber den Bibliotheken Monopolstellungen inne. Autoren konkurrieren um Platz in guten Zeitschriften. OA-Journale müssen dagegen miteinander um zahlende Autoren konkurrieren, was die Gewinne reduzieren sollte, welche nicht nur Verlage sondern auch wissenschaftliche Gesellschaften mit ihren Journalen machen.³ Dann ist denkbar, dass *peer-review*-Kosten abnehmen, weil Autoren nicht mehr nur auf die angesehensten und teuersten Journale orientiert sind. Vielschreiberei und Mehrfachpublikationen würden abnehmen. Es gibt auch Finanzierungsmodelle, die eine kostengünstigere Begutachtung dadurch erreichen wollen, dass diese online von den Lesern vorgenommen wird (Kölbel 2003; Mizzaro 2003).

Eine Reihe von Journalen – mehr als tausend sind z.Z. im *Directory of Open Access Journals*⁴ angezeigt – bieten bereits online freien Zugang zu ihren Artikeln. Das sind durchaus angesehene Zeitschriften, die auch im *Web of Science* erfasst werden (z.Z. fast 200). Von diesen stoßen in den Fachbereichsranglisten nach *impact*-Faktoren einige bis zur Spitzengruppe vor.⁵ Daneben entwickelt sich die Praxis, die noch nicht begutachteten Dokumente über das Netz frei verfügbar zu machen, was viele Zeitschriften – ausdrücklich oder auf Nachfrage – zulassen. Die Dokumente können auf Eprint-Servern (institutionellen und überinstitutionellen) oder auf der eigenen Homepage abgelegt werden.

²Dass Autoren für die Publikation ihrer Artikel bezahlen, ist bei angesehenen Journalen, wie z.B. *Physical Review*, schon lange üblich. Bei der frei zugänglichen Zeitschrift *PLoS Biology*, gegründet im letzten Herbst, müssen die Autoren für jeden publizierten Artikel 1500\$ bezahlen.

³Letztere verweisen darauf, dass ihre Gewinne der Wissenschaft zugute kommen (Giles 2004).

⁴S. <http://www.doaj.org/home>; darunter 27 Journale zur *Library and Information Science*.

⁵The Impact of Open Access Journals. A Citation Study from Thomson ISI. 2004-05-05. <http://www.isinet.com/media/presentrep/acropdf/impact-oa-journals.pdf>

Bisher haben die großen Verlage den Übergang zu *Open Access* bremsen können, er könnte eine Eigendynamik entwickeln, wenn immer mehr Autoren ihre Zurückhaltung aufgeben, weil sie gewahr werden, dass frei zugängliche Artikel stärker zitiert werden als zu bezahlende (Lawrence 2001). Die *Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen* vom letzten Oktober signalisiert die Unterstützung der großen Wissenschaftsorganisationen in Deutschland für den Übergang zu *Open Access*: “Zur Verwirklichung der Vision einer umfassenden und zugänglichen Wissensrepräsentation muss das künftige Web nachhaltig, interaktiv und transparent sein. Inhalte und Software müssen dazu frei verfügbar und kompatibel sein.”⁶

Bibliotheken würden beim Übergang zu OA-Journalen eine Aufgabe verlieren: Zeitschriftenabonnements zu organisieren. Die dafür bisher verwendeten Gelder müssten – vorausgesetzt das oben beschriebene Finanzierungsmodell setzt sich durch – den Autoren der Institution zukommen. Allerdings verbleibt ihnen die Aufgabe der Archivierung zur Sicherung des Langzeitzugriffs, die sie z.Z. für elektronische Dokumente noch ungenügend erfüllen, auch wegen der Politik der Verlage.

Seit März 2004 kann man auf den Netzseiten von *Nature* eine Debatte zu *Open Access* verfolgen.⁷ In vielen Beiträgen wird nicht bestritten, dass OA-Journale eine Zukunft haben; die Vertreter von Verlagen und Wissenschaftsgesellschaften, die Zeitschriften herausgeben, sehen jedoch eher voraus, dass wissenschaftliche Aufsätze sowohl offen als auch in zu bezahlenden Journalen zugänglich sein werden. Stevan Harnad, Kognitionsforscher und OA-Protagonist, und acht Koautoren fordern in ihrem Diskussionsbeitrag,⁸ dass Forschungsinstitute es ihren Mitarbeitern zur Pflicht machen sollten, ihre Arbeiten auf einem institutionellen Server offen zugänglich zu machen. Das wäre ein Schritt auf dem sogenannten *green road* zu *Open Access* (*golden* sind frei zugängliche Journale). Sie geben an, dass nur noch 17% von 10.000 untersuchten *peer-review*-Journalen eine solche Veröffentlichung im Netz nicht gestatten. Mehr als 4/5 der Zeitschriften sind also schon *grün*, mehr als 1000 *golden*. Viele gestatten einen offenen Zugang erst nach dem Erscheinen des Zeitschriftenheftes (*postprint*), andere auch davor (*preprint*). Die Autoren des Beitrages plädieren für den grünen Weg, weil er erprobt ist. Institutionen wie Wissenschaftler würden ihn schon deshalb beschreiten, weil ihnen

⁶http://www.zim.mpg.de/openaccess-berlin/BerlinDeclaration_dt.pdf

⁷<http://www.nature.com/nature/focus/accessdebate/>

⁸Eine vollständigere Version ihres Beitrages mit Diagrammen findet man unter: <http://www.ecs.soton.ac.uk/harnad/Temp/impact.html>

sonst im Wettbewerb um das rare Gut Aufmerksamkeit ein Nachteil entstünde. Eprints sind ideal für die schnelle Kommunikation wissenschaftlicher Ergebnisse innerhalb wissenschaftlicher Gemeinschaften. Benötigten Forscher jedoch Wissen aus Fachgebieten, in denen sie nicht zuhause sind, werden sie eher auf sicher authentifizierte Dokumente in begutachteten Journalen zurückgreifen.

2 Preprints und Eprints – wie kam es dazu?

Schon seit langem haben Autoren ihren Fachkollegen Reprints von Zeitschriftenaufsätzen zugesandt. Mit dem Aufkommen der Kopiertechnik wurde es möglich, in größerem Maße auch vor der Erscheinen des Zeitschriftenheftes Kopien des Manuskripts zu verschicken. Dadurch waren die mitgeteilten Resultate einem ausgewählten Kreis schon vor der oft erst Monate später erfolgenden Publikation zugänglich. Diese Preprints wurden von manchen Institutionen als Reihen Grauer Literatur aufgelegt. In der Elementarteilchenphysik begann dann eine neue Vorgehensweise üblich zu werden. Die Autoren oder ihre Institutionen schickten die Preprints an einschlägige Forschungsinstitute, in deren Bibliotheken sie ausgelegt wurden. Im Europäischen Forschungszentrum CERN in Genf wurden dann alle eingegangenen Preprints wöchentlich in Listen thematisch geordnet erfasst. Diese Listen konnten nicht nur die vielen Wissenschaftler am CERN einsehen, sondern sie wurden weltweit abonniert. Fand man einen interessanten Preprint in einer Liste, brauchte man nur noch die Adresse der Institution des korrespondierenden Autors in einer anderen Liste nachzuschlagen, um eine Preprint-Anforderung dorthin zu schicken und ein Exemplar zu erhalten. In der Internet-Ära tat Paul Ginsparg 1991 den entscheidenden Schritt, um dieses System online zu stellen. Er richtete den Eprint-Server *ArXiv* ein, auf den Autoren ihre Aufsätze einstellen können.⁹ Mittlerweile enthält das *ArXiv* nicht nur Hochenergiephysik-Eprints, sondern Texte aus allen Gebieten der Physik, der Biophysik, der Mathematik und der Informatik. Die monatliche Rate von eingestellten Eprints stieg seit 1991 linear und liegt jetzt zwischen Drei- und Viertausend, die Zahl der Zugriffe hat die halbe Million pro Monat überschritten.¹⁰ Inzwischen gibt

⁹deutscher Spiegel unter <http://de.arxiv.org/>

¹⁰http://arxiv.org/todays_stats

es auch für andere Fachgebiete überinstitutionelle Eprint-Server.¹¹ Die von Harnad und anderen favourisierten institutionellen Eprint-Server entsprechen den Preprint-Reihen der Papier-Zeit.

In der *Open Archives Initiative* (OAI) werden seit einigen Jahren Standards, Schnittflächen und Software zur Archivierung und zum Retrieval für online frei verfügbare Eprint-Archive entwickelt (Hitchcock et al. 2002).¹²

3 Eprint-Retrieval

Die institutionellen Eprint-Server, aber auch das *ArXiv* und *CiteSeer*, sind online frei mit *OAIster* recherchierbar (nach Titel, Autor, Erzeuger, Gegenstand und Ressourcen-Typ). Dabei handelt es sich um ein OAI-Projekt an der University of Michigan.¹³ Es zielt auf alle frei zugänglichen akademischen digitalen Ressourcen, wie Artikel, elektronische Bücher aber auch Filme, Bilder und Tonaufnahmen. Am 11. Mai 2004 waren ca. 3,2 Millionen *records* von 282 Institutionen dort erfasst, darunter auch der Dokumenten- und Publikationsserver der Humboldt-Universität zu Berlin (1294 *records*), der jedoch noch keine späteren Journal-Artikel als Eprints aufnimmt.¹⁴ Als Nachteil von OAIster gegenüber Google muss die fehlende Volltextrecherche genannt werden.

Eugene Garfield hat mit dem Science Citation Index in den 60er Jahren die Recherche nach Zitierungen möglich gemacht, sodass man nicht mehr nur zeitlich rückwärts über die Referenzen zu neuen Artikeln vorstoßen kann, sondern auch vorwärts die Wirkungsgeschichte einer Publikation verfolgen. Heute sind die Zitationsdatenbanken über das Netz im *Web of Science* (WoS) für Abonnenten zugänglich. Für Eprints wurde diese Methode wiederum zuerst in der Elementarteilchenphysik angewendet, und zwar im *Stanford Public Information REtrieval System* (SPIRES) mit der frei zugänglichen *SPIRES HEP Literature Database*, die auch Journalartikel enthält, die nicht als Eprint vorliegen.¹⁵ SPIRES HEP kann für die Nutzer auch deshalb kostenlos angeboten werden, weil Autoren in die Pflege mit einbezogen werden.¹⁶

¹¹z.B. in der Informatik *CiteSeer*, <http://citeseer.ist.psu.edu/cis>, in der Ökonomie *EconWPA*, <http://econwpa.repec.org/>.

¹²<http://www.openarchives.org>

¹³<http://oaister.umdl.umich.edu/o/oaister/>

¹⁴<http://edoc.hu-berlin.de/>

¹⁵deutscher Spiegel bei DESY in Hamburg: <http://www-library.desy.de/spires/hep/>

¹⁶<http://www-library.desy.de/spires/hep/pre-submit.html>

Die Zitationsbeziehungen zwischen Eprints (im PDF-Format) werden beim ebenfalls unentgeltlichen CiteSeer durch ein automatisches Verfahren hergestellt, das allerdings noch ziemlich fehlerbehaftet ist. Ein Vorteil ist hier, dass der Kontext jeder Zitierung sichtbar wird. Kürzlich wurde ein Zusammengehen der WoS-Betreiber (ISI Thomson) mit den Entwicklern von CiteSeer (NEC Laboratories America) bei der Entwicklung eines *Web Citation Index* bekanntgegeben.¹⁷

Auf die Zitierungsdaten im *ArXiv* und in allen institutionellen Eprint-Archiven zielt das *Open Citation Project* (OpCit). Es wurde ein kostenloses Angebot – genannt *Citebase* – entwickelt, das ebenfalls weitgehend automatisch die Zitierungsdaten von Eprints extrahiert (Hitchcock et al. 2002).¹⁸ Citebase gibt auch an, wann und wie oft der Eprint heruntergeladen wurde.

4 In welchem Maße können Eprints die Kommunikation der Forscher beschleunigen?

In einer kleinen Studie habe ich mit Studierenden im letzten Wintersemester an ausgewählten Artikeln untersucht, in welchem Maße Eprints die wissenschaftliche Kommunikation im Vergleich zu Zeitschriftenaufsätzen beschleunigen.¹⁹ Beschleunigung wird heute – weil wir unter ihr auch leiden – nicht nur als etwas gesehen, was unbedingt erstrebenswert ist. Wie Wein sollten auch Texte manchmal ablagern, von den Autoren überarbeitet werden, bevor sie publik werden. Aber Wissenschaft funktioniert als Wettbewerb um Priorität. Je schneller neue Ergebnisse bekannt werden, umso eher können andere sie verwenden, um so mehr wird Doppelarbeit vermieden.

Wir wählten für die Studie Artikel aus der Zeitschrift *Physical Review D* (PRD), die vor allem Beiträge zur Hochenergiephysik aber auch zur Astrophysik publiziert, und von deren Aufsätze wir hoffen konnten, sie überwiegend auch als Eprints im *ArXiv* zu finden. Getreu dem Prinzip, nur gleiches mit gleichem zu vergleichen, engten wir unsere Untersuchungsmenge noch mehr ein, nämlich auf Artikel, die theoretische Fragen diskutieren und auf lediglich drei Bände der Zeitschrift (58 - 60), die im zweiten Halbjahr 1998

¹⁷<http://www.isinet.com/press/2004/8217120>

¹⁸<http://opcit.eprints.org/> und <http://citebase.eprints.org/cgi-bin/search>

¹⁹Für ihre tatkräftige und kritische Mitarbeit bei der Erarbeitung der Literatur zum Thema und beim Projekt danke ich Philipp Mayr, Christina Mitrenga, Marita Pankonin, Marion Schmidt, Maria Till und Julia Vogel.

und im Jahr 1999 erschienen sind. Unsere Ausgangsmenge waren dann 211 mit SPIRES HEP (s.o.) unter dem *topic theory* gefundene PRD-Artikel, von denen 8 (d.s. 4%) nicht als Eprint ins *ArXiv* gestellt worden waren. Wir ermittelten dann die Zitierungen der Eprints in anderen Eprints (mittels SPIRES HEP) und die Zitierungen der Zeitschriftenaufsätze in anderen Zeitschriftenaufsätzen (mittels WoS, s.o.). Die drei Bände zeigten statistisch in allen Fragen ein ganz ähnliches Verhalten; ein Trend war nicht ablesbar. Die 203 Artikel bzw. Eprints (66 + 69 + 68) konnten deshalb als Einheit analysiert werden.²⁰ Die untersuchten Artikel sind keinesfalls repräsentativ, wir wollen auch nur zeigen, welche *Möglichkeiten* der Beschleunigung im Eprint-System liegen.

Die erste Frage, die wir stellten, war die nach der Zeit (in Monaten), die jeweils zwischen dem Einstellen des Eprints in das *ArXiv* und dem Erscheinen des Hefts von PRD vergangen ist (s. Abb. 1, *publication delay*). Die Artikel sind recht symmetrisch und konzentriert um Median, Modus und arithmetisches Mittel verteilt, die alle drei bei acht Monaten liegen; die Standardabweichung davon beträgt drei Monate. Die längste Wartezeit betrug 19 Monate, aber es gibt auch einen Eprint, der erst in dem Monat ins *ArXiv* gestellt wurden, als das PRD-Heft erschien. PRD-Artikel sind allerdings (für Abonnenten) schon vor dem Erscheinen des Heftes über das Netz zugänglich, manche nur wenige Tage vorher, manche zwei Monate. Der Mittelwert lag bei einer Stichprobe bei einem Monat. In der untersuchten Artikelmenge ist also der Eprint typischerweise ca. sieben Monate vor dem Journalartikel zugänglich gewesen.

Wird dieser zeitliche Vorsprung auch genutzt? Das war eine weitere Frage, die wir uns stellten. Dazu bestimmten wir den Monat der ersten Zitierung des Eprints durch einen von anderen Autoren verfassten Eprint und berechneten daraus die Zeitspannen, die Eprints unserer Untersuchungsmenge auf die erste in SPIRES HEP registrierte Fremdzitierung warten mussten (*citation delay SPIRES*). 27 (13%, eine geringe Quote) sind nicht in Eprints von anderen Autoren zitiert worden, die Verteilung der restlichen 184 nach der Wartezeit ist der Abb. 2 zu entnehmen. Abgesehen von dem fehlerhaften SPIRES-Eintrag mit einer Zitierung fünf Monate vor der Einstellung als Eprint gibt es noch zwei Eprints, die zuerst vor dem Erscheinen zitiert worden sind. So etwas kommt bei Journalartikeln auch vor. Meist ist den

²⁰27 der 203 Eprints, d.s. 13% waren nicht in einem Hochenergie-Archiv, sondern in GR-QC, bzw. 3 in ASTRO-PH.

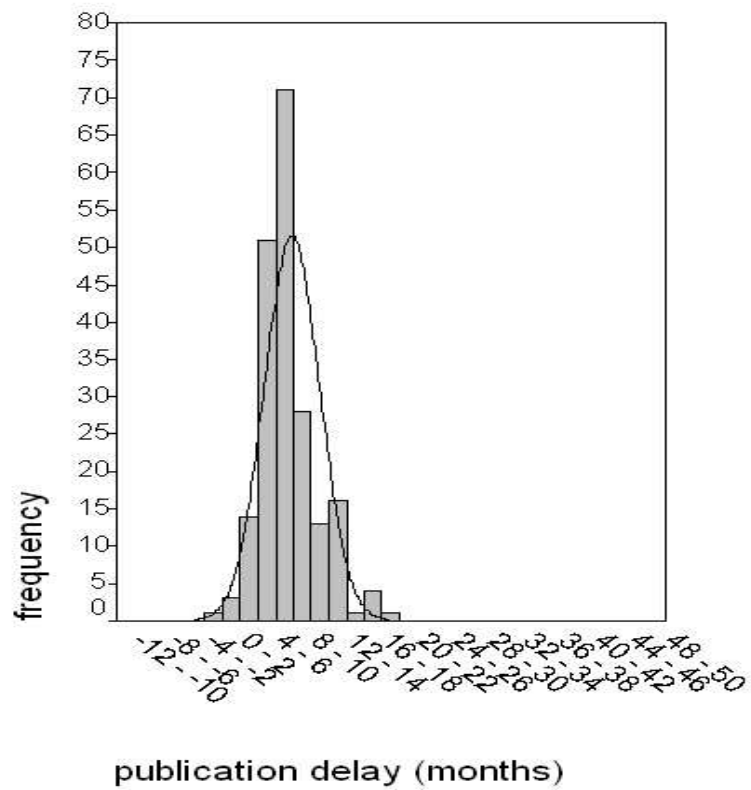


Abbildung 1: Verteilung von 203 PRD-Artikeln nach der Zeit in Monaten zwischen dem Einstellen des Eprints in das *ArXiv* und dem Erscheinen des Hefts von PRD (*publication delay*). Die Glockenkurve ist die der am besten angepassten Normalverteilung.

zitierenden Autoren der Text schon vorher privat zugekommen. Beim *ArXiv* kann die Ursache auch darin liegen, das eine spätere Version des Eprints neue Referenzen aufgenommen hat. Bei unserer Analyse entgeht uns so etwas, weil wir nicht den Datumsstempel der Version beachtet haben, sondern nur die Eprint-Nummer zur Bestimmung des Erscheinungsmonats benutzt haben.²¹ Die Abb. 2 zeigt, dass Eprints schnell zitiert werden. Nach acht Monaten sind

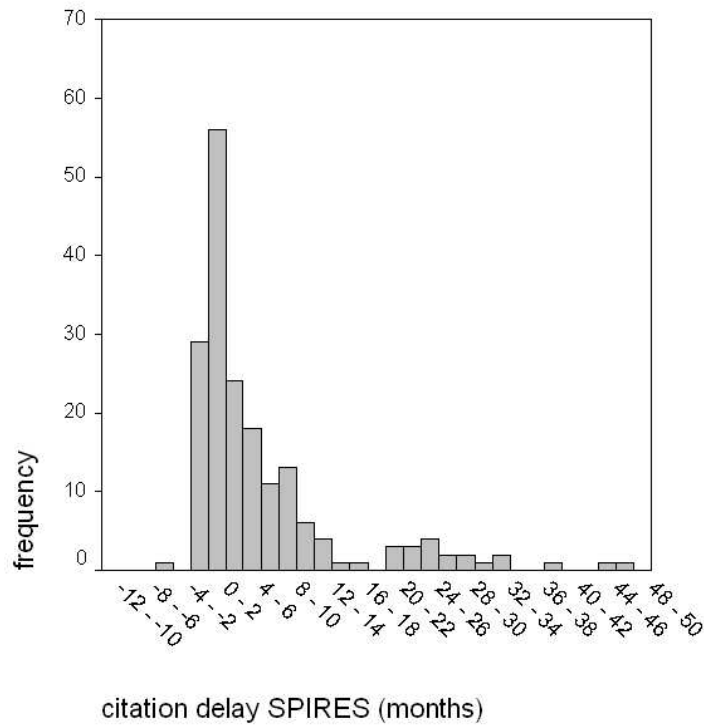


Abbildung 2: Verteilung der 184 überhaupt fremdzitierten Eprints nach der Zeit (in Monaten) bis zur ersten Fremdzitierung in anderem Preprint, ermittelt in SPIRES HEP (*citation delay SPIRES*) (−5 Monate rührt von Fehler in SPIRES, müsste +3 heißen)

²¹Sowohl im *ArXiv* wie in der *online*-Version des Journals kann der Zeitpunkt des Erscheinens eines Textes auf den Tag genau bestimmt werden. Wir haben uns hier mit einer geringeren Genauigkeit begnügt, um den Rechercheaufwand nicht zu sehr zu erhöhen. Mit automatisierter Recherche und Datenverarbeitung ist eine genauere Zeitmessung durchaus machbar. Beides lag außerhalb unserer zeitlichen Möglichkeiten.

bereits drei Viertel der überhaupt fremdzitierten 184 Eprints von anderen in ihren Arbeiten verwendet oder wenigstens diskutiert worden. Die meisten (fast drei Viertel) dieser ersten Fremdzitierungen erfolgen vor der Publikation des entsprechenden Heftes von PRD (oft Monate vorher), wie die Abb. 3 zeigt, der die Verteilung nach dieser Zeitdifferenz zugrunde liegt (*publication citation delay*). Hochzitierte Eprints werden fast alle schnell zitiert. Das

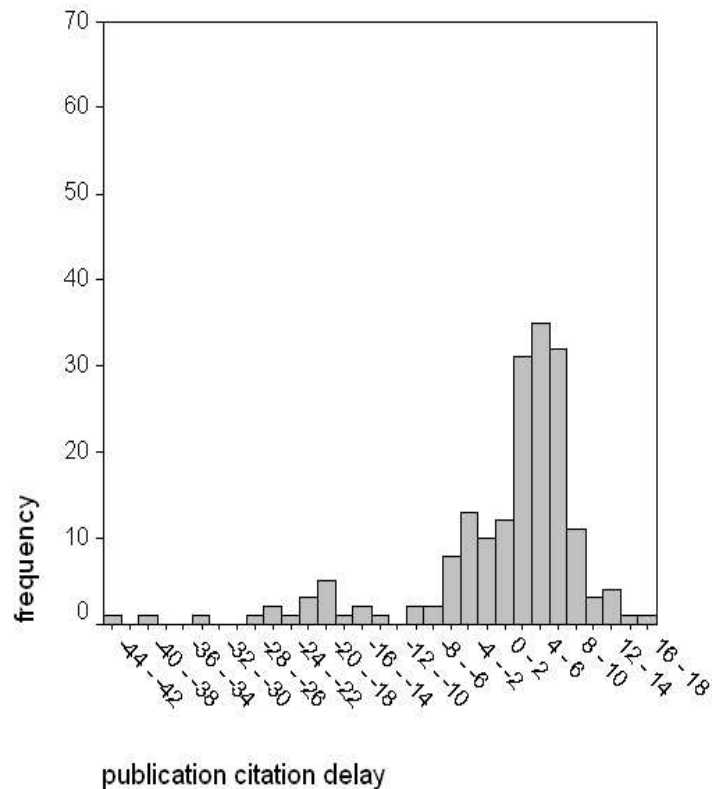


Abbildung 3: Verteilung der 184 Eprints nach dem Zeitverzug, die das Heft von *Physical Review D* gegenüber der ersten Fremdzitierung in anderem Preprint hat (*publication citation delay*)

Streudiagramm (Abb. 4) zeigt, dass nur vier der 56 (d.s. 7%) mindestens 20mal in Eprints zitierten Eprints (inklusive Selbstzitationen, Recherche im Februar 2004) länger als fünf Monate auf ihre erste Fremdzitierung warten mussten. Eine frühe Zitierung erlaubt jedoch andererseits nicht, hohe Be-

achtung der Arbeit vorauszusagen.²² Bis auf drei werden auch alle der min-

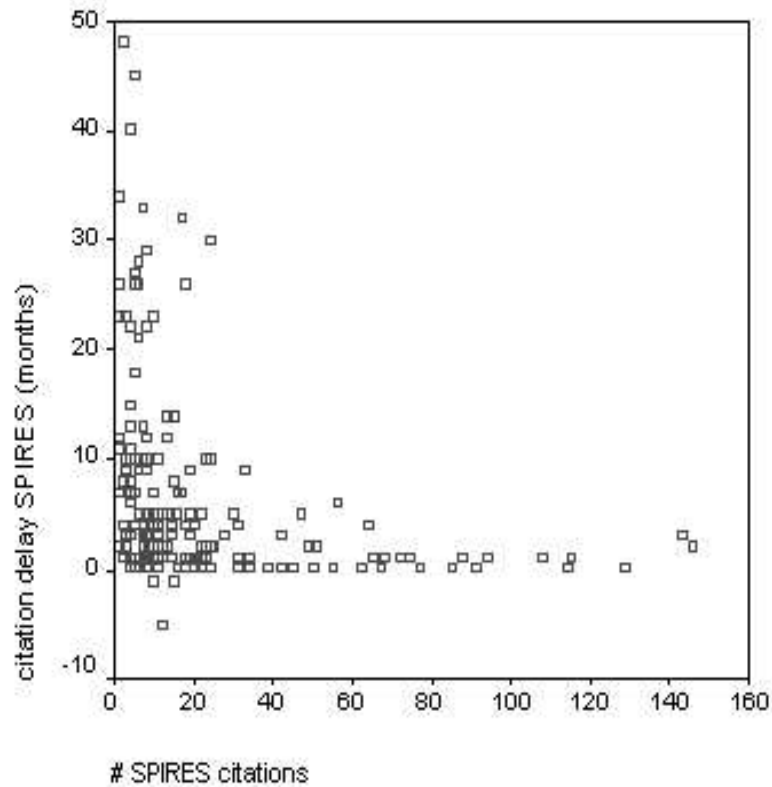


Abbildung 4: Streudiagramm von 184 Eprints nach der Zeit (in Monaten) bis zur ersten Fremdzitierung (*citation delay SPIRES*) und der Zahl aller Zitierungen (in mindesten 5 Jahren, *# SPIRES citations*)

destens 20mal zitierten Eprints vor ihrer Zeitschriftenpublikation erstmals fremdzitiert. Die schnelle Verfügbarkeit der Forschungsergebnisse über das Eprint-Archiv beschleunigt also die Kommunikation so stark, dass die überwiegende Mehrheit von Eprints die Zeitschriftenaufsätze in *Physical Review D* quasi überrunden.

²²Streng genommen sind die Zitierungen, die SPIRES HEP vermerkt, nicht alle in Eprints erfolgt, weil auch einige wenige Journalartikel registriert sind, denen kein Eprint zugeordnet ist. Die erste Fremdzitierung haben wir jedoch immer in einem Eprint gefunden.

Als letztes sollen hier noch kurz unsere Ergebnisse zu den Zitierungen der Zeitschriftenaufsätze in anderen Zeitschriftenaufsätzen diskutiert werden. Wir recherchierten hierzu im WoS, ermittelten, wie bei den Eprints, für alle 211 PRD-Artikel die Zahl der Zitierungen (in den mindestens vier Jahren bis Februar 2004) und den Monat der ersten Fremdzitierung. Zuerst wollten wir wissen, wie Zitierungszahlen in den beiden Publikationswelten korrelieren. Abb. 5 ist eine gute Korrelation zu entnehmen. Eine lineare Funktion der SPIRES-Zitierungen erklärt 95% der Variation der WoS-Zitierungen, ein beruhigendes Ergebnis.²³

Die Eprints unserer Untersuchungsmenge werden in Eprints im Mittel anderthalbmal so oft zitiert wie die zugehörigen PRD-Artikel in anderen Journalartikeln. Das rührt einmal von der längeren Zeitspanne her, die sie zitierbar waren. Dann sind nicht alle zitierenden Eprints auch im WoS erfasst, einige *noch* nicht. Eine weitere, interessante Ursache für kleineren Zitationszahlen der PRD-Artikel liegt darin, dass – wie eine stichprobenhafte WoS-Recherche zeigte – Journalartikel auch nach dem Erscheinen des PRD-Heftes weiterhin nur den Eprint zitieren.²⁴ Das bedeutet, dass Autoren, Gutachter und Herausgeber das Journal, in dem der zitierte Artikel erschienen ist, nicht beachten (was dessen *impact*-Factor absenkt). Es wäre interessant zu erfahren, ob eine Ursache für ein solches Verhalten darin liegt, dass die Autoren nicht über das jeweilige Journal verfügen.

Die Zeitspannen bis zur ersten Fremdzitierung sind bei den Journalartikeln in allgemeinen größer als bei den Eprints. Die Ursache dafür ist, dass die zitierende Arbeit den langwierigen Publikationsprozess durchlaufen muss. Von den 175 (d.s. 83%) überhaupt fremdzitierten Artikeln (bei den Eprints waren es 87%) sind drei Viertel nach 17 Monaten fremdzitiert (bei Eprints bereits nach 8 Monaten). Der Median verschiebt sich von 3 Monaten bei den Eprints auf hier 8. Es ist also eine Verdopplung der Reaktionszeit zu beobachten.

²³Korrelationskoeffizient $R^2 = 0,948$, Anstieg $a = 0,64 \pm 0,01$, Schnittpunkt mit der y-Achse $b = -0,1 \pm 0,3$.

²⁴11 von einer Stichprobe von 29 zitierten theoretischen Hochenergie-Eprints (als Eprint in einer WoS-Zeitschrift zitiert), d.h. mehr als ein Drittel, werden auch innerhalb von gut 5 Jahren nach ihrem Erscheinen als Zeitschriftenaufsatz insgesamt 28mal ausschließlich als Eprint zitiert (und 10mal zusammen mit dem Verweis auf den Zeitschriftenaufsatz). Das weist darauf hin, dass nicht selten die Zeitschriftenpublikation eines Eprints beim Zitieren ignoriert wird. Die Stichprobe wurde aus im August 1997 eingestellten Eprints gewonnen (CITED WORK: HEPH9708*). Von der angezeigten Liste von Zitierungen habe ich die ersten 29 Eprints ausgewählt (Autorennamen, die mit A bis F anfangen).

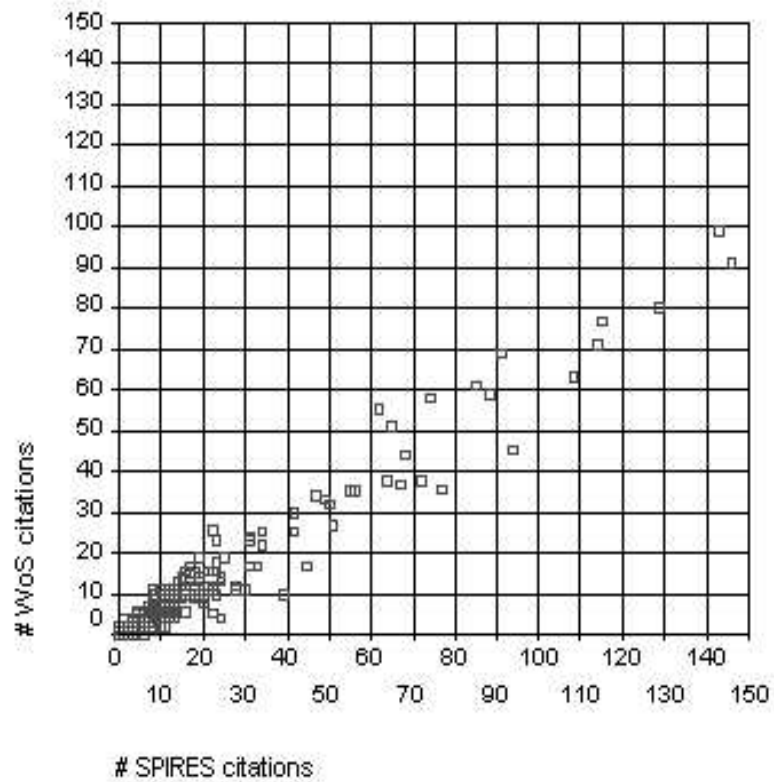


Abbildung 5: Streudiagramm von 203 Eprints nach der Zahl ihrer Zitierungen als Eprint in Eprints ($\#$ *SPIRES citations*) und der Zahl ihrer Zitierungen als Journalartikel in anderen Journalartikeln ($\#$ *WoS citations*)

Wir gingen dann auch der Frage nach, wie stark die erste Fremdzitierung des Journalartikels in anderen Journalartikeln hinter der ersten Fremdzitierung des Eprints in anderen Eprints hinterherhinkt. Bei den 166 gültigen Fällen ermittelten wir einen Median von 10 Monaten und ein arithmetisches Mittel von 14 ± 13 Monaten. Das kennzeichnet den Zeitgewinn durch Eprints bei der Beachtung von Resultaten.

5 Fazit und Ausblick

Eprints, die vor der Publikation im Journal im Netz erscheinen, können beträchtlich die wissenschaftliche Kommunikation beschleunigen. Das ist das Ergebnis der vorgestellten Studie.²⁵ Theoretische Artikel in *Physical Review D* von 1998 und 1999 sind zu 96% als Eprints in das *ArXiv* gestellt worden, typischerweise sieben Monate bevor sie online als Zeitschriftenaufsatz greifbar waren. Die Eprints überrunden die Zeitschriftenaufsätze in dem Sinne, dass 3/4 von ihnen in Eprints von anderen Autoren zitiert werden, bevor das entsprechende Heft von *Physical Review D* erscheint.

In welchem Maße und wie schnell Verlage, Wissenschaftler, ihre Institutionen und ihre Fachgesellschaften zu *Open Access* beim wissenschaftlichen Publizieren übergehen, kann heute nur spekuliert werden. Wird zunächst vorrangig der *grüne Weg* beschritten, bei dem Artikel als Eprints frei verfügbar sind, die Journale jedoch weiterhin von den Bibliotheken bezahlt werden, löst das noch nicht die Zeitschriftenkrise. Es werden weiterhin wegen fehlender Mittel Zeitschriften abbestellt werden, was die Verlage veranlassen könnte, weiter die Preise zu erhöhen. Falls wirklich die überwiegende Masse gerade auch wichtiger Artikel als Eprint verfügbar werden sollte, wird jedoch die Preisspirale die Wissenschaftler selber nicht schmerzen, jedenfalls nicht, solange davon nicht das soziale System der Wissenschaft mit seinem spezifischen Bewertungsmechanismus über anerkannte Publikationen berührt wird. Reputation muss jedoch nicht nur mit Zeitschriftenpublikationen erlangbar sein (in der Informatik verschaffen auch Konferenzbeiträge in beträchtlichem Maße Ansehen). Die Verlage müssten einen Ausweg aus der Sackgasse finden oder sich mit einem kleinen Segment am Markt wissenschaftlicher Information zufriedengeben.

²⁵Bis März 2004 hatten wir keinen Bericht über ähnlich gelagerte Untersuchungen gefunden.

Literatur

- de Solla Price, D. J. (1963). *Little Science, Big Science*. New York: Columbia University Press. dt. mit selbem Titel bei Suhrkamp 1974.
- Giles, J. (2004). Societies take united stand on journal access. *Nature* 428, 356.
- Hitchcock, S., D. Bergmark, T. Brody, C. Gutteridge, L. Carr, W. Hall, C. Lagoze, and S. Harnad (2002). Open citation linking. The way forward. *D-Lib Magazine* 8(10). The text of this article is available on <http://www.dlib.org/dlib/october02/hitchcock/10hitchcock.html>.
- Kölbl, M. (2003). FORUMnovum Dynamic Publishing. Ein Konzept für die Zukunft des wissenschaftlichen Journals. In H. Parthey and W. Umstätter (Eds.), *Wissenschaftliche Zeitschrift und Digitale Bibliothek: Wissenschaftsforschung Jahrbuch 2002*, pp. 135–142. Berlin: GeWiF. http://www.wissenschaftsforschung.de/JB02_135-142.pdf.
- Lawrence, S. (2001). Online or invisible? *Nature* 411(6837), 521. See also <http://www.neci.nec.com/~lawrence/papers/online-nature01>.
- Mizzaro, S. (2003). Quality control in scholarly publishing: A new proposal. *Journal of the American Society for Information Science and Technology* 54, 989–1005. An eprint of this article is available on the author's homepage: <http://www.dimi.uniud.it/~mizzaro/papers/EJ-JASIST.pdf>.