

Open access and institutional repositories: an evidence-based approach

Based on a paper presented at the 28th UKSG Conference, Edinburgh, April 2005

Oxford Journals are conducting experiments with partial open access publishing (*Journal of Experimental Botany*), full open access publishing (*Nucleic Acids Research*), institutional repositories (SHERPA) and subject repositories (PubMed Central). Initial results regarding open access have been encouraging, in that usage appears to have increased, but it is unclear whether open access publishing can be viable without support from institutional subscriptions. Early evidence suggests that free availability of articles through repositories also leads to increased usage but may have a detrimental impact on subscription revenues.



MARTIN RICHARDSON

Managing Director

Oxford Journals, Oxford University Press

Introduction

It seems that hardly a day can go by these days without some new policy or initiative being announced that will 'reform' the process of publishing scholarly journals, retaining the benefits of the current system that authors and readers enjoy, whilst doing away with the non-essential parts – those aspects that are both costly and low-value in the eyes of the initiators of change.

In reality, the business of publishing research articles is a complex one. As Roosendaal and Geurts described in 1997¹, any system of scholarly communication requires the 'registration' of a research finding, its subsequent 'validation', creating 'awareness', 'archiving', and finally 'reward' for the author. This is a complex process, which the current system supports – from the record of the date that the manuscript arrives with the publisher, through peer review, the publication in a journal and listing in current awareness resources, preservation of online and print versions by libraries and the prestige and/or impact factor of the journal itself, ensuring that authors are rewarded for their efforts.

Of course, the established system is not without its problems, and the enormous opportunities that the online environment offers for greatly enhanced and increased access to research information must

be developed. The challenge is to introduce changes that will build on the best aspects of the established system, combining this with innovation where it truly adds value and impact for the long term, whilst minimizing any detrimental side-effects.

Oxford Journals, as part of a university press and a department of the University of Oxford, has a particular mission to investigate opportunities to increase dissemination for the benefit of the scholarly community. Our approach is a practical one, based on experimentation. We do not pretend to have all the answers, but we are addressing the questions through our initiation of, and participation in, a number of publishing innovations. Our experiments are designed with one simple goal in mind – to discover if new business models can achieve wider, more cost-effective dissemination than existing models, while at the same time remaining financially viable.

This article follows on from a previous contribution to *Serials*, which described the background to some of our experiments.² A lot of questions remain unanswered at this stage, but we hope through this approach to encourage others to join us in collating a growing pool of statistical evidence to inform changes to existing publishing

models, so that they really are evolved for the benefit of scholarship.

The three models around which we have based our experimentation are: (1) partial open access, (2) full open access, (3) institutional and subject repositories.

1. Partial open access

Case study: *Journal of Experimental Botany*

During the latter half of 2004, the *Journal of Experimental Botany* – published by Oxford Journals on behalf of the Society for Experimental Biology – introduced an option for authors to pay a publication charge of £250 for their paper to be made freely available online. Funding from the Joint Information Systems Committee (JISC) was awarded to support this experiment, enabling us to grant a waiver to UK authors. The model was only applied to original research articles – reviews and special issues were excluded. Subscription prices for 2005 were held at 2004 levels.

The results so far this year indicate the level of interest in open access by this author community. Of the 92 original research articles accepted for publication from January to April this year, authors of 33 (36%) opted to make their papers freely available online. With JISC funding, one might expect this figure to be relatively high – for the majority of researchers working in the UK this option was free – but in fact only 18 of the 33 OA articles are from UK authors (55%) and 15 are from non-UK authors (45%).

2. Full open access

Case study: *Nucleic Acids Research*

Nucleic Acids Research (NAR) is a flagship title on the Oxford Journals list: a respected, international and highly cited journal in the field of bioscience.

We had already introduced a mandatory open access model for several special issues of *NAR* during 2003 and 2004, and we supplemented these experiments with a detailed programme of market research with our authors. Based on the overwhelming support the open access model attracted, we made the decision

to move the entire journal to an open access model for 2005. We realized that to fully test the model we needed to charge a higher amount than for other experiments we had initiated, to reflect the real cost of publication. At the same time, we also realized that it was challenging for our authors to move from one model to another virtually overnight. Until funding policies changed, and institutional budgets were better distributed across researchers and their institutions to enable them to afford these charges, we needed to offer an interim solution that would support a lower fee to authors. To this end, we developed an ‘institutional membership’ programme. To qualify, institutions simply need to maintain their print subscription for 2005 (membership is automatically included), or join our separately available membership programme (set at the price of an online only subscription to *NAR* in 2004). This entitles researchers within those institutions to pay a discounted charge of £300 for publication in *NAR*, compared with a charge of £900 for researchers from non-member institutions. This charge is also supported by JISC funding within UK HE – reducing either the full or institutional membership rates by a further £300. Free or greatly reduced publication charges are also available to authors from developing countries, and waivers are available for other authors where funding is not available. It is an essential tenet of our approach that no author is prevented from publishing within *NAR* because of their inability to pay.

Other conditions of our experimental model include simultaneous publication in the PubMed Central (PMC) archive, unlimited re-use of articles for educational and research purposes, and authors’ retention of the copyright for their papers.

Since 2002 – when we began experimenting with open access on *NAR* – submissions for publication have continued to climb. Whilst this has been a general trend for many of our journals, the growth for *NAR* has been particularly rapid – from 1,750 submissions in 2002 to 2,731 in 2004 – though we do not know if this growth is due to open access or due to other factors such as increased impact factor or decreased publication times. The number of manuscripts accepted for publication is also on the increase: from 824 in 2002 to 1,116 in 2004

(35% growth), despite a decreasing acceptance rate (47% to 41% over the same period.). The number of published pages went from 6,629 to 9,629 (45%). This trend looks set to continue in 2005 – despite the introduction of author charges – with submissions to the end of March already up on 2004 by 7%.

The availability of *NAR* has also been steadily increasing over the past few years (Figure 1), particularly as the online edition has become widely accessible through our developing countries programme and library consortia licensing agreements. Overall, the number of institutions with access to *NAR* online rose from just under 2,500 in 2002 to almost 4,000 in 2004. Regular institutional subscriptions, however, have declined by some 12% during this period, which we think may be related to free online availability with a six-month delay (see case study 3). Income from other sources, that

give *NAR* the benefit of much wider access, has been sufficient to replace the lost income from these ‘regular’ subscriptions.

It is not just accessibility that is on the increase – online usage is up, too (Figure 2). Since 2002, the number of full-text downloads of *NAR* from the HighWire-hosted site have increased from 2 million to over 5 million (250% increase). This figure is projected to reach 7 million by the end of 2005. Access via PMC is also now seeing a significant increase – from virtually nothing in 2002, when the full text was only available from the *NAR* site, to 1 million downloads in 2004, when the full text was freely available 6 months after publication. We project access to *NAR* articles in PMC increasing to 2 million downloads by the end of 2005. Combined with our HighWire site, this presents an anticipated 9 million downloads during 2005 – a three-fold increase since 2003.

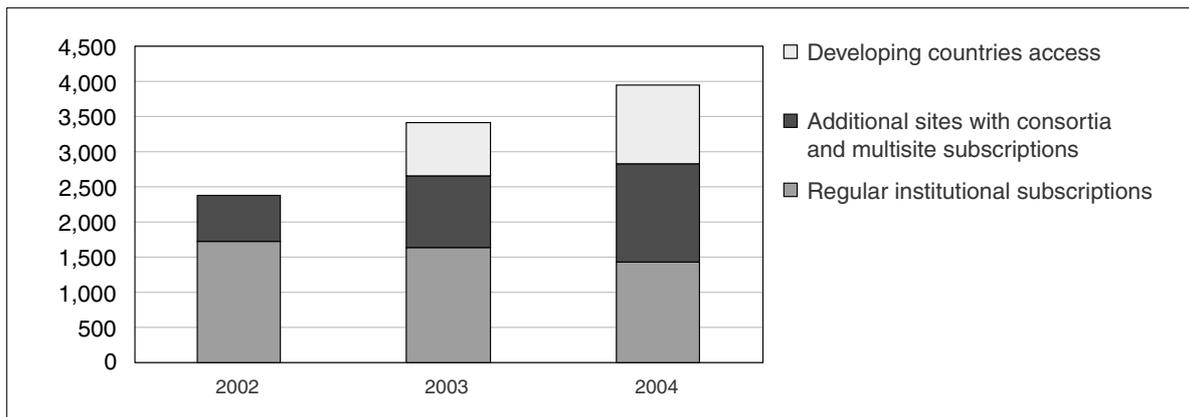


Figure 1. Institutional circulation 2002–2004

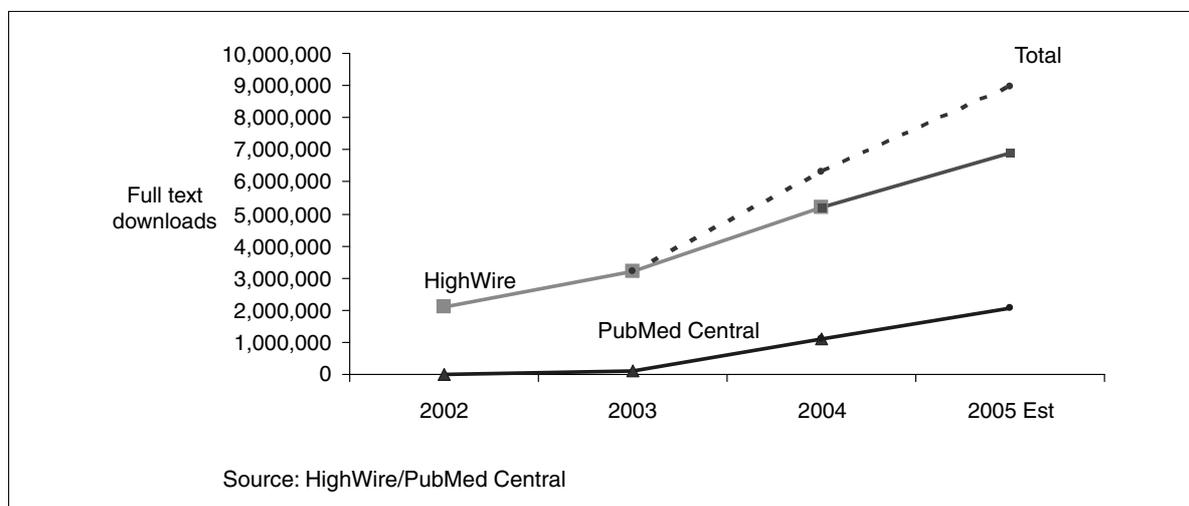


Figure 2. Online usage

It is too early in our experiment to conclude that open access is the major cause of this increased usage – we do not yet have sufficient data to analyse and the upward trend started before open access was introduced. However, we have compared open access articles published in 2003–2004 with subscription-based articles, an analysis of which does indeed show a clear increase in usage for the open access articles (by a factor of 50% in the first six months following publication). However, these open access articles were published in ‘special’ issues, so they do not provide a direct comparison with regular issues of the journal. Other factors may also have come into play – the greater profile *NAR* has received through our experiments, general trends for increased usage of online journals, the efforts of the editor and editorial board to further increase the quality of published articles, and many other factors could be contributing to this continuing growth in usage.

As well as investigating accessibility and usage of *NAR* through publishing the journal under open access, we are equally keen to monitor various cost metrics through this change in model. The prices per article, download and page have already been steadily decreasing (Figure 3) due to an increasing number of manuscripts being published, resulting in economies of scale, together with an aggressive efficiency drive which has increased speed of publication whilst reducing costs. Costs are set to decrease further in 2005. In 2002 the average price per article was £2,700, which had fallen to £2,300 in

2004, and we are estimating that the price per article will fall to £1,750 in 2005. Similar trends are apparent for price per download (£1.07 in 2002, to £0.48 in 2004, and estimated at £0.30 in 2005) and subscription price per page (£0.20 in 2002, to £0.17 in 2004, and estimated at £0.16 in 2005).

When compared with other journals in the field of biomedicine, *NAR* is below that of our median subscription price for biomedical journals (Figure 3), using information taken from the LISU study commissioned by Oxford Journals last year³, and comparing median price-per-page prices for biomedical journals across a range of publishers.

We have yet to prove that this model can be viable in the longer term – particularly without print subscriptions. Whilst it looks promising, it is worth reflecting on the fact that, in 2005, 75% of projected revenues for *NAR* continue to come from the print edition (including free institutional membership), and only 25% from author charges and institutional memberships. Without a significant shift of funding from institutions to authors, it will be difficult to prove the scalability of this model.

In order to help us further develop the model and test its long-term viability, as well as analysing the true benefits for scholarship that it offers, we plan a programme of further research. This will include a detailed review of usage under the full and partial open access models, combined with an analysis of the relationship between usage and citations; we plan to conduct

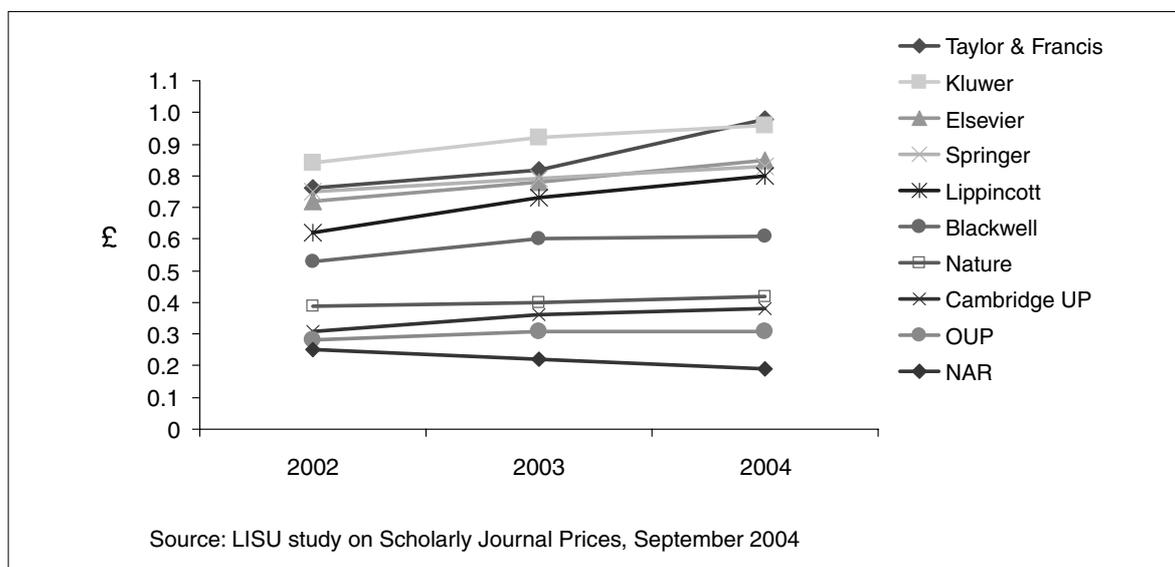


Figure 3. Median price per page 2002–2004, biomedical journals

this research in partnership with LISU and ISI. In addition we shall continue to monitor the financial implications of experimenting with these two models.

3. Institutional and subject repositories

Case studies: SHERPA and PubMed Central

Author self-archiving of articles is very much in the news at the moment, with the National Institutes of Health (NIH) having announced last year that postprints of all research articles that have received NIH funding should be deposited in PubMed Central as soon as possible following publication of the final version on a publisher's site⁴. Both the Wellcome Trust⁵ and the RCUK⁶ have also recently come out in support of this policy, which therefore looks set to gain momentum.

Such policies appear to be backed with little evidence of the real benefits for the community or analysis of the potential impact on the current publishing process. Author self-archiving will lead to multiple versions of articles being available on multiple sites. With preprints, postprints and the definitive, publisher-produced article all becoming available, how will readers be sure whether the version they are reading is the authoritative one? The effect on costs is also unknown – this policy could in fact increase total publication costs, which in turn might threaten the viability of specialist journals.

As we have demonstrated through our experience with open access, what is needed is careful experimentation to test the potential benefits and help to determine any possible pitfalls or unintended consequences. We have begun to analyse the impact of making articles freely available in institutional repositories through our partnership with Oxford University Library Services (OULS) in support of the national SHERPA project. We have provided OULS with access to 350 articles by Oxford University-based authors published in many of our journals from 2002–2004. These articles are searchable via the OULS pilot institutional repository and available free of charge to researchers across the globe.

We plan to gather information that will enable us to compare the costs, benefits

and risks of institutional and subject-based repositories. We hope to report more details on the results of this experimentation at a later date, but even at this early stage we are already seeing some increase in usage. For the OULS repository, the total number of full-text downloads from June to October 2004 was in the range of 500–700 per month, and this had reached 2,000 in January 2005; usage has subsequently remained at a similar level, since few articles were added to the Oxford Repository since the beginning of 2005. More dramatically, the usage of *NAR* content within PubMed Central (PMC) – deposited for several years now – is seeing a steep increase. The average number of full-text downloads of *NAR* articles within PMC from March 2004 to December 2004 remained steady at some 100,000 downloads a month. During that time, articles were deposited with PMC after a six-month delay. When we removed this delay period in January 2005 (when *NAR* converted to full open access), the number of downloads increased to 150,000, and by March 2005 was approaching a quarter of a million. This provides evidence that a 'no delay' policy of authors depositing postprints of their articles in PMC (as NIH and others encourage) is likely to result in such repositories rapidly becoming high-usage resources, which could in turn lead to cancelled subscriptions, as readers turn instead to PMC and institutional repositories for access to journal articles.

Our experiments with journals where we offer free back-archives provide further evidence to support the notion that subscription cancellations are likely to accelerate if free access via repositories becomes widespread. For those journals where we offer free access 12 months after publication, we saw a 2% drop in the number of institutional subscribers from 2002 to 2003. Where we offer free access after six months, we saw a much larger average cancellation rate of 6%.

If similar trends are seen by other publishers and learned societies, as repositories are expanded, it will be important to ensure that subscription revenue streams are protected if publishers are to continue to fund the process of peer review and other publication services that form the foundation of accepted practice to guarantee the quality and accuracy of published researcher findings.

Conclusion

It is all too easy to divide the world into 'goodies' and 'baddies', with the goodies being on the side of free and unrestrained access to and re-use of research information, and the baddies being intent on holding on to the old systems, collecting their revenues and turning a blind eye to the opportunities for change that our new networked world presents. The reality is, of course, that it is not as simple as this. Both factions have something to offer and a perspective to explore.

We are encouraged by the early results of our open access experiments with NAR and the *Journal of Experimental Botany*. We plan to continue with our experiments and to learn more about institutional repositories and open access, and the possible role that they could play in the future to enhance cost-effective access to research information. We look forward to sharing the outcome of our continuing experiments as further results become available.

References

1. Roosendaal, H., and Geurts, P., Forces and functions in scientific communication: an analysis of their interplay, *Cooperative Research Information Systems in Physics*, 31 August 31–4 September 1997, Oldenburg, Germany.
2. Richardson, M., 2004. Open access: evidence-based policy or policy-based evidence?, *Serials*, 2005 18(1) 35–37. Url: <http://uksg.metapress.com/openurl.asp?genre=article&id=doi:10.1629/1835>
3. Library and Information Statistics Unit, 2004. Findings taken from a LISU publication: *Journals Pricing Analysis 2004*. Electronic copies of the report can be freely downloaded from the LISU website at: www.lboro.ac.uk/departments/dis/lisu/pages/publications/oup.html
4. National Institutes of Health, 2005. *Public Access*. Available at: <http://www.nih.gov/about/publicaccess/>
5. Wellcome Trust, 2005. *Position Statement in Support of Open Access Publishing*. Available at: http://www.wellcome.ac.uk/doc_WTD002766.html
6. Research Councils UK, 2005. Available at: <http://www.rcuk.ac.uk/>

Article © Oxford University Press

■ Martin Richardson
 Managing Director, Oxford Journals
 Oxford University Press
 Great Clarendon Street
 Oxford OX2 6DP, UK
 E-mail: martin.richardson@oupjournals.org

Article received 9 May 2005; revised 9 June 2005;
 accepted 10 June 2005; published online July 2005.

To view the original copy of this article, published in *Serials*, the journal of the UKSG, click here:

<http://serials.uksg.org/openurl.asp?genre=article&issn=0953-0460&volume=18&issue=2&spage=98>

For a link to the full table of contents for this article, please click here:

<http://serials.uksg.org/openurl.asp?genre=issue&issn=0953-0460&volume=18&issue=2>