Chapter 10: Open access to the research literature: a funders perspective

Robert Terry, Robert Kiley

Introduction

In a declaration to commemorate the publication of the first draft of the human genome, UK Prime Minister Tony Blair and US President Bill Clinton commented that, "unencumbered access to this information will promote discoveries that will reduce the burden of disease, improve health around the world and enhance the quality of life for all human kind" (quoted in BBC, 2000).

One of the major funders of the human genome project was the Wellcome Trust[1], an independent charity that funds research to improve human and animal health. And, having been at the forefront of the decision to make the genome sequencing data freely available, it was perhaps inevitable that this funding body would lead the way in advocating free access to the research literature. If, as the Wellcome believes, it makes sense for scientists to have free access to raw, genomic data – to help realise the promise of this research – then it makes equal sense for scientists to be able to access the outputs (journal articles), to enable this research to be built on and developed.

This chapter considers the issues around open access from the perspective of a research funder.

Background

The Wellcome Trust first began to look at issues of access to the research literature following concerns raised by its own *Wellcome Library* Advisory Committee during 2001. In response, the Trust commissioned two reports from economic consultants SQW in order to understand the structure and economics of scientific journal publishing (SQW, 2003; SQW, 2004). These reports highlighted a number of issues.

One was the rising cost of paying to read the research literature. Subscriptions to science, technical and medical (STM) journals had risen by over 200% in the period 1990 – 2000, at a time when inflation was in single figures (from Blackwell's Periodical Price Indexes 1990–2000, quoted in SQW, 2003). This, coupled with the knowledge that certain publishers were taking large profits out of the system (House of Commons, Science and Technology Committee, 2004a), suggested the market for publishing scientific research is, from our perspective, a failing market.

Secondly, researchers, as authors, give their research papers to publishers for free and are mostly unaware of the cost of reading the material - when they are readers - because subscriptions are dealt with through the library budgets. As such, there is no real link between the user of the journals and the cost of obtaining access to those journals and consequently, no downward pressure on prices (see Figure 10.1)

[1] The Wellcome Trust is an independent charity funding research to improve human and animal health. Established in 1936, and with an endowment of around £11 billion (2005), it is the UK's largest non-governmental source of funds for biomedical research. See: http://www.wellcome.ac.uk/

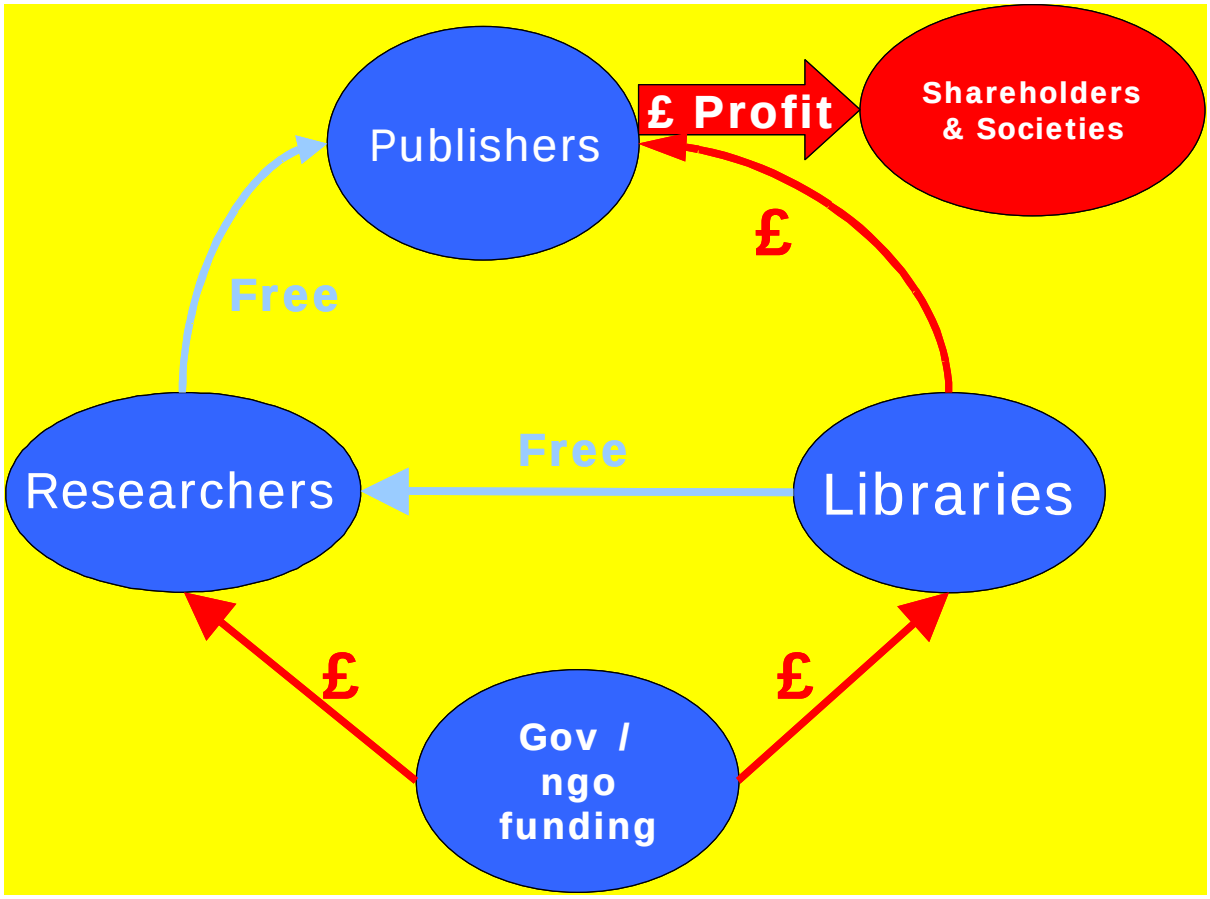

Figure 10.1 - The economic cycle of scientific research publishing

As journal costs were increasing at a far greater rate than library budgets, the total number of journal titles that researchers had access to actually fell. Data from the Association of Research Libraries (*ARL Statistics and Measurement Program*) showed that in the last decade of the 20th century the number of serials titles purchased by large academic research libraries fell by 5%.

Against this background – often referred to as the 'serials crisis' – and fuelled by the internet revolution, a new method of distributing peer-reviewed research papers was born, the Open Access (OA) model.

Defining open access

For the Wellcome Trust, an OA publication is one that meets two conditions, as defined by the Bethesda Statement (2003). The Bethesda statement, drafted in 2003 following a meeting organised by the Howard Hughes Medical Institute, is important as it marked the first occasion when funding bodies discussed the issues pertaining to OA with a number of publishers, learned societies and universities (see Bailey, Jr., this volume).

The Bethesda Statement (2003) defines access, reuse and archiving as important elements of an OA work:

> (a) The author(s) and copyright holder(s) grant(s) to all users a free, irrevocable, worldwide, perpetual (for the lifetime of the applicable copyright) right of access to the

> work, and a licence to copy, use, distribute, perform and display the work publicly and to make and distribute derivative works in any digital medium for any reasonable purpose, subject to proper attribution of authorship, as well as the right to make printed copies for their personal use.
>
> (b) A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in a suitable standard electronic format, are deposited immediately upon initial publication in at least one online repository that is supported by an academic institution, scholarly society, government agency, or other well-established organisation that seeks to enable open access, unrestricted distribution, interoperability, and long-term archiving.

In practice this means that the Trust supports two routes to making the research it funds freely available. One is to publish the original research paper in an open access publication (such as journals published by the Public Library of Science, or BioMed Central); the complementary approach is to publish in any journal that allows deposition of a copy of the final manuscript into *PubMed Central* (PMC) or, once established, UK PubMed Central (UKPMC) [see below].

Wellcome Trust: realising the goals of OA

The Trust was the first – and to date, the only – research funder to introduce a grant condition that *requires* grantees to make their research papers freely available to all.

Specifically, Trust grantees are required to deposit in PMC, electronic copies of any research papers that have been accepted for publication in a peer-reviewed journal, and are supported in whole or in part by Wellcome Trust funding. All deposited papers must be available for free no later than six months after the official journal publication date (Wellcome Trust, 2005b). Recognising the significant change this introduced to research publishing, the Trust phased in this condition over a 12-month period. It applied to all new papers arising from grants awarded after 1 October 2005, but does not come into force for existing grants until 1 October 2006. In addition, the Trust has provided grant holders with additional funding to cover the costs of page processing charges, levied by publishers who support the open access model.

The Trust is also working with the *National Center for Biotechnology Information* (NCBI) to establish a UK version of PubMed Central (UKPMC). The aim is to create a stable, permanent, and freely accessible digital archive of the peer-reviewed biomedical research publications.

Initially this service will simply mirror the data held in PMC, but longer term the objective is to create an independent resource that can process new content, offer enhanced linking and searching capabilities and be configured to meet the specific requirements of both UK researchers and funders. We anticipate UKPMC to be available early in 2007.

The driver for OA: a funders perspective

For a funder such as the Wellcome Trust, providing open access to the literature it has funded is attractive for a number of reasons.

Firstly, it is a fundamental part of our charitable mission to ensure that the work we fund can be read and utilised by the widest possible audience. Unrestricted distribution, via the internet, currently offers the best available method to do this. In contrast, the current 'reader pays' model significantly restricts access. For example, a survey undertaken by *BioMed Central*

found that found that fewer than half of the articles resulting from NHS research grants are accessible online to NHS employees (Cockerill, 2004).

Secondly, providing open access to the research literature enables these outputs to be linked and integrated with other resources. Research papers that are tagged in a standard, uniform way – such as the *NLM Journal Archiving and Interchange DTD* – can be read and searched by computers (as well as people), thus enabling context-sensitive links to be made to other online sources, such as gene and chemical compound databases. Figure 10.2 shows how an article in *PubMed Central* is linked to a number of online resources, including *PubChem* and *OMIM*.

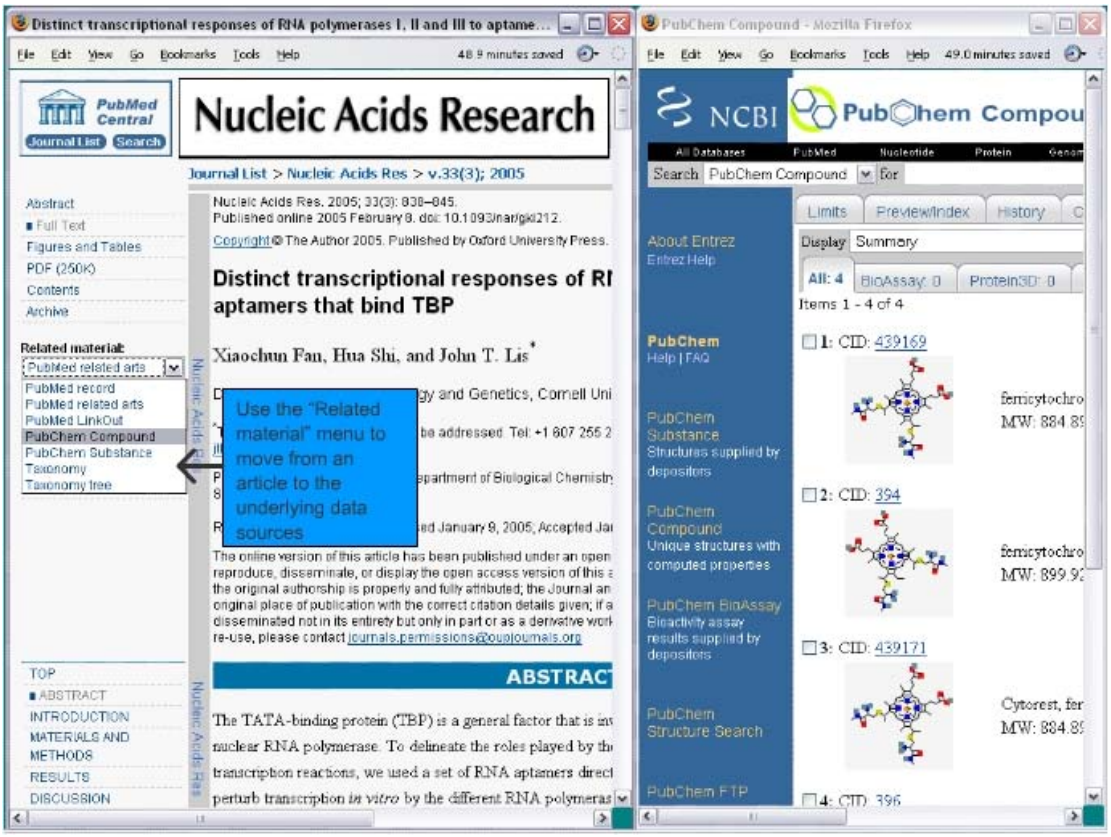

Figure 10.2 – Context-sensitive linking from PubMed Central to other online resources

Over the next few years we will start to see new types of search facilities being developed – based on data mining techniques – which will create new knowledge by linking research papers that previously had not been seen as being relevant to each other. The creation of this capability will be a hugely significant development in the life sciences as massive data sets, such as gene sequences are linked with environmental data sets and patient records in cohort studies. For the development of new drugs, compound databases can be searched and matched with references in the published literature to help discover previously unrelated activity or side effects.

Making research outputs freely accessible also helps funding bodies evaluate the research they have funded. In a recent exercise conducted by the Trust it was found that only 10% of papers in which the Trust was acknowledged as a funder were available without a charge online. This makes any systematic analysis of the value of the research we fund, using the

web, very difficult and costly. Once all Wellcome-funded research is available 'under one roof' (in PMC, or UKPMC) it will be possible to examine the effectiveness of our funding strategy and re-align it as appropriate.

Finally, by mandating its grantees to make all research outputs accessible through PMC/UKPMC, the Trust is helping to ensure that the digital record of medicine can be preserved. All papers that are added to the PMC/UKPMC repository are marked-up to the *NLM Journal Archiving and Interchange DTD*. Mapping documents to this standard, non-proprietary format should ensure that future generations will be able to read these digital files, irrespective of developments to either hardware or software environments

Cost of open access

The effect of the author (funder)-pays model is that it realigns the market, making publishing a research cost, rather than a library cost. In addition the volume of funds available to publish should remain in line with those research budgets. It will be a more transparent system enabling a clear comparison between the charge made to publish and the service offered by the publisher.

Though the OA model provides free access to the literature for the reader, there are costs associated with this approach. For example, managing the peer-review process, and copy editing the final manuscripts are value-added services that incur expenses.

However, looking at the costs (see Table 10.1) levied by both open access publishers and those publishers that have introduced a hybrid model, these, from the perspective of the Wellcome Trust, are affordable.

Publisher	Example title	Cost per article (assumes no discount)
BioMed Central	Arthritis Research and Therapy	£750
Blackwell – Online Open	Journal of Physiology	£1250
OUP – Oxford Open	Rheumatology	£1500
Public Library of Science	PloS Medicine	$1500 (approx. £850)
Springer	Journal of Human Genetics	$3000 (approx.£1700)
Average cost per article (across these 5 publishers)		£1210

Table 10.1 - Typical costs (Jan 2006) of publishing in an open access journal

In a typical year the Trust is acknowledged in approximately 4000 original research papers. If every single one of those papers was published as an open access article, and taking the average cost calculated above (£1210 per article), the total cost to the Trust would be £4.84 million; just over 1% of our annual research budget.

It is also worth noting that the Trust is rarely the sole funder of a research team, and more than 80% of papers that acknowledge our support also acknowledge the support of one or more other funders. In time these costs will be spread throughout the research budget and fall below the 1% figure estimated here.

Much of the debate around OA has focused on the traditional publisher and whether this model will bring about their decline. However, though the rhetoric has, at times, been apocalyptic in tone, what evidence there is suggests that open access publishing *can* provide a means by which publishers can continue to meet costs and turn a profit. Springer, for example, a traditional for-profit publisher, believes that it can generate a profit through open access; Jan Velterop, Director of OA Publishing at Springer, commented "we are absolutely convinced that with open access we can have good profit margins" (Velterop, quoted in Pincock, 2005).

Further, a study (Waltham, 2005) that looked at the viability of the open access model for learned societies concluded these publishers could *continue* to deliver the average surplus to their societies, by introducing an OA fee of £1166 per article (2004 costs). (See also Waltham, this volume.)

In support of these arguments are the experiments by a number of established publishers with a hybrid model of publishing (author-pays open access articles) within a subscription journal. This appears to be a low risk way of exploring the viability of the open access model, both by journal title and subject discipline. The two models, subscription and author-pays, should be able to co-exist for as long as they are both needed.

Conclusion

Providing free and open access to all research outputs provides real benefits to both the research community, and the funding bodies. Over the next few months it is likely that the Research Councils UK (RCUK) will mandate its grant holders to make their research papers freely available – either through institutional or subject-based repositories.

In the US, the National Institutes of Health (2005) Public Access Policy has – through the creation of *PubMed Central* and the supporting manuscript submission system (*National Institutes of Health Manuscript Submission*) – provided the infrastructure for improved public access. However, the supporting policy of encouraging (rather than requiring) NIH grantees to deposit research papers in PMC can be described as a failure. At a meeting in November 2005 of the NIH Public Access Working Group[2] (2005), it was reported that less than 3% of unpublished manuscripts authored by NIH investigators were being deposited in PMC (See also Suber, this volume). It will be interesting to see whether this failure to make a critical mass of research available to the American taxpayer will encourage the Congress to require the NIH to adopt a stronger, mandatory approach.

It should surprise no-one that many funding bodies from around the world are making a commitment to increasing open access to the research they fund, for example by signing the Berlin Declaration (2003). Funding bodies have clear objectives, in the case of the Wellcome Trust to improve human and animal health, so they are obliged to explore and use the most efficient and modern practices to achieve those objectives. Making the research and data they

[2] The NIH Public Access Working Group is a subgroup of the National Library of Medicine's Board of Regents to review the impact of the NIH Public Access policy

fund accessible to as many people as possible, for free, via the internet, offers a significant advance in the research process and should form a key aim for all research funders.