Future Professional Communication in Astronomy
(Eds. A. Heck & L. Houziaux, Mém. Acad. Royale Belgique, 2007)

PUBLISHING IN THE NEXT FEW YEARS: A COMMERCIAL PUBLISHER'S PERSPECTIVE

HARRY (J.J.) BLOM
Springer New York
233 Spring Street
New York NY 10013-1578, USA
harry.blom@springer.com

Abstract. Over the past 15 years, internet technology changed the ways of publishing tremendously. It is truly revolutionary that both fresh and historic science publications are so much easier to search and find. This revolution has not been completed and all parties involved in science publishing are continuously adjusting their activities to the new rules and opportunities. From a commercial publisher's perspective, I will extrapolate what happens today to predict what happens in the next few years with journal subscriptions, book publishing, marketing, production and other steps in the publishing process.

1. Changes

Before the internet really took of, driven by the success of *NCSA's Mosaic* Web browser in 1993, the publishing market was quite straightforward. Publishers were getting most of their income from selling printed books and subscriptions to printed journal issues to individual libraries. However, at that time, a crisis was in full swing: library budgets could not keep up with the growing production of new titles and expanding journals. Librarians needed to make more and more cuts in their serial collections, cancelling subscriptions and standing orders. Publishers were subsequently selling their books and journals to a shrinking customer base. In order to keep up healthy returns, prices were increased significantly above inflation rates, increasing the pressure on those who kept paying, which triggered new cancellations. The scenario of a vicious circle.

Then came the internet, the growth of electronic publishing and library consortia. Gradually it became possible to share access to electronic sub-

scriptions without the time delays of inter-library mailings of printed issues. A familiar business model is the so-called *Big Deal*: a "one price, one size fits all" package which includes all journals of a particular publisher. It allowed libraries to team up into consortia, add all funds for purchase, and then combine collections.

Not that budget problems are history, but the result has been spectacular. For example, niche journals like *Solar Physics* had about 300 subscribers to the printed edition in 1998. Today, print is disappearing but its electronic articles can be accessed by thousands of institutes, basically taking away pressure on the publisher to allow free access via ADS, a few years after publication. Indeed, almost any researcher in this field is now able to download back-files and new articles from the publishers platform because his/her institute's library is part of a consortium that bought access to an entire package of physics and astronomy journals. This example illustrates much of the current situation for many fields of science.

1.1. LARGER SCALES

Commercial companies (but also many other not-for-profit organizations) strive for growth in order to reach higher revenues and stay in business, fending off competition. The pressure to grow increases in a rapidly globalizing world. The desire to gain negotiation power, to capture new markets, and reap up the benefits of larger scales, drives further consolidation of industries. So-called *horizontal mergers* (between companies in the same industry) and acquisitions have become the norm in publishing over the past 3 decades (see Mary H. Monroe's interesting timelines for the main academic publishing houses[1]). The most recent example is the takeover by John Wiley & Sons, announced on 17 November 2006, of Blackwell Publishing (Holdings), Ltd.

In the first two quarters of 2007, the global value of mergers and acquisitions (all industries) reached US$ 2.7 trillion (The Economist, July 2007). This unprecedented level of economic activity, combined with continuous GDP growth rates around 10% for the two most populous countries in the world (India and China), are two indicators that further consolidation will happen in order to be ready for the bigger scales in larger new markets. The growing importance of internet and database technology for publishing may trigger a *vertical merger*: e.g. a large publisher teaming up with an IT company. But so far, publishers appear to be buying smaller high-tech start-ups to acquire new technology.

Driven by much lower salaries that are, for the time being, the norm in emerging markets, the outsourcing of labor-intense activities to countries

[1] `http://www.niulib.niu.edu/publishers/`

like India, China and the Philippines is in full swing. For about the past five years in academic publishing, printing, type setting and customer services are moving to these countries. This trend will continue, not least because a different local market of end-users (scientists that are doing high-quality research and who start to publish locally) is emerging. They will request more often the proximity of other than the more basic services of a publishing house. Indeed, all big publishers are rapidly opening new Asian offices that include an editorial staff.

1.2. LIBRARY CONSORTIA

Just like publishers, libraries followed the same logic of joining forces and have set up consortia and associations, growing rapidly in the 1990's. Today, most of the larger ∼18 000 research libraries in the world are part of a few hundred consortia. While surrendering independence and flexibility, the gains in negotiation power (towards publishers *and* funding agencies) and increases in value-for-money by sharing collections have been significant. Many consortia of academic libraries are organized by state (e.g. OhioLINK in Ohio, PALCI in Pennsylvania) or by country (e.g. Russia, Canada, UKB in The Netherlands). This trend will continue, with national consortia crossing borders to become international.

1.3. DEVELOPMENT OF STANDARDS

When a small number of companies and organizations are dominating an industry, the push for standards increases. There is a true sense of opportunity to get things better organized, perhaps combined with a wish to control the business. Whatever the motivations, standards indeed do appear in publishing. I am giving three examples below, the third not yet fully operational:

There is *CrossRef* for citations[2]. *This is an independent membership association founded and directed by publishers. It operates a cross-publisher citation linking system that allows a researcher to click on a reference citation on one publisher's platform and link directly to the cited content on another publisher's platform, subject to the target publisher's access control practices.* Digital Object Identifiers (DOIs) to any piece of scholarly information are registered by CrossRef. From then on, the DOI will serve as an internet address to find this information online.

Another familiar standard is *COUNTER* for online usage (**C**ounting **O**nline **U**sage of **N**e**T**worked **E**lectronic **R**esources[3]). *This is an interna-*

[2] `http://www.crossref.org/`
[3] `http://www.projectcounter.org/articles.html`

tional initiative designed to serve librarians, publishers and intermediaries by facilitating the recording and exchange of online usage statistics. The use of online information resources has been growing exponentially and it is widely agreed by producers and purchasers of information that the use of these resources should be measured in a more consistent way. Building on a number of important, existing initiatives, COUNTER has set out to achieve this.

With mergers and acquisitions, journals change ownership. To prevent too much trouble with lost subscription lists, changing formats of publishers platforms and many other transition problems, publishers are now working on a *Transfer* protocol that may become operational in the next few years.

2. Trends in Journal Publishing

2.1. GROWING OUTPUT

The number of scientists in the world is growing. A popular claim is that the number of scientists doubles every 15 years (De Solla Price, 1986) and it is therefore no surprise that the total scientific output is increasing. Given that scientists publish more: ∼0.5 million articles in 1988 to ∼1 million articles in 2006 (in ISI ranked journals[4]) the number of articles read per scientist also doubled on average over that same time period (e.g. Tenopir & King, 2003).

Where is the strongest growth coming from? There is a strong correlation between the scientific article output and Gross Domestic Product (GDP), which is partly related to the allocation of higher budgets for research. It is indeed observed that fast-growing economies have the highest growth in scientific output (Rowlands and Olivieri, 2006). Mara Hvistendahl (2005) writes for example: *between 1988 and 2001, article output grew by a factor of five in China. In the same period it only increased by 10% in the U.S. According to research done by Rice University, by 2010, if current trends continue, more than 90% of all scientists and engineers in the world will be living in Asia, and many of those in China.* Clearly, the strongest growth is currently coming from Asia.

2.2. CHANGING BEHAVIOR

The amount of time available to do research, to write, and to read is not keeping up with the growing output in science, so researchers seek and find efficient ways to handle information. Specializing in ever smaller subfields helps. But if we focus on publishing, we see that no longer visiting the physical library saves time, using powerful search engines and databases like

[4] `http://scientific.thomson.com/isi/`

ADS saves time. Also, ignoring articles that are of no immediate interest saves time. Indeed, people no longer *browse* through printed information. Scientists search for specific information and are often satisfied with a few *accessible and trusted* search results (hits) that provide the requested details. The push for harvesting knowledge more efficiently will continue as long as the scientific output keeps growing exponentially. Perhaps only a revolutionary knowledge database like Rudolf Albrecht is suggesting (this volume) could break this trend completely.

2.3. ELECTRONIC PUBLISHING CONTINUED

The growing scientific output combined with the speed and ease of electronic publishing predicts the end of *printed* journals. Indeed, subscriptions to print still decrease by about 5 – 10% per year. If paper disappears, *pages*, *issues* and *volumes* become meaningless and online article-based publishing remains.

Despite resistance among both librarians and researchers who have become attached to the idea of page numbers, this trend is unstoppable. It is already hard to refer to multi-media using page numbers, but one could still point to the pages where they are first discussed. But it is the push for online publication immediately after an article is ready that is driving this trend. It takes too much time to wait for the complete volume contents to be ready such that page numbers can be assigned. Perhaps only for topical issues the coherence and order of articles still has additional value. But for general science articles, aggregation into issues and volumes has become a redundant step. The end of page numbers and paper means a restructuring of reference formats. Article numbers with a time stamp (month, year), or DOIs will take over.

2.4. MULTI-MEDIA AND LINKING

Most electronic articles can still be printed without losing any information. However, an electronic article can include multi-media features such as a movie, high resolution images, or a dynamical graph that allows for different parameter values. Linking references directly to literature, but also words and terms to dictionaries, observations to databases, and objects to catalogues can enhance the value of an electronic article significantly. Such links can be added and updated continuously as more related information becomes available online. The connected layers of information present a "3-D" article which can no longer be printed on 2-D paper without losing information. Despite questions about proper peer review of multi-media items and about the safe long-term archiving of such articles, this is a trend that will continue.

2.5. OPEN ACCESS

The serials crisis discussed in paragraph 1, triggered initiatives to remove the grip of the publishers on journal content, who are blocking access to articles for those who cannot (or will not) pay for subscriptions. Again, the internet came to rescue, making it easy to post articles online, freely available (*open access*) to any person with a connection. The pre-print archives and postings on personal home pages were the first examples. This was followed by the launch of subject-specific repositories. Indeed, in physics and astronomy, the problem of not having access was quickly solved by the arXiv pre-print server set up in 1991 by Paul Ginsparg at Los Alamos National Laboratories, NM, and the launch of ADS in 1992 at the Harvard-Smithsonian Center for Astrophysics in Cambridge, MA.

In other fields of science, however, the lack of access to information was perhaps felt a bit stronger. An Open Access movement emerged, stating that publicly funded research should be publicly available, or the "tax payer should get what he paid for". This led for example to the launch of peer-reviewed open access journals by the Public Library of Science (PloS) in biology and medicine in 2003. The UK-publisher BioMed Central started publishing with open access in 1999. These early initiatives, while perhaps not yet financially stable, have clearly been successful in attracting top quality articles that are frequently cited. Some research funding agencies (e.g. the Wellcome Trust, NIH, CERN) demand that science results they have sponsored should appear online, freely available to everybody, at least after some period of time. Traditional publishers have responded to such desires by setting up both hybrid (e.g. Springers Open Choice$^{\mathrm{TM}}$) and fully open access models for specific journals. Nowadays, almost one new open access journal is launched every day, many of which will not survive...

The business model of open access is based on article fees paid by the author, or the funding agency who sponsors the research of the author. It is therefore, somewhat incorrectly, called the *author-pays* model. Critics of this model mention the barrier to authors without funds (in developing countries) and not accounting for the peer-review costs of articles that are, or need to be, rejected. In addition, productive institutes may be paying much more than a traditional subscription with free publication would have cost them. But most discussions focus on the proper amount that needs to be paid per article to cover all costs, allowing for a surplus to stay in business. The current fees range roughly between €1000 and €3000. It will depend on an increasing number of factors how much one will pay, and a wide range of article fees will likely emerge.

3. The Future of Books

3.1. EBOOKS

The internet allows access to complete electronic journal article collections. The most important research journals have now been back-digitized to volume 1, issue 1, and many libraries have bought access to back-file packages to complete their collections (sometimes catching up again were they had to break serials when cancelling subscriptions). Nowadays, it is mostly the book content that is still "locked up" in print-only format, so digitization projects have turned to books.

Making knowledge accessible to the public (or to research communities) is popular and valuable, so government and private funds are now stepping in to create and sponsor electronic books, or eBooks (see for example Project Gutenberg[5]). This is an exciting development and complements the purposes and idealism of Wikipedia. It is a no-brainer to predict that the publicly available eBook collection (fiction and non-fiction) will grow exponentially over the coming years.

But commercial eBook initiatives are taken up (again) as well. As an employee of Springer, I witness this from within a company that owns copyright to the largest academic book collection in the world. The first serious business models to sell books electronically (eBooks) were introduced about five or six years ago. In fact, this introduction was not very successful for various reasons (unnecessary and expensive file encryption is one). But now that Google points so easily to the proper book when people are searching for particular phrases, the demand and usage of eBooks goes up *and* (very important) drives the sales of traditional print books. It now becomes more profitable and less risky to sell eBooks, especially when sold in packages to big library consortia, which often allows a lower price per title compared with the traditional print list prices.

So, will printed books follow the fate of journals and disappear in the next couple years? Probably not. Apart from the often greater coherence of information between back and front cover, it is the emotional and cultural value of a well-produced tangible book that will keep the printed book on the shelves. Not so much because paper makes a more comfortable read. Younger generations seem to have less trouble reading larger parts from screen. The eBooks serve a different and complementary need; specific pieces of information are much easier to find electronically and selected parts are printed for further use or reference. This is similar to how journal articles are used. We will therefore see that science eBook chapters will start to resemble journal articles. They will have their own abstract

[5] `http://www.gutenberg.org/`

and references, which are no longer collected after the final chapter. Peer review may be introduced for academic book chapters to further enhance credibility of science books. This is also driven by better online tools, in which article submission, peer review and production workflows are much faster and better linked.

3.2. PRINTING ON DEMAND

Besides the growing importance of electronic books, there are also developments in the production of traditionally printed books (and journal issues). For a number of very specialized titles, off-set print-runs have gone down to less than 400 copies, reflecting a rather low global demand. This forced publishers to seek cheaper alternatives for printing niche books. Digital printing at a fixed unit cost, albeit initially at a lower quality, provided a solution. But now that the quality of digital printing (including color and half-tones) is rapidly improving, other big advantages are looming on the horizon. What if a new book can be printed on the day it is ordered, at a printer close to where the buyer lives? That would eliminate the warehouse and its large stocks, reprint delays, mailing costs, and long waiting times for shipment. It also allows for flexibility in correcting mistakes in the master PDFs. Digital printing (and fast delivery) on demand will become huge in the near future.

4. Changes in Promotion

In STM publishing, book and journal promotion has been a relatively small-scale (and perhaps undervalued) activity. The traditional goals of attracting more articles and book authors has been the job of active editors and increasing sales has been the job of dedicated sales reps, or intermediary agents. Many publishing houses didn't go much beyond printing flyers and brochures that were mass-mailed to librarians and scientists (the classical *push-marketing*). The "target" addresses were (and are still) bought from e.g. learned societies that keep their membership lists tidy. The results of a mailing (and other) actions were often disappointing. The very lukewarm response to deep-discount codes, or even free books, leads time-and-again to the conclusion that flyers are apparently mostly going straight to the paper waste basket upon arrival. Now this is a somewhat exaggerated picture and print promotion has not disappeared. But the lack of being able to measure and prove direct success limits the appetite for allocating healthy promotional budgets. This is changing, thanks again to e-mail and internet combined with a new goal: reaching a high *usage* (i.e. many PDF downloads).

Nowadays, it is key that publications are easy to find online. Apart from

some unique scientific discoveries, most information, even at an academic level, is available in various formats written by several different authors. A search on Amazon using the phrase "X-ray astronomy" returns more than 700 book titles. If you are the publisher of a book on X-ray astronomy, better make sure your book is shown on the first page with search results. Book promotion has become an art of anticipating how people perform online searches. Just like the brick-and-mortar library, the physical bookstore is losing customers. This is especially true for academic books. A particular piece of information can be need-to-have for research and the purchase of a book that shows it is what we call a *destination sale*. Your local bookstore may still thrive on *impulse buys*, books bought because you were looking for random fun stuff to read, or to give away as a present. But people who are looking for less popular, particular information, will go and search online.

Online marketing techniques are therefore rapidly evolving. The so-called *pull marketing* becomes much more important. Here the communication starts with the customer (who is searching). One should make it easy for the customers to *first* find your company and publications when a need for information arises. For example, book information online should provide unique but relevant titles, clever and accurate key-words and links to related topics. Try and search for your own publications online. It can be both embarrassing and pleasantly surprising.

The traditional *push marketing* is shifting to e-mailed newsletters and table-of-content (ToC) alerts. These go to people who have specifically requested such information, with some degree of choice in message format, content and frequency. Technology allows an accurate measurement of how these e-mails are treated by the receivers, but not at the individual level to protect privacy. Open-rates, click-through rates, forwarding rates and unsubscribe rates, provide immediate feedback on the success of a mailing. Fine-tuning your messages and online information have become key factors for the successful promotion of publications. As marketers are now learning to follow the shifting attention to the online world, I predict that the battle to appear on the *front Google page* has only just begun.

5. Measures of Quality

It is evident that the continued growth in scientific output, combined with the increased accessibility to both new and archival information can lead to information overloads. Questions like "which new article is most relevant?", "which article has highest quality?" and even "which scientist is currently producing the most interesting results?" become more imminent. The days that one could read *and browse* through the latest journal issues and then be fully up-to-date on all that matters in astronomy, are gone.

Top-quality journals with highest *impact factors* are still the first ports of call for astronomers and physicists. Peer review is still regarded as an essential filter against the publication of incorrect or redundant results. But changes are coming. In particular in cross-disciplinary fields like biophysics, scientists are becoming extremely selective and do not want to waste time by checking the table-of-contents of journals. In high-energy physics, many new articles are no longer submitted to a journal, because peer review takes too long. A posting in arXiv is deemed sufficient. So if the name of a researcher, an institute, or a top journal is no longer a reliable measure, what else is?

The answer may come from community developments that are enabled by the internet. For example in biology, where the *Faculty of 1000* [6] a large group of top scientists, recommend the best articles that they have read. This selection is faster than waiting for many citations to an article in other articles. Some journals have online blogs attached to their home pages, in which readers comment on latest results.

A different type of immediate measure could be the download intensity (or *usage*) of an article. Even though "bad" science may get a lot of (temporary) attention as well, an article's download intensity over time (peak and decay) could be an interesting indicator of quality. Usage is already the new currency for librarians who want to measure the relative value of their journal subscriptions (see the COUNTER protocol described in 1.3). Publishers would have to prevent artificially high downloads, but it is my prediction that such new measures of quality will emerge to support the needs of the modern impatient researcher.

6. Conclusions

When internet technology became a tool for the masses, electronic publishing revolutionized the market for disseminating (academic) information. This revolution is still ongoing with companies and organizations growing bigger to gain power to operate effectively on larger scales. This is driven by economic growth, especially in emerging markets and causes further globalization. As a result, business models, distribution, production, and marketing activities all change rapidly to adjust to the new opportunities of the global publishing market of tomorrow.

Acknowledgements

I am thanking Prof. André Heck and Prof. Léo Houziaux for organizing a very interesting workshop, bringing together a unique group of people

[6] http://www.f1000biology.com/home

to discuss the future of publishing in astronomy, and for inviting me to present my thoughts. I am thanking my colleagues at Springer for discussions and presentations that taught me some of the ideas and terminology I am presenting in this article.

References

1. De Solla Price, D.J., 1986: in Little Science, Big Science . . . and Beyond, Columbia University Press, New York
2. Hvistendahl, M., 2005: What happens if Science is Made in China?, Seed Magazine, Oct/Nov issue, http://www.seedmagazine.com/news/2005/11
3. Rowlands, I. & Olivieri, R., 2006: Overcoming the Barriers to Research Productivity, Publishing Research Consortium, London
4. Tenopir, C. & King, D., 2003: http://www.dlib.org/dlib/october03/king/10king.html