

Open Scholarship and Research Universities

Malcolm Getz

Vanderbilt University

Nashville TN 37235-1819

Malcolm.Getz@Vanderbilt.edu

Presented at Middlebury College, April 30, 2005

May 12, 2005

I want to thank the participants in the Middlebury College conference for comments that helped improve this essay. I also appreciate the continuing support of The Cornell University Project on *Creating an Open Access Paradigm for Scholarly Publishing*.

Research universities are intellectual powerhouses whose primary reputations derive from the published scholarship of their faculty. Intellectual reputations are central to recruiting faculty and students, generating grant funds for research, and attracting philanthropy. Legislators and governors value the intellectual reputations of their public universities, particularly when research is linked to prominent local industries as with agriculture, computing, and biotechnology. Over the last 50 years, competition among all institutions of higher education has become more intense, inducing institutions to put increasing emphasis on attracting and sustaining faculty who publish influential, original ideas.

Commercial publishers and professional societies own and manage many of the academic journals that publish the articles that build faculty reputations. Most publishers rely on revenue from subscriptions as their primary source of finance. The publications, however, bear increasingly high prices, sometimes exceeding \$1,000 per year per title.¹ Many publishers have developed journals with sufficient reputations for quality so as to be able to push subscription prices far beyond the cost of producing the journals. Indeed, some publishers package dozens or even hundreds of journal titles in an essentially all-or-nothing bundle that forces universities to pay quite high prices or forego access to significant elements of scholarship, even scholarship produced by the university's own faculty. Many academic journals are marketed almost exclusively to academic libraries so that the institutions' payments are the dominant source of finance for the journals. The number of journals also increased over the decades with journals splitting or new journals developing in new areas of scholarship. The growth in both the quantity and price of journals is well recognized.(Parks 2002; Colander and Plum 2004)

The average library system at a research university in the US spent \$5.4 million for serials in 2002, buying more than 10,000 of the world's 21,000 academic journals plus a variety of other serials.(Getz 2005) In addition, it spent \$800,000 in processing its serials, that is, in selecting, ordering, receiving, binding, and paying for them. It also spent about \$300,000 to build and maintain the shelf space to store the added volumes of printed journals. The total cost of serials for a large library then is about \$6.5 million per year. Academic journals probably account for about \$3.7 million of the purchases and \$500,000 of the processing and storage, a total of \$4.2 million, roughly \$420 (=\$4.2M/10,000) per academic journal title per library per year. If journals average 40 articles per annual volume, the cost per library per article is about \$10.54 (counting both publisher and library expenditures). Of course, this average encompasses a considerable range from less than \$1 to over \$25 per article per library. If, on average 500 libraries subscribe to each academic journal, the cost to the whole system of publication may average more than \$5,000 per article overall including both publisher and library costs. Table 1 gives a breakdown of costs as discussed below.

An important question for the universities, and indeed scholars everywhere, is whether the Internet and associated digital tools will transform scholarly publication both to enhance the intellectual reputations of the universities and to reduce the cost of supporting publications.² More deeply, improvements in publication ought to allow scholars to be more productive, that is, generate more new ideas, and to have more influence with

works that are more widely consulted.³ This essay explores how digital publication might move publishing toward these goals.

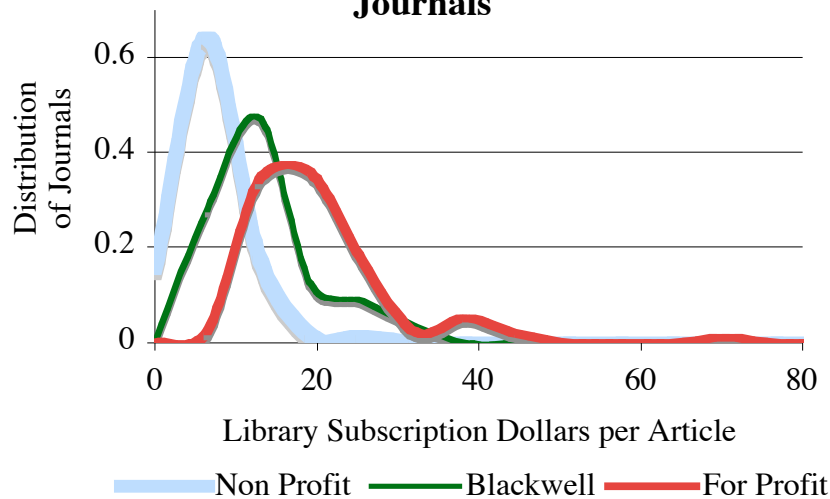
Table 1 Average Cost of Publishing Per Article Per Library

Distribution	\$0.80	Reduce with digital
Accounting	0.60	Reduce with open access
First Copy	4.60	No change
Profit	3.23	Reduce with reorganization
Total Publisher	\$9.23	
Processing	0.93	Reduce with open access
Storage	0.38	Reduce with digital
Total Library	\$1.31	
Grand Total	\$10.54	

Source: Based on Getz (2005) Assumes the average US academic research library subscribes to 10,000 journals each with an average of 40 articles per year.

Journal prices vary widely, particularly by type of publisher. Theodore Bergstrom's information on the price of economics journals provides a convenient way to show the dispersion. Figure 1 summarizes the library subscription rate per article in economics in 2004; assuming each article has 16 pages. The relative frequency curve on the left shows the distribution of library subscription rates for 78 economics journals produced by non-profit publishers with a mean of \$4.06 per article. The curve on the right shows the distribution of library subscription rates for 176 economics journals published by commercial publishers other than Blackwell, with a mean of \$17.15. (It also omits two journals from the Emerald group with prices over \$200 per article.) The middle curve shows the distribution for 57 journals from Blackwell with a mean of \$10.39. In addition to publishing some of its own journals, Blackwell has various contractual relationships with professional societies that may set prices somewhat independently of the publisher. Blackwell is an intermediate case and important enough to be viewed separately. The weighted average of the library rates for economics journals per article is \$9.71, reasonably close to the \$9.23 for all disciplines given in table 1.

Figure 1: Price per Article for Economics Journals



Source: Authors calculation from data at Theodore C. Bergstrom,
<http://www.econ.ucsb.edu/%7Etedb/Journals/pricing.html>

A move to digital publication could induce a change in how publishing is organized and financed. One notion is to move from subscription fees to open scholarship wherein access to all readers on the Internet is without charge. The cost of publication would be borne directly by universities and through author fees. Academic journals, like a number of broadcast services, for example, television and Internet search services, would be offered without charge to users.

Several questions arise in thinking about a move to digital scholarly publication. One question is the extent to which digital publication might lower costs. A second question is whether digital publication will improve the intellectual reputations of the universities. A third question is how a shift to digital distribution might reorganize the scholarly publishing industry so as to lower prices.

Digital Tools and the Cost of Publication

The shift from print to digital is well underway. Nearly 12,000 academic journals are available digitally, although most continue to be available in print.(Getz 2005) Digital distribution saves the expense of printing ink on paper, the cost of postage, and also the cost of storage in libraries. Digital distribution adds some expense in managing digital storage and retrieval systems, however, the expense to support digital distribution need be incurred only once for the Internet rather than being replicated in each library.

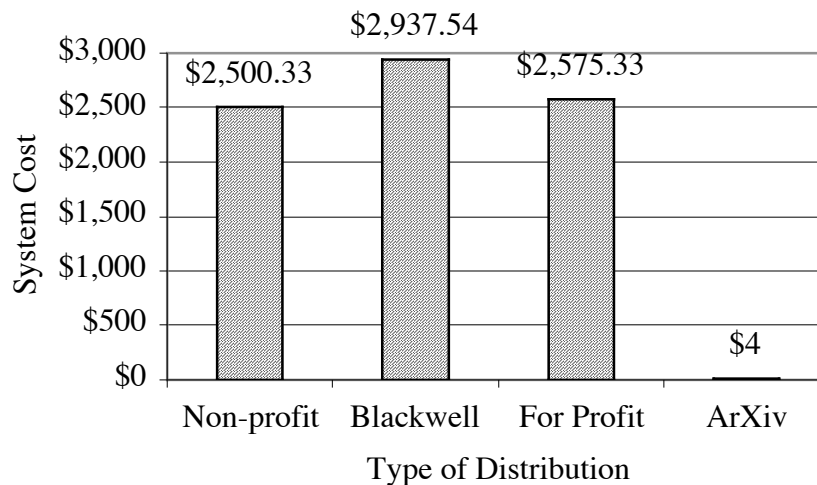
A shift to digital-only distribution would save about \$300,000 in paper, print, and postage and another \$100,000 in library storage costs for each library. Assuming 40 essays in each of 10,000 journals, the savings would be about \$0.80 per article per library.⁴

The remainder of the cost of publication is primarily the first copy cost (\$4.60 per article per library), that is, the cost of selecting, editing, and design that turns manuscripts into polished written presentations. The publishers also incur costs in billing and accounting for subscription revenues (about \$0.60 per article per library) and they earn a profit (about \$3.23 per article per library).

Open Archives

Open archives are a significant initiative to take advantage of the Internet as a low cost method for distributing scholarly works. The arXiv service that began at Los Alamos in 1991 invites authors to post essays without charge in physics, math, and computer science in a standard format. Cornell University now hosts the service, providing about \$200,000 per year in direct support.(Hickerson 2004) ArXiv now receives new postings at an annual rate of about 50,000, a flow equal to about 1,250 journals.(arXiv 2004)⁵ The full cost of the service is about \$4 per article with access to all via the Internet without charge. ArXiv often experiences many more than 12,000 connections per peak hour at its home site. Seventeen sites around the world mirror the original. (The cost of mirror sites is unrecorded.) In effect, the \$4 per article cost of arXiv compares to the \$10,000 per article cost of conventional print subscription journals deposited in 1,000 libraries, except that the arXiv service is fully automated and requires neither selection nor editing.

Figure 2: Cost with Library Distribution



Source: Same as Figure 1. Uses the following count of libraries from OCLC records (then inflated by 2.43 to account for the libraries who do not post their serials holdings on OCLC): 191 libraries noted on OCLC for the average non-profit journal, 103 for the average Blackwell journal, 57 for the average for profit journal, and none for the archive. Only about 41 percent of libraries post their serials holdings on OCLC, estimated from 1,841 reported holdings of the American Economic Review on OCLC to the 4,482 library subscribers reported in the May, 2004 AER.

Figure 2 illustrates the differences in costs by multiplying the cost per article times the number of libraries who subscribe to each category of journal, on average, as

shown in OCLC files. The OCLC record of library holdings is incomplete and the chart understates the differences. The important point is that cost of library subscriptions appears to be several orders of magnitude more expensive (with editing) than the once-for-the-whole-Internet arXiv-style service (without editing). Note that the quality of access is best with the Internet Archive and worst for the for-profit publisher because fewer libraries subscribe at their higher rates.

Authors may post works in progress as well as post-prints, that is, digital versions of essays that have been published elsewhere. A number of groups promote universal posting of works by authors.(Harnad 2005) In February 2005, some 55 institutions in many countries agreed to require all faculty to post their latest research in open archives. Open posting is to become part of the system for monitoring the research performance of faculty. The National Institute of Health in the US adopted a rule that expects all works funded by NIH grants be posted to an open archive, PubMed Central, within one year of original publication.(National Institutes of Health 2005) The British government considered requiring general use of digital archives among universities in the United Kingdom but rejected the idea, presumably because of opposition by publishers. (House of Commons 2004)

Institutional repositories offer services related to open archives. Recently, more than 150 universities around the world have installed software to support decentralized open digital archives on their campuses. Faculty may submit digital documents of a wide variety including preprints. A modest global system harvests meta-data (author names, titles, subject headings, even full-text) from all of the campus archives and allows the repository content from potentially hundreds of campuses to be searched as though they were a single entity.(Carnevale 2004; Foster 2004) The decentralized archive system allows each campus to bear most of the cost of archiving its own scholarship while creating integrated, global open access. The campus-based digital archive provides a platform from which a variety of digital services may launch. A large file of biological images, geographic data on climate, digital video files of subjects from psychology, and more might be readily managed via the digital archive and made available in an organized fashion to the Internet. For example, Cornell University launched its Internet-First University Press in 2004.(Cornell University 2005) The institutional repository system may support open journals as discussed below.

The cost of distributing scholarly works via the Internet is quite low, including even full-text searching of an integrated scholarly database of substantial size. Distribution, however, does not include the selection and editing of raw works needed to turn them in to the polished gems that build reputations.

Open Journals

Open journals add the evaluation and development of scholarly works that help them add credibility and reach a wider audience. Open journals can use the tools of open archives to manage the flow of raw materials into the editorial process and to distribute the final product. The content of open journals is then articles that carry the imprimatur of the editors of the journal and justify more attention from readers than works in an open archive. The content of open journals are available free to all on the Internet.

The Directory of Open Access Journals (Lund University Libraries 2004) listed 1,442 open journal titles in February, 2005, 358 of which are searchable at the article level. (Some number may not yet have produced a first issue; the list may not be complete.) The number of open journals is more than sufficient to demonstrate the concept.

The Public Library of Science used philanthropic support to launch open journals in biology and medicine. Its blue ribbon editorial board assures high intellectual standards and its expenditure on copy-editing and design create appealing articles. The Public Library of Science expects author fees and institutional support to allow it to become financially self-sustaining within a few years.(Twyman 2004)

The cost of launching an open journal may be minimal. A campus with a standard open archive may install software to manage the flow of materials through an editorial process. The Open Journal System software (Willinsky 2004) is in the public domain and requires relatively straightforward local implementation.⁶ The editorial software maintains a database of articles in process, keeps track of each in the editorial flow, and sends timely queries to nudge the process. The upshot is that any group of faculty members who want to edit a journal would face little expense in managing the editorial process or distributing the final product.

For example, Vanderbilt launched an open journal, AmeriQuests (2004), as part of its Center for the Study of the Americas. The University Library supports an open archive system and installed the Open Journal System software to support the creation of journals. The new journal has no business plan because its costs are embedded in the library's infrastructure and the editors serve voluntarily. The incremental cost of launching a new journal with voluntary editors is whatever the venture might spend in promoting its journal and little else. In short, volunteer editors can launch new open journals on most campuses that support open archives. The barriers to entry are minimal.

In the print world, the proliferation of journal titles adds cost. Libraries who once sought comprehensive collections can no longer afford them. In an open journal world, every library, indeed every researcher and student, will have access to all the journals. If we view journals only as the output of intellectual work, we give emphasis to judging quality. In this view, there can easily be too many journals publishing works of limited value. Proliferating journals might look like junk mail. Journals, however, are also inputs to intellectual work and provide a forum for active scholars to exchange ideas much as conferences, workshops, and open archives do. Indeed, before published journals appeared, scholars like Benjamin Franklin simply exchanged letters with others they thought would be interested in the topics they were contemplating. Early journals were accumulations of such letters. That tradition will continue in open access journals with voluntary editors who emphasize the exchange of ideas among active scholars. Different kinds of journals play different roles, as discussed below, with some putting more effort than others into careful selection and development of works. Smaller journals with volunteer editors might be thought of as development journals where a peer review process provides a validity check and some measure of the degree of interest. In the limit, a publisher might make a judgment about whether a small journal amounts to junk mail.

The usefulness of small development journals with volunteer editors seems to contrast with the hard-publishing constraint advocated by (Colander and Plum 2004). Put differently, the small open journal might be half way between posting an unevaluated essay in an open archive and publication in a large journal with high-cost editing, a kin to the multi-tiered publications in Colander and Plum.

New journals require time and effective editors to develop reputations. Open journals haven't been available long enough to produce citation rates or other evidence of intellectual influence. The Public Library of Science's first journal, PLoS Biology, appeared in the October, 2003 and is publishing ten to twelve articles per month in addition to other features. Such open journals should achieve significant reputations for intellectual influence within a few years.

Many commercial and society publishers pay scholars with well-established reputations to be editors, to attract, select, and develop high quality and interesting works for publication. They may also engage copy editors and designers to turn the authors' manuscripts into polished products. The expense of the editorial process can be considerable. Although some reputable subscription journals function with volunteer editors, open journals that aim to compete against journals with paid editors are likely to have to pay for editorial services as well. The Public Library of Science intends to finance its operations with \$1,500 per article author fees and institutional support.

The American Economic Association commits \$601,000 per year to salaries for the editors (Siegfried 2004) and staff who produce the American Economic Review with about 100 articles, that is \$5,700 per article. Adding the cost of offices, supplies, and computers, the cost per article approaches \$10,000. This is considerably more than the voluntary editing and contributed office support that sustains early entrants to the open journal arena. Open journals will need to attract financial streams sufficient to support high quality editing in order to compete with quality subscription journals. Of course, the culture of volunteerism and rates of compensation offered editors differs by discipline. Open journals are likely to be successful first in disciplines where volunteer editing is a tradition.

In the aggregate, the cost of producing the first copy as distinct from the cost of distributing the journals may be half of the expenditures on journals. Inexpensive journals use volunteer editors and lower cost production methods and yet have small distributions with the result that first copy costs are a substantial share of total cost. Expensive journals produced by not-for-profits may apply higher levels of effort to editing and production but have large distribution with the result that first copy costs are also a high share of the total. Expensive for-profit journals, however, have a significant third component, namely a significant level of profit. For expensive, for-profit journals first copy costs are well below half of subscription rates.

Open journals could serve scholars without the intermediation of libraries. Open publishers, however, should include in their financial plans support for permanent archiving of their products. JSTOR and other vendors may offer such services with costs borne by publishers. Methods for archiving are evolving and will probably include mirror

sites for reliability, off-line storage for permanence, and systems for migrating digital files to new media and software as computing evolves. Ultimately, the aggregate cost should be much less than the perpetual storage of paper in 1,000 libraries. A credible archiving program should play a role in attracting authors to open publications. Without permanence equivalent to replication of paper across many libraries, open journals offer an inferior product.

Open journals might benefit from additional software to help create more systematic meta-data. For example, searching large digital files would yield more precise results if every author's name had a unique representation. Libraries use name authority files that use the full name, date and location of birth to distinguish each Jane Smith from every other Jane Smith. Suppose that an open journal offered a webpage for an author to submit her essay for publication. The webpage might offer a field for entry of the author's name that would link to an online authority file. The webpage would allow the author to indicate which Jane Smith she might be or to create a new entry with birth date and place. The open journal would then add a unique digital name code that would distinguish the author. Web search engines might then offer access to the same online file of authors to allow a search of a specific author. Similarly, an open journal might present a menu of subject headings for its discipline and invite authors to choose appropriately. The search engine might offer a choice of discipline and within a discipline, a choice of subject headings. For example, "elasticity" might yield different results in physics than economics.

Name and subject authority links may be relatively straightforward additions to a web interface for submissions to an open journal that would impose little burden on authors. Editors and reviewers would have a chance to verify the labels, at least casually. A considerably greater challenge would be to introduce meta-tags to the components of each reference. The notion would be to invite authors to use software like EndNote that tags each field in each reference with the resulting tags being embedded in the documents. Search engines might then key on the tags and yield more precise linkages.

A move from subscription journals to open access journals would reduce the publishers' accounting costs, about \$0.60 per article per library for 10,000 journals. It would also reduce the library's costs of processing, ordering, receiving, shelving, and binding journals by \$370,000 per library per year for 10,000 journals. This is a rate of about \$0.93 per article per library. (Getz 2005)

Open Indices

The third leg of open scholarship is open indices. For example, Google Scholar launched in beta version in November, 2004. Scholar allows searching full-text files of scholarly works without charge to all on the Internet. When copyright allows, it will deliver the page image or a full-text file to the reader. In other cases, it may link to a commercial site where access may be purchased. In still other cases, it links to a library that holds the item.

Going beyond journals, Google has contracted with a number of significant libraries to scan and digitize many of their books and allow the book content to be

searchable by Google Scholar. Harvard plans to have Google digitize “a substantial number of the 15 million volumes” in its libraries.(Harvard University Library 2004) Michigan intends to have Google digitize all its seven million volumes.(Delaney and Trachtenberg 2004) Stanford, Oxford, and the New York Public Library will also allow Google to scan large numbers of books.(Keller 2005) Although Google has not disclosed all of the components of the content of its Scholar service, it appears to encompass open scholarly archives of working papers, the content of many journals, and more ephemeral policy papers from government and not-for-profit agencies.

Google Scholar has a number of strengths. Unlike a standard Google search, a Scholar search is limited to scholarly materials, albeit of broad definition. It encompasses nearly all disciplines, relying on searches for words and phrases to identify items for retrieval. Scholar uses information about linkages and references among documents to establish a priority ranking and returns the highest priority items first. It provided a count of items “cited by” and a list of those citing the item. The count of links may indicate the importance of the identified work.⁷ In many cases Scholar provides ready links to ways of accessing the items found on the web. In many ways, it dominates the performance of conventional, discipline-specific indices that rely on less sophisticated search engines.

Google Scholar’s weaknesses are not catastrophic and are likely to be addressed as the service matures. A comprehensive database with an explicit description of what’s available in the database would allow a user to anticipate what may be missing. The beta version of Scholar does not document its content and although the content is large, it doesn’t seem to be complete in any particular domain. The opportunity to limit searches by discipline and to use a controlled vocabulary for searching would support more powerful searching. The ability to limit searches to works that have been published in reputable journals would allow a reader to focus on more important works. Scholars read about 200 works per year, on average.(King 2004) They will read working papers and ephemeral materials in a narrow area of their own active scholarship but limit more general reading in their discipline to works published in better journals. Google may seek collaborative relationships with libraries, scholarly societies, and publishers to address the shortcomings of Scholar.⁸

The three elements of open scholarship, namely, archive, journal, and index, provide a powerful, low cost, highly automated architecture for making scholarship freely available to all on the Internet. An open index like Google Scholar will make the success of open journals more likely. Open journals ought to be instantly covered by open indices and thereby allow its content to be found by searching the open index. The success of open journals enhance the prospects for open archives because an open journal will face no loss of revenue when its content is made available in other ways. Data generated in “cited by” counts will inform scholars and universities about the influence of individual journals, track the growth of reputation of successful editors, and indicate in subtle ways the tiers of a multi-tier publication world. The hard-publishing constraint in Colander and Plum seems less important when reputations can be traced quickly by citation information for journals, authors, and articles.

The market for indices could change radically as a consequence of the entry of Google Scholar. The Google search engine that can return hits in priority order based on number of references and linkages dominates search engines that return items in alphabetic or chronological order. In some disciplines, even the beta version of Scholar appears to have sufficient coverage to be widely useful. Scholar returns a list of “cited by” references. A similar feature appears in the ISI Citation Index, but Scholar reaches across most disciplines with a wider set of works. Scholar builds its index with automatic web crawlers that would benefit from more carefully structured tagged fields in the original documents. As Scholar matures from its beta form, it is likely to offer more advanced searching functions. Conventional discipline-based indices seem likely to lose market share rapidly. There is some risk that Google will gain market power from a near monopoly position. The service may have attributes of a natural monopoly and Google is a for-profit company that could come to exploit its position.

Citations and Access

Open access scholarship may cause scholarship to be more useful. There is some evidence that articles published in the open are more frequently cited, given quality, than those distributed by subscription. In addition, open publication may allow more readers from beyond the immediate research community who produce similar work to read and be influenced by the work. For some purposes, searching digital files makes them more valuable than the same works in a paper format. Digital files can be more readily, searched, sorted, and repurposed than paper. Open digital publication may enhance scholarship in a variety of ways.(Gass and Doyle 2005)

Recent evidence from Thomson Scientific demonstrates that open access journals are sometimes achieving significant intellectual impact within a decade of launch. McVeigh (McVeigh 2004) uses the Web of Knowledge database to compare an impact factor and an immediacy index for 239 open access journals in six areas of knowledge to 8,818 subscription journals selected for indexing by Thomson Scientific. The impact factor reflects citations with two and three years of publication and the immediacy index reflects citations in the year of publication. Of 14 journals ranked among the top ten percent in impact in their field, five are open access journals launched as open access journals in the last decade and another two moved from a subscription model to making the full current and back content open. Among 20 journals ranked in the top ten percent in immediacy, five were launched as open access in the last decade, and two have become fully open. The most influential open journals appear more often in medicine and physics. The newest quite influential open journal is Malaria Journal, launched in 2002. It would be interesting to know how age, the financial support for editing, and the institutional ties of the publisher influence the degree of influence and the time required to achieve it. Nevertheless, the evidence demonstrates that some open journals achieve high levels of intellectual influence within a few years of launch.

Other evidence on citation rates derives primarily from disciplines represented in arXiv. Citation rates for essays that appeared in the open are double or more than those that had not appeared in the open.(Kurtz 2003; Brody 2004; Harnad 2004) Reputations then would seem to build faster and higher among authors who publish in the open.

Because the quantity of material that a reader can read is limited, the rate of use of material not published in the open may be declining.

There is a tipping point in the use of scholarly materials that occurs once a substantial number of scholars come to use given tools. Libraries saw a point after the introduction of the card catalog when materials not well represented in the catalog would rarely be used even if they were of compelling interest to scholars. Once the electronic catalog became the norm, materials a library owned that were not well represented in the electronic catalog became essentially invisible even if represented in the card catalog. Electronic indices to the literature have a similar point such that materials not listed in the index are invisible and little used. It is possible that open publication will reach a similar tipping point such that readers seldom turn to materials not available in the open on the Internet. Absence of appearance in Google Scholar may come to be associated with lower rates of use. Such a point, of course, may occur at different times in different disciplines.

Reputation often builds top down. Leaders are opinion makers who have significant influence on others. A scholar who wants to build a reputation will put primary emphasis on attracting the attention of leaders. Open journals will succeed when opinion makers look to them for better quality materials.

At the same time, the intellectual enterprise depends on support from many wider constituencies. Students, legislatures, and philanthropists are drawn not just by outstanding scholarship but also by evidence that the scholarship influences the larger world. Law schools whose graduates become judges, legislators, and corporate leaders will be more valuable than a law school that simply trains legal scholars. Schools of agriculture that influence production and marketing decisions will have more sway in the local legislature than one that seeks only to impress other agricultural scientists. Even astronomy, a science with little practical value, has adherents among amateur astronomers and draws enough public fascination to be of interest to general readers.

Although individual scholars may sense little gain in intellectual reputation from making their works openly available to broader constituencies, their institutions may well see significant advantages in attracting political and philanthropic support as well as in attracting students. (Aligning individual behavior with collective interests is a common problem in many arenas.) Open publication is significantly more available beyond the walls of academia than subscription materials. For this reason, academic leaders in research universities may find significant value in promoting the development of open journals on their campuses. This motive might cause open journals to appear first in disciplines with sophisticated external constituencies like law, medicine, theology, business, art, and music. After the National Library of Medicine put its MedLine index in the open, use increased sevenfold. More than 30 percent of use is by people other than health care professionals. (National Library of Medicine 2004)

Articles published in the open are more readily used in instruction. Many campuses support course management software systems that allow instructors to post materials for students. Students may post items as well. A syllabus can readily provide

links to open articles. Subscription-based articles require copyright permissions, a process that discourages use.

Articles published in the open more readily link to associated digital content. Readers can reach materials cited by URL. Digital objects that are not facsimiles of printed essays can be linked to text. For example, numeric data sets, software algorithms, animations, sound and video files all may be referenced and reached in the open but require transactions and delay in a subscription environment. Of course, more sophisticated digital documents raise complexity and cost in providing search and permanent archives, costs that will be weighed on their own merits. Open articles themselves may enable but do not require association with other digital documents.

Open publications give prospect of becoming the norm for scholarship because they will be more widely cited and will reach a wider audience. The critical issue is who will be the entrepreneurs and how will the financial streams come into place to cause open publications to gain significant influence?

Finance, Profit, and Reorganization

Subscription journals are a significant and growing problem because many of them cost too much. Some commercial publishers earn extraordinary profit by offering good journals at high prices.(Bergstrom 2001) A rough calculation identifies about one-third of the average research library's expenditure on scholarly journal subscriptions as profit, about \$3.23 per article per library per year.(Getz 2005) We can recognize that the for-profit publishers generally offer about 8 percent more pages per journal title and their journals generally have lower circulation, perhaps one-third the circulation as noted in Figure 2.¹ Their journals also tend to have lower citation rates, about 25 percent less in economics. Taking account of these forces on prices yields an estimate for economics journals that for-profit publishers (other than Blackwell) offer library rates per page that are about 60 percent higher than the non-profit journals. (Blackwell prices are about 36 percent higher but not statistically different than the non-profits). This compares to a 54 increase in cost over publisher cost with profit compared to the cost without profit shown in table 1.² While the estimated price differential for for-profit journals is less than the simple comparison of means noted in discussing figure 1 above, the differential is substantial and inline with the magnitude shown in table 1.

¹ Author's analysis based on data concerning economics journals by Theodore Bergstrom at <http://www.econ.ucsb.edu/%7Etedb/Journals/pricing.html>.

² A regression of the log of price per page for 247 economics journals (that omits those with zero price, those with incomplete information, and two that are priced at over \$200 per page) yields the following result with adjusted r-square of 0.603 (t-ratios shown below):

$$\begin{aligned} \log(\text{price}) = & 0.657 - 0.044 \log(\text{library circulation}) - 0.490 \log(\text{pages}) \\ & (3.68) \quad (-1.64) \quad (-6.43) \\ & + 0.084 \log(\text{citations}) + 0.597 \text{ for profit binary} + 0.363 \text{ Blackwell binary.} \\ & (2.79) \quad (16.3) \quad (8.46) \end{aligned}$$

In recent years, some publishers have come to offer their suites of journals as near all-or-nothing packages.(Carnevale 2003) Reed Elsevier may ask as much as \$1.7 million from a single library for its Science Direct service with online access to 1,200 journals. An individual university may have little interest in many of the journals but a significant need for some. The all-or-nothing package affords the university little reduction in expenditure, even if it cuts many individual titles.

The commercial publishers achieve significant market power by recruiting good editors and spawning titles to fill intellectual niches. Once a journal has developed a reputation for quality, universities that support study in the domain of the journal will want to subscribe. For example, could a medical center library drop a journal in a discipline in which it offers instruction or supports research? As a consequence of the cumulated reputation, a publisher can charge a price well above the cost of production. A new journal in the same intellectual domain will be a poor substitute for an established journal. A new journal will need to become established with a strong editorial group and a record of publishing interesting, good quality articles to gain the confidence of authors in submitting essays and readers in deciding to read.

To compete with quality subscription journals, open journals need quality editing and presentation. Although voluntary efforts can launch and sustain a journal, they are unlikely to be sufficient to sustain a journal of sufficient quality to vie for author and reader attention with quality subscriptions journals. When editors are to be paid, an open journal needs entrepreneurship, start-up capital, and the prospect of a reliable revenue stream to sustain the venture. In short, it needs a publisher.

The publisher will attract the capital, recruit and sustain the editorial group, and keep the enterprise on track to meet intellectual, production, and financial goals. Established, subscription journals have publishers who recruit and sustain editors. A publisher will have experience with the production process and can manage it effectively. A publisher may generate surpluses from established titles that may be used to launch new journals. A publisher may split a journal with a growing audience into two or more new journals.

A university might choose to publish open journals in order to enhance its intellectual reputation much as Cornell supports arXiv and many universities support unique, comprehensive library collections and museums. They do so not just for the faculty and students currently on their campus, but as part of a broad mission to promote intellectual inquiry for the ages. Universities build their reputations by sustaining unique and significant tools for inquiry. Publishing open journals fits readily into the missions of great research universities.

The subscription fees universities pay for journals that are essentially only purchased by university libraries amount to a form of institutional support that the university provides the publisher. When a single publisher receives more than \$100,000 per year from a university for a suite of titles, often offered as an all-or-nothing package, the university has a deep commitment to the publisher with little voice about editorial direction, quality, quantity, or format. The university has little possibility of

differentiating itself, that is, gaining reputation compared to other universities, as a consequence of its subscriptions.

Deploying similar magnitudes of funds, a university could publish open journals. The advantage of financing open journals with institutional support is that the university can retain control and gain reputation from its unique association with its intellectual products. A few subscription journals today carry a university's name in the title, for example, the University of Chicago Journal of Accounting Research. At present many journals with a university name in the title are student publications. The university, then, could commit sustained institutional support to fund the publishing and editorial functions for its open journals. As publisher, a university might choose an editorial board that includes, but is not limited to, its own faculty. It might recruit a faculty member to fill the role of editor. Clearly, the cost of the editorial function will vary by field and will reflect a university's aspirations for the intellectual impact of its journal. More impact is likely to come from investing more in editing. The notion of universities as publishers who make decisions about financial commitments to the editorial process is related to the solution proposed by Colander and Plum. Colander and Plum propose that a university financed editorial-like review be completed on each campus as part of faculty promotion decisions. The advantage of university support for the editorial process in open publishing is that the benefit of the editorial review accrues to all scholars. Personnel decisions may then be informed by the editorial process as well as by citation information.

In considering launching a journal, a university leader would ask why should the university forego subscription income to underwrite its venture? Even the University of Chicago Journal of Accounting Research is published by Blackwell and has a \$507 per year subscription fee for print and premium online.(Blackwell 2005) A first answer for the university publisher might be that the open journal is likely to have a considerably larger intellectual impact than a subscription journal. A second answer is that the open journal will have somewhat lower total cost of production because it avoids the expense of accounting for transactions with subscribers and maintaining subscriber lists. A third is that any profit that the journal earns will accrue to the university as publisher rather than to a third party. Nevertheless, the direct outlay for a given journal will be higher to the university that provides institutional support than if it sought subscription revenue.

One could imagine a consortium of universities who each agree to support a journal in a different area of a discipline such that each supports its own title with the expectation that others will support theirs. Law journals appear to follow this pattern with heavy subsidies from the law schools. Because no money need change hands, issues of governance should be minor. Indeed, once a base level of open journals succeeds, the consortium may no longer be needed. Entry and exit might occur with little cost to others. Launching new journals on a subscription basis or converting an open journal to a subscription basis might then be unattractive because it would both be more costly, lower intellectual impact, and threaten reprisal.

Author fees might provide an additional source of revenue once journals are attracting more high quality submissions than an editorial group can comfortably

evaluate. Author fees could include submission fees as well as page charges. Author fees have something of a checkered history because the incentives facing an individual scholar may differ from those of the university or a funding agency. Nevertheless, they have their place.

Transition

The open architecture could allow reorganization. Currently, independent societies and commercial publishers own most scholarly journals. The societies earn revenue from memberships and library subscriptions. Societies may generate a surplus on publishing operations that they use to help underwrite other ventures including headquarters, lobbying, meetings, and unprofitable publications. Some societies offer dozens of publications. The commercial publishers rely primarily on library subscriptions alone with profits going to shareholders. Many offer dozens or even hundreds of publications across disciplines. Reorganization might change the cross-subsidies and profits and lower costs to universities and scholars.

Open journals are likely to co-exist with subscription journals for some time, perhaps even permanently. Journals with large numbers of personal subscribers provide a service directly to readers at modest cost per reader. The accumulation of revenue from many thousands of readers may support editorial services that universities may not find worth challenging by supporting competing open journals.

Open journals should quickly replace subscription journals that rely exclusively on voluntary editors. The advantages of the open architecture may be so compelling that existing journals might switch to the open format in order to save costs and increase intellectual impact. Such publications would be based at universities that support the infrastructure for open archives and open journals that conform to open standards. The universities would need to implement a cost-effective strategy for archiving its digital journals.

In the intermediate ground where subscriptions are primarily sold to libraries and editors are paid by commercial publishers, leaders of research universities will want to consider investing in open journals. They are likely to support as infrastructure the open-source software tools for sustaining open archives, open journals, and a system to assure permanence. They will then want to behave as publishers in launching and sustaining open journals. They will be strategic in investing in new open journals in disciplines with strong faculty on campus and where the university recognizes a significant off-campus constituency. They will seek to associate the host university prominently in the promotion of the journal. Perhaps an associate provost for research will play an entrepreneurial role in recruiting editors, overseeing budgets, and evaluating performance. The research universities, then, will shape the speed and direction of the development of open journals as they consider where to invest.

Conclusion

The Internet is changing how scholars work and how universities build their reputations. The movement of academic publications to an open architecture of archive, journal, and index is underway. The shift to open scholarship may cause scholarship to be of wider influence in our society and across the planet. Researchers are likely to be more productive and students will learn more by using open scholarship.

The cost of scholarly communication will fall modestly with a shift from paper on library shelves to digital communication via the Internet because much work need be done only once for the Internet rather than on each campus and also because open publishers need not keep track of transactions with individual readers or libraries. A shift to open publication may also bring control of journals to the universities and recapture some of the rents from scholarship for scholarly purposes.

The initiative is with university leaders. Instead of paying a few individual commercial publishers' very large bills for journals, a university might invest thoughtfully in the infrastructure and in the editors needed to launch open journals. Each of 100 research universities might invest carefully in ten new excellent quality open journals in strategically chosen disciplines. Each might also see another dozen smaller scale journals develop with voluntary editing. As a result of such ventures, the scholarly landscape would show significant results within five years. Within ten years, open journals are likely to dominate scholarly communication.

References

- (2004). AmeriQuests. R. F. Barsky, Vera Kutzinski. Nashville, TN, Vanderbilt University. 2004. <http://ejournals.library.vanderbilt.edu/ameriquests/index.php>
- arXiv (2004). arXiv.org e-Print archive. Ithaca, NY. 2004: Online electronic archive to scholarly essays in physics and more. <http://arxiv.org/>
- Bergstrom, T. C. (2001). "Free Labor for Costly Journals?" Journal of Economic Perspectives 15(4): 183-98.
- Blackwell (2005). University of Chicago Journal of Accounting Research. Oxford UK, Blackwell Publishing. University of Chicago Journal of Accounting Research,
- Brody, T., et al. (2004). The Effect of Open Access on Citation Impact. <http://www.ecs.soton.ac.uk/~harnad/Temp/OA-TAadvantage.pdf>
- Carnevale, D. (2003). "Libraries With Tight Budgets Renew Complaints About Elsevier's Online Pricing." The Chronicle of Higher Education 50(17): A33.
- Carnevale, D. (2004). "6 Institutions Will Help Fine-Tune a Popular New Archiving Program." The Chronicle of Higher Education 49(23): A36. DSpace at MIT is tested.
- Colander, D. and T. Plum (2004). "Efficiency, Journal Publishing and Scholarly Research—A Discussion Paper." Middlebury College Economics Discussion Paper 04(14). <http://ideas.repec.org/p/mdl/mdlpap/0414.html>

- Cornell University (2005). Internet-First University Press. Ithaca, NY, Cornell University Library. 2005. <http://dspace.library.cornell.edu/handle/1813/62>
- Delaney, K. J. and J. A. Trachtenberg (2004). 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: D1.
- Foster, A. L. (2004). "Papers Wanted, Online archives run by universities struggle to attract material." The Chronicle of Higher Education 50(42): A37.
- Gass, A. and H. Doyle (2005). "The Reality of Open-Access Journal Articles." The Chronicle of Higher Education 51(24): B13.
- Getz, M. (2005). "Open-source Scholarly Publishing in Economic Perspective." Journal of Library Administration forthcoming. dspace.library.cornell.edu/handle/1813/62
- Harnad, S. (2005). "The Implementation of the Berlin Declaration on Open Access." D-Lib Magazine 11(3). <http://www.dlib.org/dlib/march05/harnad/03harnad.html>
- Harnad, S., Tim Brody (2004). Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. D-Lib Magazine. 10. <http://www.dlib.org/dlib/june04/harnad/06harnad.html>
- Harvard University Library (2004). "A Unique and Compelling Opportunity: Harvard's Pilot Project with Google." Harvard University Library Letters I(7): 5.
- Hickerson, T. H. (2004). Project Euclid and the ArXiv: Complimentary and Contrasting Elements for Sustainability. Ithaca, NY, Cornell University Library. 2005. <http://hdl.handle.net/1813/266>
- House of Commons (2004). Science and Technology - Tenth Report. S. C. o. S. a. Technology, House of Commons. II (HC 399-II). <http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/39902.htm>
- Keller, M. (2005). Stanford University and Google Mass Digitization Project. Palo Alto, CA, Stanford University Library. http://library.stanford.edu/about_sulair/special_projects/google_sulair_project.html
- King, D. W. (2004). Remarks. Cologne Summit on Open Access Publishing, Cologne, Germany, Open Scientific Publishing Forum. <http://www.openspf.de/space/start>
- Kurtz, M. J. (2003). Restrictive Access Policies Cut Readership of Electronic Research Journal Articles by a Factor of Two. OpCit The Open Citation Project. Highfield, Southhampton, UK. <http://opcit.eprints.org/feb190a/kurtz.pdf>
- Lund University Libraries (2004). Directory of Open Access Journals. Lund, Sweden, Lund University Libraries. 2004. <http://www.doaj.org/>
- McVeigh, M. E. (2004). Open Access Journals in the ISI Citation Databases: Analysis of Impact Factors and Citation Patterns. Thomson Scientific White Paper Series. Philadelphia, PA, Thomson Corporation. 2005. <http://www.thomsonscientific.com/promo/openaccess/>
- National Institutes of Health (2005). Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research. Notice Number NOT-OD-05-033. Office of Extramural Research. Bethesda, MD, Department of Health and Human Services. <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-05-022.html>

- National Library of Medicine (2004). "Digitizing Back Issues of Journals." NLM Newsline 55(4). <http://www.nlm.nih.gov/pubs/nlmnews/octdec00/2000CIM.html>
- Parks, R. P. (2002). "The Faustian grip of academic publishing." Journal of Economic Methodology 9(3): 317-335.
- Siegfried, J. J. (2004). "Report of the Treasurer." American Economic Review 94(2): 495-8.
- Twyman, N. (2004). "Launching PLoS Biology - six months in the open." Serials 17(2): 127-131.
- Willinsky, J. (2004). Public Knowledge Project Open Journal System. Vancouver, BC, University of British Columbia. 2004: open source software for journal management. <http://www.pkp.ubc.ca/>

Endnotes

¹ Among 150 economics journals with subscriptions paid at Vanderbilt's Library in 2001/2002, 3 had prices over \$1,000. In addition, the library received 83 journals in economics from Elsevier as a package across many disciplines.

² Even if universities are currently at a ceiling amount of attention available for all universities, those who promote open publication may gain reputation at the expense of universities that eschew open publication. In such a world, if all participate in open publication, the rank order of reputations may remain the same but the costs of sustaining rank will be less. However, open publication may expand the total influence of academia when people outside academia have ready access to academic works.

³ Faculty who are already committing as much time to intellectual work as is humanly possible, may find that with better tools they can use the same effort to produce more or better quality work. Those who do not embrace the best tools will have lesser intellectual impact, on average. Readers who already commit a ceiling amount of effort to reading each year may find reading works produced with better tools to be more valuable and therefore read fewer works produced with lesser tools. When works are produced with better tools, the time spent reading will yield more and deeper ideas per unit of effort by the reader.

⁴ The cost savings refer to newly created digital documents not the conversion of existing print documents to digital formats.

⁵ For 2000 through 2004, 24.6 percent were from US "edu" domains, 9.9 percent from Germany, 6.3 percent from Italy, 5.8 percent from the UK, 5.7 percent from Japan. And 5.6 percent from France. ArXiv is a genuinely global service. arXiv (2004). arXiv.org e-Print archive. Ithaca, NY. 2004: Online electronic archive to scholarly essays in physics and more. <http://arxiv.org/>

⁶ Cornell University Library is developing DPubs as open source software to serve a similar function. See <http://dpubs.org/>

⁷ The count of “cited by” items is similar to a feature of the ISI Citation Indices. Such counts may inform appointment and promotion decisions for professors.

⁸ Internet search engines continue to evolve and new entrants may eclipse current leaders. Clusty.com, for example, is a meta-search engine that uses linguistic and other approaches to cluster search results by common elements rather than in a one-dimensional ranking.