

## **Three Frontiers in Open-Access Scholarship**

Malcolm Getz  
Vanderbilt University  
Nashville, TN 37235-1819

[Malcolm.Getz@Vanderbilt.edu](mailto:Malcolm.Getz@Vanderbilt.edu)

January 13, 2005

Draft

## **Three Frontiers in Open-Access Scholarship<sup>1</sup>**

There are three important frontiers in moving from subscription-based scholarly publications to delivery of scholarly works to readers without charge via the Internet. First are automated archives of preprints and post prints that do not require formal editorial review before posting. The arXiv service, now at Cornell, is emblematic of this frontier. (arXiv, 2004) Second are the quality-assured journals that are distributed on an open-access basis. The Public Library of Science initiative in launching journals in biology and medicine is emblematic of the second frontier. (Public Library of Science, 2004) Third are open access indices to the scholarly literature. Google Scholar, launched in beta version in November 2004, is emblematic of this frontier.(Google, 2004) Each frontier advances the prospect that the best scholarship will be readily available to all via the Internet. Our goal here is to identify where each frontier is today and how it may evolve.

There are related frontiers in open access to educational materials as at MIT and in many initiatives to make information products available with little or no charge on the Internet as with Linux and open-source software. The focus here is on the original works scholars produce to establish their ideas.

The goal of new formats for scholarly communication is clear. Open-access archives and journals might make scholarship more useful to more people, while lowering total costs. All of the participants in the scholarly process may see their roles change.(Herbert Van de Sompel et al., 2004) An effective editorial process to winnow lesser ideas and develop and promote the most valuable ideas will remain of vital importance. Good ideas gain value from being shared, developed, and celebrated. Digital archives and open journals have a considerable advantage in being shared over the

---

<sup>1</sup> I appreciate comments from Paul Gherman, Kenneth M. King and John J. Siegfried. The Cornell University Project on *Creating an Open Access Paradigm for Scholarly Publishing* commissioned this work.

Internet. Creating a process for development and celebration in the digital arena is the larger challenge.

### **Open Archives**

The concept of an open-archive of scholarly works with no required peer review for quality assurance dates to the creation of arXiv in 1991. The service plays a significant role in the disciplines it serves. The frontier for open-archives has two parts. One is to expand the content to a comprehensive set of scholarly works in a discipline and the second is to extend the scope to more disciplines. Thinking about how this frontier may expand depends on understanding costs as well as anticipating the value of a broader service.

### **Cost**

The basic cost structure of arXiv demonstrates that a comprehensive digital archive for most disciplines would entail modest costs. Extending the coverage and scope of open-archives is both technically feasible and well within the range of costs universities might find affordable.

Cornell University supports the original arXiv budget at a cost of under \$200,000 per year in 2004, making some allowance for management and overheads. (Thomas H. Hickerson, 2004) Submissions have grown linearly since August 1991 ([http://arxiv.org/show\\_monthly\\_submissions](http://arxiv.org/show_monthly_submissions)) and exceeded 4,000 per month in October 2004. Connections to the main service exceed 12,000 per hour in peak hours. In addition, there are 17 mirror sites around the world with costs and connections not counted in the figures for the primary site at Cornell. This cost encompasses processing new items into the database, storage and maintenance of the files, and retrieval on demand. The cost structure provides no financial support for an editorial process but does include a full-text index and retrieval capability.

The total cost at the primary site per submission is just over \$4 with zero cost allocated to readers. [ $\$200,000 / (4,000 \times 12)$ ] This total cost carries the automated intake as well as a significant flow of current connections.

ArXiv cost might also be considered relative to connections. Suppose there were 200,000 connections per day for about 240 working days per year (a low estimate) and that each connection involved using one item. This would mean 48 million “uses” per year (one essay per connection per hour). If essays are used only in the first year of submission, then the 48 million uses divided by the 48,000 submissions gives 1,000 uses per submission as a rough average (and probably an understated estimate). That puts the cost per use at \$0.004 (with no cost allocated to submissions) with use meaning rapid search of a full-text database with delivery of the article in a few seconds to any reader’s desktop everywhere the Internet goes.

A system of charging either authors or readers could be implemented in this environment but the cost of accounting would probably dwarf the cost of the service itself and would change the character of the service. If the goal is to recognize, develop, and celebrate great ideas, transaction-based fees are an impediment.

The alternative to reader fees or author submission fees is institutional support. Cornell chooses to fund arXiv. Consortia with membership fees would be possible and would be similar in character to offering journals on the basis of subscriptions to academic libraries. If those who don’t join the consortia do not gain access to the arXiv either as readers or authors, then the service will be less comprehensive and less valuable.

A better regime for extending the scope to other disciplines may well be one in which each campus bears a share of costs as a part of its commitment to the community of scholarship without expecting compensation. Whether voluntary finance by major universities that seek to serve the world would be sustained is an open question. In announcing support for making some of its libraries’ collection available without charge on the Internet, Harvard University’s President Larry Summers said: “We have always thought of our libraries at Harvard as being a global resource.” (John Markoff and Edward Wyatt, 2004) Sustaining information resources for the planet is part of the central mission of research universities.

The key to arXiv’s low cost is that all its processing of scholarly works is automated. Authors must conform to standard formats in preparing materials for

submission. Similarly, readers must provide their own Internet search tools that comply with (now nearly universal) standards. The submission format allows capture of metadata that describes the submitted document. The full-text file is added to a searchable index. The system allows sophisticated searching and delivers page images. This regime assures low cost and scalability. Use of simple systems that are fully automated is a key to successful open scholarship.

New protocols allow a physical open-archive to be distributed over many campuses with metadata harvested from each site to build a common index so that the whole can be searched as though it were at one site. Universities in the United Kingdom proposed to pursue the distributed architecture.(Alma Swan, et al., 2004) This approach would allow the open-archive to grow across disciplines with each campus bearing its own share of costs with modest expenditure in common facilities. The government in the UK rejected the proposal in November 2004. Nevertheless, the UK proposal is a useful point of comparison for developments in the US.

If an arXiv-like total cost were born by 100 campuses (this is the approximate number of universities in the US whose libraries are members of the Association of Research Libraries), the cost of arXiv would be \$2,000 per campus. ( $\$200,000/100$ ) Extending the scope to 25 disciplines would put the cost per campus at \$50,000 per year. This cost would provide immediate access to about one million new essays per year, roughly the annual content of 10,000 journals. ( $48,000 \times 25 = 1,200,000$ ) An average research library in the US now subscribes about 10,000 current refereed scholarly journals. An open digital archive, however, provides no method to finance processes for recognizing, developing, and celebrating ideas in the manner provided by well-edited journals.

## **Benefits**

A good open archive of substantial scale makes content instantly searchable with essays delivered immediately to anyone on the Internet. The content must be large enough with a sufficient proportion of work of good quality that is relevant to a given reader so that the reader will enjoy enough success to turn to the archive frequently. An

archive that is too small or has too little quality in content will attract few readers. The search facility must be relatively easy to use and yield desired works with a high probability and yield unwanted works with a low probability. The work delivered to the reader must have a clear source, be easily readable, and allow printing on demand. The tools for creating such a service are now relatively inexpensive. Inducing authors to submit their works has been difficult except in a few disciplines like those in arXiv.

A number of investigators compare rates of citations to published articles that appear in arXiv as e-prints to those that do not appear in arXiv. Brody, for example, finds citation rates in nuclear and particle physics to be roughly 2.5 to 5 times higher with arXiv posting than without. (Tim Brody, et al., 2004) The method of analysis, however, does not account for other possible sources of differences in citation rates, for example, the author's and journal's reputations. Lawrence compares citation rates for conference articles in computer science that are openly available online to those that are not. (Steve Lawrence, 2004) Comparing citation rates for a given conference in a given year and limiting the analysis to the top 20 conferences yields a 284 percent higher rate of citation for open articles. Harnad notes that over time more publishers are allowing authors to post their post prints in open archives: 34 percent in 2003, 52 percent in 2004, (a recent communication says 92 percent). (Stevan Harnad, Tim Brody, 2004) As more literature is readily in the open, the likelihood of use of each open article increases. As a consequence, more readers turn to open sources more frequently, disadvantaging material not openly available. The analysis of citation rates are generally limited to fields included in arXiv, making it difficult to distinguish the effect of arXiv from other attributes of the disciplines studied. The size of the effects, however, is large.

The convenience of open access doesn't necessarily increase the total amount of reading. A student of libraries and scholarship, Donald King, notes that university scholars read about 200 articles per year on average. (Donald W. King, 2004) Open access may not increase the total amount of reading. Instead, as open access materials are used more frequently, materials not available in the open may be used less frequently. As open access makes more materials more convenient, scholars may choose what to read more carefully and be more likely to read material that is better suited to their needs.

More sophisticated investigation will be needed to reveal whether use of items not in the open is declining, whether the total rate of reading is changing, and whether users of open access find materials that better match their interests or simply settle for what happens to be more convenient.

Open access via a large digital archive offers the benefit to an author that work reaches readers faster, even in a matter of minutes, and that more readers have ready access. I have not seen an estimate of the proportion of new works in the arXiv disciplines that appear in arXiv but I take it to be a substantial share, but well under 100 percent. (The figure in (Steve Lawrence, 2004) might indicate 50 percent of the literature in computer science is in archive.) It would be of interest to know what fraction of different kinds of scholars who read work in physics, math, and computer science use arXiv. Casual observation suggests the rate is approaching 100 percent.

Institutional repositories supplement the discipline-based services like arXiv. EPrints is in use on nearly 150 campuses, DSpace on nearly 70 more and Digital Commons at a dozen with other systems also in play. (DSpace, 2005, EPrints, 2005, ProQuest Digital Commons, 2005) Institutional repositories allow authors to post pre and post prints in local archives in a manner that allows information from the documents to be harvested centrally to support Internet-wide searches. Institutional repositories allow decentralization of most of the cost of archiving while providing universal access. With cooperative effort, they could provide coherent meta-data structures that would allow them to function across many disciplines with the coherence that arXiv provides for its disciplines. Institutional repositories, however, do not in themselves provide methods for identifying the quality of specific works.

The open-archive concept has been slow to spread beyond the arXiv disciplines. Indeed postings on arXiv have grown linearly, suggesting a process other than the exponential growth rates of a contagion. Perhaps there is a barrier to first use, even a psychological one, which discourages authors from posting their works in open archives. If open archives are to become significant in more disciplines, more authors must post their works. Authors can use posting an essay to an open archive as a method of

establishing ownership because the archive provides a permanent record of date of authorship. (Preservation and permanence are discussed below.) One impediment to posting may be a concern that essays posted openly will be more difficult to publish in formal journals.

The UK considered requiring faculty at public institutions to post their works.(Alma Swan, et al., 2004) Government regulation requires universities in the UK to provide faculty curriculum vitae with copies of published works to allow external examiners to evaluate research. An open archive of all faculty research would facilitate the evaluation. The NIH in the US proposed that authors who receive grants from the NIH provide electronic copies of their works so that NIH could post them at PubMedCentral after six months of publication. (U. S. House of Representatives, 2004, Rick Weiss, 2004) Other government agencies and private philanthropies that fund research could require work they support to be offered openly. Most agencies and philanthropies that sponsor research require that the results of the work be published. Requiring that works be posted in open-archives derives from the same view that ideas are most valuable when they are shared. However, some publishers have responded by offering some published works publicly on the Internet but object to allowing all of their publications to be freely available.(Bernard Wysocki Jr., 2004) Other publishers including Nature invite authors to post articles in open archives six months after they appear in Nature.(Nature Publishing Group, 2005)

Faculty, however, want their works to be shared, developed, and celebrated. The open-archive is a pathway for sharing but provides little in the way of development and celebration.

A significant proportion of journals allow authors limited rights to post their published essays in open-archives as post-prints. However, open-access to post-prints may become a substitute for journal subscriptions and thereby undermine the revenue streams that support the editorial process that turns raw work into attention-getting articles. If the substitution becomes significant, subscription journal publishers will face

significant choices about limiting post-print publication rights, finding alternate sources of finance, or ceasing publication.

In promoting open-archives, then, one must weigh the value of the editorial process. The risk of requiring that all post-prints be made available in an open archive is that such a policy may undermine financial support for the editorial process. Open-access journals address this concern. Although open archives are inexpensive and widely used once the content reaches a critical mass, they are only likely to thrive when quality journals are also open access.

### **Open Access Journals**

The second frontier for open-access scholarship is the growth of open-access journals. On November 16, 2004, the Directory of Open Access Journals (Lund University Libraries, 2004) reported 1,357 open access titles of which 331 were searchable at the article level. The UK proposal mentioned above (Alma Swan, et al., 2004) reported 1,148 titles in July 2004. The difference indicates about a 50 percent annual growth rate. The rapid growth in titles may indicate in part the fact that the cost of the infrastructure needed to mount a journal (exclusive of editing and design) has dropped to near zero. Note, however, that counts in the Directory may be misleading because as many as 20 percent of the titles may simply be URLs with no content. (Sally Morris, 2004) (See comments on journal qualification below.)

Vanderbilt's new Center for the Americas has launched a journal, *AmeriQuests* (2004a) It is a multi-language, peer-reviewed journal designed to reach audiences across the Americas. It is open-access to broaden its reach with a fee-based print option. Institutional support provides the finance. It uses journal management software that is available to any non-profit organization without charge. The Vanderbilt Library has installed the Public Knowledge Project Open Journal System software (John Willinsky, 2004) and will consider migrating to journal management software expected from Cornell next year. Vanderbilt supports an electronic institutional repository; the journal management software layers on top of it. The library has no business plan for the

individual journal because it sees no incremental cost. (Private communication from the Library Director.)

The library appears ready to support other journals that arise within the University with no incremental charge. The faculty group that launched the Americas research center is responsible only for the selection and editing; they face no production costs. The incremental cost of launching a new journal at Vanderbilt (and by implication at many research universities) is simply the commitment of editors to edit and promote the journal. The fixed costs of software (no license fee, just local support costs) and hardware (the cost for a journal is too small to worry about) allows these services to be treated as infrastructure to be supported in common for all with no direct charge to users, either authors or readers. Of course, when the volume of submissions is high as with widely cited journals, the process of screening involves a level of effort unlikely to be offered voluntarily. There is more discussion of the editorial process below.

The number of open-access journals is likely to continue to grow rapidly using voluntary editing and reviewing. This set of journals might be integrated into Internet data with a fully integrated index as discussed below.

As an open journal gains reputation, it may choose to introduce submission fees and use the funds to induce reviewers to respond promptly. When submissions are many times the number of acceptances, a fee will reduce congestion in reviewing both by discouraging less promising submissions and encouraging reviewers to act in a timely manner. Another possibility is that open journals will operate at smaller scale within a tighter community of scholars where peer pressure for performance will be more effective. Page charges at publication also become possible when the quality of an open journal is high. Indexing services can help establish quality. Investment is likely to be larger and come more quickly to activities where use is growing exponentially because the rate of return will be higher compared to services where use is growing only linearly.

## **Open Index**

In November 2004, Google, Inc. launched Google Scholar, an open index to scholarly content on the Internet. (Google, 2004, Jeffrey R. Young, 2004b) In concept, Google Scholar may provide an open-access, fully integrated index to all scholarly literature that would put the content of open journals on an equal or better footing than subscription journals. Understanding how an automated index performs will be useful in forecasting the prospect for both the open index itself and the value of an open index to the success of open journals. Because the service is new, the discussion here provides some details.

Although the announcement of Scholar was vague about the scope of its content and its method of indexing, in broad terms, Scholar applies a Google search engine to a substantial body of scholarly essays and books. An announcement earlier in 2004 showed that Google was developing a search service with MIT and other research university libraries that would provide access to the full content of digital archives on the campuses. (Jeffrey R. Young, 2004a) Another announcement described a relationship with OCLC that would allow Google to provide access to the bibliographic records of books in a significant number of libraries, perhaps ultimately catalog records to all 54 millions titles in the OCLC database. (Jeffrey R. Young, 2004c) Google will digitize the full-text content of the 7 million volumes of the University of Michigan Library and has agreements with many other libraries to digitize significant parts of their holdings. (Kevin J. Delaney and Jeffrey A. Trachtenberg, 2004)

Although the initial release is a beta version likely to be improved over time, the service has sufficient merit to become a potent tool for academic work. Its weaknesses are also apparent but are not fatal and are amenable to improvement.

### **Names**

Three brief tests of Google Scholar demonstrate its power, highlight its limitations, and point to changes that might improve its performance. First, typing in an author's name should return a list of the author's works, in effect the publication list on

the author's curriculum vitae. Typing into Scholar the name of a scholarly author with a unique name will yield something close to the person's life publications presented in order by a proprietary scoring that tends to put frequently cited works first. For example, entering the name of a prominent economist, author: "Zvi Griliches" yields 689 references, the first of which shows that it has been cited 391 times that Google Scholar could identify. Clicking on the "cited by" button returns the list of the 391 sources that cite the Griliches' article. In contrast, the standard index to the literature of economics, ECONlit, returns 209 references for Zvi Griliches (with the ERLWEBSPIRS5 search engine) of publications in a less valuable chronological order by publication date. Comparing the Scholar result with a general Google search is also instructive. A general Google search of "Zvi Griliches" returned 16,700 references to Zvi Griliches that led with a homepage, an obituary, and a paid link to Amazon for books by Griliches.

Common names cause more difficulty. For example, entering into Scholar author: "Robert Jones" returns 306 items on widely varying subjects unlikely to have been authored by the same person. At least nine of the first ten items appear to be by different people. In contrast, a search for author "Robert Jones" in ECONlit (with the ERLWEBSPIRS5 search engine) returns 43 references with nine of the first ten referring to the same person. The focused character of the discipline database and the use of full author names reduce ambiguity. A general Google search on "Robert Jones" returned 500,000 references with little in common.

The standard library method for allowing names to be treated in a systematic way is to verify names against a name authority file. OCLC with its thousands of library members maintains a name authority file, a unique listing of full name with birth date and place for each author. As each new work by an author enters OCLC's catalog database, a cataloger verifies the name against the name authority file and uses the established name listing. A new author gets an entry in the authority file. The open-archive proposal in the UK (Alma Swan, et al., 2004) included the idea that author's names would be verified against the name authority file maintained at OCLC but the plan involved handwork by a cataloger with consequent increase in cost per item of about \$10. Google could not have

created a large index in a few months if all the author's names were to be verified against an authority file.

To maintain the very low cost of an automated service like arXiv and Google Scholar, it is essential that tasks be automated. Perhaps open journals will find a way to allow authors to identify themselves in the OCLC name authority file and to enter a unique name-authority code as meta-data that Google Scholar (and other search engines) could use to focus author searches. The notion would be that when an author is submitting an essay to an open journal (or any journal), the author would look up his or her name in the online name authority file and enter it into the electronic submission in a meta-data field. An author may need to do this only once per lifetime. After publication, the search engine would use the unique code to distinguish one name from another. In this way, the search engine could distinguish all works by one author from all other works. Of course, references would not list authors' names uniquely and therefore Google's index would reflect errors in attribution used in building the "cited by" lists.

## **Titles**

Google Scholar is much better at identifying known items by title. A search of words from a title returns a list of candidate items in a priority order that has a high probability of putting the desired item first, given that it can be found at all. As a trial, I searched for the 64 items listed as references in two articles in the *Journal of Economic Perspectives*, Summer 2004 (Perold and Stulz). Google Scholar returned a valid hit in 80 percent of the items. Non-hits were primarily more than 30 years old, not in academic journals or books, or were electronic sources with URLs in the original citation. In a few cases, the valid hit was not first but among the top five. In a few other cases, entering the author's last name with the words from the title helped identify the item. An eighty percent hit rate is very good for a service in beta version, a few weeks old. Google Scholar provided a link to the articles themselves for about a third of the items. The books could be located in libraries via OCLC's Open WorldCat, a feature of Google Scholar.

In 75 percent of the test references, Google Scholar provided a count and a link to a list of items that referenced the identified work. The average number of citations listed was 350 (median was 121). The forward-reference feature is a powerful way to identify relevant, related items that can be overlooked with word searches alone. Scholar uses the counts of linkages to establish priority order for returns from its word searches. Indeed, the index may weight each link by a count of links to it.

In ten percent of the title searches, Scholar returned more than one citation to the same item, usually because of variation in the spelling of an author's name or because there are multiple editions of a book.

In about one third of the cases, Scholar provided links that could return the page images (or HTML version) of the item itself. Typically, a reader must be on a campus that subscribes to JSTOR or other subscription supplier or be willing to pay per look in order to see the item. Open journals offer a clear advantage here.

Scholar retrieves brief references that identifies the authors but sometimes just with first initials, the year of publication and the journal title but not volume, issue, and page numbers (for books, the publisher and city of publication may be lacking). The Scholar record may link to a source with more complete information but many citations do not have links. Some links go to files of uncertain origin, for example, PDF documents that do not necessarily begin with title pages. A reader may be left to puzzle about whether an essay has been published, updated, or even repudiated by the author. (Einstein recanted one of his published essays.) The provenance may be unclear.

Journals might enhance their offerings to allow Scholar to address the completeness and authenticity of the documents it references by including a systematic labeling of the origin of the reference. Specifically, open journals might provide meta-data flags to indicate origin and evaluation. Google Scholar approaches validation by providing counts of citations. Refereed materials might also gain some priority. Publication in refereed journals provides authenticity, particularly when more complete tagged date and publication information is part of the record.

For comparison, ECONlit identified the reference by title of 22 of the 64 items in the references lists in the two essays, missing many early essays by Nobel Laureates, recent working papers, articles in journals and other publications not indexed in ECONlit, and several books. ECONlit's 34 percent success rate in identifying sources in these two essays compares with an 80 percent success rate in Scholar. ECONlit's references were detailed with volume, issue, page, and date.

A generic Google search for the same 64 titles led to a reference (among the top ten links returned) to the item in 60 percent of the cases (compared to 80 for Google Scholar). The references found in the generic Google search, however, were more difficult to use. Many were course syllabi or reading lists that have to be scanned to locate the desired reference. The references in Scholar are immediate. Books in the generic search linked to Amazon or other bookseller, compared to the library links in Scholar. Just over ten percent of the titles in the generic search linked directly to the items themselves (compared to a third in Scholar counting paid or subscription access). References with URLs including government documents went directly to the documents. Some working paper databases are available in both Scholar and the generic Google. Neither Scholar nor a generic Google search identifies newspaper stories.

## **Subjects**

The third strategy for finding material is to search by subject. Searching Scholar with very precise terms tends to yield a well-focused and appropriate set of references. For example, "Stolper Samuelson Theorem" returned 1,030 references with the first listed item having been cited 58 times and the 30<sup>th</sup> item having been cited one time. (The original 1941 essay that enunciated the theorem was not included.) All of the first ten items were items from the economics literature concerning the theorem. For comparison, ECONlit yielded 93 references. The ECONlit list was tightly focused on the professional economics literature. (ECONlit provides a subject index with the "Stolper Samuelson" entry leading to 53 references.) The Scholar search encompassed a wider range of literature. Its "cited by" lists provide a method of access to the literature that is not available in ECONlit. Because ECONlit is a service offered through the campus library,

it provides an easy link to library holdings including digital versions of articles when available. Scholar provides links to electronic sources but only links to the library for books. Libraries may seek ways to allow Google Scholar to provide links to their subscription-based ejournal services and other holdings information.

A search in Scholar on broad terms not readily identified with a discipline often yields many extraneous references. A search on “open access journals” yielded 235,000 references, however, only three of the first ten concerned journals (#7, #8, and #10). Other references concern transmission of natural gas, genetic algorithms, and the harvesting of North Atlantic Whales. ECONlit yielded no references for this search. A general Google search on this phrase yielded 4.75 million references with one paid link. The first listed reference was to the Directory of Open Access Journals and many other of the top listings were to specific open journals. Scholar identified sources that discuss the topics.

Subject searching might be improved with the addition of a formal subject structure. The open-archive plan in the UK (Alma Swan, et al., 2004) proposed the use of the Dewey Decimal System numbers to label subjects. OCLC maintains a list of subject entries and a staff person in the proposed UK archive would have assigned a subject number. OCLC would provide a mechanism for adding new subject headings. The cost of handwork in maintaining systematic subject entries, however, is prohibitive for an open service and out of character for Scholar.

An alternative might be to have authors propose one or more subject labels. As with the name authority, the author would consult an online read-only listing of Dewey Decimal subjects or other subject labels and select appropriate subject numbers. On submission to the journal, the automated system would verify that the subject numbers were valid and the editor would give visual confirmation. The use of universal subject numbering would allow a journal index to span all disciplines yet readily allow a reader to limit searches by subject. A universal index would more easily allow a reader to find relevant material in tangential disciplines.

Google Scholar or another open-journal index might support other author-initiated, controlled vocabulary content labeling as well, invoking labels for empirical content, visual information, time period (for example, century), language, and continent. OCLC or another entity might maintain the controlled vocabulary on a read only basis. The author would look up and copy the appropriate label and enter it into a page of meta-information that is part of the essay being submitted. The automated journal system would verify that the terms used are valid. The editor would provide a visual check. An editor might decide what labels would be required in a given journal. History journals may require labels for time period and continent. Each discipline may have its own set of controlled content labels. For example, the subject labels used in the ECONlit might be readily used as a required subsidiary label with automated control in open-access economics journals. Surprisingly, the ECONlit search (at least with the ERLWEBSPIRS5 search engine) does not support searches by the official subject headings that are a vestige of its print forbearer.

The general Google search screen provides a link to “advanced search” that allows searchers to search by word proximity or for the exact phrase and to limit searches by language, date, format, and domain. An advanced search in Scholar (if such should become available) might allow a searcher to specify a discipline and to use name and subject codes that are common among journals in that discipline.

We may have some concern that Google will treat Scholar as a proprietary service that, once established, could be exploited for extraordinary monopoly profit. Consider, for example, the possibility that Google, Inc. encounters financial distress a few years hence and decides to sell its under performing assets to a major commercial publishing firm. Various pricing strategies might then come to exploit the market power inherent in a dominant service with significant economies of scale. Witness the once intense competition among firms producing word processing software and the current market concentration and extraordinary profit.

## **The Changing Role of the Journal**

With the very low cost of launching open journals and the automatic indexing in Google Scholar or the equivalent, we can expect journals to proliferate rapidly. In such a world, journals will differentiate themselves in three ways. First, of course, is the quality of the editing that recognizes, develops, and celebrates good and relevant new work. Second is the prescription and monitoring of metadata to enhance searching. Third is a commitment to preserve the journal's content for perpetuity.

## **Editing**

Launching a new open journal at a research university with now common database tools for archiving documents (DSpace or the equivalent) has essentially zero out of pocket costs for launching a new journal given that the digital archive is in place. Therefore, would-be editors can readily launch journals. Many journals (I don't know how many) run with volunteer editors, often working within the context of a community of scholars, a scholarly society, or other group. There are, however, limits to what voluntary editors and reviewers can achieve. An editor has few inducements to offer associate editors and reviewers to encourage timely decisions.

The critical issue is to understand the value and cost of editing. What induces people to become editors? How important is the quality of editorship? Great journals have great editors. Some professional societies bid for distinguished scholars to edit premier journals. Conventional wisdom is that commercial publishers and some scholarly societies pay editors well. The publishers' willingness and ability to do so is an essential ingredient in the publisher's success. Information about the specific payments publishers make to editors is likely to be proprietary, making empirical analysis difficult. We may recognize that successful editors need some special talents, some organizational skill, and contacts with many scholars to recruit reviewers and associate editors and to solicit manuscripts, and a lot of breadth of perspective and tolerance of various methodologies found in their disciplines. Some scholars who are offered prominent editorial posts decline to accept the roles. It would be interesting to learn the pattern of compensation of editors and to see how compensation is correlated with journal citation rates or other

measures of quality. The pattern is quite likely to vary by discipline with some more hierarchical (higher peaks) than others.

Premiere journals involve significant editorial cost. The American Economic Association reported \$601,000 of employee compensation for the *American Economic Review* (AER) and \$325,000 for the *Journal of Economic Perspectives* (JEP). (John J. Siegfried, 2004) These expenditures are for the editorial staff of each journal, including senior editors and co-editors, copy editors, and support staff. (Most administrative functions are in separate lines.) Most of these editorial costs are not specific to a particular format and would presumably be necessary to produce journals of comparable quality whether on an open or subscription basis. The AER published 106 articles in 2003, putting the editorial compensation at \$5,700 per article. (Ben S. Bernanke, 2004) With space and other office expenses, the editorial cost would be more than \$10,000, much more than is likely be paid as author fees. JEP published 44 articles and features in 2003, putting editorial compensation at \$7,400 per article. (Andrei Schleifer, 2004)

### **Design and Copy Editing**

In addition to direct compensation of the scholars who husband the intellectual content, editing may also involve a variety of tasks in preparing works for publication, including copy-editing and design. How much does expenditure on document preparation fall within the journal's editing budget? Some journals like the JEP mentioned above provide extensive rewriting services to create a journal that is readable by a wider audience.

The Public Library of Science is a publishing venture with significant philanthropic support for its start-up. (Public Library of Science, 2004, Nick Twyman, 2004) Tracking its usage rates, prices, and financials will be a significant benchmark for open-access journal publishers. Tracking the experience of other open-access peer-reviewed journals will be helpful in understanding the long-term prospects for self-finance of open-access journals.

What is the value of editing? The editorial process has a number of dimensions. A better editor may attract better submissions, a role that is more important for less well-established journals. An editor may credibly solicit manuscripts from interesting scholars. An editor may influence authors to reshape or even redirect their work. Ultimately, editors may select articles for publication from among many that reviewers judge as excellent. In a top journal, reviews may qualify 25 percent of submissions as well worth publishing. The editor may then select less than half of these for ultimate publication. The editor may be selecting 100 articles for publication per year from more than 200 that have been praised by reviewers, a significant task.

Ultimately, editors are arbiters of good taste. An editor may select from among well-crafted essays those he or she thinks are most interesting for the journal's readers. Because interesting journals attract readers, the editor may choose essays that address novel topics or use clever methods. The editor is a kind intellectual entrepreneur, looking for good ideas to promote.

Entry of effective new journals depends critically on the ability to recruit and sustain effective editors. PLoS began by hiring editors from *Cell* and *Nature*. (Nick Twyman, 2004) Even if the incremental cost of acquiring and publishing manuscripts is zero, a plan for a journal must address what editorial functions the journal will support and how to finance them. It is also appropriate to contemplate the incentives editors face. With subscription finance, a for-profit publisher (owner) looks to the rate of change of subscriptions as an indication of editorial success. A not-for-profit may also look at submissions. Is the journal attracting the best essays within its intellectual scope? When sales or submissions lag, it's time to look for a new editor.

How would editors be recruited, evaluated, and rewarded for an open-access journal if the open-access publisher will compete with the commercial publishers for editorial talent? Open-access journals will need support from universities through author fees, support for editors, or direct subvention as with a university press. The issue here may well be whether universities are willing to compete for editorial talent? In time, the fully integrated journal index, whether Google Scholar or a similar service, would readily

summarize citation rates for open-journals and give a score of successful editorship. Of course, such scores are not substitutes for good judgment about intellectual success. Will universities value the intellectual contributions made by editors?

Faculty may be reluctant to value service as an editor in recruitment and retention. Although research, teaching, and service are valued, original intellectual contributions that are recognized by others are central. Achievement in research has enduring value. Success in teaching is less enduring and gets less emphasis. Service is in third place. As an alternative, universities may come to pay specific salary supplements to its faculty who edit open journals, much as they engage faculty as department chairs and in other time-limited managerial roles.

Premier journals in each discipline may continue to be financed by subscriptions and make significant expenditures in editing in order to serve a wide audience. Open journals will proliferate in increasingly smaller intellectual niches with editing financed through university salaries, subventions tied to editors, and perhaps page charges to authors. Journals in the middle with high costs but without strong reputations may be under increasing financial pressure as library budget constraints limit purchases of subscriptions. Articles that require a subscription or pay-per-look fee may be at an increasing disadvantage in being less likely to be read and cited except where journal titles have strong reputations.

### **Journal Certification**

Assume that many readers prefer to limit most of their reading to peer-reviewed works that have been selected for publication by a careful editor. The index might then allow a reader to limit a search to published items. A conventional index establishes a threshold of quality for treating a journal as an active, referred, scholarly journal, because the decision to index the content of another journal involves a commitment of continuing cost by the index. A journal that is not indexed in the standard index for its discipline will remain invisible and insignificant.

In contrast, an open index can readily index in full-text all essays whether published or not as long as they are posted in an open archive. Some secondary process will certify the bona fides of a journal to distinguish unpublished materials from those that have passed through a significant review process. Although some observers expect journal titles to disappear in favor of direct access to articles via an open index, many recognize that editors play valuable roles as noted above. The journal title and the reputation of the editor should differentiate content but to be valuable in an open index, the designation of “published” ought to rest on some process of certification. A university’s willingness to support an open journal will depend on its own sense of the journal’s success, an informal certification.

The open index itself might rank the journal title based on the count of “cited by” links to its content. A certified journal might be one that has published, say, at least 20 articles per year for two years and has achieved a median “cited by” count of citations per published article that exceeds a suitable threshold. More careful analysis of the cited by rates for journals would establish suitable thresholds that might vary from one discipline to another. A new journal would need to pursue a sufficient level of effort to achieve certification and to sustain a level of effort to maintain certification. After initial certification, open journals may well come to achieve higher levels of intellectual impact as signaled by citation rates. The sponsor of an open journal should be readily able to monitor the performance of its journal.

Readers may turn to new issues of an open journal to track developments in an area of scholarship and elect to receive an email notice of the table of contents of each issue, as is now common for subscription journals. When searching for essays on specific topics, a reader might choose to limit searches to the content of certified journal titles along with limits on dates, names, and subjects. A scholarly group that contemplated launching an open journal would anticipate a level of effort sufficient to create a certified journal so as to attract the attention of readers who limit their search to certified journals. Of course, Google Scholar’s ranking of items from a search goes some way toward certifying individual works. Journal certification should be a low cost method of recognizing the value of editorship.

## **Metadata**

A second way an open journal might differentiate itself is by requiring and monitoring metadata from authors to support higher quality searching in the open index. As suggested above, an open journal might require authors to include an authenticated version of their names, located, say, in an OCLC name authority file, as data in a tagged field in an opening block of metadata at the beginning of the digital version of the essay. With authenticated names available to the search engine, the author's name can become linked to all works by the author in the open index. Enhanced access by name makes the author's work more accessible and therefore more valuable to the journal and to the author. This enhancement is only practical if it can be achieved in a highly automated and therefore inexpensive way. Perhaps OCLC or similar service might provide a name registry for a modest initial fee that would allow an author to provide the needed information (for example, full name, birth date and place) to be affirmed by one or two other people to assure authenticity.

Similarly, the journal might include its title, volume, issue, and page numbers as metadata so that the open index can provide complete bibliographic information with each reference. Inclusion of a permanent electronic reference tag is also useful so that location is not dependent on transient URLs.

Finally, the open journal might require and monitor metadata to support structured subject searching. Journals in specific disciplines might identify their discipline in a way that would allow the open index to restrict searches to items that are self-identified as within a discipline. Journals might also require and monitor author use of controlled subject headings. For example, the American Economic Association publishes headings for use in its index that are widely used by journals and working paper services in economics and could be used by open journals. With some effort, the Dewey Decimal system might be adapted. Again, this enhancement is only practical if it is achieved with a high level of automation and offers very easy use by authors.

Open journals that support even a modest number of metadata fields would provide better access to their content in the open index. Subject tags might also form the

basis of an email service that would email notice of new essays to registered readers by subject area. Reliable metadata might allow articles in open journals to be more readily found in an open index that includes many disciplines and preprints and other documents as well as quality-assured works.

### **Evaluation of Scholarship**

Active scholars gain reputation and thereby earn higher salaries and promotions as a consequence of the success of their publication programs. The process of selecting, developing, and celebrating works through publication plays an important role in the lives of active scholars. As open journals replace subscription journals, the information available for evaluating an individual scholar's intellectual contributions will change. Citation rates as reported in Scholar will gain importance because they are universally available and because the lags in the appearance of citations will be shorter than those in conventional citation audits. As open scholarship generates higher rates of citations, scholars will be drawn to post materials in open archives and to publish in open journals. Of course, the thoughtful evaluation of scholars who work within the scholar's individual domain will remain important. Word and citation counts may inform the process but won't replace reasoned judgment. The judgments of journal reviewers and editors will continue to play a role. The possibility of blind reviews, however, disappears when all working papers are posted in the open. The reviewer may remain anonymous, but not the author.

### **Preservation**

Open-journals might also differentiate themselves by providing reliable, durable, permanent access to their content. The electronic medium is inherently volatile, making the issue of preservation problematic. The LOCKSS (Lots of Copies Keeps Stuff Safe (2004b)) is one approach. ArXiv's mirror sites enhance reliability and survivability. More technical approaches from JSTOR and others are likely. (Kevin M. Guthrie, 2001) Deployment of an effective preservation strategy may help attract authors. The finance of preservation will likely fall on university hosts and will await a low-cost, highly automated system. The preservation of journals with their relatively simple digital

formats (compared to video and other file formats) may be affordable and a service authors will find of value in an open journal.

## **Speed**

Although subscription journals that are distributed digitally can publish materials very quickly, an open journal might experience some advantage in speed of publication. With well-established standard protocols for electronic submissions, the open-access editor can dispatch candidate essays to reviewers instantly. Perhaps editors will devise better methods to encourage speedy turn-around from reviewers. As noted above, once an open journal has achieved a reputation for publishing frequently cited works, it may have enough appeal to authors to allow a submission fee from authors to provide payment to reviewers that would encourage faster review.

We might expect that open journals would publish more frequent, smaller issues, say five essays per month rather than 15 per quarter. A smaller scale, more focused journal would require less effort from its editor than editing a much larger journal and allow responsibility for a journal to fit more easily into an scholars life. Many prominent scholars might find that editing a small open journal in their specialty is an intellectually rewarding way to provide intellectual leadership on a par with participating in workshops and conferences. More frequent publication would allow new works to appear more quickly. PLoS Biology averaged 4.9 articles per month for the first eleven months of 2004 with a range for two to ten. (Public Library of Science, 2004) Once accepted, an essay can be published rapidly and even be posted individually when ready. It would be interesting to compare the time from submission to publication across journals and particularly across disciplines. Open journals may have better prospects in disciplines where existing journals have long publication lags. Given the ease of entry and low economies of scale in producing a journal, there may be many more open journals. Each may serve a smaller intellectual community than conventional print journals have served. Smaller intellectual communities may create peer-pressure for speedier reviews.

## **The Scope of Open-Access Journals**

To understand how open-access journals may evolve, it will be useful to characterize the disciplines of new titles and judge whether open access is more likely in some than others. New open-access journals may be more likely in disciplines with large and effective open-access archives. The disciplines in arXiv are one set. Medicine might plausibly be a second with PubMedCentral as e-archive and free access to Medline as front-end. Journals can be indexed in Medline, in the fully integrated journal index, in Google Scholar, and elsewhere. PLoS Biology did not appear to be indexed in Google Scholar in November 2004.<sup>2</sup> It might also be of interest to observe whether entry of new journals is more likely in disciplines served by many journals from commercial presses that charge high prices as compared to disciplines without such price pressure. Ordinarily, one would expect that high prices and high profits would be more likely to attract new entrants. Patterns of entry of new open access journals might be the goal of a study. It will be of particular interest to learn whether open access journals are more likely to appear in disciplines where journal prices are high.

Law reviews may also be a promising arena for entry of open journals. Law reviews are usually heavily subsidized by their law schools and seek to reach an audience that extends well beyond academia. Open access law review journals might have low cost and allow law reviews to become more widely read. Use of Medline increased sevenfold when it became freely available on the Internet. Perhaps law review articles would see a similar jump in use.

## **Time to Maturity**

Given a credible plan to launch an open-access journal, how much time will be needed for the journal to gain stature? A new journal typically will attract submissions from its editorial board and their close associates. As it gains reputation, it will attract manuscripts from a wider group. If it is to become self-financing, at some point the revenue streams must be sufficient to cover continuing costs. Call this the breakeven

---

<sup>2</sup> Title searches in Google Scholar for several articles in PLoS Biology got zero hits on December 3, 2004.

level of operation. Time to maturity might then be thought of as the number of years from the date the first issue appears until the journal is financially self-sufficient. The PLoS business plan expects financial self-sufficiency within five years.(Scott Allen, 2004)

In the open-access arena, maturity might mean financial self-sufficiency if revenue streams are to cover significant costs including particularly payments to editors. However, if editing is voluntary and the incremental operating costs are minimal or if universities pay for editorial talent directly, we might want to define maturity in terms of intellectual impact. Citation rates are the most likely guides. It should be possible to track citation rates by age of journal over time.

Open indices like Google Scholar mean that the content of open journals could be indexed overnight, including the full text. The links to references in the “cited by” files will also be updated instantly. In this fashion, some part of the delay in the accumulation of citations may decline. Some of the items in the reference lists used in the searches reported above were visible in Scholar but not in ECONlit because they were published within the last two years or were referenced as working papers. The mean time from publication until first citation (call it the citation lag) may drop by a year or two.

If financial viability is important, then the number of years to reach self-sufficiency will define the amount of start-up capital needed to give the journal a permanent place. If financial costs are not at issue, viability may simply reflect the editors and their university hosts’ sense that the journal is an intellectual success. By either criterion, there will be a period of time from launch to achieving financial and intellectual viability. With open access, that time is likely to be shorter than with conventional journals and conventional indexing.

### **The Scholarly Society**

Open journals will challenge scholarly societies to rethink how they sustain themselves. Scholarly societies have played a central role in publishing subscription journals. Membership and library subscription fees are significant sources of revenue. Although some societies offer only meetings, many publish the proceedings of their

conferences and many publish more than one journal. A few societies also produce indices to the literature in their disciplines. Societies often may use a surplus of revenue from one activity to subsidize other activities. For example, a surplus from established journals might underwrite the expense of launching new journals or support a lobbying effort.

The growth of open scholarship threatens the revenue streams that sustain scholarly societies and other publishers. Open access to pre and post prints reduces the demand for secondary journals by individuals. As open journals gain stature, demand for subscription journals may decline. As open indices gain scope and sophistication, the demand for narrow proprietary indices will fall. Lower demand means lower gross revenue, reduced possibilities for cross subsidy, and pressure to find new revenues.

A society might respond by embracing the technical change to promote open scholarship. A society may license its name and provide a certification service to universities who host open journals in their disciplines. Certification would involve approval of peer-review policies and consideration of intellectual impact. A society may come to certify dozens of open journals within its intellectual discipline. An author may first post his or her working paper to an open archive. Once he or she has vetted his or her work through seminars and conference presentations, he or she may submit the essay to an open journal. Open journals may be frontline journals, where works appear once they satisfy reviewers and editors of their originality and importance. Editors of the frontline journals may then recommend the best essays published in their journals to the editors of the society's celebration journals. The editors of the celebration journals may provide a second level of review and select a limited number of essays for publication for a wide audience. They may impose a significant fee on authors and provide copy-editing, rewrite, and design services that turn frontline work into the flagship presentations that draw wide and enduring attention. The celebration journals may continue to be published in print that is available by subscription and be offered to members of the society as a part of membership. Celebration journals may carry more advertising than is common now. The effort required to select essays for publication in celebration should be much less than is common now because celebration journals might give preference to the

publication of works published in a journal the society licenses or works its editors commission, as with review essays and essays designed for a wider audience. The copyright of all works published in the open are likely to remain with the authors who assign non-exclusive rights to publish to the open archive, the open frontline journal, and to the celebration journal.

Alternatively, a society might seek more revenues from its proprietary products while pruning products that yield insufficient revenue. The pricing policies of commercial publishers reveal that societies could earn more revenue by raising library subscription rates. Offering current full-text electronic access exclusively through an integrator who accumulates a significant proportion of all journals in a discipline would give significant market power to the integrator and increase the society's revenue. The society would block publication of material that is available in the open and thereby limit the growth of open archives. Membership rates and fees for conferences would increase. The open index would then point to proprietary websites that would be available at premium rates.

Either of these paths leads to a smaller role for scholarly societies. The latter involves higher prices and the cost of the transactions needed to sustain subscriptions and access fees along with greater limits for access to scholarship. The open path will have lower cost and significantly greater access to scholarship. The open path, however, involves lower revenues and greater financial risks and may be less likely to be chosen. For open scholarship to succeed, universities will need to be successful in launching and sustaining effective open journals. The growth of open journals will limit the possibility of significantly higher prices and other restrictive practices by scholarly societies.

## **Implications**

Nearly everyone involved in the scholarly process will see new possibilities and significant changes.

1. Authors will see an advantage to making their preprints and post prints available in open archives that are readily searchable by Google Scholar. The advantages of

- ready access by full-text search will allow works to be seen and used more widely, more quickly, and with all likelihood, greater rates of citation. Materials not found in Scholar will be less convenient and less likely to be used.
2. Leading universities will want to provide the infrastructure to support open journals. By promoting the introduction of open journals, universities can promote research at modest cost.
  3. Universities will want to reward the intellectual work their faculty provide in editing open journals. Endowments for journal editors might provide quite visible naming opportunities. Universities may want to discount service in editing journals that are not open, reflecting the likely decline in influence of such journals.
  4. The magnitude of effort is considerable. If each of 100 research universities in the US launched one journal in each of 12 departments, programs, and research centers on average, with 40 articles being published per year in each open journal, the set would publish 48,000 articles per year. (This volume would match the current flow of submissions to arXiv.) An activity level ten times higher at US research universities that was matched at other US institutions and abroad would likely allow open journals to dominate scholarship in most disciplines.
  5. Open journals will seek to be accessible with full-text searches in Google Scholar. Ready access through Google Scholar will allow essays published in open journals to be more conveniently available than essays published conventionally. Google invites email queries from authors and publishers at [scholar-publisher@google.com](mailto:scholar-publisher@google.com). Experimentation with automated methods of producing more reliable metadata may enhance an open journal's success in attracting readers and authors.
  6. Publishers of traditional subscription indices will want to rethink their business plans. Google Scholar may dominate many of the services they provide.

7. Scholarly societies will see a change in role. Many will attempt to raise prices on proprietary products, prune underperformers, and seek new revenue streams. Some may support open journals. They may continue to produce premiere celebration journals on a subscription basis. Scholarly societies are unlikely to be the driving force in creating open scholarship.

## **Summary**

The open-archive frontier continues to grow at a modest pace in building archives in disciplines beyond arXiv's set. Some funding authorities are advancing the idea of requiring scholars to post their work in open-archives but the consequences are unclear and the government in the United Kingdom rejected such a proposal. The NIH proposal has generated intense lobbying by publishers as a counter. There is the threat that the finance of journals from scholarly societies and other publishers will be undercut without the creation of effective substitutes.

The open-journal frontier is moving rapidly to create new titles probably most often with volunteer editors. The content of these open-journals should be indexed nearly instantaneously in Google Scholar and equivalent open indices. The challenge for open-journals is to achieve a threshold level of quality so that the journals are effective in developing and celebrating our best ideas. The built-in "cited by" counts in Scholar will be used as one measure of value in judging the impact of open journals. Universities that want to expand their intellectual influence may choose to support faculty who edit journals by offering monetary inducements to editors. Editing may become an integral part of faculty life.

Among the thousands of open-access scholarly journals that are likely to be launched in the next decade, we can hope that a significant number will achieve intellectual distinction. Universities might value and reward the important role that editors play. Editors will create budgets for copy-editing and design that allow them to sustain readable, quality open-access products. A highly automated, fully integrated index that includes the content of open-journals will enhance the value of the new journals. The Google Scholar service may play a significant role in causing readers to

turn to open journals and, therefore, draw authors to the open journal format. A commitment to careful review with appropriate metadata, some commitment to preservation, and timely publication decisions will give open journals an additional edge.

The success of open journals will come from the greater rate of citation for materials published in quality journals in the open and their greater level of accessibility in open indices. Scholars who publish in the open are then likely to see their work cited more frequently and faster and therefore to build stronger reputations.

The success of open journals should enhance the prospects for the growth of open archives. When the final products are open, an important reason for avoiding open-archives for works in progress evaporates. Open archives and open journals are complementary. In fact, open archives and the shared systems that support open journals with the associated indices and preservation seem likely to become part of the standard infrastructure of universities. As journal editing becomes a part of the intellectual activity of faculty, the open journal may a vital role in developing and promoting our best ideas.

### References

- "Ameriquests," R. F. Barsky, Vera Kutzinski, Nashville, TN: Vanderbilt University, 2004a. <http://ejournals.library.vanderbilt.edu/ameriquests/index.php>
- "Lots of Copies Keeps Stuff Safe LOCKSS," Palo Alto, CA: Stanford University, 2004b. <http://lockss.stanford.edu/>
- Allen, Scott. "An Upstart Medical Journal Debuts," The Boston Globe. Boston, MA, 2004, np. [http://www.boston.com/news/science/articles/2004/10/26/boston\\_hospital\\_ends\\_missed\\_surgeries/](http://www.boston.com/news/science/articles/2004/10/26/boston_hospital_ends_missed_surgeries/)
- arXiv. "Arxiv.Org E-Print Archive," Ithaca, NY, 2004, Online electronic archive to scholarly essays in physics and more. <http://arxiv.org/>
- Bernanke, Ben S. "Report of the Editor, American Economic Review." American Economic Review, 2004, 94(2), pp. 501-03.
- Brody, Tim, et al. "The Effect of Open Access on Citation Impact," 2004. <http://www.ecs.soton.ac.uk/~harnad/Temp/OA-TAadvantage.pdf>

- Delaney, Kevin J. and Trachtenberg, Jeffrey A. "Google Goes to College, Search Engine Plans to Scan Books in University Libraries, Making Them Available on the Web," Wall Street Journal. New York City, 2004, D1.
- DSpace. "DSpaceinstances," Cambridge, MA: MIT Libraries, 2005.  
<http://wiki.dspace.org/DSpaceInstances>
- EPrints. "Eprints.Org - Self-Archiving and Open Access (OA) Eprint Archives," Southhampton, UK: University of Southhampton, 2005. <http://www.eprints.org/>
- Google. "Google Scholar," Mountain View, CA: Google, Inc., 2004. scholar.google.com
- Guthrie, Kevin M. "Archiving in the Digital Age: There's a Will, but Is There a Way?" E. R. N. D. p. 56-65, EDUCAUSE, 2001.  
<http://www.educause.edu/ir/library/pdf/erm0164.pdf>
- Harnad, Stevan, Tim Brody. "Comparing the Impact of Open Access (OA) Vs. Non-OA Articles in the Same Journals," D-Lib Magazine. 2004.  
<http://www.dlib.org/dlib/june04/harnad/06harnad.html>
- Hickerson, Thomas H. "Project Euclid and the Arxiv: Complimentary and Contrasting Elements for Sustainability," Ithaca, NY: Cornell University Library, 2004, Presentation at Workshop on Sustainable Models for University-Based Scholarly Publishing at Columbia University, June 1, 2004. <http://hdl.handle.net/1813/266>
- King, Donald W. "Remarks," C. Koltzenburg, Cologne Summit on Open Access Publishing. Cologne, Germany: Open Scientific Publishing Forum, 2004.  
<http://www.openspf.de/space/start>
- Lawrence, Steve. "Free Online Availability Substantially Increases a Paper's Impact," nature webdebates. nature, 2004. <http://www.nature.com/nature/debates/e-access/Articles/lawrence.html>
- Lund University Libraries. "Directory of Open Access Journals," Lund, Sweden: Lund University Libraries, 2004. <http://www.doaj.org/>
- Markoff, John and Wyatt, Edward. "Google Is Adding Major Libraries to Its Database," New York Times. New York, 2004.  
<http://www.nytimes.com/2004/12/14/technology/14google.html?pagewanted=1&th&oref=login>

- Morris, Sally. "Remarks," C. Klotzenberg, Cologne Summit on Open Access Publishing. Cologne, Germany, 2004. <http://www.openspf.de/space/start>
- Nature Publishing Group. "License to Publish," Basingstoke, Hampshire UK: Macmillan Publishers Limited, 2005. [http://npg.nature.com/pdf/05\\_news.pdf](http://npg.nature.com/pdf/05_news.pdf)
- ProQuest Digital Commons. "Build Your Custom Institutional Repository the Quick and Easy Way," Ann Arbor, MI: ProQuest Information and Learning, 2005. <http://www.il.proquest.com/umi/digitalcommons/>
- Public Library of Science. San Francisco, CA: Public Library of Science, 2004. <http://www.plos.org/journals/>
- Schleifer, Andrei. "Report of the Editor, Journal of Economic Perspectives." American Economic Review, 2004, 94(2), pp. 518-19.
- Siegfried, John J. "Report of the Treasurer." American Economic Review, 2004, 94(2), pp. 495-8.
- Swan, Alma, et al. "Delivery, Management and Access Model for E-Prints and Open Access Journals within Further and Higher Education," The Joint Information Systems Committee, Loughborough, Leicestershire LE11 3TU: EPIC (The Electronic Publishing Innovation Centre), 2004. [http://www.jisc.ac.uk/index.cfm?name=funding\\_eprints](http://www.jisc.ac.uk/index.cfm?name=funding_eprints)
- Twyman, Nick. "Launching PLoS Biology - Six Months in the Open." Serials, 2004, 17(2), pp. 127-31.
- U. S. House of Representatives. "Conference Report to Accompany H.R. 4818 - Consolidated Appropriations Act, 2005," C. o. Rules, U.S. Congress, 2004, 104-5. <http://www.house.gov/rules/stmgrh4818ttext1.htm>
- Van de Sompel, Herbert; Payette, Sandy; Erickson, John; Lagoze, Carl and Warner, Simeon. "Rethinking Scholarly Communication, Building the System That Scholars Deserve." D-Lib Magazine, 2004, 10(9). <http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html>.
- Weiss, Rick. "NIH Proposes Free Access for Public to Research Data," Washington Post. Washington, DC, 2004, A21.

Willinsky, John. "Public Knowledge Project Open Journal System," Vancouver, BC: University of British Columbia, 2004, open source software for journal management. <http://www.pkp.ubc.ca/>

Wysocki Jr., Bernard. "Biomedical Publishers Plan More Free Articles, Industry Aims to Head Off Government Effort to Make Materials More Accessible," Wall Street Journal. New York, 2004, D6.

Young, Jeffrey R. "Google Tests Search Engine for Colleges' Scholarly Material." The Chronicle of Higher Education, 2004a, 50(33), pp. A36.

\_\_\_\_\_. "Google Unveils a Search Engine Focused on Scholarly Materials." The Chronicle of Higher Education, 2004b, 51(15), pp. A34.

\_\_\_\_\_. "Libraries Aim to Widen Google's Eyes: Search Engines Want to Make Scholarly Work More Visible on the Web." The Chronicle of Higher Education, 2004c, 50(37), pp. A1.