Chapter 11: Business models in open access publishing

The Wellcome Trust supports both 'green' and 'gold' open access. In the following two chapters the focus is specifically on open access publishing ('gold'), and illustrate one of the key debates around open access: is there a viable open access business model for those currently publishing journals? In brief, Matthew Cockerill (BioMed Central) argues that yes, there is, whereas Mary Waltham is more circumspect.

Matthew Cockerill, in this chapter, draws a parallel between open access and open government, noting that only by making information freely available can we hope to build the kinds of systems and services that are already useful, and that will be essential in the future.

Chapter 11: Business models in open access publishing

Matthew Cockerill

Introduction

The move towards open access in scholarly publishing is a curiously two-stranded affair. Theoretical debate about the feasibility and desirability of open access continues, as it has for years, with little sign of movement in entrenched positions on either side. Meanwhile, the 'facts on the ground' in science publishing have been changing rapidly, with sustained growth in the number of new open access journals, and with ever more existing journals introducing open access options or switching to a fully open access model.

This chapter provides an open access publisher's viewpoint on the economics of scholarly publishing, laying out the case for open access as a viable and clearly preferable alternative to the traditional model. Whilst no arguments are likely to convince those who have a reason to defend the *status quo*, I hope to at least provide a useful perspective on the changes that are undeniably happening in the world of scholarly publishing.

What is an open access business model?

The defining characteristic of an open access business model for scholarly publishing is that it should not depend on restricting access to the published research in order to recoup the inherent costs associated with publication. Since the subscription and pay-per-view models used by most traditional scholarly publications inherently depend on restricting access, other models are clearly needed.

The most well known model for open access, often misleadingly referred to as 'author pays', is based an article processing charge (APC), generally paid by the author's funder or institution, which covers the cost of publication. Hundreds of journals now offer this type of open access, either for all articles, or as an option for the author. The APC model is simple and has the benefit of making authors aware of the cost of publication, something that is unfortunately opaque in a subscription-based system. Given that page charge and colour figure charges have long played a role in covering at least some of the publication costs for many traditional journals, APCs introduce no fundamentally new issues.

Are journals encouraged to publish any old rubbish by page charges or article processing fees? Open access opponents frequently imply that this be the case, but it is evidently not true. Journals only attract submissions if they can convince authors that they will achieve kudos and credibility from publication in the journal. If a journal cannot offer such credibility then it serves no purpose for an author to publish in it. As a result, all journals, including open access ones, are strongly motivated to ensure that they do not harm their reputation by publishing poor quality research. There is abundant evidence that open access journals offer every bit as high a standard of peer review as their toll access competitors. *BMC Bioinformatics*, *PLoS Biology*, and *Nucleic Acids Research* are just three examples of fully open access APC-funded journals that are at the forefront of their respective fields in terms of citation impact and quality.

Another increasingly common business model is for a journal to be directly funded, by a society, a foundation, or a research organisation, as part of the mission of that organisation. This results in the

journal then being free at the point of use for both authors and readers. Examples of *BioMed Central* journals funded in this way include the *Beilstein Journal of Organic Chemistry* and *Chinese Medicine.*

It is sometime inaccurately claimed that such journals are 'subsidised', in contrast to traditional journals that are 'commercially self-sustaining'. The distinction is illusory however. Profitable journals do not make their profits from thin air – the majority of revenue for scholarly journals comes from academic library budgets, and thus the scholarly community is 'subsidising' traditional journals too, even if those journals appear profitable. There is nothing inherently unsustainable about either form of subsidy. Scholarly publication is a service that must be paid for somehow and for the most part it is already the scholarly community that pays for it.

Having established what constitutes an open access business model, it is worth taking a step back to look at the fundamental objectives of scholarly publishing to see why open access publishing provides a model that is better suited to the needs of the research community. To do so, it is informative to contrast scholarly publishing with the other types of publicly funded material.

Open scholarship vs open government

If you are a UK citizen and you want to know what your elected representative has been doing in Parliament, it is impressively easy to find out. *Hansard*, the official record of parliamentary proceedings, is freely available via its official website. Furthermore, anyone can get permission to reuse and/or redistribute the material in *Hansard*, just by filling out a *Click-Use form*, to take out a '*Click-Use license'*.

What does this freedom of re-use mean in practice? *TheyWorkForYou.com* is an independent, volunteer-run website which reorganises parliamentary material by bringing all the latest speeches and written answers by an MP together on a single web page, organised by date and topic. It thereby makes the material vastly more informative and accessible for constituents. In this case, it is clear that freedom of reuse made possible innovation that has furthered the democratic goals that are the reason for publishing *Hansard* in the first place.

Similar possibilities exist for the re-use and enhancement of the scholarly literature. For example, *Google Scholar* mirrors *TheyWorkForYou.com* in that it too takes information available on the web (in this case, scholarly articles), and provides navigational enhancements. Specifically, Google Scholar makes it possible to find articles that cite another article, and to search for the most highly cited articles on a given topic. Until recently, this kind of citation information was so labour intensive to gather, and so complex to manage that it was only available through expensive services such as the Institute for Scientific Information's *ISI Web of Science*. In contrast, Google Scholar, by using modern technology to automatically extract this information from articles already available on the web, is able to deliver a similar service that is free to all.

A major limitation of the service, however, is that since much of the scholarly research literature is currently 'owned' by publishers, Google Scholar can only index what publishers allow it to index. At the time of writing, for example, Elsevier, which controls more than 20% of the scientific research publishing market, does not allow any of its articles to be indexed by Google Scholar. As a result, Google's Scholar's coverage is significantly incomplete and skewed.

It is not only commercial innovators such as Google who are held back by lack of access to the scientific literature – academics too are frustrated by restrictions as to which scholarly articles that they can access and reuse. For example, taxonomists and biodiversity researchers are keen to create a comprehensive database that would provide authoritative information about all the world's biological species. Unfortunately, however, most descriptions of new species are published in journals that claim exclusive rights to the articles concerned, As a result, species descriptions cannot be included in such a database without complicated negotiations for licensing rights (Agosti, 2006).

The motivation for scholarly publication – not cash but communication

Like parliamentary publishing, scholarly publishing operated for hundreds of years on the basis of print economics – those who wished to have access to the material had to pay for a printed copy. However in both cases the purpose of the publishing activity *from the point of view of the producer* was not, fundamentally, to maximise financial return, but to ensure that the material concerned should achieve wide visibility and readership at reasonable cost.

The arrival of the web fundamentally changed in the economics of distributing such information. The cost to distribute additional copies suddenly became negligible. As a result, when wide distribution and readership is the key objective, universal access suddenly becomes the most obvious and natural model *from the point of view of the producer*. Making content freely available via the web reduces or eliminates both printing and subscription-management costs, while maximising access.

The alternative chosen by most traditional science publishers, however, was to charge for online subscriptions just as they did for print subscriptions. If we consider applying this model to parliamentary information, we can imagine that *Hansard* might be turned into a subscription-only website, charging for an annual subscription. This would certainly generate additional revenue, and it could be argued that anyone who truly wanted to know what had been going on in Parliament could subscribe to find out. But it seems clear that closing off free access to the parliamentary record in this way, and thereby obstructing both casual users and websites such as *TheyWorkForYou.com*, would be a retrograde step for open democracy. Similarly, the scholarly community currently loses out because access to scholarly literature is not as free and open as the web allows it to be.

The UK government has no great reputation for technological innovation, but it is to be credited with recognising that the best model for information distribution in the age of the web may not be to simply translate the print model into its online equivalent.

Scientific funders, such as the *National Institutes of Health* (NIH) in the United States and the *Wellcome Trust* in the United Kingdom, have made a similar realisation. They understand that their research budgets would be used more effectively if they could ensure that the published results of the research were universally accessible so that researchers could build on them. Both the National Institutes of Health (2005) and Wellcome (2005a) have introduced policy statements to that effect. However, although funders ultimately pay the bills, they do not directly control all aspects of the publishing process. Librarians, authors and publishers (both commercial and non-commercial) are involved in a complex web of relationships and dependencies. As a result, shifting towards a more open model of publishing is not something that will happen overnight - it will require coordination between the various players concerned.

Publication and dissemination as the last part of the research process

In 2007, the Large Hadron Collider (LHC) at CERN is expected to begin collecting experimental data. The LHC is probably the most ambitious and complex physics project ever brought to fruition. Vast public sums have been invested in the project over several decades. As such, it provides a very clear example of the unsatisfactory nature of conventional science publishing. After vast investment by publicly funded research agencies, the research articles which constitute the final payoff from that investment would traditionally end up owned and controlled by publishers whose contribution has been, in the grand scheme of things, quite miniscule.

CERN's Director General, Robert Aymar, is well aware of the undesirability of this situation, and is taking steps to address it. In his own words, “The next phase of LHC experiments at CERN can be a catalyst for a rapid change in the particle physics communication system. CERN's articles are already freely available through its own web site but this is only a partial solution. We wish for the publishing and archiving systems to converge for a more efficient solution which will benefit the global particle physics community” (European Organization for Nuclear Research, 2005).

Several funding organisations (such as the Wellcome Trust) have similarly noted that seeing the publication and dissemination of research as the last part of the research process, and paying for it upfront, inherently makes more sense than surrendering control of research to publishers and then paying for access (Walport, 2004). By paying for the cost of publication, research funders not only ensure that their research is universally accessible, but also ensure that funding for journal publication naturally keeps pace with the amount of research that is being funded. The Wellcome Trust estimates that in biomedical research the cost of publication is typically only around 2% of the cost of doing the research in the first place (SQW, 2004). Currently this cost of publishing is paid indirectly through library budget. Paying it directly from research funding is equally feasible and need not lead to any overall increase in expenditure, seen from the perspective of the system as a whole.

The myth of ‘sustainability’

Of all the objections raised against the open access model, one of the most prevalent and yet misguided is the suggestions that it has not been proven whether open access publishing is ‘sustainable’.

Questioning the ‘sustainability’ of open access journals misses the very basic point that in macroeconomic terms it is obvious that both an open access model and a toll access model are equally ‘sustainable’ (that is, affordable to the scholarly community). At most, an open access model has the same costs as a toll access model. Funders such as NIH and Wellcome have recognised this point, for example, see Wellcome’s report on Costs and Business Models in Scientific Research Publishing (SQW, 2004). However, it is steadfastly ignored by those who oppose open access, who choose instead to speak ominously of the potentially "disastrous consequences" (Royal Society, 2005) of switching to an open access model of unproven sustainability.

Traditional toll access scholarly publishing is 'sustainable' (that is, profitable) because the academic/research community funnels large amounts of money into it through library budgets. It is not 'self-supporting'. The fact that journals 'break even' (and in fact often make large profits) in an environment such as this simply indicates that the community is currently choosing to pay for these journals from library budgets. Looking at the system as a whole, the same funds can clearly cover the cost of open access publishing (since opening up access introduces no new costs). This is a choice that the community can make, but given the intertwined roles of authors, libraries, funders and publishers in the process, it will require coordinated action. Meanwhile, it is a particular unhelpful form of circular for organisations to try to block the policy changes and budgetary shifts that are a necessary part of the process of funding open access, on the basis that open access has not been 'proven' to be sustainable.

In fact, the large publisher profits that are regarded as an indicator of the 'sustainability' of the traditional toll access publishing model are better seen as one of its greatest pathologies. Librarians are prepared (however reluctantly) to pay inflated sums in order to get access to the research literature for their users. This is entirely understandable, as their users *need* access to the latest research. The problem is that authors, through historical circumstances and inertia, have given away the rights to the research articles concerned and the research is now owned by publishers. Publishers (even so called not-for-profit publishers) therefore have libraries over a barrel – and journal pricing inevitably reflects this. The willingness of libraries to pay whatever it takes to get access to research is not an indication of greater 'sustainability' for traditional publishing so much as it is an indication of a failure of the market mechanism. The problem is that you cannot have a truly effective market for access to research articles because you cannot substitute one research article for another. If you need to follow up a citation, you really do need *that* article – no other can serve as a substitute.

Under an open access model, on the other hand, there is a choice and therefore a genuine market. Journals can compete to offer authors the best quality of service and value. As open access evolves, market forces will naturally ensure that the open access publishers who are efficient and provide a good service will be able to recoup their costs and make a reasonable profit. Those that are inefficient or provide poor service will not. A move towards open access is therefore not likely to leave the scholarly publishing landscape unchanged, but rather it will be a positive force for change and improvement in the services offered by scholarly journals.

What is a reasonable amount for an open access publisher to charge?

One of the great virtues of open access publishing is that it offers a far greater degree of pricing transparency than the traditional model.

When an author publishes in a traditional journal, the scientific community is paying the publisher for the service of publication and dissemination, just as with open access publication. However, in the case of traditional publishing the service is paid for indirectly through large numbers of subscriptions, and so it is generally unclear how much is paid by the community for this service, per article published.

In a few cases, the figures are available for comparison, however. For example, the American Physical Society is considered by many to offer reasonably priced subscription-only journals. Figures presented by Martin Blume (2005) indicated that the society received US $30 million of annual revenue from its journal publishing activity, and published about 16,000 articles, meaning that their revenue per published article from the subscription model was around

$1900. This suggests that open access charges in the $1000-$1500 range, as charged by many open access journals such as those from *BioMed Central* and *Public Library of Science*, are certainly not excessive. These open access publishers are providing universal access, for a total cost per article published than is lower than even relatively inexpensive conventional journals.

The future for open access

Given the benefits of open access to the scientific community, it might be asked why open access has not already taken over completely from traditional publishing. One might point to the necessary and proper conservatism of science, which ensures that trusted paradigms are not discarded overnight. On the other hand, one might also note the recent survey by the Deutsche Forschungsgemeinschaft on attitudes to open access among its researchers, which found that younger researchers were consistently both more aware of open access, and more enthusiastic in their support for it, than their elder peers (Over, Maiworm and Schelewsky, 2005). Change is never easy, and no doubt many obstacles remain, but the enthusiasm of a new generation of researchers for open access, and the ongoing expansion of activity in this area by funders and policy-makers, together mean that open access to the results of scientific research could soon be the norm rather than the exception.