Re-modeling Scholarly Communications (a parallel tale from the U.S.)

National Scholarly Communication Roundtable, no. 12
The Structures of Scholarly Communication
Canberra, Australia
August 2001

Ann Okerson
ann.okerson@yale.edu

My contribution to this roundtable is practical rather than theoretical and takes its departure from a survey of recent history. My own involvement in this history has given me a series of vantage points from which to view that history: first as a research consultant, then as program officer with the Association of Research Libraries in Washington, and finally for the last six years as Associate University Librarian at Yale. But I take my point of departure from a study I wrote a dozen years ago, at a moment when it was still difficult and in some ways dangerous to a empt to understand and describe the realities of the serials market place in research libraries. Much of what I said then is now commonplace, but it is worth tracing the history from that time to this.

Ancient History

The study was compiled in 1989 on contract from the Association of Research Libraries (ARL): "Of Making Many Books" it was called, with a quiet bow to the prophet Isaiah, and in it I set out to explain the "serials crisis" of the time – the experience of librarians that the costs of providing journal subscriptions to their users were running far ahead of inflation and of their ability to keep up within typical academic library budgets.[1]

That study described the serials crisis from three points of view: from a library view, that is to say, taking the question as a consumer/marketplace problem (namely, one of high prices); from a scholars' view, that is to say, taking a view of the overarching systemic dilemma (fed by the information explosion, competitiveness among scholars, the culture of academic tenure and promotion, and the expectations of researchers); and from an economists' view, that is to say, looking at a situation of natural monopoly characterized by high fixed (or "first copy") costs and low marginal costs and depending on limited but relatively inelastic demand. The upshot of the study was a tripartite set of recommendations.

First, the report advocated a library action agenda. Most directly, consumer actions could impact the situation by raising consciousness of the pricing policies of publishers, by publicizing widely the pricing issue, and by organizing concerted protest against extreme cases of pricing increases. More strategically, the report called for creation of a research agenda to study usage, costs, and the system as a whole in greater depth. Finally, it urged libraries to work to strengthen the not-for-profit sector (focusing support on cost-effective publishers, on government agencies, on building cooperative collections, and developing consortia).

Neither was the report silent on emerging new electronic technologies:

> "a serials publication system which simply collects all research results into an an enormous research database would not by itself be adequate, even if it were free."

The above statement arose from an assessment of the genuine value added by all the traditional elements of the scholarly publishing system, including editorial intervention, peer review, and systematic formalization of the published body of research.

I was, of course, much younger then, and very idealistic, and so I concluded the report by urging that librarians work with the rest of the academic community to encourage change in the criteria by which research grants and academic tenure are awarded. Quality, not quantity, of publications should prevail, the report urged. I was neither the first nor the last to suggest a modification in publication criteria, nor the only voice to be ignored by a system that still depends, in large measure, on quantitative criteria.

Actions Taken and Lessons Learned

1. Librarians can, to some extent, control their own spaces

The first lesson we can take from this excursion into late 1980s history is that it was the recommendations for action within the control of librarians that were acted on most effectively. Many of the issues identified in the report (such as those of pricing and access) were then and remain today of broad concern to the entire research community; of course, they affect first of all the librarians who are charged make a large array of scholarly materials available to their user communities. While consciousness about pricing and access to the scholarly literature has broadened over the last decade, the fact that there is not more widespread interest in the issues is surprising, perplexing, and troubling – a phenomenon that might well be in itself the subject of research study.

But – librarians certainly did all they could. Over the last decade, a series of ARL interventions (some of which I was involved with in the early 1990s) have culminated in the "SPARC" initiative, which combines advocacy with focused attempts to create new journals from the not-for-profit sector, some on non-traditional business models and plans. More broadly, in the last decade the library point of view has been well and widely publicized. It is fair to say that all librarians have become believers in the importance of scholarly communications issues (and some faculty have joined librarians in awareness as well). The library community certainly has engaged more broadly than ever before with not-for-profit publishers and learned societies. From the dawn of the history of e-journals (late 1980s), the ARL was active in fostering the creation of such journals and raising awareness about them. In 1991, we published the first *ARL Directory of Electronic Journals* and, from 1993, I have been co-moderator of *NewJour*, an on-line alerting service with archive that reports new electronic journals as they come to our attention.

Where librarians controlled the spaces – organizing ourselves for resource sharing, organizing consortia for bargaining over prices and licenses – we as a community have had considerable success. Where we sought to build partnerships with administrators, faculty, publishers, and (in particular) funders, we have found that those groups all have their own priorities and agenda, on which the issues of serials prices and access to serials information do not loom as large as we could like.

2. Some strategies did not work and are unlikely to work

Some strategies the library community has tried in the last decade years simply have not worked. Here are some of the forms of advice that have fallen flat:

· "You should publish less"

Likeliest answer: "When donkeys fly: my dean and the National Science Foundation may not read my papers, but they do count them."

· "Give self-publishing a try"

Likeliest answer: "Publishing is not my job: I do the research and write up the results: I want to be able to submit my work to a robust and reliable publishing system and not have to worry about managing the publication of my research in the future."

· "Take control of your copyrights and manage your work"

Likeliest answer: mostly incomprehension. Though there are individuals who manage to do this to some extent, the ambiguities of copyright, especially in the shifting sands of Internet life and the actual and proposed changes of the last year, seem to make this impractical for most scholarly authors.

· "Are you sure you need this journal subscription? It's costly!"

Likeliest answer: "Yes, I need it." Scholarly and scientific information is rarely fungible. That is to say, there is rarely a case when an article in another, cheaper journal will meet the researcher's needs.

· "Let's start more, cheaper journals"

Likeliest answer: "There are too many journals already." (This is not, in fact, a bad answer.)

· "We'll marginalize the big for-profit scholarly publishers"

Those who single-mindedly pursue such a strategy have not so far experienced success. It might be said to be an unrealistic ambition.

3. Along the way, the unexpected happens and requires strategic re-adjustment

The world, moreover, had more than its share of surprises for everyone in the last dozen years. One need not be very old at all, for example, to remember the days when one imagined that the Internet was born and would remain forever a noncommercial space, where any form of advertising would be flatly illegal and commercial transactions would unthinkable. But the commercialization of the Internet is now an axiom of the economy of the twenty-first century and every organization that had a leading position in the "old economy" is aggressively pursuing Internet strategies. In our context, this means that precisely the same journals and journal publishers (both for-profit and not-for-profit) that were powerful in the paper world, continue to be very influential in the electronic world. Even more important, some of the influential information providers have used their economic power and influence with governments to further strengthen protections on intellectual property in cyberspace. Though the ultimate goal of such increased protectionism is most often the support of mass market intellectual property (such as popular books, film, video, and their derivatives), the big publishers of electronic information for scholars and scientists benefit as well. Although increasing numbers of scientific publishers, both not-for-profit and for-profit, say publicly that the prevailing legal copyright protection periods of "life plus 50 years" or "life plus 70 years" make no sense for scholarly works and that the protection period for such works needs to be much shorter, the reality is that boundaries between scholarly and commercial works are large and very blurry. It is unlikely that we will develop national and global legislation that can make these kinds of distinctions.

4. The marketplace for scholarly information can be made to work much more effectively in the electronic environment than it did in the print world

The best news to counter protectionist shifts in favor of the information sector has been the discovery that aggressive negotiation in licensing library resources, and in particularly through consortial arrangements, can provide libraries with at least one significant antidote to excessive copyright and database protections: it turns out that almost everything, even scholarly journal access, is negotiable these days. Much has been said and written in other venues about the increasing effectiveness of librarians in large-scale negotiations with the information sector.

5. Economic fluctuations have large global repercussions

Of course, all were surprised by the booming economic times of the last decade. The sense of constraint and constriction that were felt through the economic turns of the 1980s subsided somewhat, even for academic libraries. On the other hand, the relative purchasing power of the wealthy countries further distorted prices for others and the gap between haves and have-nots has probably increased.

Since the late 1980s, the research economy of higher education has mostly boomed and the research effort overall has continued to expand. The Internet economy exploded – with it came the discovery that electronic publication would not miraculously or immediately reduce costs over those of print publishing – but would in fact prove far more expensive than predicted. Furthermore, the costs and

implications of long-term archiving for electronic publications are likely to be considerable and have, as yet, been largely unaddressed.

6. Electronic information resources – an unstoppable phenomenon

Nothing is clear than that the all forms of publication, including indexing services, reference works, full text of all kinds, and traditional print journals in electronic form became incredibly popular very very quickly. From what could easily be a thousand datapoints, here are just some samples:

· Usage at JSTOR participating sites more than doubled between November 1998 and November 1999. Overall, usage of the JSTOR database is increasing at a rate of roughly three times per year. Part of this growth comes from the addition of new sites, and of course one must control for that. In any case, total 1998 usage was 5,920,398 accesses with 432,714 articles printed; 1999 usage was 17,311,453 accesses and 1,224,400 articles printed. (Data provided by Kevin Guthrie, President of JSTOR and cited in the U.S. National Academy of Sciences Report called *LC21*, chapter 1, note 35)

· The Florida Center for Library Automation reported that its Elsevier server's articles were accessed 58,842 times in 1998 and 207,006 times in 1999. The American Society for Cell Biology reports that the number of hits for its e-version of *Molecular Biology of the Cell* in the two years since its release increased about 500 percent. By contract, the number of subscribers to the print journal increased about six percent in the same time. (*LC21*)

· Academic Press usage at Yale sees a dramatic rise every year (illustrative PPT slide was shown).

7. Electronic resources – a costly phenomenon

But the prices for online electronic resources are surprisingly high, especially given that electronic information distribution enjoys very low marginal costs, compared to print information distribution. Sometimes the electronic version of a journal is priced as a surcharge on top of the print price, while sometimes it is only a little cheaper than print (90 – 95%). Moreover, absent long-term archiving assurance, libraries have been slow to decrease their acquisition of the print editions of the very journals they are buying in electronic form. *Some* price savings can be harvested by libraries that cancel print and retain only the electronic versions of journals, but my suspicion is that long-term archiving will bring its own costs to offset those savings at least in part.

Moreover, electronic journals positively cry out to be accessed through "gateway" services that seem to be even more costly than the journals themselves. A number of our key indexing and abstracting gateways cost my institution anywhere from US $20,000 - $125,000 peryear. These are staggering prices.

It would be reassuring to know that usage of these resources justifies the prices, but there are some real obstacles to good usage data. Many online information providers do not supply usage data at all,

even after we have made repeated requests and they have made repeated promises to supply it. None provide data in ways that are compatible with that of other information providers. We receive reliable data from only about 29 vendors at this point. JSTOR does the best job, Silver Platter and Elsevier Sceince do pretty well, and about 60% of those that do supply usage data do so in a way that ranges from mediocre to unusable. The numbers we do have suggest high rates of use and increasing use, but we all still know far too little – surely far too little to use such data to help us make choices about our subscriptions for the future.

Where Things Stand Today

I cannot recount this story without pausing to ask why we are really discussing scholarly publications issues today in many of the same ways we did in the 1980s. What seemed a dozen years ago like a situation spiraling out of control has remained somehow, vaguely, in control. We have made some compromises – reduction in the purchase of monographs to provide funds to pay for serials is chief among them – and we have survived

The apparent question, the one that brings us to the table, is this: How do we remake the communications infrastructure to take best advantage of new technologies? When we ask that question, however, the subtext question is much more pointed. It goes to *who* will remake the infrastructure – that is to say, who has power in this new scholarly communications environment? What we as librarians fear is that we will be rendered relatively powerless and irrelevant. The anxiety that drives our questions is the fear about our own future: do we have one? Will information providers price their resources to address directly the end-user or "retail customer" and thus "disintermediate" the library as information service provider? Fear and hope live side by side, but fears of this kind are rarely either simply confirmed or simply refuted. The good news is that power is still widely distributed in the scholarly communications world and bids fair to remain so.

Of course the publishers have power. The strong STM publishers like Elsevier ScienceDirect and Springer's SpringerLink, as well as the large scientific societies continue to be unavoidable and powerful players. The existing gateway providers as well, such as ISI, have roles that continue and will continue. I have no idea what the next bullet point, retained here in large font, may mean.

- Collaborations by publishers
  - DOI standard
  - CrossRef functionality

To a degree we did not imagine a decade ago, researchers have taken pieces of power themselves. The mixed success of various preprint servers should not be underestimated. There are and will remain fields of research such as high energy physics (http://xxx.lanl.gov changing to ??> where this kind of direct scholar-to-scholar communication will be powerful enough to obviate the need for intermediation by publishers. Librarians, interestingly, do retain a role in this area, inasmuch as libraries do not simply link high end researchers to each other but also mediate the results of research to users who are not part of the privileged initial community.

There are other initiatives today that need to be tracked carefully in this area. The Electronic Society for Social Scientists <www.elsss.org.uk> may be an unlikely contender in this field, while the Public Library of Science (heatedly debated, most recently in the American Congress: http://www.publiclibraryofscience.org> is a stronger contender for success and recognition. All these issues are lively subjects of debate in the United States, and I would commend readers to look and listen to some of the controversies.[2] ELSSS is a new non-profit-making society whose mission is to improve scientific communication in the social sciences, especially by the provision of electronic publications of high quality, wide diffusion, and low cost for the direct benefit of the the academic community (and indirectly for the taxpayer and general public). It aims to:

1. increase competition in the academic journal publishing marketplace
2. introduce a fairer and more efficient way of producing, distributing, and consuming academic journals whereby the large surpluses currently being earned by some commercial publishers are redistributed to the individuals who make journals.

Governments have also recognized their power in this arena in the last decade and undertaken various initiatives. The PubScience initiative in the US has been threatened and seems at this moment to be in reprieve (http://pubsci.osti.gov), while the more ambitious PubMed initiative (http://www.ncbi.nlm.nih.gov/PubMed) that would require scientific research results funded by government grants to be placed on the web for free distribution within a very short time of first publication, remains the subject of much contention. PubScience is a free gateway service developed by the U.S. Department of Energy, to facilitate accessing and searching peer reviewed journal literature in the physical sciences and other energy-related disciplines, as of October 1, 1999. Abstracts and citations are offered; a hyprerlink is provided and available if the organization has a subscription to the journal. Sources of content include participating publishers and the database of journal citations maintained by the Office of Scientific and Technical Information, one of the largest in the world. PubMed is a service of the NLM; providing access to over 11 million citations from MEDLINE and additional life science journals. Includes links to many sites. PubMedCentral is a digital archive of life sciences journal literature, managed by the National Center for Biotechnology Information (NCBI) at the U.S. NLM. It is not a journal publisher. Access is free and unrestricted. Currently 8 journals; 11 more promised. Government-supported consortia in other jurisdictions, such as JISC in the UK, OhioLINK in the US state of Ohio, and CNSLP (?) are ventures at somewhat higher risk.

Not surprisingly, the strong not for profit players in the scholarly publishing arena remain in play as well. BioOne, for example, (http://www.bioone.org) is an example of consortial e-publishing. It focuses on the biological, ecological, and environmental sciences. 2001 is first active subscription year. It offers an aggregation of the full texts of about 40 small bioscience research journals, mostly from small societies and a few not-for-profit publishers. Aim to dramatically reduce costs to libraries. Operating expenses are kept as low as possible in a virtual organization with low overhead. It depends on in-kind and cash contributions from "Founding Organizations" (Big 12, SPARC), Amigos, other strategic partners and draws on a high level of volunteerism. As the number of subscribers increases, prices will drop (a positive feature of not-for-profit publishing in particular). The only real question is whether this level of idealism can be sustained.

Also in the not-for-profit sector is HighWire, based in the Stanford Univeristy Library. As of 7/27/01, it has one of the two largest archives of free full text science on earth, comprising 287,663 free articles (free after 6 months or a year from date of original publication) and over one million total articles. Approximately 132 or half of the publishers make access available in this way.

There are some new providers of free or inexpensive journals. The World Health Organization (http://www.who.int) and the Soros Foundation (http://soros.epnet.com/eifl_description.html> have missions that run well beyond the domain of scholarly publishing and can include free publishing activities in their activities in support of that strategic purpose. HighWire Press also includes free material, as indicated above, as part of its overall project.

A different approach and view can be taken if we concentrate on the power that technolgoy itself has to change the way people work . The Open Archives Initiative (http://www.openarchives.org) is based on technology rather than ideology and builds interoperability standards that facilitate:

Author self-publishing/archiving of article
Online retrieval of those articles
Effectively a global virtual archive

The underlying concept is a powerful one that can be applied to other digital formats (besides articles and can be applied to paid content.

But libraries are far from powerless in this world as well. HighWire and OAI register the power of innovation that has come from within our community. SPARC, as outlined above, has been powerful at telling its story and is having at least the first stages of success in creating new journals: whether those become important pieces of the system remains to be seen. Less ostentatiously, the power of the library community to organize together in consortia that cooperate to share resources and in particular to license electronic information resources for readers, have a record of real achievement for their users and for the scholarly community as a whole.

It is in just such cooperation that libraries find their strongest new resources. In traditional formats, cooperation was difficult: physical materials are site-specific and need to be moved around with great difficulty and cost. Zbut with electronic formats, materials can be available to all equally. Given the propensity of publishers to make electronic information available under licensing agreements, electronic information is particularly suited for consortial arrangements bringing together many library customers in a single purchasing group. The alrger the consortium, the greater the benefits of joint licensing. Such consortia can be defined in many ways, including close relationships among a specific group for a continuing period or alternately entirely ad hoc and temporary. With electronic journal package deals, small increments in investment can bring huge benefits: we've seen this at Yale with Elsevier particularly.

Let me give a detailed example, arising from the experience of the Northeast Research Libraries Consortium (NERL) with the Academic Press familyi of journals. If our 23 schools had subscribed independetly to the print versions of AP journals in 1999-2000, it would have cost us $2,165,517, but the consortial cost to the whole group was $1,988,839. This does not seem to be a great saving, but in the print regime, the average institution in NERL was subscribing to 71 journals before our license

arrangement began. But of the X00 AP journals, the average institution in NERL was in fact using 167 on average after the licensing arrangement began. Thus for a slight reduction in cost, we gained a huge increase in functionality overall. Schools with higher prior print holdings benefitted less (about a 19% increase in number of journals accessed), while smaller institutions (such as RPI and Skidmore) saw huge increases (some 60+%).

My conviction arising from this set of observations is that librarians still have in fact considerable power to influence our and our institutions' futures. The individual librarian is now already what fashion speaks of as a "knowledge worker". We are close to our users and we meet their content and service needs, without primarily being concerned with format or means of access. Our key roles are evolving as ones of selection, access, and support: places where we add and continue to add real value. We will undeniably have a part o play in the evolution of e-archiving: we have the independence to give credibility and assurance to archiving enterprises, and we have the motivation to consolidate our low-use print collections and replace collections that take up substantial space and costs with more functional resources. We are supporting those roles by the way we are devloping our business and legal savvy, developing the technological expertise to supoprt our needs, and continually updating our skills. That is a true and valid model for the future of librarianship.

So let me close with a few prophecies – or platitudes, true and important, I think, for giving us the inspiration and direction we need.

First, ours is and remains a rapidly expanding and evolving information universe. In that world we must all hang together, or we will surely all hang separately. We will work with other libraries to establish patterns of sharing, with publishers to develop business approaches, and with our universities in the political arena. We can and will icense our resources globally and serve our readers locally. Information will never be free, or cheap: but knowledge workers will never be obsolete.

****

Here let me just append a list of select readings that can support and develop some of my themes or perhaps provide the reader with ways and means of quarreling with what I have to say.

- Paul Ginsparg: "Creating a global knowledge network," <http://arXiv.org/blurb/pg01unesco.html>
- National Academy of Science, "LC21: A Digital Strategy for the Library of Congress," found at <www.nas.edu>
- Ann Okerson, "Of Making Many Books..there is no end,"" <http://www.library.yale.edu/~okerson/making/>
- Tom Sanville: "Use of Electronic Journals in OhioLINK'S EJC," <http://www.ifla.org/IV/ifla67/pprog-e.htm>
- Alicia Wise: "Can evaluation get us to the heart of learning in the electronic age?" <http://www.ifla.org/IV/ifla67/pprog-e.htm>
- FOS Newsletter: <http://www.topica.com/lists/suber-fos>

[1] http://www.library.yale.edu/~okerson/making

[2] http://www.nature.com/nature/debates/e-access
and http://www.npr.org/ramfiles/totn/20010622.totn.02.rmmm