Open access: reflections from the United States

Based on a paper presented at the UKSG seminar 'Scientific Publications: Free for all?', The Geological Society, London, Tuesday 23 November 2004

This article offers a perspective about open access and the current 'churn' in scholarly journal publication from a collections development and scholarly communications specialist librarian in the United States. The differences between the US and UK higher education systems are not insignificant, the former being more extensive as well as more diverse in sources of funding support, with less direct involvement by national government. Thus, while librarians in both countries very much aim to improve various aspects of scholarly communications, their approaches and expectations may be somewhat different. The author suggests that much more experimentation with business models is needed before we can well understand the sustainability of open access, the models that can effectively support it, and to what extent. Early evidence suggests that there are no panaceas that will result in close to free journal publication; under an open access publishing system, the costs to research-intensive universities may well increase.

ANN OKERSON
Associate University Librarian, Collections & International Programs
Yale University

Introduction

For the purposes of this article I am going to use the metaphor of *The Good, the Bad and the Ugly,* which is a classic 1966 film about the search for hidden treasure. Even the 'Good' is a badly flawed person. All three are inter-dependent yet do not trust each other. See if the search for ever-improving access to information bears any resemblance to this tale.

The US Context

In the United States, there are up to 4,000 very diverse institutions of higher education of all types: two-year colleges, technical schools, undergraduate institutions, and research intensive, graduate level universities. We have an extensive, diversely funded, very well developed public *and* private institutional higher educational system in the States. So in many ways the intervention of government and funding councils is not as much of an option as it is for the UK. We also have a very high proportion of students enrolled in higher education. The population in the United States is around 290 million souls, and as many as 15 million of those are enrolled in post-secondary education. According to a recent search result on the web, the European Community population is around 450 million, and 13 or so million are enrolled in post secondary education programs. So for better or for worse – and whatever one may think the quality of US education – on average we have proportionally close to twice as many students enrolled in higher education as in Europe.

Even the largest state (public) institutions receive very little funding from their states. Let me give you the example of the University of Michigan, where I taught briefly. The University of Michigan is a major and highly respected research institution in the US (with 40,000 or so students) and has documented that it receives under 20% of its funding from the State of Michigan. The fact is that institutions of higher learning in the US are funded in various ways, through their state legislatures, through often substantial tuition and fees (which are particularly important for the private schools), donations, endowments, federal grants and contracts, numerous corporate sponsored projects, and

subsidised student loans from one sector or another, such loans being at least indirectly a way of supporting higher education.

We have 50 states with their own state university systems, scores of library consortia, and 60,000 librarians who are members of the American Library Association. Many have very different views about all kinds of library issues. However, one of the things that Americans do seem to agree on is that we do not want the government involved very much in our lives.

The paradox of open access

Open access can be a highly emotional topic, and yet it turns out that no one is opposed to it. In fact, almost everyone is sure that they are contributing towards open access in some way or another, whatever the method or the reason. One reason might be that they are the Good; another is that open access is at least to some degree a natural consequence of the ease of information access on the Internet: the way in which such information 'leaks.' Another reason is that there has been a fair amount of marketplace push-back by librarians and others against the traditional journal subscription models, particularly of the higher-priced publishers, and this pushback has made an impact. These days, we see more competition in the publishing arena, which leads to more rapid change than we have been used to. Still, the theoretical agreement about the benefits of wide access disintegrates rapidly in meetings as people opine about who and what is Good, or Bad, or Ugly, as if effective scholarly communications did not need the interdependence of scholars, publishers, librarians, and other players in the communiations chain.

The Good

Librarians think of themselves as Good – indeed, that is why I became a librarian, to become part of the mission of bringing knowledge and information to users. Librarians have a long tradition of providing on-site access to information. We circulate materials widely and we do this not only in our own institutions but also broadly for users we don't even know, in completely different locations, through inter-library loan and document delivery, whether through high tech or standard, traditional means. In the United States, librarians work very hard at monitoring government legislation related to higher education and particularly to copyright developments. We scrutinize developments in database legislation and push back at proposed legislation that would get in the way of users' education and scholarship. We support library- and user-friendly legal briefs in major court cases, circuit courts, state courts and also the Supreme Court: wherever a decision might restrict access or re-define authorship. There is a high level of legislative and political awareness by librarians in the US.

Librarians have also embraced and transformed our services using electronic technologies and the Internet. The library community has been supporting alternative publishing ventures and models, in order to test their effectiveness and sustainability in a time when technology has transformed so many publishing-related activities.

I also believe that publishers can be called Good. Many journal publishers have long traditions of distributing free or cheap print subscriptions to institutions and countries that cannot afford to pay in Eastern Europe, Africa and other parts of the world. Electronic versions of journals can now be delivered free or inexpensively to countries that cannot pay, and some of this is being done through global public-private partnership projects such as the World Health Organization's *HINARI* and the FAO's *Agora*, as well as numerous other independent projects, such as HighWire's journal access activities.

Publishers have had to shift their journals from print to electronic publishing mechanisms and platforms at considerable expense, with extensive planning and re-tooling. Many publishers have been willing to participate and experiment with new business models, whether print, electronic bundles or consortial, and we are now seeing open access experiments, along with library and consortial payment models to support them. This experimentation is extraordinarily beneficial.

Most publishers have loosened up a great deal from their earlier fears that library site licences would lead to numerous copyright violations. A report on RoMEO (Rights Metadata for Open archiving) here in the UK stated that approximately 90% of publishers do not require exclusive copyright transfer as a condition for article publication and allow authors to license their articles

or give extensive re-use rights such as the right to mount their articles on institutional repositories and/or on their own home pages. That is an exceptional change in less than ten years. This is also Good.

Today's government interest in STM publishing is also Good. Agencies and funding bodies have generally been very detached from scholarly publication, what it costs, and who has access, in the belief that a hands-off policy expresses proper neutrality about the doing of science. That is changing with various sectors now asking for agency and government involvement from the UK House of Commons and the US National Institutes of Health (NIH). We are suddenly awash in studies, recommendations and suggestions for policy change.

However, none of the foundations or government agencies that have explored open access and recommended change have proposed to alter their own dissemination practices. For example, the US National Science Foundation should publicly identify all the proposals it has funded, linked to any progress reports and final reports from the scientists who do the research. Likewise, the Wellcome Trust, the NIH and others should do the same. Instead these agencies are targeting the downstream articles that arise out of their funded research, articles to which publishers have added value. Why is that?

The Not So Good

There seems to be agreement about only one thing in the open access discussions (apart from the fact that everyone supports the concept): no matter how it is to be achieved, open access is not going to be free. Someone somewhere somehow will need to pay for the process of managing, reviewing, editing, producing, electronically distributing and hosting the journals or articles – as well as delivering numerous value-adding features. After that, the disagreement begins. Open access players, whether pro, con or neutral, are at odds about many things. Some of the many areas of disagreement devolve around exactly what open access is:

- Does open access require all articles to be free to all readers from the moment of publication?
- What version would be satisfactory to open access advocates?
- Does delayed free access, such as free access after six months, still count as open access? … Or is it a pathetically weak substitute, a kind of sop to the readers?
- Do the numerous developing nations' initiatives count as open access?
- How about hybrids in which articles prepaid up-front by authors are open access and articles that cannot be paid for up-front are not open access?

Many open access proponents criticize the 'imperfect' compromises, and many librarians are hypercritical about a large amount of wonderful change and progress in a very short time.

Another set of questions in play these days relates to cost of open access:

- Just how much does or should open access cost?
- Who is going to pay for it?
- How will we sustain this access?
- Will open access require entirely new business models such as 'author pays'/'funder pays'?
- Will it be done through institutional repositories, and what is the role of government, if any?
- Will libraries save money under open access?

(I am somewhat skeptical as there is very little hard data to answer most of these questions, and there is a great deal of speculation.)

The big deal

Sometimes called 'bundling', the 'big deal' in the US is a contract made by collectives of libraries for access to large packages of electronic journals. There may be very little local library choice about whether it will join and pretty much no choice about which titles will or will not be included, as the package includes everything – or a large subset – produced by a publisher. If open access means the ability of readers to access and use articles without payment or toll barriers, then perhaps for many that deal is facilitating open access.

Think about national and multi-national site licences. The Canadian National Site Licensing Project has brought huge amounts of content to all institutions of higher learning in the country. The eIFL project of the Open Society Institute has resulted in licences for free access to huge numbers of journals by readers in dozens of developing nations in the world. Some of the users who participate in those projects have told me that this looks

and feels exactly like open access and they have said: "You rich westerners should go away and solve your rich people's problems. We are now starting, thanks to the publishers and the web, to get the access we need."

Institutional repositories and open access

Institutional repositories (IRs) are sometimes thought of in the same breath with open access, because so far they have been conceived and set up as the freely accessible and searchable sites of not-for-profit organizations such as universities. IRs are intended to provide opportunities for institutions to care for and disseminate their authors' creations and allow a wide range of faculty and staff to gain experience in using information technologies. The question often arises: will such institutional repositories provide access to works that have so far been organized and distributed by publishers? Will IRs displace or even replace publishers, especially STM journal publishers as we have known them? Will IRs be *the* or at least *a* path for achieving the open access ideal?

In September of 2004, prior to speaking at an STM meeting at the Frankfurt Book Fair, I surveyed informally 40 of my colleagues in the largest research libraries in North America and received 16 responses. Of those 16, six have in place some kind of institutional repository, and the other 10 plan to have one within the next five years. Frequently expressed concerns were that:

- Institutional repositories are costly to create.
- Institutional repositories and placement of materials in them are a hard sell on my campus.
- Long-term sustainability of the IR is a daunting issue.
- So is a compatible infrastructure, cross-searchable with other IRs around the world.
- One librarian wrote: "Personally I am intimidated by the burden of preservation of archive materials, considering the wide variety of formats that are likely to be deposited."
- Several said that the different levels of rights and access issues loomed complicated and very large.

At a SPARC (Scholarly Publishing and Academic Resources Coalition) and SPARC-Europe sponsored institutional repository symposium in Washington DC, one panel addressed the topic of IRs and rights management. One of the biggest concerns academic authors have about university IRs is that access be provided to their materials in ways that are comfortable or appropriate for these authors. Clearly, levels of rights can be complex, and that those who are responsible for these repositories are spending a great deal of time handling rights issues, agreements, templates and the campus discourse about rights and copyrights (much as publishers do)!

In response to my survey question about the kind of materials that would be placed in the IR, the highest priority for many were faculty and student scholarly electronic works and databases, as well as many electronic research and teaching/learning projects on campus: projects that have no other permanent home. Other categories included 'grey literature' (informal publications) created on campus and content the library staff digitises or creates, such as images (which might not be able to be made freely accessible to the world). Less than half of the respondents (7 of 16) will target articles that their faculty submit to journals, especially in STM fields. So, will IRs in the US be in a position to replace journal publishers? At the moment, it seems that while repositories will expand access to a number of articles, the coverage will be far from complete or comprehensive, as compared to articles that continue to be produced in more formal publishing outlets.

My take-away from the survey, as well as from US discussions in general, is that we will indeed invest significant resources in IRs on our campuses. We will focus less on STM journal articles than on the urgent need to care for our institutional creations, such as faculty and student databases, data sets, teaching materials and research projects since those do not have any other outlets, whereas journal articles do. It is unlikely that under this kind of scenario in the US, scattered local versions of STM articles would compete effectively with the completeness or the value that the publishing community adds.

Savings for libraries?

The costs of journal publishing are not all that well understood, and pricing for open access is understood even less well. Open access as advocated these days requires that all costs associated with initial article/journal creation and long-term

hosting, value-adding and migration need to be collected 'up front', as articles are unlikely to yield any additional revenues in the future. (See the BioMed Central e-mail signoff, 'All Use is Fair Use.') Happily, several important publishers, among them the US National Academy of Sciences and Oxford University Press, have begun open access experiments, from which we all can learn a great deal. Let me briefly describe a couple of examples that have me far from sanguine that the costs to my library will decrease.

1. Oxford University Press's *Nucleic Acids Research* (*NAR*). The current subscription price for this important journal is $2,855 for paper plus an online edition. In 2005, this journal will become open access and the library subscription will remain at the current level, while faculty will also contribute $500 per submitted article. We have estimated that about 22 articles would be published in *NAR* by Yale authors in 2004. If exactly the same numbers hold true in 2005, Yale University would pay approximately $11,000 for faculty submissions. (If we did not have a library subscription, our authors would pay $1,500 dollars an article.) Thus, a crude estimation of the price that Yale University will be paying for worldwide open access to *NAR* next year will be just short of $14,000. That is not an increase for the library, but it is a significant instiutional increase and something to think about.
2. The National Academy of Science's journal *PNAS* (*Publications of the National Academy of Sciences*). Our library's current electronic subscription costs approximately $3,000. As of 2005 this journal will convert to open access. Authors will be charged an optional $1,000 dollars per article, because a survey of those authors said they were not willing or able to pay any more at this time. That $1,000 is certainly well less than the cost of publishing that article, and will be charged in addition to the library's subscription cost.

In short, at this time no reduction in library subscription prices is expected, and there will be increased costs university-wide. Of course, at some future time, if OUP and *NAS* secure sufficient author revenue, they may adjust the library subscription price downward. Nearly a year ago, I made some back-of-the-envelope calculations, aimed at assessing whether the Library would be ahead or behind in an entirely open access publishing environment. We made some rough approximations of Yale author output, using PubMed Central and ISI's *Web of Science*, from which we estimated that the University's authors had published as many as 4,000 STM articles in the previous year. We had expended in that year somewhere between $3.6–4 million for STM journals. Had we no library subscriptions for STM journals, and if our authors each paid $1,000 per published article, we would more or less spend as much money in the open access environment as we do today on our subscriptions. However, note that most publishers estimate the cost of article production as somewhere between $800 to $5,000 with many clustering at the $2,500–$3,000 level. So, a per-article cost to Yale of only $1,000 does not seem realistic. Nor does it take into account the many other kinds of subscriptions and licences we purchase.

Cornell University Library staff conducted a similar study – only far more meticulously – with surprisingly similar results. My conclusion from these two examples (admittedly limited in number) is that savings to large research libraries under an open access model are unlikely, unless substantial production cost reductions can be realised by many categories of publisher. I am not sure where those would come from; but one hopes there is potential for at least some reduction. Also, I have described (in the *NAR* and *PNAS* examples) potential and significant cost shifts away from the library to other parts of the university, which raises a number of institutional policy issues. One could ask: if there is such a shift, and if other parts of the university are paying for STM publications, could not budgets for libraries be commensurately reduced?

The National Institutes of Health recommendation

The NIH recommended in July 2004 that PubMed Central should host all articles that result from NIH-funded research, after a maximum lag of six months after journal publication. The NIH and any funding agencies are well within their rights to ask this as part of the condition of providing research grants. Yet, it is not clear to what extent access will be improved thereby. The NIH currently funds about 20–25% of US biomedical research, which though high is far from 100%. HighWire Press asserts that it already offers in its free content (its

publishers generally exercise an embargo period or 'moving wall' of their choice, anywhere from six months to two years) about three times as many free articles as PubMed Central (PMC).

Publishers are asserting that there may be better ways than the NIH recommendation to improve access, and that different versions of the same article between journal and PMC are likely to cause confusion for readers. Many have asked that the NIH instead permit links to publishers' own journal sites and have requested as well to choose the time after which publication will be available for free via PMC. We should know by December 2004 what the NIH has decided, and only time is going to tell us what the impact of that initiative will be upon formal journal publications. [Author's Note: As press time, 2/1/2005, the future of this recommendation continues to be uncertain.]

What is a librarian to do?

Our first commitment is to our users, to meeting their research, teaching, and learning needs. Next, our commitment is to describing the scholarly communication system to them, engaging in clear discourse, so that librarians and users understand each others' needs. As librarians, we need to engage with university administrations and other partners on campus to share the likelihood of significant cost shifts within our universities, costs which are beyond the purview of the library to mandate or pay for. Finally, we have a commitment to the scholarly communication system itself. We should treasure its achievements, which have evolved over several hundred years. It isn't a bad system, yet it could surely be better. Thus, we librarians want to be well informed, to support our colleague publishers' experiments, to be curious and open-minded, and to examine evidence that emerges from those experiments. We must continue to push back in our e-negotiations with publishers, in order to make clear what both our users and we need and can afford to pay. One of my mantras during these times of churn and change is: "Expect no windfalls, expect no miracles, expect no dazzling budget relief; be civil and turn down some of the rhetorical temperature."

My final point is neither good, nor bad, nor ugly. Let me draw our attention to a couple of recent important articles. Several authors suggested last fall that open access might be misplaced as the total centre of our attention. As I read their arguments, they resonated with me a great deal. In the September 2004 issue of *D-Lib Magazine* on the web, Herbert von de Sompel and colleagues published 'Re-thinking the Scholarly Communication'. Their thesis is that open availability is only one small dimension of scholarly communication. We must look, they reason, at the rapidly changing nature and pace of research, think about all the new media that are coming before us, re-think models totally, and develop goals that hasten discovery and enrich information for our readers. It is apparent that these authors do not think that the scholarly journal is the centre of all scholarly communications activity: it is one piece. In October, the *Charleston Advisor* published an editorial by the UK's David Worlock, called, 'Open Access – Free for All?' He asserted, perceptively, that today's dialogue needs to be about many things: about broken economic models, the transformation of scientific research, the journal, access, and the power of the web.

Let me conclude thus: let us work hard for the freest possible access to information, while recognizing that as librarians, publishers, and other key players in the scholarly environment, we have other, even larger and more encompassing challenges to meet.

■ **Ann Okerson**
Associate University Librarian
Collections & International Programs
Yale University, Sterling Memorial Library
P.O.Box 208240 (postal), 130 Wall Street (courier)
New Haven, CT 06520-8240, USA
Tel: 203-432-1764
Fax: 203-432-8527
E-mail: ann.okerson@yale.edu