

Towards a Vision of Inexpensive Scholarly Journal Publication

ANN OKERSON

Yale University Library, Yale University, New Haven, Connecticut, USA

The dream of free and easy scientific and scholarly publishing on the Internet antedates the introduction of the World Wide Web. The last ten years have seen that dream dampened but not extinguished by the growth of the online for-profit journal, but multiple voices continue to speak for new models less tied to high prices and restricted access. This paper reviews the main models now in play, to identify

their characteristic features, their strengths and their weaknesses. It reviews the criteria that will determine success and failure and concludes that management of costs and assurance of academic quality are the two critical factors. As a result of these factors, an altered, high-quality system of scholarly communications does not show promise of being truly inexpensive.

A brief history in time: not-for-profit prophecies

My mission today is to squint into the crystal ball and foretell the future. My credentials for this exercise are fairly straightforward: I've done it before. Nobody would suggest that I've done it *successfully* before, but that's a different matter.

Now, the best way to look ahead is first to look backwards. We are entering the second decade of coherent prophecy about the future of scholarly and scientific communication in an Internet environment, and so even prophets have a history. We will begin by reviewing some of that history.

Ten years ago, the beneficial transformation of scholarly communications seemed so easy and obvious. A number of individuals in academia and libraries, at least in the sciences, imagined that future publication, by utilizing electronic distribution, would happen quickly, by authors themselves and without mediation of publishers as we had known them. *Voilà!* Thereby would vanish the problems of high prices of scholarly

journals and of slow dissemination time. In fact, a handful of us believers in this vision gathered in Washington, D.C. on May 20th, 1992, to imagine what a free "e-print" initiative might look like. Such a model for non-commercial dissemination was already in hand, in the long-standing practices of scientists distributing their photocopied preprints around the world to like-minded researchers. (We were told that the mathematics department of one of our greatest universities employed a half-time clerk whose only job was to photocopy paper preprints as they arrived and to distribute them to faculty mailboxes.) And in the previous year, 1991, physicist Paul Ginsparg, then at the Los Alamos National Laboratory, had already led the way by facilitating the electronic distribution of preprints in the field of high-energy physics.

We believed – and it is important to recall that this moment of vision occurred in a room in which no one had yet heard of the World Wide Web – that it would only be a matter of time before that kind of model replicated itself to the world and the existing system of high-priced

Paper given at La Communication Scientifique en Quatre Dimensions Montreal, Quebec, June 5, 2003.

Ann Okerson is Associate University Librarian for Collections and International Programs, Yale University Library, PO Box 208240, New Haven, CT 06520-8240, USA. Homepage: <http://www.library.yale.edu/~okerson/alo.html>. E-mail: ann.okerson@yale.edu

STM journals would be hounded into irrelevance and disappear.

Some things did happen as we imagined they should. Paul Ginsparg's "e-print" system (he liked to call it "desk-bottom publishing" because the server for his system lurked under his desk at Los Alamos National Laboratory) took off quickly and has flown high and well. Ginsparg has since moved to Cornell University, a hotbed of initiative in the areas of free access to scientific information and he won a MacArthur "genius" fellowship in the bargain. Today, in a number of fields, the e-print is a norm for scientific communication.

That success is deservedly famous, but it has a dark side that is seldom noticed: in some sense, it seems as if little has changed. That is, in various disciplines the e-print system has replaced the paper preprint system, but the costly formal journal system remains in place – perhaps irrelevant to many scientists in need of the most current information, but no more and no less irrelevant than it was in the old order of photocopied preprints.

Meanwhile, in April of 1993, the first graphical Internet browser was released: Mosaic took the world by storm. Hardly more than a year later, Netscape swarmed into view and took over the burgeoning world in a matter of weeks. And several years later, Bill Gates noticed the Internet, created Internet Explorer, lost a massive antitrust suit for his anti-Netscape practices, and succeeded in seizing nearly universal domination of the browser market. There is a logic to all this, even if no one understands what it is.

But the availability of graphical browsers revolutionized the Internet. What had seemed a pipe dream when Al Gore and Bill Clinton spoke in the 1992 campaign of the "Information Superhighway" became a cliché and then a stock market bubble. In 1993, colleagues and I began the *NewJour* alerting service for new electronic journals on the Internet [1]. By 1995, there were 250 journals in this archive. By 1997, there were four times that many, Jeff Bezos had invented Amazon, and the traffic in rare Beanie Babies had begun in earnest, shifting a year or so later to the extraordinarily efficient and transparent marketplace of e-Bay. The world was changing before our eyes.

And every scientific journal publisher from here to eternity got the message pretty quickly.

For example, in October 1995, at the Frankfurt Book Fair, were found some of the more venturesome, the managers of Academic Press, who were full of ideas about making an access arrangement for all their journals in electronic form under a series of consortial licenses. For the next several years, librarians invested a huge effort in creating the business of making licensing arrangements with both for-profit and not-for-profit publishers. In 1997, the *LIBLICENSE* project [2] went live and became the place where serious students of the art of this kind of deal (both librarians and publishers) meet to identify emerging issues and announce important next steps. Six years and nearly 5,000 moderated *liblicense-l* messages later, that is still where much of the e-journal action is.

But our prophecies went wrong, as prophecies will. Distribution costs must fall, we all believed – and we argued intently in those days whether distribution amounted to more than 30% of the cost of journal publication. When the dust of 1995–1997 had cleared, we found that the old familiar faces of STM publishing, the giants, were still the giants. If anything, some publishers had swallowed others – and Academic Press, among the first movers, was among those swallowed. By 2003, it is fair to say that there is no existing scientific, technical, or medical journal of global significance that does not have a Web version available to its subscribers. And the economic model of today is simple: *not* print minus 30% (to allow for the most conservative notion of savings that e-distribution would bring) but more like print *plus* 10% or 20% or 30% – because few enough libraries could or would give up the print version of journals as e-versions became available.

The percentage of library acquisitions budgets allocated to e-versions of existing print titles shot up, at Yale, for example, multiplying by 8-fold over the last seven years. After the revolution, libraries and institutions of higher learning were paying more, not less, money to publishers whose pre-eminence the activists had thought to overthrow.

Keep the flame burning

Yet, throughout late 1990s, voices continued to speak of the possibility for more transparent and

affordable distribution of research information. Here and there, some stakeholders (scientists and librarians, in particular) continued to believe and to enact the belief that a more efficient and economical mode of distribution of scientific information was possible, desirable, and indeed inevitable.

Thus, over the last five years, a few proactive initiatives were born: a number of additional “e-print” services [3], a few cheap or for-free alternative scholarly e-journals [4], ARL’s SPARC (Scholarly Publishing and Academic Resources Coalition) [5], the PLoS (Public Library of Science) [6], the BOAI (Budapest Open Access Initiative) [7], and the latest darling of many, the institutional repository [8]. Behind the potential success of a number of these initiatives lie some important technological innovations and aspirations (the most promising being the Open Archive Initiative [9], which will foster cross-searching of open e-publications sites), building tools that would make it easier for scientists to distribute the results of their work easily and efficiently and easily.

In the past year or two, the term “Open Access” has gained in visibility and has become the watchword and vision of a growing number of scientists and librarians. Simply defined, Open Access would make scientific articles available for free (i.e., without financial barriers to any readers anywhere). To achieve that freeness, the system would be subsidized “up front” by a variety of players, such as universities or laboratories via “memberships,” governments, private foundations, other patrons, and/or advertising. No subscription fees would ever be charged to anyone for access to such works. In fact, the definition of Open Access is somewhat complex, as there seems to be some disagreement among insiders about exactly what it is and is not – more of that below.

The Open Access Revolution

There are several important features of this incipient and so-far under-fulfilled Open Access revolution.

The great person phenomenon

Today the sources of the most widely publicized initiatives are a few voices expressing vision and

a core of start-up money, for example, the Association of Research Libraries (ARL) membership, the U.S. National Institute of Health’s and now Sloan-Kettering’s Harold Varmus, and the philanthropist George Soros, although we see also some high-visibility smaller, for-free journal startups.

The article and the journal will not die

Remarkably, most of the new initiatives have also accepted the fundamental model of scientific communication that goes back centuries, i.e., the individual article on a single topic, produced in one intellectual context (for example, the research laboratory) but distributed in another (the journal). Of course, not all of the new startups reproduce the traditional journal, but all assume that the production of information is crucially separated from the dissemination of information. There are relatively few, for example, research projects distributing their own research results in a systematic manner on a prestigious website of their own making. Instead, the model of granular production and then prestigious placement remains in place.

A time of ideological and economic war

This revolution plays itself out against, very much *against*, the landscape of existing models of scientific publishing. Relations between the advocates of new models of distribution and the existing publishing community have, if anything, deteriorated over this time. Scientists are put in the position of making an ideological and economic choice in deciding where to place their articles.

The dog that didn’t bark in the night

What is not happening? Most agree, however reluctantly, that the “distributor” (to use a neutral term) or “publisher” (to use a less neutral one) has a key role in the future scientific information food chain. The most aggressive new models for article distribution, ones that imagine dis-intermediation, still envision some kind of known logo or environment, even the institutional repository – where the university or research institute begins to take on the role of the traditional

publisher in archiving and providing metadata for material published by researchers from within their ranks. So it is that in 2003, we have come to this point: there exist “old” journal publishing models and “new” models, but they differ perhaps less than one would think.

The issue that all journals share is economic – what will they cost, and who will pay the piper for the dance? Where they may differ is in the identity of those who pay and possibly the price for the payers. The traditional subscription models assume that users (or their libraries) pay for the costs of publication, supplemented by occasional page charges from creators of information, i.e., scientists (or their grants). As noted above, new models attempt to relocate the payments away from the end-users or their institutions, or else to eliminate or reduce subscription prices, under certain conditions:

Up-front author charges

For example, the Public Library of Science journals (to begin publication in the fall of 2003), expect to charge authors “up front” about \$1500 per accepted article [10], which sum should support the article through a long lifetime online and make it available “for free” to all readers.

Up-front institutional memberships (or other revenue sources)

The model of BioMedCentral’s [11] Open Access journals (not all of BMC’s publications are Open Access; a number are funded by subscriptions) is to charge \$500 to authors per accepted article; or to offer their universities or laboratories annual “institutional” memberships at a flat rate per type and size of institution, this flat rate covering all articles that those authors will submit for publication during a given year. That per-article amount intends to care for the article during its lifetime.

Universities as aggregators, i.e., institutional repositories

Still other “new” models depend on learned institutions to establish structures to make available their own authors’ works widely in the Internet, relying on technological developments such as

the OAI and other open source tools to provide the necessary search and retrieval oomph. The presumption with repositories such as *dSpace* [12] is that content will be free to all readers, though it is likely that over time there will be a wide spectrum from freeness to costliness.

Free access delayed

Some new models rely on libraries or individuals to pay regular journal subscription fees, but the journals may choose to make their content available for free under certain conditions. The leader among these is probably HighWire Press, [13] many of whose publishers have established a practice of making journal content available for free after a year or two (or even less) after initial publication.

Free access to those who cannot pay

Other publishers give away content for free to developing nations, for example through the World Health Organization’s HINARI project. [14] Still others make very inexpensive license arrangements with poorer nations. [15]

Cheaper infrastructure, cheaper subscription prices

Other “new” publications are less new in the charging model than in the amount being charged: by taking advantage of Internet-only publication management and dissemination and not needing (yet) to support a large organizational structure, they can charge a small annual subscription fee.

Non-peer-reviewed distribution

Some authors are delivering their articles for free through their own university sites or home pages, or through disciplinary e-print services. These tend, so far, to be free of charge.

What we have above is a brief but emerging listing of improved (more affordable, at least for some) access options. Not surprisingly, for the array of “newer” options, there is some disagreement about the boundaries. Where does Open Access stop? In a kind of war of ideologies, some assert that any journal article that requires a sub-

scription charge is by definition not "Open." By that pure type of definition, only e-articles that charge *no* library subscription price or membership can be called Open Access – which would exclude journals that make content available for free after a delay, journals that charge subscribers in richer countries and not in poorer ones, and journals that charge a low subscription price. Thus, one could reason that where a library or university pays an institutional membership fee (for example for BioMedCentral), that library/university's readers are not benefiting from Open Access – rather, they are supporting, through a subscription-type fee (though otherwise named) readers who cannot afford to pay. It is certainly reasonable that those who can afford to pay should help subsidize those who cannot, whether by paying institutional memberships or creating institutional repositories, but I am troubled by the fact that the costs in the proposed Open Access systems seem to be, by their advocates, insufficiently acknowledged. Where Open Access proponents accept the existence of such costs, they may not acknowledge the range of these costs, often dismissing them as very cheap or unnecessary – or to be borne later by others. It is likely that the concept of Open Access is far more complicated – and costly – than any of its advocates have so far acknowledged.

What continues to remain unclear is whether relocation of charges can effectively reduce costs and prices paid by subscribers – and whether relocation results in a fairer or more efficient system. For example, one of the values of the current subscription system that has evolved over some centuries is that it offers an economic incentive for journals of quality. By this I mean that subscribers will eventually stop paying for journals that are of low quality (even though the university and library communities have not shown themselves strong-minded in this regard). A for-free system is not discriminating in that regard, being one in which a thousand flowers can bloom, but the distinction between weeds and perennials is largely obscured.

Voice of prophecy

So perhaps now I can become a prophet. What are the key facts of the present landscape that will shape developments over the next ten years?

We can identify these trends without reducing them to a single model future, for the trends are at variance with each other, and the future will be shaped when we know which of them proves to be more powerful than the others. The following are some trajectories.

The power of rhetoric

The Open Access movement does not need to define its success in terms of number of subscribers for non-traditional journals or other sources. A canny innovator can change the behavior of the object of its attentions by staging a select number of well-targeted initiatives and by winning the battle for the hearts and minds of a relatively limited part of the population that cares about these journals. Simply by persistence and by demonstration models – that need not themselves survive for the long term – the movement is rapidly changing – and becoming part of – the discourse for publishing. This is a very good thing. No longer do we talk solely about subscription price (as we did from the late 1980s to the late 1990s); now we talk about the fundamental model of access – and that discourse reduces itself to principles on which many scholars and scientists agree, and which are disadvantageous for many existing, especially for-profit, publishers. The developing nations initiatives are so far quantitatively small (though qualitatively and morally significant), but the argument made that *some* readers deserve free and easy access regardless of ability to pay is proving remarkably effective.

The law of unintended consequences

Unintended consequences are the wild cards of history. What could happen as a result of the Open Access movement that would confuse expectations or hopes? What would success look like?

I mean to be provocative here, but consider this scenario. Suppose that existing largest publishers see that their business is in fact being impacted by the Open Access models and by agitation for a change in the system – and suppose they respond? What happens if the major publishers say, "All right, you are right: we *have* been too aggressive in creating myriad new journals in every sub-sub-discipline of science. We are going

to respond by downsizing, rightsizing, and outsourcing. We will let *you* have the commercially marginal fields for your e-print services, we will cut our commercially marginal titles, and we will concentrate on adding such persuasive value to our core collection of major titles that they will be unavoidable.”

If that happens, then suddenly the not-for-profit community will have succeeded in wresting back the unprofitable publishing outlets for itself – and be faced with the costs of paying for them. With no subsidies or economies of scale from more successful publishing ventures, the marginal scientific journal will fall into the plight of the university press of the 2000s: too small a market, too high a price, too great a need for cost-recovery from not-for-profit entities.

Quality quality quality

All the dreams of frictionless dissemination of information assume that the business of disseminating scientific and scholarly information is all about the business of, well, disseminating information: a better mousetrap will take over the mousetrap business. But it has also become clearer than ever in the last decade that a significant part of the business of which we speak has little to do with the dissemination of information per se. It does not even have to do with the scholarly validation of that information: peer review is essential to the existing system of publishing and needs to be reproduced in any new environment, no question.

But there is another related and more elusive part of the system that we have not dealt with in sufficiently direct ways. The hidden function of the system of scholarly and scientific publishing that we inherit is its ability to create, value, and preserve a unique form of capital: academic prestige. Many readers see an article published in a given journal and read it attentively and carefully to learn what it has to say. But I venture to suggest that for each such reader there are several more who observe the *fact* of publication and come to a conclusion about the author, his or her institution, and the quality of the work done there. The author looking to place an article does not simply mean to reach the people who are best qualified to *judge* the work, but to influence the widest body of opinion-shapers in and around

the field of study – and many of those will not actually read the article but only observe its existence and its prestige.

What the existing journal community has, in addition to a strong position in the actual dissemination of information, is a near stranglehold on the awarding of prestige-for-publication. The genuinely open question for the next decade, it seems to me, is this: when and how will non-traditional outlets for scientific and technical information find a way to confer upon their contributors prestige congruent with that provided by traditional journals? I believe that question is answerable, but I do not believe that it can be predicted confidently that it will be answered successfully.

Costs, costs, costs

In the end, the consuming issue is and will be costs. The topic of costs can be divided into two:

(a) Does the system of scholarly communication (especially STM) really need to cost very much at all? This is of course the question that divides the community, some of whom argue for the adequacy of a distributed literature, posted by authors on their Web sites (which are supported by their research institutions anyway) and retrieved by ever-improving search engines and manipulated by ever more sophisticated softwares. Others in the scholarly community (or one should better say “communities”) point out that the need for high-quality peer review, the functionalities increasingly demanded by readers, and the integration, consistency, and reliability that save time and distinguish utility from chaos, are and will continue for a long time to be costly. I find myself, even if at times reluctantly, in the second camp, believing that at least the STM journal literature represents a careful set of users, unlikely quickly to change their practices or values.

(b) Who will pay that piper? I place the question at this point in my survey because it is in fact consequent on the last one. Not-for-profit institutions are not in a position to dispose of significant financial resources in the name of a potential revolution. Yet, if the Open Access movement wishes to succeed, it should be putting every effort into making the circumstances for institutional support as persuasive as possible.

Are there alternatives?

Of course there are, and the most visible one is government support. Scientific, technical, and medical research benefits from substantial governmental support in many countries. What does not directly benefit is the economy of publication. One can argue that indirect costs awarded to the researcher's institution pay for some part of the library subscription budgets – but when the dollar amount of grants go up and associated indirect costs go up, no institution I know passes many of those dollars to libraries. In short, it could be argued that there exists a potential subsidy that is not making its way through the pipeline. Could governments be persuaded to contribute a greater share to the economy of scientific and scholarly communication? I am, frankly, skeptical; but this remains the greatest wildcard for our future.

Guarded optimism

I come to the end of these remarks with very guarded optimism. That optimism arises because so many individuals and groups have a common interest in the successful nurturing and maintenance of a system of scientific communication. To allow that system, built up at great cost and with great care over the last half century, to decline into even temporary chaos is unacceptable to all. A more affordable system for distributing scholarly results is a hugely important and laudable goal, one worthy of our pursuit.

We will probably muddle through this time of transition to a better world, even though there may be some disruptions. I can observe that many of today's STM publishers are fiddling while Rome burns, I can believe that we will need to conceive something very different from today's unintegrated, piecemeal journal systems with their high prices, and I can also believe that Open Access is far from free. That means I know what the problems are. What are the solutions? Or at least what are some steps we could take towards solution? These are less clear, but some of the new models for journal publishing have been outlined earlier in this talk.

The real question is not what we will call a new system, so much as how we will achieve consensus, consensus that will enable us to bring

enough players together who agree about the pursuit and work together to find ways to effect change. It seems to me that we need to do the following, for a start:

1. Continue to build awareness on the part of the leaders of research and research institutions, as well as government agencies.
2. Support new journals with good models –no matter who publishes them
3. Turn down the heat in the debates among parties to the conversations
4. Remember that there are no free or easy answers.

Notes

1. For the current archive, over 12,000 titles strong, see: <http://gort.ucsd.edu> [viewed 2 September 2003]
2. The *LIBLICENSE* project offers many features for students and practitioners of academic library content licensing, including definitions, bibliographies, live links to resources, sample licenses, an academic library "Model License," developed over two years with much input from the library and publisher community, a downloadable software that enables the user to create his or her own content license, and much more. See: <http://www.library.yale.edu/~llicense> [viewed 2 September 2003]
3. According to Tim Brody, University of Southampton, in a PowerPoint presentation called "The Research-Impact Cycle," shared with me by a colleague, the number of e-print archives is about 70; apart from the three largest e-print archives, these average about 120 papers per archive. Mr. Brody also reports the existence of 23+ institutional repositories or full-text archives, with a mean number of 225 papers. For more information, Mr. Brody's e-mail address is: t.d.brody@soton.ac.uk
4. Perhaps a good source of information about Open Access journals today is the recently launched (May 12th, 2003) Lund University Libraries' *Directory of Open Access Journals* <http://www.doaj.org> [viewed 2 September 2003], supported by the Information Program of the Open Society Institute and SPARC, The Scholarly Publishing and Academic Resources Coalition. According to the press release, this Directory contains information about 350 open access journals; the service will continue to grow as new journals are identified
5. For comprehensive information about SPARC, which recently celebrated its first energetic five

- years, see the organization's Web site at: <http://www.arl.org/sparc> [viewed 2 September 2003]
6. The Public Library of Science is deeply committed to "making the world's scientific and medical literature a public resource." For complete information about this movement and project, see: <http://www.publiclibraryofscience.org/> [viewed 2 September 2003]
 7. The BOAI was launched by the Open Society Institute (of George Soros) in 2001–2002. Exhaustive information about the BOAI's goals can be found at: <http://www.soros.org/openaccess/index.shtml> [viewed 2 September 2003]
 8. Institutional repositories, being created by universities to house the works of their faculties and students, are becoming interesting to many institutions. For a very good summary explication of the concept, see: Raym Crow, "The Case for Institutional Repositories," in the ARL Newsletter # 223, August 2002, available online at: <http://www.arl.org/newsltr/223/instrepo.html> [viewed 2 September 2003]
 9. The Open Archives Initiative "develops and promotes interoperability standards that aim to facilitate the efficient dissemination ..." The OAI's home page is at: <http://www.openarchives.org/> [viewed 2 September 2003]
 10. According to the PLoS web site, the project is launching a journal publishing venture aiming to produce high-quality, high-profile information. The startup is funded with \$9 million from Gordon Moore (of Intel fame) and in-kind contributions by the Howard Hughes Medical Institute.
 11. See: <http://www.biomedcentral.com/> [viewed 2 September 2003]
 12. Comprehensive information is available at: <http://www.dspace.org/> [viewed 2 September 2003].
 13. On May 31st, 2003, HighWire offered 551,163 free journal articles. For more information about HighWire's program, see: <http://highwire.stanford.edu/lists/largest.dtl> [viewed 2 September 2003].
 14. For links to a number of publishers' access programs for developing nations, see *LIBLICENSE*'s link at: <http://www.library.yale.edu/~license/develop.shtml> [viewed 2 September 2003]
 15. The Open Society's eIFL project makes affordable and effective licensing arrangements for over 40 less-affluent countries. See their web site at: <http://www.eifl.net/> [viewed 2 September 2003].

Editorial history:
paper received 2 June 2003;
accepted 10 June 2003.