Open Access to Journal Content as a Case Study in Unlocking IP

Roger Clarke* and Danny Kingsley†

Abstract

The Internet has brought with it both means to disseminate and access content, and an enhanced expectation that content will generally be readily accessible. This has threatened entrenched for-profit activities, which have long prospered on closed, proprietary approaches to publishing, facilitated by anti-consumer provisions in copyright laws. The ePrints and Open Access (OA) movements have been complemented by the emergence of electronic repositories in which authors can deposit copies of their works.

The accessibility of refereed papers published in journals represents a litmus test of the extent to which openness is being achieved in the face of the power of corporations whose business model is dependent on the exploitation of intellectual property (IP). A specification of the requirements for “Unlocking IP” in refereed papers is presented and applied, leading to the conclusion that a great deal of progress appears to have been made. The copyright arrangements applied by most publishers enable authors to self-deposit PrePrints of their papers on their own web-

* Assistant is gratefully acknowledged from John Houghton, Arthur Sale and Colin Steele generally, and from John Lamp particularly in relation to the analysis of Information Systems Journals. The OA, ePrints and repository communities are responsible for most of the progress that has been achieved, particularly Steven Harnad, Peter Suber and John Willinsky, and staff at the Universities of Southampton and Nottingham.

† Principal of Xamax Consultancy Pty Ltd, Canberra; Visiting professor in the Cyberspace Law & Policy Centre at the University of New South Wales; visiting professor in the E-Commerce Programme at the University of Hong Kong; visiting professor in the Department of Computer Science at the Australian National University.

† Manager, Scholarly Communication and ePublishing, Australian National University.
sites and in open repositories; and in many cases authors can also self-deposit the PostPrint, i.e. the author's copy of the final version.

The theoretical success of the OA, ePrints and repositories movements has not – or at least not yet – resulted in success in practice. This is because only a small proportion of papers are actually self-deposited, and a large proportion of refereed papers continue to be accessible only through highly-expensive subscriptions to journals and journal-collections controlled by for-profit publishers. The unlocking of IP in refereed papers is therefore still very much a work-in-progress. Moreover, the gains may be ceded back to the for-profit publishing industry, unless concerted efforts are made within academe.

[This article was presented at the 'Unlocking IP' conference held in New South Wales on 16-17 April 2009.]

DOI: 10.2966/scrip.060209.234

© Roger Clarke and Danny Kingsley 2009. This work is licensed under a Creative Commons Licence. Please click on the link to read the terms and conditions.
1. Introduction

Access to content is vital, for a variety of cultural and economic reasons. The Internet ushered in a new era of accessibility, and in the process challenged existing publishing technologies and the business models that had grown up around them. Fifteen years into the revolution, a few market segments are sufficiently mature that retrospective analysis is appropriate. This paper performs a stocktake in relation to one such segment: refereed journals and journal articles.

While the journal is the primary formal publishing channel in many disciplines, in some it is not. The analysis in this paper is broadly applicable to some other collections of refereed papers, particularly refereed conference proceedings. However, the analysis does not encompass scholarly books, or works in the depictive and performance arts. The simplifying assumption is made that papers have a single author, in order to avoid the complexities of syntax that would otherwise arise.

It is useful at the outset to distinguish several broad categories of publishers:

- “Mutuals” – these operate as not-for-profit organisations, and in many cases are not even incorporated. Revenues from these operations are low or even non-existent, and most activity is subsidised, primarily by individuals and universities;
- For-profit corporations – these have become dominant in many segments;
- Segments of “learned societies” – this term encompasses disciplinary and professional associations and other forms of institutionalised scholarly communities that transcend the boundaries of universities. In some segments, these are dominant;
- Universities – a small number of journals are run by universities (particularly by law faculties).

The cost-profiles of these various categories were assessed by Clarke.\(^1\)

Learned societies may budget to run their journal publishing activities at a loss (by subsidising them from other revenues, such as membership subscriptions), at break-even, or at a profit. If they make a profit, the excess is available for allocation to other activities. For the purposes of the analysis undertaken in this paper, for-profit publishing includes not only publishing by for-profit corporations (referred to where necessary as “commercial publishers”) but also profit-making publishing by learned societies.

The paper commences by reviewing the nature of the journal prior to the widespread availability of the Internet. It highlights key steps in the interwoven processes of the

---

open access and repository movements and the adaptive behaviour of for-profit journal publishers.

A framework for the analysis is then provided, by identifying the key aspects of the unlocking of intellectual property (IP) necessary to facilitate open access to content (OA). The scope for genuinely value-adding activities by publishers is then considered.

The paper then provides a review of the current scene, focussing firstly on the OA credentials of journals as identified by the SHERPA/RoMEO undertaking, and secondly on an empirical study of journals in the information systems discipline. This is complemented by consideration of three specific cases. Finally, an assessment is provided of the extent to which the theoretical openness of access to refereed papers is being exploited in practice.

2. The Changed Nature of Journals

This section provides a brief review of journal publishing patterns in the period prior to and shortly after the electronic publishing revolution that was unleashed by the widespread availability of the Internet.

2.1 The Scene in 1995

In Clarke’s study, the interests that authors have in publishing papers in journals were distilled down to the following:

- Gaining the recognition of one's peers;
- Gaining citations in relevant literature;
- Developing personal reputation;
- Enhancing the likelihood of having research grant requests approved;
- Enhancing recognition by potential employers; and
- Contributing to organisational reputation.

Contrary to the assumptions sometimes made, most academics have little interest in wide dissemination of their work. What matters to them is that it is discovered, seen, and valued by the small group of targeted readers who the author believes can understand it, and who actually matter to the author.

The interests that consumers have in journal papers are quite different from those of authors. They were summarised in the same paper as follows:

- Gaining access to relevant information and ideas;
- Gaining access to exemplars;

---

• Avoiding barriers, such as the delay and inconvenience involved in negotiating a copyright licence;
• Sustaining a sense of community; and
• Gaining access to precise citations and references.

Journals need to service both the supplier and the consumer groups. Clarke & Kingsley\(^3\) described the functions of a journal in the period 1950 to 1995 as combining:

• A quality assurance (QA) mechanism;
• A publication channel;
• An archival mechanism; and
• A discovery mechanism.

The QA of individual papers has been (and continues to be) performed almost entirely by academics, although in some cases publishers have provided some amount of editorial assistance, directly and/or by funding administrative support staff working adjacent to editors. Archival aspects have been performed by libraries, primarily in universities and the public sector. Discovery of relevant articles has been to some extent facilitated by publishers but to a large extent by third party collection-consolidators and libraries. The primary contribution of publishers has been to the publication aspects of the undertaking.

Moreover, the QA, archival and discovery functions have been performed with little or no direct recompense by any party, i.e. primarily by subsidy from individuals, universities, governments and to a limited extent corporations. Publisher contributions have been small. Yet all revenues and profits have accrued to the publisher. (In a minority of cases, some profit-sharing has occurred, to the benefit of learned societies.)

During the period 1950 to 1995, the number of journals exploded. The number of scientific papers published annually has been doubling every ten to fifteen years for two centuries.\(^4\) To accommodate this, journals have increased – to some extent in size – but to a considerable extent in number. In recent decades, the sustained proportional increase has been on an ever-larger base, so the number of new journals launched each year or decade is now very high.

At the same time, the penetration of for-profit corporations into the field also exploded. Editors and learned associations had less understanding of the significance of IP ownership than did their service-providers. Legal control of a great many journals was ceded by the academic sector to the publishing sector. In the field of Economics, for example, by the late 1990’s, two-thirds of journals were controlled by

---


commercial publishers. As the industry matured, consolidation naturally occurred, and the concentration became so great that, by 2005, the top eleven publishers in the sciences were publishing more than 70% of the journals.

From about the end of the 1970’s, that dominance was turned into super-profits. Publishing companies used their monopoly advantage to progressively raise prices. They also bundled journals in such a manner as to maximise the revenue that they could extract from subscribers. This strategy was highly successful, and achieved of the order of 60-70% market penetration. Between 1984 and 2002, the price of science journals increased by a factor of nearly six, or over five times the increase in the Consumer Price Index.

The key subscribers were – and remain – libraries; particularly university libraries. They have felt enormous pressure on their budgets. They have also lacked the power to do anything about it, however, because universities have generally played quite limited roles in journal publishing, and access to journals by their staff was a quite fundamental requirement.

2.2 The Internet Era and Open Access

The Internet created a great many new possibilities for the dissemination of, and access to, publications. With it came high expectations, reflected in the term “electronic frontier” – as famously argued by Barlow. (For a review of the

---


7 Bergstrom, see note 5 above.


aspirations, see Clarke.)\textsuperscript{12} Critical among the expectations was greatly enhanced freedom of access to information, summed up by the call-to-arms “information wants to be free” – which was actually a corruption of the original, more complex statement “information wants to be free; it also wants to be expensive”.\textsuperscript{13}

Technical disciplines (particularly computer science and physics) were well-advanced in exploiting the opportunities by the time that the Internet became commonly available from about 1994. Since about 2000, all disciplines have been distributing material electronically. Many new “pure eJournals” have emerged, some existing journals have converted from print-only to electronic-only, and almost all of the remainder are available in both forms. In the Information Systems (IS) discipline, for example, the count of electronic-only journals grew from 102 to 131 between April 2007 and March 2009 – double the growth-rate of paper-only and paper-and-electronic journals. (See Exhibit 1.)

What had once been referred to by such terms as “working papers” became “ePrints”, defined as “the digital text of a peer-reviewed research article ... before and after refereeing”.\textsuperscript{14} Two kinds of ePrints are commonly distinguished, and a third, related form is usefully juxtaposed against them:

- “A PrePrint” – this is a stage of a work-in-progress. Multiple versions may exist, e.g. prior to and after one or more presentations, and prior to and after one or more rounds of refereeing;
- “The PostPrint” – this is the author’s own copy of the final version, which has been accepted by the journal and forwarded to the publisher. In practice, the PostPrint may be (or at least have the same content as) the last PrePrint, because it was accepted without change;
- “The Publisher's Copy” – this is the version that appears in the journal, and which incorporates the publisher's investment in presentation, production-editing and branding.

Authors may perform what is often misleadingly called “self-archival”, but is better described as “self-deposit”. Self-deposit can be in the author's own repository. For long-term accessibility, one or more institutional repositories are preferable, by which is meant here a repository run either by a university or some other research-oriented employer or by a learned society (as that term was defined earlier in this paper). This has been supported by the ePrints movement.\textsuperscript{15}


Uptake of the opportunity to self-deposit into institutional repositories has been slow, however. There have been some barriers to self-deposit – to some extent in the imaginations of authors – but also to a considerable extent real. The OA movement emerged to address the impediments. What is usefully referred to as “core OA” requires “free and unrestricted access’ to content without ‘price barriers’ or ‘permission barriers’”.

Automated tools have been developed to support management of the review process, in particular the Open Journal System (OJS). Discussions took place about business models to support “content commons”.

It took until about 2005 before governments and research funding bodies finally began to address the accessibility issue – and, indirectly, the monopoly pricing problem – by requiring self-deposit of ePrints into an appropriate repository.

2.3 The Responses of For-Profit Publishers

The renewed emphasis on openness represented an apparent threat to the entrenched position of for-profit publishers. They have adopted various measures to shore up the high levels of profitability that they had achieved.

Some responses have been constructive, particularly in the form of enhanced services such as search facilities, notification/alert services and auto-generation of hotlinks to cited works. These have offered value, but only within the particular publisher’s range of journals. From the perspective of each scholarly community, such boundaries are artificial. Moreover, some of these advantages are being neutralised – in part by for-profit companies in other industries such as Google, and in part by the high level of sophistication achieved by open source software and open access services.


Some responses have, on the other hand, been destructively competitive. The American Association of Publishers (AAP) hired a public relations consultant,\(^\text{20}\) which led to the launch of the Partnership for Research Integrity in Science and Medicine (PRISM). This partnership promoted to policy makers “the very real threat to peer review that ill-considered government interference represents, and [the need] to explore the ways in which we can safeguard peer review as a critical component of scientific integrity”.\(^\text{21}\) This was a seriously misleading assertion, because open access does not in any way threaten peer review, and peer review is in any case performed not by the publisher, but by the academic community.\(^\text{22}\) PRISM failed to publicly declare who it was, but it appears that AAP, Wiley, Elsevier and the American Chemical Society were key players.\(^\text{23}\)

A further response was the adoption from learned societies in the sciences of what is misleadingly called the “author pays” approach. This results in what is sometimes referred to as a (or the) “hybrid model” for journal-publishing. The choice is offered of: paying an up-front fee – permitting open access to that article; or paying no fee – in which case it is only accessible by subscribers. Springer launched the first such scheme by a for-profit publisher in 2004, at the price of USD$3,000.\(^\text{24}\) Other publishers have subsequently set a very wide range of prices. The SHERPA/RoMEO website offers a catalogue of publishers that offer the “author pays” form of OA, including their prices. The range displayed on 12 March 2009 was US$150 to US$5,000.

It is open to conjecture that an important motivation for introducing the “author pays” option was to deflect attention away from the primary postulate of the OA movement – that research papers should be much more readily and inexpensively available. The adoption level of “author pays” in commercial journals appears to have been very low, however.\(^\text{25}\) Some journals have resorted to an “editor’s choice” mechanism to ensure that at least some issues include at least one paper that is openly accessible. If the intention was to dissipate the energy of the OA movement, the attempt has had little success to date.

---


\(^{23}\) Giles, see note 20 above.


However, there are some indications that publishers are accepting the inevitable moves towards change. One initiative, Publishing and the Ecology of European Research (PEER), supported by the European Union, is a collaboration between publishers, repositories and researchers. This project, running from 2008 to 2011, is attempting to investigate the effects of the large-scale, systematic depositing of authors’ final peer-reviewed manuscripts on reader access, author visibility, and journal viability, as well as on the broader ecology of European research.26

Following a decade of desperate concern about the impacts of OA and repositories, it appears that for-profit publishers may be gradually coming to accept that the scene has changed. Openness is making inroads, but this has not to date threatened their existence, and they have sustained their high prices and high levels of profitability.

3. The Nature of “Unlocked IP”

This paper is a contribution to the research domain described by the term “Unlocking IP”. The focus of researchers in this domain is to contribute to understanding of policy formation concerning the accessibility of materials of all kinds. This section considers the properties of “unlocked IP” as it applies to refereed papers.

For content of any kind to be accessible, a number of conditions need to be satisfied. One of them is that the business models under which production, publication and consumption are conducted are sustainable. Most of the production and consumption activities have been, and continue to be, subsidised by individuals, employers and the state. Further, as discussed in Clarke,27 the technical difficulty and the costs of publishing, discovery and access have all decreased dramatically.

Despite this, some journals and publishers may struggle to survive if they suffer a substantial reduction in revenue. The problem is particularly serious for journals with high cost profiles. These arise variously because of inherently high costs (e.g. for clinical testing), and bloated organisations and inefficient practices arising from decades of assured revenue – including, during the last 25 years, super-profits arising from the successful exploitation of monopoly advantages.

Given a sustainable business model, the requirements of “core OA” were noted above as being “free and unrestricted access” to content without “price barriers” or “permission barriers”. In the terms used by free software pioneer Richard Stallman, the need is for content to be “free as in air”, but not necessarily “free as in beer”. However, such price as may apply must not be prohibitive for any relevant category of reader.

Analyses have been previously undertaken of the needs that the various actors have of copyright licences for PrePrints28 and PostPrints29. The broad licence specification was summarised as follows:

27 Clarke, see note 1 above.
28 Clarke, see note 2 above.
• The object may be copied in whole, but not in part, and each copy must carry the copyright notice with it, including the means of discovering the terms of the licence;
• Reproduction is permitted, but not for commercial purposes;
• Dissemination of copies is permitted, but not for commercial purposes;
• Adaptation is not permitted;
• There are no limitations on who can be a licensee;
• The licence is not time-limited, and is irrevocable;
• The copyright-owner provides no warranties or indemnities, other than any that may be imposed by law;
• The licence is gratis;
• The power to grant a licence is retained exclusively by the copyright-owner (i.e. no power to sub-license is granted).

For all categories of PrePrint, adoption was recommended of the least liberal of the Creative Commons licences, “Attribution - Noncommercial - No Derivative Works”.

The following is a proposed operational definition of the necessary and desirable characteristics of Unlocked IP, in a form that facilitates assessment of the policy statements of journal publishers.

3.1 Ownership of the Copyright in Each Collection of Papers

The term “collections of papers” is used here to encompass journal volumes and issues and such similar concepts as are emerging in the digital era, but also collections such as proceedings, books of readings, and festschrift.

_The terms of ownership must be such that the collections become and remain readily accessible._ This can be achieved by the organisation that performs the QA retaining ownership, or by the publisher acquiring ownership subject to conditions that ensure publication and long-term availability.

3.2 Ownership of the Copyright in Each Paper

For IP to be “unlocked”, _the author must retain sufficient rights in relation to the work._ This can be achieved in two broad ways:

• By the publisher acquiring copyright from the author, but granting a licence back to the author that is sufficiently liberal that ongoing accessibility is enabled; or

By the author retaining copyright, i.e. not assigning or transferring copyright in the paper to the publisher, but rather granting a licence to the publisher that is non-exclusive, but sufficiently substantial (and hence constraining on the author) that the publisher is able to operate an effective business (whether for-profit or not).

Further details on relevant aspects of such licences are provided in the following subsections.

### 3.3 Accessibility of PrePrints

*The terms of the agreement between author and publisher must enable PrePrints to be accessible.* By this is meant that prior versions up to, but perhaps excluding, the final version accepted by the journal and forwarded to the publisher, must continue to be generally accessible.

In practical terms, this needs to include the right to place PrePrints, and to leave PrePrints, in all of:

- The author's own electronic repository/ies;
- Repositories of the author's employer(s) or affiliated institution(s); and
- Repositories of the author's learned society/ies.

On the other hand, it would be entirely reasonable for the publisher to impose a requirement that the author link closely-related PrePrints forward to the Publisher’s Copy.

### 3.4 Accessibility of the PostPrint

*The terms of the agreement between author and publisher must enable the PostPrint to be accessible,* i.e. the author must be free to self-deposit the final version that was accepted by the journal and forwarded to the publisher into any and all of the categories of repository listed in the previous sub-section.

It is necessary (or at least highly beneficial for the publisher) for the author to accept a number of constraining terms:

- The author is to display an appropriate copyright notice (preferably with a link to the licence and/or further information about licensing). This may involve changing the identity of the copyright-holder;
- The author is to provide accurate citation of the published work; and
- To facilitate access to the published work, the author is to provide a link to at least the journal, e.g. the contents-page of the issue in which the paper appears. It is highly desirable that a reliable identifier be provided (such as the DOI) and/or that a link be provided to the published work, whether or not there are “price” or “permission” barriers that have to be negotiated in order to achieve that access.
3.5 Accessibility of the Publisher's Copy

Provided that the conditions specified in this section are satisfied, it is only desirable and not essential that the Publisher's Copy be freely accessible without “price” or “permission” barriers.

Useful though it may be, it is not even essential that the Publisher's Copy be freely accessible after the elapse of some period (such as six or twelve months after publication). If PrePrints and PostPrints are openly accessible, then market conditions can be left to determine whether or not access to the formal version of the work has a price-tag placed on it, and, if so, how much and for how long.

3.6 Copying

The terms of the agreement between author and publisher must permit PrePrints and PostPrints to be copied by any party.

On the other hand, copying may be subject to constraints such as limited numbers made by any one person, and no commercial use (e.g. for sale or use in marketing).

Provided that the conditions specified in this section are satisfied, it is only desirable and not essential that the Publisher's Copy of the paper be copiable, and that it be copiable after the elapse of some period (such as six or twelve months after publication).

3.7 Republication

The terms of the agreement between author and publisher must enable the PostPrint to be republished by any party in a subsequent collection, without any right of veto on the part of the original publisher.

This may be subject to reasonable constraints, such as proper citation, display of copyright ownership and where to seek a licence, and a fee – provided that it is sufficiently modest to not represent a barrier to publication.

4. Measures of Progress

The remainder of this paper applies the requirements of unlocked IP, as defined above, in order to assess the extent to which open access to journal content has been achieved to date. The first sub-section below considers progress in the enablement of open access through the adaptation of copyright arrangements, and the second sub-section presents multiple measures of the extent to which the new possibilities have been exploited.

4.1 Progress in the Unlocking of IP

No one measure is sufficient to define how far the liberalisation of copyright arrangements has gone. This sub-section commences by summarising the available information about the policies of publishers and journals generally, and then presents an analysis of progress within one particular discipline. Those “macro” views are then
complemented by three case studies, one each of a learned society which publishes journals as a service, of a for-profit corporation, and of a learned society whose journal-publishing is dependent on revenue from subscriptions and sales.

4.1.1 The OA Credentials of Journals Generally

For some years, the OA movement has documented the degree of openness of the copyright policies of journals generally. This is recorded in the SHERPA/RoMEO website. This currently categorises journals into four groups:

- Green – allows self-deposit of both pre-prints and post-prints;
- Blue – allows self-deposit of post-prints but not pre-prints;
- Yellow – allows self-deposit of pre-prints but not post-prints; and
- White – does not allow self-deposit.

The SHERPA/RoMEO site appears not to include an auto-generated summary. At the level of journals, the Open Access Newsletter reported in January 2008 that 91% of 10,000 surveyed journals support self-archiving. The ePrints site presents a statistical summary, which, in March 2009, indicated that, of 523 publishers: 51% were “green”; 12% “yellow”; and 37% “white” – but many “white” publishers were small. Hence, of about 10,000 journals: 63% were “green”; 32% “yellow”; and only 5% “white”. Even allowing for the mechanism being based on self-reporting by publishers, for some uncertainties and ambiguities in definitions, and for the total population of journals being much larger (somewhere between 30,000 and 50,000), it appears that, at a macro level, OA has achieved strong penetration.

4.1.2 The OA Credentials of Information Systems Journals

Within the first-named author’s primary discipline of IS, a comprehensive catalogue of relevant journals is maintained by John Lamp, at Deakin University in Australia. A previous publication included a summary table showing the then penetration of electronic accessibility and OA. Exhibit 1 below presents an expanded version, including more recent data.

Some inferences from the data are that, among IS journals:

- Paper-Only journals still constitute 7% of the total (unchanged), and are (and have to be) subscription-funded;
- Journals available in both electronic and paper forms remain 90% subscription-funded;

---


33 Clarke and Kingsley, see note 3 above.
The number of electronic-only journals grew about three times the rate for other categories (nearly 30% in two years, cf. 11%);

Of electronic-only journals, 77% are OA (up from 72.5% two years earlier); and

OA journals continue to gain ground on subscription journals, but:
  o Growth is slow (1.5 percentage-points p.a.); and
  o OA journals are mostly in “new” and “fashion” areas, and 90% of the longstanding, large and prestigious journals remain behind “permission and price barriers”.

Viewed from the macro perspective, there has been considerable progress in the adaptation of IP arrangements to enable greater openness, although the progress may be much less in the case of well-established, mainstream journals than in newer journals addressing recently-emerged sub-disciplines and research-domains.

It is therefore useful to complement the general impression provided by the macro-data with more specific assessments of the policies of particular publishers. A purposive sample was selected for this assessment. The first and third are learned societies with somewhat different approaches to their journal-publishing activities, and the second is the largest of the for-profit corporations active in journal-publishing.

*Exhibit 1: Penetration and Growth in OA in the IS Discipline*

<table>
<thead>
<tr>
<th></th>
<th>April 2007</th>
<th></th>
<th>March 2009</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Counts</td>
<td>Percentages</td>
<td>Counts</td>
</tr>
<tr>
<td>Paper-Only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper-Only</td>
<td>OA: 1</td>
<td>Sub: 39</td>
<td>Total: 40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic-Only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic-Only</td>
<td>OA: 74</td>
<td>Sub: 28</td>
<td>Total: 102</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>OA: 41</td>
<td>Sub: 382</td>
<td>Total: 423</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OA: 116</td>
<td>Sub: 449</td>
<td>Total: 565</td>
</tr>
</tbody>
</table>
4.1.3 IP and Proceedings of the US National Academy of Sciences (PNAS)

The US National Academy of Sciences (NAS) publishes several series of refereed papers, directly and through affiliated organisations. One of these is the Proceedings of the National Academy of Sciences (PNAS). As a “learned society” (in the sense used in this paper), it has a strong orientation towards open access.

From 1915 until 1992, PNAS did not require transfer of copyright. Indeed, PNAS only commenced registering copyright in collections (a peculiarity of US copyright law) in 1978. From 1993 to 2009, PNAS policy was to require transfer of copyright (equivalent to assignment under Australian law), but to provide a liberal licence back to the author.\(^3\) In comparison with the requirements identified above:

- Copyright ownership in each collection satisfied the requirements;
- Copyright ownership in each paper was largely satisfied, subject to two qualifications underlined in the following bullet-points;
- Accessibility of both PrePrints and PostPrints was supported, in any repository, subject to the proviso that the presentation version invested in by NAS could not be used;
- Accessibility of the Publisher's Copy was supported, in the form of a PDF of the paper, which could be made available in the author's own repository. The policy appeared to be silent about whether it was acceptable to self-deposit in one or more university and/or disciplinary repositories. (That is understandable, because the policy was established before repositories became a significant part of the landscape);

• Copying of PrePrints, PostPrints and the Publisher's Copy was permitted, for “personal use, including classroom use”, and for attendees at a meeting or conference at which it is presented. Commercial use required separate permission to be sought, and could require compensation; and

• Republication was permitted, although the only categories mentioned were “a printed compilation of [the author's] works, such as collected writings or lecture notes”. That left open the question of republication in books.

PNAS changed its policy again with effect from the beginning of 2009. Authors now retain copyright, and provide a licence to NAS. The effect of this was to further liberalise the terms, as follows:

• Self-deposit of the author's own “manuscript” (i.e., in the terms used in this paper, the PostPrint) is now expressly permitted in any “preprint servers such as arXiv” and “in [the author's] funding body's archive or designated repository”. This appears to deny self-deposit in other repositories (e.g. those of the employer and the author's discipline), unless specific funding was provided by them; and

• Republication for commercial purposes (phrased as a “request from commercial companies”) continues to require a formal request, and by implication NAS retains the rights to set fees and to veto republication. On the other hand, most authors would probably be very comfortable to depend on NAS to not let commercial considerations unreasonably override the benefits of republication.

• The new PNAS policy also referred to its “author pays” option (currently US$1,200 in addition to the base charge of $70 per page). It also expressly stated that it had considered the Creative Commons licence (presumably the US version of the least liberal one recommended). It cited as the reasons for not adopting it: first, the lack of clarity as to what “commercial use” means; second, the uncertain consequences of “unrestricted redistribution”; and, third, the lack of any means or even mechanism to “vindicate the author's rights if they have been violated”.

4.1.4 IP and Elsevier

The second case considered here is the world's largest journal-publisher, which has long earned massive profits from the business division. It might be anticipated by many casual observers that the conditions applied to authors would be among the least liberal of any publisher.


36 Clarke, see note 2 above.

Clarke, see note 28 above.
The company’s policy requires transfer of copyright, but it grants a relatively liberal licence back to the author. Much of the wording mirrors that used by PNAS, suggesting a common source for some aspects of the policy. The comparison with the requirements identified earlier is essentially the same as that for PNAS and presented in the immediately preceding section. The exceptions are as follows, with areas in which the policy may not satisfy the requirements highlighted by underlining:

- In relation to PrePrints, permission is granted to post “on Internet web sites including electronic-pre-print servers, and to retain [them] indefinitely”; and
- In relation to the PostPrint, permission is granted (and according to the reference has been since May 2004) to self-deposit to “the author's personal or institutional web-site or server [subject to citation and linking to the Elsevier site]” (emphases added). Elsevier clarified that:
  - “institutional” means “university”, i.e. a PostPrint can be deposited to a university repository but a PostPrint cannot be deposited to a subject or disciplinary repository without prior permission from Elsevier;
  - The “or” is intended to be inclusive, i.e. and/or, i.e. either or both of a personal web-site and a university repository;
  - The permission is not limited to a single university repository, i.e. a visiting professor with more than one university affiliation is permitted to place PostPrints in several university repositories;
  - Copying and republication for “commercial purposes” is defined to include posting for use by customers of a company, associating advertising with the PrePrint, charging for distribution or access, and distributing to parties other than known colleagues, e.g. via email lists; and
  - PDFs of the published version of the paper may not be posted to public web-sites.

In short, the Elsevier licence terms conform, to a very considerable extent, to the requirements defined earlier in this paper, but fall short in the specific and important area of self-deposit of PostPrints in learned society repositories.

4.1.5 IP and the Transportation Research Board (TRB)

The US has a National Research Council (NRC), which is jointly administered by the National Academy of Sciences (NAS, whose copyright policy in relation to PNAS was discussed in section 4.1.3, above), the National Academy of Engineering, and the Institute of Medicine. The Transportation Research Board (TRB) is one of six major divisions of the NRC.

---


38 Personal communication, 30 March 2009.
During the incubation period of this paper, TRB presented itself as an opportunistic case study within the learned society classification. The first-named author of this paper was a co-author of another paper, which was submitted to a TRB journal. Based on the TRB’s formal statements and correspondence, its policies are less liberal than PNAS in one key aspect (even though NAS actually provides publishing services to TRB). The explanation for this is that “TRB is a non-profit educational institution, and these activities require funds. ... The purpose is ... to protect TRB’s expenditures to make TRR Online possible. TRB is not making a profit from TRR Online; the goal is to make the service sustainable – that is, to break even”. 39

In comparison with the requirements of Unlocked IP identified earlier in this paper:

- Copyright ownership in each collection satisfied the requirements;
- Copyright ownership in each paper is transferred to TRB, and the conditions in relation to PostPrints fall short of the requirement, as highlighted by italics below;
- In relation to PrePrints, “[a]ny other versions of the TRR paper, before or after these versions submitted to TRB, can be published or posted as the author or copyright holder wills”, and hence may be placed in, and remain in, the author’s, institutions’ and learned societies’ repositories;
- The PostPrint is not permitted to be placed in any repositories. The same applies to the Publisher’s Copy. They may, however, be “posted to a restricted-access website; or the abstract may be posted to a publicly accessible website, with a link to the TRB Publications Index”. An exception is made where the source of funding mandates open access (which is common with research funding provided by US government agencies);
- The policy appears to be silent on copying, and hence to technically preclude it (although the practice might be more liberal than that); and
- Republication requires permission, but TRB’s policy is “to approve promptly any request from an author to reprint or reuse material published by TRB ... [and] if that book or journal is then published openly on the web, TRB will not refuse permission”.

This constitutes a “yellow” policy in the terms used by SHERPA/RoMEO, i.e. it is reasonably liberal, but not “green” (or “unlocked IP”) as is the case with PNAS and Elsevier. This more conservative position reflects concern that author self-deposit of PostPrints may threaten the publisher's business model, which is based on running at break-even based on subscription revenues.

4.2 Progress in the Adoption of OA

Broadly speaking, and with many specific qualifications at levels of detail, a great deal of progress has been made in “Unlocking IP” in refereed journal papers. In a preponderance of cases, copyright arrangements permit at least PrePrints to be

39 All quotations from personal communication, 27 March 2009.
published in any repository/ies ("yellow"), and in many cases they permit self-deposit of the PostPrint as well ("green" / "Unlocked IP"). Policies in relation to copying and republication also appear to be generally facilitative of ready access to research outputs.

Appropriate copyright licence terms are a precondition for OA and Unlocking IP. But the real test is whether papers are readily available. Indications of the current state of play in relation to the exploitation of liberalised copyright policies can be gained from several different tests.

4.2.1 PostPrint Volumes

One measure of adoption is the proportion of articles that are being deposited into repositories. This includes both self-deposit and deposit on an author’s behalf (e.g. through automated processes within a university or by the publisher). The repositories may be run by universities to reflect the output of their employees and associates, or by learned societies in a discipline or research domain (sometimes referred to as "subject repositories"). Authors’ own web-sites are commonly either ignored or deprecated in this literature.

Two recent studies of deposit-rates provide insight. Of Australian doctoral theses completed in 2005, only 12% were self-deposited into university repositories. Meanwhile, a much wider study of the open availability of journal articles published in 2006 came up with an estimate of 11.3%.

The deposit-rates vary enormously across universities, and across disciplines. Few universities have yet achieved high deposit-rates. In some areas of physics – particularly high-energy physics – arXiv has achieved a very high penetration-rate. In health-related disciplines, a high volume has been achieved by PubMed Central, run by the US National Institutes of Health (NIH); but NIH has had to resort to mandating self-deposit in order to lift its penetration-rate above 15%.

4.2.2 ePrint Availability

A second indicator of the success of OA is the proportion of the accumulated literature that is available in open repositories. This was informally investigated by means of searches in Google Scholar, in order to determine the extent to which copies of PrePrints or PostPrints were openly accessible rather than all copies being locked up behind publishers’ “price and permission barriers”.

The work is laborious, and hence a small sample was used. Searches using the names of authors were avoided (because authors vary in their practices). Instead, several topic areas with which the author had some familiarity were selected. The method used was to search Google Scholar for the term, and inspect each hit for which links

were provided. Where the entry offered additional links (of the form “all ‘n’ versions”), these were also inspected. The locations in which the publications were available were classified according to the scheme reported below.

The first-named author had recently conducted a citation analysis within the IS discipline.\textsuperscript{42} Two previously used terms were selected. One was a term of relatively long standing. The other was of more recent origin. The rationale for this was that a term of recent origin could reasonably be expected to only locate relatively recent publications, and a larger proportion of more recent publications, especially of paper (rather than book)-length, might be expected to be openly available.

The term of long standing that was selected was “information systems failure”. The first fifty hits on Google Scholar were inspected (some of which were citations only, without any links, and were ignored). The outcomes were:

- Forty-three publications with a total of 2,432 citations (This ignores one publication from outside the discipline, and that appeared to be a spurious result in any case, which had 1,979 citations);
- Three were unavailable in electronic form, including two major books (published in 1993 and 2002), with, in all, 454 citations, or 19% of the total;
- Of the forty articles, seven were available from an open source and thirty-three were only behind barriers;
- The seven open publications accounted for 571 citations, an average of eighty-one (but dominated by one publication with 380 citations, with the others averaging thirty-two each), and the thirty-four closed publications were responsible for the other 1,761 citations (averaging fifty-two each);
- Of the seven open publications:
  - One was open-only; and
  - Six were behind barriers but an open version was also visible.
- Of the seven occurrences of open publications (three of which were available from two different categories of sources):
  - None were on authors’ own web-sites;
  - Three were in a single disciplinary repository (Citeseer);
  - Three were in formal university repositories (at Brunel, Kingston and Loughborough, all in the U.K.);
  - Four were in informal university repositories (on instructors’ web-pages), in two cases more than one; and
- None were in OA journals.

The seven open publications were published in 1995, 1998, 1999, 2000, 2001 and 2002. Publications in the topic-area were primarily during the 1990’s and into the early 2000’s, suggesting that the sub-sample distribution is not markedly different from the sample distribution as a whole.

The term of recent origin that was selected was "reintermediation". The date of the oldest publication appeared to be 1997. The outcomes were markedly different:

- Twenty-eight publications with a total of 733 citations;
- Fourteen were available from an open source and Fourteen were only behind barriers;
- The fourteen open accounted for 485 citations and the fourteen closed for 248 citations;
- Of the fourteen open:
  - Three were open-only; and
  - Eleven were behind barriers but an open version was also visible.
- Of the eighteen occurrences of open publications (four of which were available from two different categories of sources):
  - Eight were on authors’ own web-sites;
  - Six were in a single disciplinary repository (Citeseer);
  - None were in formal university repositories;
  - Three were in informal university repositories (on instructors’ web-pages); and
  - One was in an OA journal.

The Association for Information Systems (AIS) operates an eLibrary that contains AIS journals and AIS and AIS-affiliated conference proceedings. The eLibrary is not formally open, although a login ID is provided to members (for a membership fee of US$125 p.a.). However, even within that small catchment, the eLibrary finds and provides access to forty-three papers that contain “reintermediation”.

Any generalisations from a tiny convenience sample from a single discipline must of course be expressed extremely cautiously. The following very tentative inferences are drawn:

- There are signs of a moderate degree of openness for publications since the late 1990’s, and probably at a level far greater than was the case prior to the mid-1990’s;
- In the case of older publications, there is little evidence of retrospective capture or opening up of previously captured but closed publications;
- The impact of two forms of institutionalised OA (OA journals and formal university repositories) may be less than another form of institutionalised OA (disciplinary repositories); and
• At this stage, formal OA may be little more significant than informal OA (authors’ own web-sites and instructors making copies available on web-sites intended for their students, but not subject to any permission barriers).

4.2.3 Journal Prices

A third measure of successful OA would be reductions in the prices of journal access, i.e. of subscription fees for individual journals, or for the bundles sold to libraries. Some journal-publishers, in announcing the availability of “author-pays”/hybrid journals, indicated that the subscription-price of journals would be reduced according to the number of OA articles that appear in the issues.\(^{43}\) To date, one report of a change of this nature has been seen: Oxford Journals claims that the online-only price increase for 2008 was only 1.7%, compared with the average increase for all journals of 6.9%.\(^{44}\) It may, however, take some time before it is apparent whether any widespread reductions will arise.

4.3 Progress in Australia

The Unlocking IP project has been conducted largely within Australia, and hence it is appropriate to consider not only progress generally, but also specifically within that country.

To the extent that OA is associated with journals and learned societies, most of the analysis needs to be conducted at the level of the world as a whole, because most journals and the most influential learned societies are universalist in their outlooks. The legal discipline is more parochial than most, however, because a great deal of legal analysis focuses on specific jurisdictions. In order to assess progress in law in Australia, the research method that was developed above and applied to the IS discipline needs to be applied to the Australian law journals indexed at the Australasian Legal Information Institute (AustLII).

There has been a recent surge in development of institutional repositories in Australian universities due to specified funding relating to government reporting requirements. Indeed, by 2009, all but one of the (\textit{circa} 40) Australian universities claim to have developed a repository.\(^{45}\) However the majority of these repositories have experienced a slow population rate, mirroring the experience in many other countries. The sole institution that has made considerable progress is the Queensland University of Technology (QUT), which identified its Digital Repository as a strategic factor in 2003, and published a policy in January 2004 mandating self-deposit. This


mandate did not have penalties for non compliance, however by February 2008, the repository had accumulated about 10,000 items.46

5. Conclusions

This paper has assessed the OA credentials of journals against an operational definition of the requirements, and found that the terms of copyright licences are currently relatively liberal. The expectations of openness stimulated by widespread availability of the Internet, reinforced by the ePrints, OA and repositories movements, have to a considerable extent succeeded in creating the appropriate legal context for open access and unlocked IP.

On the other hand, the exploitation of the opportunity has lagged, because of impediments to adoption, especially the lack of any positive incentive to self-deposit, and downright apathy. The outcomes to date are disappointing for proponents of OA and Unlocking IP. Only a small proportion of the literature is readily available, academics continue to be primarily dependent on the formal versions, academics-as-consumers continue to be uninformed and apathetic about self-deposit, and libraries continue to pay inflated prices to enable academics-as-producers to gain access to the papers that they collectively wrote and that they collectively quality-assured. There are limited signs of the adoption process speeding up sufficiently to deliver significant results. OA and Unlocking IP in the area of journal articles are at serious risk of being still-born.

The nature of journals, and perhaps to some extent of papers, is in some flux. The possibility of de-construction into separates is one possible direction. Hybrid models are being experimented with, in endeavours to sustain cash flow in an open access context.

One example, the “author pays” approach, can be seen as an attempt to diffuse the message conveyed by the OA movement, and sustain publishers’ high charges and profitability by collecting equivalent revenue at a different point in the industry chain. A range of other possibilities exist,47 including the Berkeley Electronic Press (bepress) notion of “quasi-open access”.

For-profit publishers might be seen as being under pressure to reduce their prices, and perhaps to enhance their value-add as well. Unless they can create new barriers to the open economy, the era of gross monopoly profits might be coming to an end, and the


new forms that are emerging might offer the benefits of the digital era at much more reasonable cost.\textsuperscript{48}

That outcome is far from assured, however. If author self-deposit of PrePrints fails to become the norm, then access to papers will continue to be dependent on access to for-profit publishers' unduly expensive services. On the production side, many mutuals and learned societies may lack the commitment and sustained professionalism necessary to perform the publishing function. If so, enough of the new e-journals may be vacuumed up by for-profit publishers, enabling them to resume their highly-profitable business as usual. Given the limited evidence of any price falls to date, perhaps they may even sustain their super-profitability in the interim.

A natural positioning step by for-profit publishers will be the offering to universities and learned societies of outsourced repository services. That will have the effect of denuding institutions of the technical capacity that they, and government research funding agencies, have invested in. But, faced with low adoption-rates and uncertainty as to whether the service is their 'core business' anyway, many may well give the business away \textit{gratis}, under the assumption (probably mistaken, but at best valid only in the short term) that they will thereby reduce their costs.

The period 1995-2010 has seen considerable progress in establishing the preconditions for Unlocking IP in journal papers, and to some extent in journals as a whole. Whether those gains are permanent depends on ongoing commitment by universities and learned societies, and stimulation of a far higher level of adoption of self-deposit by authors.

---