Introduction

The Internet is rapidly changing scientific publishing. Scientists have both the equipment and skills to produce and access material in electronic form at their workplace and have in fact been forerunners in using email and the Internet for knowledge sharing. This paradigm shift is discussed by, among others, Guedon, who stresses the importance to society and the global scientific community of the shift from print to electronic dissemination. Borgman makes the assumption that the relationship between publishers and libraries has become unbalanced with the advent of electronic publishing. The print paradigm involved libraries organizing the distribution of publications to clients and ensuring archiving. Now that the publishing process is dominated by large electronic site licenses for journals, and as scholars can exchange data with colleagues around the world quickly and in a convenient way through peer-to-peer networks, the role of the university library has become less obvious. However, the traditional refereed journal article, albeit in electronic form, is still the essential element in disseminating research results and the role of the libraries is mainly in handling subscriptions, licensing and access.

The shift from print to electronic format immediately affects the distribution process and reduces reproduction and delivery costs. It also affects the business models of scholarly journals. Within mainstream publishing, where journals now typically are produced both as paper and electronic versions, the traditional single title subscription is supplemented by bundled electronic licenses for multiple journals and pay-per-view access to single articles. The electronic distribution channel is eminently suited to both of these. The trend towards very comprehensive subscriptions (so-called 'big deals') has

The open access scientific journal: an empirical study

Turid Hedlund, Tomas Gustafsson and Bo-Christer Björk
Swedish School of Economics and Business Administration, Helsinki

© 2004 Turid Hedlund, Tomas Gustafsson and Bo-Christer Björk

ABSTRACT: The open access (OA) model for journals is compared to the open source principle for computer software. Since the early 1990s nearly 1,000 OA scientific journals have emerged – mostly as voluntary community efforts, although recently some professionally operating publishers have used author charges or institutional membership. This study of OA journals without author charges shows that their impact is still relatively small, but awareness of it is increasing. The average number of research articles per year is lower than for major scientific journals but the publication times are shorter.

Turid Hedlund

Bo-Christer Björk
caused much debate as to how advantageous these are for libraries and end-users.3

Among librarians there is a perception that the subscription prices of journals have risen faster than inflation for a couple of decades, thus causing 'the serials crises'.4 The assumption is also that this 'big deal' trend will continue because of the title monopolies held by a few large commercial publishers and the lock-in situation of customers.5 The cost structure of new journal pricing models has been discussed by Cox,6 and the market trends and pricing of electronic journals in mathematics have been reported by Kingma.7

Even before mainstream publishers began to produce electronic versions of their journals, pioneering scientists and groups of scientists seized the opportunities offered by the Internet for unprecedented new modes of producing and delivering scientific publications. These efforts can roughly be divided into two groups: electronic-only peer-reviewed journals, and subject-specific repositories where authors can upload manuscripts in different stages of the publication lifecycle, in order to disseminate the results more efficiently than through traditional channels. The best-known example of such a repository is the arXiv e-print server for manuscripts in high-energy physics and related fields, which already contains over 200,000 papers (http://arxiv.org/). Most of the repositories founded in the 1990s were subject specific. In the last couple of years there has been increasing interest in setting up institutional repositories run by individual universities or research organizations, since these are believed to offer substantial benefits for long-term sustainability.8 Recently the label ‘open access’ (OA) has been attached to both journals and repositories that offer universal free access to research publications. Currently there is an on-going debate among proponents of OA as to which route is more promising, primary publication in OA journals or parallel publishing in OA repositories. For the latter route to be successful on a large scale, issues of copyright and of indexing via search engines dedicated to scientific publications need to be resolved.

What is the current impact of the OA model on the totality of scientific publishing? In 2003 the cataloguing of OA journals was improved by the DOAJ directory of OA journals from Lund University. In February 2004, the DOAJ directory contained 745 journal titles, of which 100 titles were from BioMed Central. Scielo is a regional directory of OA journals in South America and is not yet included in DOAJ, so the number of journals registered today is close to 900 and the actual figure may well be nearer to 1,000. One of the problems with all the directories except Scielo is that they mainly concentrate on journals published in English and/or published in the Anglo-Saxon world.

The essential aspect of OA is that it is a model that would have been impossible, because of the considerable costs of print and distribution of marginal copies of publications, in the earlier paper-based world. It offers free electronic access to primary scientific knowledge, not only to the research community within the university but to society as a whole. The user is able to read, print and distribute the publication for non-commercial purposes, without payment. In order to study the effects of this change more emphasis should be put on the total life-cycle economics of the publishing process, including activities as research, publishing, archiving, indexing, retrieval and use.9,10

In this study the aim is to characterize the OA scientific journal in the framework of the changing publication model.

The structure of this article is as follows. In the next section we introduce the framework of the changing models for organizing journal publishing. We then present the research questions and the research settings for the empirical study of OA journals. The results of the study are given, followed by discussion and conclusion.

The changing models for organizing journal publishing

During the latter half of the 20th century scientific publishing evolved to its present stage. A number of commercial publishers, which were able to answer the rapidly rising demand of authors for publication outlets,
established major positions in journal publishing.

In recent years scientific journal publishing has, like so many other businesses moving to e-business models, undergone radical changes. It is indeed an area eminently suited for e-commerce, since the product itself is divisible into small packages of 10–20 pages of text, which can be read on the computer screen or printed out. The central issue is that the Internet offers a new type of technological platform with opportunities for innovation in the business model. (For discussions of such models, see for example Turban et al.11 and O’Brien.12) It also offers opportunities for non-commercial grassroots movements in the form of peer-to-peer networks where commercial intermediaries can be bypassed altogether. The term disintermediation has been used to describe this phenomenon, where whole stages in the earlier value chain can be bypassed.

In the early days of OA, most journals and repositories were organized on a voluntary basis by small groups of scientists, in a manner closely resembling open source programming projects, such as the one resulting in the Linux operating system. Since scientists had been used to working without pay as editors and reviewers of journals in the service of commercial publishers and society, it was not difficult to adapt to this production mode. And since these journals only appeared in electronic form, the activities involved with the printed journal only could be avoided.

Figure 1 shows the value chain for delivering a scientific article to its potential readers. The illustration contains both the steps of the traditional print/electronic model as well as of the OA model, and is based on a much more elaborate model developed in the European SciX research project.13,14

The essential differences between current mainstream publishing and the OA models as practised by most independent OA journal publishers are:

- OA publishing is usually much faster; the delays in publication to meet a regular issue schedule are usually avoided.
- Traditional publishing relies to a large extent on commercial indexing services for spreading information about an article to potential readers.
- OA publishing has until now mainly relied on general search engines as a means of ‘marketing’ their content to readers.
- In traditional publishing there is a need for an intermediary between the publisher and the readers in setting up subscription arrangements; this need is perhaps greater in the electronic environment when the library consortia are involved.
- In traditional publishing there is an...
opportunity cost to society in the form of many potential readers who do not get access to research results that would have been useful to them – either because of high subscription prices or the amount of extra effort needed to get access to a publication in another library.

Although the OA mode solves this last problem, potential readers may fail to find out about interesting OA articles because these are marketed efficiently only to a select community of researchers and because general search engines are rather inefficient tools for finding relevant and quality assured material.

The overall objective of using formal modelling tools for modelling the value chain of scientific publishing is to broaden the analysis of alternative models from the narrow question of subscription fees alone to the other cost and quality effects of process change.

After about a decade of experiments it has become evident that OA can produce successful journals and repositories, but also that it is very difficult to change the habits of academics, in particular as submitting authors. The obstacles for a major shift towards open access are formidable, as discussed by Björk, and the traditional subscription based model still dominates the picture. The success of the OA model is, according to Parks, dependent on finding the right incentives for the actors in the publishing process to move to it.

Recently a different type of OA journal has started to emerge. Thus publishers such as BioMed Central (http://www.biomedcentral.com) and the Public Library of Science (PLoS, http://www.publiclibraryofscience.org/) operate on an OA basis, yet they have the stable source of revenue needed for financing large-scale professional publishing activities. This new type of business model is based on the fundamental realization that the paying client of a scientific publisher is not the reader (and his/her representative the library), but the author striving for recognition, dissemination and long-term preservation of his/her scientific results. Publishers that provide these services can finance the operations through author charges or revenues from the organizations that finance the authors – universities, research funding bodies, etc. Even in the traditional subscription model one can argue that authors in fact do not give away their manuscripts for free to the publisher, but rather trade them in exchange for these kind of services. Some publishers have in fact also earlier levied page charges from authors for publishing their work.

Table 1 contains a classification of the business (or revenue) models that can be used to sustain the publishing of a refereed scientific journal. In some cases the business

<table>
<thead>
<tr>
<th>Paid access</th>
<th>Paper</th>
<th>Electronic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per article</td>
<td>Document delivery</td>
<td>Pay-per-view</td>
</tr>
<tr>
<td>Per journal</td>
<td>Traditional journal subscription</td>
<td>Electronic journal subscription</td>
</tr>
<tr>
<td>Bundled</td>
<td>‘The Big Deal’, e.g. Science Direct</td>
<td></td>
</tr>
<tr>
<td>Hybrid access</td>
<td>Delayed</td>
<td>The ALPSP journal Learned Publishing</td>
</tr>
<tr>
<td>Limited functionality</td>
<td>Read-only possibility</td>
<td></td>
</tr>
<tr>
<td>Individual article basis</td>
<td>The Oxford University Press journal Nucleic Acids Research</td>
<td></td>
</tr>
<tr>
<td>Open access</td>
<td>Community service</td>
<td>Majority of small OA publishers</td>
</tr>
<tr>
<td>Advertising</td>
<td>British Medical Journal</td>
<td></td>
</tr>
<tr>
<td>Grants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author charges</td>
<td>Public Library of Science Biology</td>
<td></td>
</tr>
<tr>
<td>Institutional membership</td>
<td>BioMed Central</td>
<td></td>
</tr>
</tbody>
</table>

...
The open access scientific journal: an empirical study

Research questions and research settings

In the study we aimed to answer three important research questions on OA journal publishing. (i) How prevalent are online OA refereed scientific journals compared to all online journals of the same type? (ii) What are the main characteristics of OA journals, e.g. which are the subject areas covered, what is the number of published articles, what are the main sources of funding and are the journals indexed in subject-based and citation indexes? (iii) To clarify the question of the cost structures of OA journals and, in the long run, the implications those structures have in the framework of the value-chain of a publication.

Data collection methods

The empirical data for the cost study consists of (i) a study of the prevalence of OA refereed journals, and (ii) a web survey sent to editors of OA journals.

The study on the prevalence of OA journals was performed in the summer 2002 in order to establish the number of OA refereed scientific journals on the Internet. The survey resulted in a list consisting of 317 OA journals and information on web addresses, subject and country of origin. The sources of information were two databases for journal information (Ulrich's web (http://www.ulrichsweb.com/ulrichsweb/), ARL (http://db.arl.org/index.html) and a directory of journals provided by Alison Wells (http://panizzi.shef.ac.uk/elecdiss/edl0001/index.html). The study also included a check if the journal was still active in 2002.

The survey to editors of OA journals was conducted in May 2003 as a web survey. The initial contact was an email sent to 300 of the editors on the list of OA scientific refereed journals established in the earlier part of the study. Due to missing and incorrect web addresses the whole list of 317 could not be used. Sixty answers to the questionnaire were collected; thus the percentage of responses amounted to 20% of the population.

The questionnaire contained five parts. The first part provided general information about the journal and the publisher. The
second part provided information about the average labour time the publisher spent (i) on general management and (ii) on the processing of an article. The third part provided answers to six statements regarding publishing, queue time and rate of publishing. The fourth part concerned intellectual property rights, such as copyright and copyright transfers. The fifth part was general comments by the editors on the questionnaire and on OA publishing.

Results of the study

Prevalence of scientific refereed OA journals

In the beginning of the study collecting basic data about OA journals was important, and 317 refereed scientific OA journals were identified from the sources used. The main research question was: what proportion of scientific refereed journals are available as open access? The proportion of OA journals in the most comprehensive database of journals, Ulrich’s, was established. This proportion was 1.5%.

The proportion of OA journals of all new refereed scientific journals founded during a particular year is higher if we look at the interval 1993–2002 (see Figure 2). The proportion of refereed scientific OA journals increased year-by-year over that time. Of the new scientific refereed journals started in 1993–2002, there is a steady rise in online distribution, related to the success of the World Wide Web. The number of new refereed OA journals as a percentage of all new refereed journals for the period 1993–2002 is shown in Figure 3. For the last three years 2000–2002 BioMed Central data is shown in the figure. The percentage of OA journals for 2002 is 32%, including BioMed Central; otherwise 22%. The source data are collected from the Ulrich’s web directory of journals in April 2004.

One concern regarding OA publishing is the sustainable development over time for journals. In a study by Gustafsson,17 the number of lasting titles of OA journals was established by comparing the years 1998 and 2002. For the situation in 1998, Gustafsson used data from the study performed by Wells.18 Of 387 journals identified as active by Wells in 1998 only 193 or 50% were still active in 2002. The results can be compared to a study by Crawford,19 which found that of 86 OA journals active in 1995, 49 (57%) were still active in early 2001.

The country of origin of scientific refereed OA journals reflects a clear predominance for the United States (53%). However, the United Kingdom and Germany as European countries represent together around 11%.

The subjects covered by scientific refereed OA journals are shown in Figure 4. The journals in the area of social science form around 40% of all journals, followed by arts and humanities, physical sciences and life
sciences with around 17% each. Technology forms around 8% and the category 'general' is 1%. Of single subjects, most journals were found in medicine and mathematics, with 36 journal titles in each.

**Characteristics of OA journals**

A common assumption is that scientific OA journals receive a large amount of funding from the publisher’s institution. The work of editing and publishing is to a large extent looked upon as voluntary and accepted as a part of the normal work of a professor engaged in research activities. As shown in Figure 5, the assumption was supported by the survey data. Grants are also one important source of funding. In some cases professional societies provided funding.

Advertisements, member fees and author charges were the main sources of funding in just a few cases. In the survey the participants were able to give more than one main source for funding.

The number of published articles is one of the quantitative measures of the importance of a journal to the research community. In the survey we asked the editors to provide data from the year 2002 on published articles. The large spread of the answers to this question was expected. The minimum number of published articles was 3 and the maximum 110. The average number of articles was 20 and the median was 16. Compared to the number presented in a study by Tenopir and King, where the average of traditional publishers is 123, the number of articles is relatively low. The average figure for BioMed Central for 2003 is 14 articles.

The number of articles is not, however, a measure of quality. The additional measure ‘rate of acceptance’ or the percentage of submitted articles that are actually published gives a more complete picture of a journal. For the OA journals participating in the survey the rate of acceptance was 50%.

The quality of a journal is also measured by the impact factor, where the impact factor calculations are based on the amount of citations to the articles in the journal. Impact factor calculations are produced by the Institute for Scientific Information (ISI). For a journal editor the impact factor is an important prestigious measure of quality and therefore it is vital to be indexed by the ISI.

Based on data provided in the survey only 10% (6 of the participating 60 journals) were indexed by the ISI. However, OA journals are new and the procedure to be included in the ISI index takes time.

Subject-based indexes are a measure of visibility as they are very widely used by readers to identify or locate articles. Of the participating journals, 57% (34 of 60) were indexed in one or more subject-based indexes.

Regarding the question on how a journal handles the long-time preservation of articles, the option of making their own arrangements was favoured by more than half of the population (58%), while 13% chose some other alternative, mostly mirror

![Figure 4 Subject categories of scientific refereed OA journals.](image1)

![Figure 5 Main sources of funding for scientific OA journals.](image2)
sites. Preservation arranged by the national library and joint arrangements with a library or disciplinary-based repository were chosen by only 10% (see Figure 6).

Cost structure of OA journals

Since OA publishing in its current form is presumed to be a voluntary activity, to some extent at least, we knew that it would be very hard to obtain exact cost data covering the whole publishing process. The difficulties in providing cost data were also mentioned and regretted by several of the participants in the ‘free comment’ section. Editing a journal is considered part of the ordinary work at a university and a measure of idealism is also a driving force for carrying out the editorial work.

The participants had three options for giving cost data; they could use one or a combination of all three. (i) They could give the information as direct expense numbers or as budgetary cost. (ii) As an alternative, they could estimate the time spent on a task. The last alternative was used in most answers. It is also the most relevant measure and best reflects the characteristics of OA journal publishing. The editor puts in a lot of his own time and this is not calculated as direct or budgetary cost. (iii) Instead of giving information on individual tasks such as administration, marketing, etc., the participants could estimate the total time spent running the journal.

Fifty participants of the 60 that took part in the survey provided cost data in some form, mostly as an estimation of the time spent on a task. Direct expenses and budgetary cost figures greater than zero were provided by relatively few, for direct expenses less than ten and for budgetary costs less than five. For most general management tasks the reported value is zero, which means that the editors state that they have no direct expenses or budgetary costs. If they could not answer the question the editors were asked to leave the answer field blank. The median values for time spent on general management tasks per year are shown in Table 2. The editor’s general estimate of the total time spent on general management is considerably higher. This option was used when the editor was not able to provide numbers for the individual tasks.

Due to the low number of observation of direct expenses and budgetary costs, the ‘time spent’ unit is used in the reporting. The time spent processing the average article is 22 hours (see Table 3).

Editors’ answers to statements regarding OA publishing and their free comments

In order to get a picture of how the editors regard central questions on OA publishing, six statements were made to which the editor either could agree or disagree. The statements and the answers are shown in Table 4.

A general assumption is that electronic publishing is a means to speed up the

Table 2 General management, measured by the unit ‘time spent per year’ for OA scientific journals

<table>
<thead>
<tr>
<th>Task</th>
<th>Median time (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration</td>
<td>50</td>
</tr>
<tr>
<td>IT-infrastructure</td>
<td>40</td>
</tr>
<tr>
<td>Planning issues</td>
<td>50</td>
</tr>
<tr>
<td>Marketing to authors</td>
<td>20</td>
</tr>
<tr>
<td>Marketing to readers</td>
<td>3</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>163</td>
</tr>
</tbody>
</table>

Editor’s approximation of the total time spent on general management per year: 250
publishing process from submission of the article to the actual publication date. In traditional publishing the queue time, i.e. the time elapsed from the time of acceptance of the article to the publishing date can be very long. The editors were asked to anticipate the time from the submission of an article to the publication date. The average was 5.1 months and the median 4 months, with a minimum of 1 month and a maximum of 18 months.

The question of copyright has been discussed mainly in the case of institutional repositories for pre-prints and post-publishing of articles. The transfer of economic rights of an article is normally not needed in the OA environment. On the other hand, for the protection of the authors’ rights a copyright statement of some kind on the website of the journal is needed. In the survey, formal copyright transfer was required by 32% of the journals and 68% had some kind of statement regarding copyright on the journal pages, thus protecting the authors’ moral rights. The use of, for example, copyright licenses like those developed by the Creative Commons were not asked about in the survey. An example of a copyright statement is as follows:

Authors that publish their articles in . . .

<table>
<thead>
<tr>
<th>Table 3 Average time spent processing an article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Article-specific activities</td>
</tr>
<tr>
<td>Editor’s review</td>
</tr>
<tr>
<td>Peer review</td>
</tr>
<tr>
<td>Technical editing</td>
</tr>
<tr>
<td>Placement of article</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4 Statements regarding open access publishing and the editors’ answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement</td>
</tr>
<tr>
<td>Most submissions are by the authors on their own initiative</td>
</tr>
<tr>
<td>We mostly invite authors to submit articles</td>
</tr>
<tr>
<td>Most submissions are to special issues of the journal</td>
</tr>
<tr>
<td>It is easy to receive submissions</td>
</tr>
<tr>
<td>Most received submissions maintain a high standard</td>
</tr>
<tr>
<td>It is easy to get reviewers to produce quality reviews in time</td>
</tr>
</tbody>
</table>
the help of editors and referees. The department supplies the hardware and software used. We buy some admin time.

**Analysis and discussion of the results**

In this study the aim has been to characterize the OA scientific journal and the first step was to discover how prevalent it was. The impact of this type of publishing is still small compared to all scientific publishing. However, the general awareness of this type of business model is rapidly increasing due to several national and international conferences on the subject as well as the emergence of pressure groups. The emergence of professionally operating OA publishers, such as BioMed Central and the PLoS, has also forced mainstream publishers to look at new business models. However, the critical mass in attracting enough high standard submissions and reviewers is not easy to achieve. The general awareness of OA publishing and the implications on savings in university budgets is a prerequisite for scientists to choose this channel for publishing.

The second research question was to describe the characteristics of OA journals. The typical OA journal is mainly produced as a single journal by an editor or publisher and is mainly funded from the institution of the editor or the publishing institution. The number of articles published per year is low compared to major scientific journals. However, the rate of acceptance for submitted papers, which is a qualitative measure of a journal, is 50%. Open access journals also cover quite a large area of science, representing medicine, mathematics, education, law, sociology, computer science, economics, history, biology and information science.

The third research question was intended to identify a cost structure and its implications in the framework of the value-chain of a publication. The constraint regarding this question is that a typical OA journal does not have a budget and actual direct expenses or out-of-pocket costs are reported in only a few cases. The unit used in this study was therefore the time spent working on a task. The data collected were grouped into the main categories of general management and article-specific activities. In order to compare production costs of OA journals with those of traditional journals, the measure ‘time’ could be transformed into monetary value by calculating the salary costs per hour for the editor and editorial assistants. The calculation of the overhead costs for using the space and computer equipment and network of the publisher’s or editor’s institution is an alternative to the direct expenses and budgetary costs of a traditional publisher. On the other hand, the time spent by the editor and the reviewers on the review process is in most cases not counted as a cost in traditional publishing. The elements of the value-chain, reported in Figure 1, queuing time for publishing and the subscription and access rights handling, are lifecycle costs that can be bypassed in the OA model. The main immediate impact of OA publishing would be in the reduction of subscription prices to library budgets.

The key question for OA publishing is whether it can be scaled up from a single journal publishing model with relatively few articles published per year to a comprehensive major journal with of the order of 50–100 articles annually. One has to remember, however, that most average estimates are based on counts in the journals included in major indexes like Ulrich’s and the ISI; if we include minor journals not listed in those indexes it is likely that the average number of articles published each year would be much lower.

The editor-in-chief has a very central role in OA publishing, in many cases acting as a pioneer. The continuation of the journal relies very heavily on the personal involvement of the editor and is as such a risk to the model. Employing staff to handle, for example, management, layout and copy-editing tasks, is a cost-increasing factor that also is a threat to the model.

The OA model in which an author charge or an article-processing fee is applied is a promising business model combining open access with a stable income for the publisher. Author charges can be compared with the page charges used by some journals in addition to subscriptions. BioMed Central publishes around 100 titles in biomedicine,
with a target author charge of US$500. The PLoS is an initiative in the United States where the author charge is calculated as US$1,500. Six of the OA journals participating in the study used author charges as a form of funding.

Institutional membership is considered as a replacement of author charge or article-processing fee in OA publishing. Institutions like universities or research centres take the responsibility to pay author charges for their researchers. This is the case for BioMed Central, where, for example, the Helsinki University has agreed to pay a yearly fee covering the costs for papers from its faculty. The Budapest Open Access Initiative of the Soros foundation has agreed to finance the fees for authors from developing countries.

In Europe there are numerous regional or national scientific journals published in some European language or in English, published with small budgets and public subsidies. This is a type of journal that would benefit from being published as open access but they would need support with IT infrastructure and advice.

The marketing of OA journals is also largely neglected. Directories of OA journals such as the Lund University Directory (DOAJ, http://www.doaj.org/) can be helpful in drawing attention to a journal and its articles. Branding is also extremely important from a marketing viewpoint. A key issue for marketing is proper indexing in subject-based indexing services and the successful implementation of the OAI protocol (http://www.openarchives.org/) for harvesting metadata.

References


Turid Hedlund, Tomas Gustafsson and Bo-Christer Björk

Swedish School of Economics and Business Administration
Helsinki, Finland
Email: turid.hedlund@hanken.fi