# Challenges to the role of publishers

**Mary Waltham**

ABSTRACT: *What future roles will publishers need to build on and develop given changes driven by the online system for the publication of research-based information? By reviewing some of the key issues currently faced by publishers online, this article is an attempt to provide a constructive framework for working through both immediate tactical pressures and priorities, and longer-term plans on product refinement and market needs. The areas of publishing affected by online availability are identified and contrasted with those that are not. Success rests on the ability of publishers to listen to and observe online users, to be open to the new opportunities that arise, and to be willing to adapt publications to provide value to the research community and libraries by making full and sound use of online as it becomes the dominant route for accessing research information.*

## Why does online publishing change everything?

The primary market for research information is dominated by individual researchers who through their publishing activities are both producers and consumers of information. Individual researchers are customers of the research information system in two different ways: they supply material to it in the process of disseminating research results and, at the same time, they use it as a means of learning of the research of others. This identity of supplier and user is unusual across business, and with the almost universal availability of the tools and medium for publishing online, the intermediary role of the scholarly publisher in distributing and packaging information may well be markedly reduced. Providing an independent mechanism for peer review, on the other hand, is a role publishers often stress and arguably becomes more important online, as does filtering and selecting information in a consistent and non-partisan way.

The opportunity to group content online to meet both the individual requirements of users and their information needs changes the traditional role of the publisher. It also promotes the publishers' need to co-operate online with other information suppliers who are in direct competition for the attention of the user.

Online access is always on in many countries[1] with the result that there is a growing requirement and impatience for prompt and accurate customer service round the clock. The high-quality consumer web businesses have set high standards already with the result that specialist publishers of research information are forced to continually improve their service levels in a medium that is both demanding and unfamiliar.

Online functionality such as linking and searching add value to the user well beyond

*Mary Waltham*

print, and at a basic level both are trivial to put in place. Unlike print references, hyperlinks, which are critical sources of value, do go dead – and key ones need regular checking for accuracy to avoid utterly frustrated users who will be unwilling to come back to the site. Unlike print, searching for information online is anything but absolute. Users all want comprehensive yet delimited online search but as with print, you never know what you are missing. Current search engines are poor for most research fields although Research Index™ [2] provides an excellent model for the computer science literature. There are some interesting new entrants to the specialist search engine market and in due course the accuracy and granularity of online searching will doubtless improve.

Online publishing systems are increasingly scalable – so that more content and functionality can be built into most systems – enriching the user experience. The cost of this continual growth and innovation is proportionately higher for small systems and while some large publishers have developed in-house solutions – remember publishers used to own printers – many medium and small publishers turn to a technology partner to provide this support. In the future it seems most likely that publishers' proprietary systems will evolve to deliver more neutral content formats to a wider range of platforms distributed in a broad variety of ways. Users want common interfaces and fewer different front ends through which to access content.

**Why is the online information market different?**

*Authors and readers*

The needs of authors and readers of research information are distinct and may conflict even though over time the majority of research readers are also authors. However, most publications have more readers than authors and so it is important for publishers to understand the priorities of each in building a coherent online strategy.

Authors want all their ideas, no matter how well formed, to be published in full and recognized as such. Online publishing makes

this an easier ambition to achieve and as a result pre-print servers thrive in some, but not all, disciplines, and increasingly research papers are posted on individual websites before or during submission to a journal. Open archiving initiatives encourage authors to common standards and broad open access.[3] None of this is new – it is simply easier and more convenient online.

Readers want to select and access distilled and well-written arguments. Publishers clearly have a role here in supporting and enhancing high-quality and clearly sorted communication formats. The need for this added value is exaggerated online due to the large volume of information that is neither of quality nor sorted.

Both readers and authors want an open system with few/no financial barriers to access that *they have to deal with*. This does not mean all information must be free – but that accessing it must be as simple and seamless as is possible. So what we see is the already rapid migration from print to online as the medium of choice for finding information quickly and efficiently. The large databases of journal articles have been particularly successful here and evidence from both the PEAK project[4] and Ohio Link[5] show that providing users with access to more online information (than they had in print) means that more online information is downloaded, although an article downloaded does not equal an article read. In the longer run, wider access to online information must have an impact on citation patterns and does put pressure on smaller publishers to be visible through larger gateways. (See Table 1)

*Free access to information*

What sorts of information are free? How does that challenge the role of a publisher?

Bibliographic information and increasingly research articles are free online. Consider HighWire Press at Stanford University which now has 450,000 articles free online out of some 1,300,000 articles in the database, so 35% of the database is available free. Pre-prints are free online – examples include the physics and mathematics pre-print server http://arxiv with some 198,000

**Table 1  What people want from research publications**

|  | Research authors | Research readers | Non-research readers | Librarians | Publishers |
|---|---|---|---|---|---|
| Quality | Y | Y | Y | Y | Y |
| Retrievability | Y | Y | Y | Y | Y |
| Permanence | Y | Y | ? | Y | Y |
| Speed | Y | Y | Y | ? | Y |
| Breadth |  | Y | Y | Y | Y |
| Relevance |  | Y | Y | Y | Y |
| Wide distribution | Y |  |  | Y | Y |
| Citability | Y | Y |  |  | Y |
| Low cost |  | Y | ? | Y | Y |
| Credit | Y |  |  |  |  |
| Online searchability | Y | Y | Y | Y | Y |

Adapted from Morris, S. Learned journals and the communication of research. *Learned Publishing* 1998:11, 253–258.

pre-print articles.[6] Course literature for 36 courses is now free online from MIT within the Open Course Ware initiative (and for 500 courses by September 2003). Everyone is trying to give away information online – so where does that leave publishers?

If pre-prints and research reports are free to the reader on major websites or on the author's website, then the value of that information for a publisher of research who is interested in selling the information falls, no question. This issue is particularly pressing as technologies enable smart searching for all free sources of an article. The trend also is speeding up under pressure and action from the author market and it forces publishers to review their fundamental strategy, and purpose.

One vivid recent example of this is the Public Library of Science group – who wanted publishers to make their articles available free on an open database six months after publication. The key point about the Public Library of Science was not in my view the pledge letter or the number of signatories, but the level of insurrection it started to brew afresh within research establishments at the same time as librarian groups were presenting coherent alternatives and tactical challenge – in the form of SPARC,[7] Create Change[8] and Declaring Independence.[9] Note, however, that signed-up members of the Public Library of Science group continued to publish in the highest impact and most visible journals even if – as

most – these journals did not agree to deposit their articles in an open database after six months.[10] Individual behaviour and competition for visibility and impact clearly override consensus here.

### Copyright and online postings

Attitudes to copyright in the online world are driven by authors who want their work widely distributed and recognized (see Table 1). Most print copyright assignments by authors to publishers do not permit the widespread online posting by the author of the work in question. The publisher used to be entirely responsible for content distribution when that was entirely print based, but the simplicity of online posting of an article and authors' desire to have their articles visible wherever they may get attention changes viewpoints on copyright. Certain universities have come out strongly to stop their faculty assigning copyright to publishers.[11] This has been brought to a head by research authors posting their latest key article online and publishers being thereby forced to state their position – which until recently has been 'take the article down'. Much of this appears to be bluster from the publishers. Publishers never had copyright assigned by government employees or from a substantial portion of book authors; did that damage the print publishing business? The question comes back to that faced by newspapers and magazines: what publishing rights must publishers have to run the business

*universities have come out strongly to stop their faculty assigning copyright*

effectively? Copyright is a 'nice to have' for most journal publishers not a 'need to have'. An exclusive license to first publication, with a non-exclusive right thereafter, is a good alternative solution for journals as it is for newspapers and magazines. Copyright assignments can be written to give both authors and publishers what they want and need.[12]

### Aggregation and linkage of online content

Publishers have tremendous power to aggregate content but many have been slow to deploy it. The initial entry cost of creating and sustaining a collection of online articles was high and tended to be limited to proprietary databases from the larger publishers. Fear of a perhaps too rapid transition to online and the economic consequences of inevitably losing print subscriptions was and still is also a factor especially when online business models are in flux. To address this, some rigorous analysis of print and online cost bases is required in order to understand which pricing models will work and which will erode the surplus from a typical publishing business as it makes the transition to online.

*online information… is purchased by readers in different and smaller units than print*

Trends to note here are for both more aggregations of information and more articles to be available from a number of different sources. Single-publisher sources are generally just not adequate in breadth and depth since they cannot possibly provide the comprehensive search mentioned earlier. Using single-publisher databases is not efficient for information seekers, who as a result are forced to move between databases to cover any one search topic. Cross-database search tools are being developed and these will ultimately help solve the user's problem. Of course abstracting and indexing services that are more comprehensive and do link into single-publisher sources are another solution. The CrossRef initiative is another promising one – but its true success from the users' perspective will rest on how many of the core publishers in a given field join CrossRef. For example, at the time of writing, physical sciences – an area dominated by a few large commercial and not-for-profit publishers – are well covered by CrossRef.

Coverage of subject areas with many small publishers is much less complete.

### Pricing models

Broadly, three markets exist for scholarly published information online:

- The institutional/organizational site license sold to libraries
- Individual or member subscriptions
- Pay-per-view article sales

The price of each of these has an impact on the others.

Libraries have until now tended to buy large packages of information – journal subscriptions and books, and now the entire publisher database ('The Big Deal') – yet readers are interested in smaller units, individual articles. Since the market is taking the lead in professional publishing, development of the single-article economy is surely the next clear step. Online information becomes more granular and is purchased by readers in different and smaller units than print.

There are no signs of stability in current publisher pricing models as many different experimental approaches are tried out and evolve year on year. Even in 1998 Stephen Rhind-Tutt[14] reported over 50 commonly used pricing models for electronic resources. Some recent trends are, however, clear.

Publishers of research information are beginning to see the savings from online-only publishing but only when their print runs start to shrink and often only then after they are challenged to save costs – by a governing or management board. First-copy print production costs are high and for most journals already published in print will obviously remain so until the publisher can declare with confidence that online only will in future be the format.

Publishers of relatively high circulation titles such as *Nature* and *Science* are dependent on advertising revenues to sustain their current business model. In print this rests on a low price for individual or member subscriptions and so a large number of such subscriptions is required. These in turn drive advertising revenues because of high circulation (advertising exposure) in a well-

targeted market. Loss of print subscriptions to this group of publishers presents a threat to revenue. Advertisers have not migrated their spending intact to the Web,[15] where in general there are too many pages and not enough advertising. Current online advertising formats simply do not achieve the impact of print.

Pricing of journals is flipping from a price for *print*, with online as part of that price or for an additional sum, to *online* as the base price with print at an optional additional cost.

Pay-per-view was initially too high-priced and clumsy for the individual user. The so-called 'non-pecuniary' costs of pay-per-view – passwords and credit card inputs – dramatically reduce participation.[16] However, the pay-per-view process online is becoming simpler for users, and some publishers have reduced the costs of pay-per-view to reflect more accurately a per-article cost. Given these changes, article sales can and and will evolve into a more sustained revenue stream and enable publishers to reach a broad range of irregular users who will pay for single articles but not entire journal subscriptions, and whose need for immediate information puts them off the often long-winded and time-consuming article supply systems. Individual article sales online allow publishers to reach readers (and librarians) with what they are interested in – the individual article – without having immediate access to the publisher's unit – the entire journal. The opportunity to reach new markets with such access is clear, but in such an article economy readers will be less aware of brand and journal title. A strategic consideration for publishers is how crucial brand name and journal title each are in maintaining the quality of their publications through author and editor recognition and how it is possible to sustain this within the so-called article economy.

Global pricing for online products is the expected norm with no distinction based on country of residence. Some publishers are pricing by country, even online. On a positive note there are numerous initiatives firmly in place to enable lower-cost online access for countries where gross national income is low.[17]

## Archiving and the journal of record

Publishers are driven to put all of their published content online and develop a formal archive policy because:

- the electronic library is a primary goal for many institutions;
- print-only materials are and will become invisible to the next generation – if not this; and
- the journal of record is the online version.

Only with a well-articulated and sturdy archive policy in place are librarians willing to abandon print entirely and so publishers are faced by a number of strategic questions, such as:

- Who should be responsible for the online archive – the publisher or a third party?
- Should everything published in the print journals be archived?
- What level of functionality should the archive have – complete reference linking?
- How far back should the archive extend – is citation half-life relevant?
- Who should have access and at what price?
- The cost of creating and maintaining an online archive is substantial – can that investment be recovered?

With a number of exploratory projects on archiving underway funded by the Mellon Foundation in the process of reporting their findings, it is too early to propose here a best practice position for publishers. However, the arrangement between the American Physical Society and the Library of Congress is one that meets many of the criteria librarians prefer for an effective archive.[18]

## Partnerships and alliances

As the volume of scholarly material available online is set to continue to grow because more publications are made available and more archived material is converted into online formats, most publishers find they want and need to build business relationships well beyond their own website.

Pressure from the market is encouraging some formerly unlikely partnerships. Cross-Ref – an alliance of some 160 publishers – is

*the pay-per-view process online is becoming simpler for users*

not one that would have been possible even six years ago (in 1997). Subscription agents such as Ebsco (as Ebsco Publishing) have evolved to become online content aggregators. Technology partners and service providers are evolving to become suppliers of additional services, which only become relevant and possible with collections of information. For example, HighWire Press has launched a life science portal to include all the journals hosted at HighWire. Ingenta has launched a number of vertical portals with additional information sources associated with each. Within both the HighWire and Ingenta sites former arch-competitors sit alongside each other on one site, with one interface and one search engine – because that is what users want.

### What stays the same online?

- *Information overload* characterizes the professional world of researchers. Access now to more information online increases the need for reliable resources that can provide depth and breadth and each of which can be controlled and quite tightly prescribed by the individual user.
- *Authors want* all their ideas out in their community as fast and as fully as possible.
- *Readers want* a more summative experience with well-written and concise information, which is presented clearly. Daily site updates are not a reader-driven feature, but the ability to tell the reader what has been newly posted since they last visited the site is useful.
- *Filtration* of information as characterized by a brand has always been the strong element in publisher identity and this becomes even more essential online for selecting sites to use and trust.
- *Quality content* means the publisher is adding value by selection, peer review and careful editing. Where relevant, publishers also often commission and publish additional material – commentary, analysis or review articles – which sets the research in context. This is in sharp contrast to the e-print server where very little screening is done before an e-print is published.
- *Commercial support* from advertisers and

*rumbles and roars of discontent from researchers and librarians*

sponsors who want the niche professional publishers can assemble around their content persists and this support will carry through to online versions if there are compelling reasons to do so and if the formats evolve. However, for smaller publishers this is likely to mean their content is part of an aggregated cluster of content since small sites cannot deliver the visibility and impact required.

### What changes in the transition to online?

- *Business models* are changing to more hybrid versions. Revenue streams are possible from any and all of: author submission fees, advertising and sponsorship, subscriptions – site licenses, individual subscriptions and individual document transactions.
- *Aggregation of content*, which is aimed at professional communities and integrates different types of documents or sources of information, such as books, monographs, journals, maps and charts, and so-called grey literature. Integration enables the user to follow a research thread across and between different sources through linking of the documents online.
- *Electronic-only products* are increasing in number. The successful ones are essential to the user because they provide support for research activity or do something online that is not possible in print.
- *The journal of record* is rapidly becoming the online version. In fact as a rule of thumb this clearly has already happened to any publication in which information that can be cited is available online only.
- *Insurrection* is the term I use to cover the rumbles and roars of discontent from researchers and librarians who want publishers to speed up and become more cost-effective. The Public Library of Science and the Budapest Open Archive initiative are both recent examples. But there are numerous others across virtually all disciplines. Some say these have always existed and they will melt away. From where I sit there now seems to be much more visibility and potential power accruing to the insurrectionists even if they stumble in their initial steps. They will not

be blown off course and have already changed the behaviour of some high-profile publishers.

- *More free access to* information. Time-scales differ for free access – free so long after publication (*Science*), or even free on publication (Institute of Physics) for a period, or free on the author's website. This trend points to a fall in transaction fees as research reports become a commodity available from a number of different, often competing sources. As search engines become more sophisticated at identifying sources rather than just sites, and the volume of openly accessible information continues to grow, the pace of this change will accelerate.

### Changing the role of scholarly publishers

Some of the ways in which publishers can evolve their role and continue to provide outstanding value – both to the research community and to libraries – can be articulated by addressing the questions that follow.

#### What do your customers want and how do you know?

Publishers need to get much closer to their customers – get to know them and interact with them as intuitively as Amazon.com or e-bay might, to find out what they really want and value, how they use the online product – not how they say they use it. A key question here is what are customers now online unable to do for themselves and so what must publishers now online do for them? The current burst of activity on reporting usage statistics[19] is one example. Looking beyond this basic service: do all customers want the same product or will different versions[20] at different prices provide a closer match between a non-elastic information budget, user needs and publishers' wishes for deeper market penetration?

#### Where do you add value to the publishing process and is this where value is perceived by customers?

A crisp analysis of where you add value is essential to compete as part of routine business management. The ability to confront where the value added is no longer necessary or even desirable online will in most cases be revealing and will encourage further change in how publishers operate and present themselves to their author and reader communities. The essential questions are tough ones. For example, how important is your copy-editing to your authors and readers? Given the development of PC-based spelling and grammar programs, are you in fact spending time and money where it yields less value than you perceive?

#### Which sales channels and markets will you cover yourself and which will you choose not to?

Only large publishers have sufficient resources to cover all their global sales potential. With the expansion of accessibility to new, underdeveloped online sales channels, publishers need to decide which to do well themselves and which to make other competent arrangements for. When did you last seriously analyse each channel and each market and take an informed position on where growth was either possible, or achievable with some additional support? Can digital delivery of your publications to your customers reduce costs and speed delivery as it has for Technology Review.[21]

#### What new products are you uniquely poised to create and develop?

If you succeed in honestly answering the questions here, you will gain insight into where your new product development lines may lie.

Many of the changes reviewed in this article imply a firm shift in the traditional publishers' focus to become primarily a service provider. This is plainly not an easy shift to make but it is likely to distinguish success from failure.

### References

1. See www.netsizer.com for current data on internet usage according to almost any parameter.
2. The Research Index site developed at NEC is at http://www.researchindex.com/.
3. The Open Archive Initiative is at http://www.openarchives.org.

*publishers need to get much closer to their customers*

4. Peak project articles are mostly at
   http://www.lib.umich.edu/retired/peak/.
5. Experience from OHIOLink is summarized in Tom Sanville's paper presented at IFLA in August 2001 at http://www.ifla.org/IV/ifla67/papers/118–141e.pdf.
6. Data on, among other measures, the number of pre-print articles submitted to the arXiv server are at http://arxiv.org/show_monthly_submissions.
7. Information about SPARC is at http://www.arl.org/sparc/home/index.asp?page=0.
8. The Create Change home page is at http://www.createchange.org/home.html.
9. The Declaring Independence home page is at http://www.arl.org/sparc/DI.
10. Update on the Public Library of Science: http://chronicle.com/free/2002/05/2002051601t.htm.
11. See 'To publish or perish' in *Policy Perspectives*, March 1998 (http://www.arl.org/scomm/pew/pewrept.html) and for the view promoted at the California Institute of Technology for a while see 'A provost challenges his faculty to keep copyright on journal articles', *Chronicle of Higher Education* 1998: 18 Sep.
12. For the ALPSP model 'Licence to publish', which allows authors to retain copyright in their own name and also provides publishers with the necessary rights to publish and to grant licences to publish to other entities, see http://www.alpsp.org.uk/grantli.pdf.
13. See CrossRef at http://www.crossref.org.
14. Rhind-Tutt, S. What a tangled web we weave. *Against the Grain* 1998:Feb, 24–28.
15. Article from *The Economist* summarizing the state of online advertising: http://www.economist.co.uk/business/displayStory.cfm?Story_ID=513133.
16. An article by R.S. Gazzale and J.K. Mackie-Mason on the impact of 'non-pecuniary cost' on the willingness of users to buy individual articles is at http://www-personal.umich.edu/~jmm/papers/gazzale.pdf.
17. Many of the initiatives to provide low or no-cost access to developing countries are listed at http://www.library.yale.edu/~llicense/develop.shtml.
18. The American Physical Society archive policy is at http://librarians.aps.org/archivepolicy.html.
19. The intense level of activity to provide usage statistics can be seen by following links from http://www.projectcounter.org/other_initiatives.html which gives the seven other current active initiatives about measurement of online usage.
20. Varian, Hal. Versioning information goods. Publications or 'information goods' have high fixed costs and small variable costs, so Varian argues that cost-based pricing makes little sense and value-based pricing is much more appropriate. He discusses the main principles of versioning publications for different markets at http://www.sims.berkeley.edu/~hal/Papers/version.pdf.
21. Tedeschi, *New York Times*, 5 Aug 2002, section C, 5. Technology Review is enabling subscribers to download exact replicas of the magazine that they can read at their leisure and is targeting foreign readers first since digital delivery allows publishers to get copies to distant readers without making them wait weeks for postal delivery.

**Mary Waltham**
*184 Springdale Road*
*Princeton, NJ 08540*
*USA*
*Tel: +1 609 430 0897*
*Email: mary@marywaltham.com*
*Website: www.MaryWaltham.com*

Mary Waltham founded her own consulting company in 1999 to help international scholarly publishers confront the rapid change that the networked economy poses to their traditional business models and to help develop new opportunities to build publications that deliver outstanding scientific and economic value.