

# The Impact of Electronic Publishing on the Academic Community

## Session 1: The present situation and the likely future

### The impact of information technology and networks: new perspectives for scientific, technical and medical publishing

Arnoud de Kemp\*

Springer-Verlag, Tiergartenstrasse 17, D-69121 Heidelberg, Germany

\*For correspondence or a bibliography of further readings, please contact: [dekemp@adkathome.de](mailto:dekemp@adkathome.de).

©Kluwer Academic Publishers, 1997.

[Copyright Information](#)

---

#### Abstract

This contribution can only be a small collection of ideas and experiences from my (personal) publishing point of view. The subject area assigned is very generic and there are indeed many developments going on, so I had to be very selective and restrictive, while avoiding redundancies as much as I could. I have included some metaphors and paradigms, such as the shift from print publications to electronic information. The traditional role of publishers in the information chain is compared with the new opportunities that electronic publishing is offering now and may offer in the foreseeable future. From desktop publishing to document-type definitions, unplugged and unbundled information, linearity, appropriateness, packaging and customizing, filters, intelligent agents, quality, integrity and authenticity are some of the items hidden in the text. The overall conclusion is that the Internet still can learn a lot from print.

#### What publishers do?

Most scientific, technical and medical (STM) publishers are small organizations or small units in much larger media companies, but publishing as a whole is a big

business with a long tradition and an interesting historical background. The publishers in Europe originally evolved from typesetters and printers, who imitated Gutenberg's invention, and in keeping their stock, became booksellers and later publishers.

Many societies in this part of the world publish in close cooperation with professional publishers, whereas in North America society publishers are more common. The American Association of Publishers (AAP) divides publishers into school, higher education, trade, professional and scholarly publishers. In this contribution I will concentrate on the last of these categories, also known as scientific, technical and medical publishers.

There is an International Association of STM Publishers, registered in the Netherlands, affiliated with the International Publisher's Association (IPA) in Geneva. They jointly support the International Publisher's Copyright Council (IPCC). These organizations and the AAP have been instrumental in developing and adopting rules, e.g. the SGML (standard generalized markup language) rules for manuscript preparation. One of the most important activities in 1997 has been the introduction of uniform information and/or document identifiers, in addition to ISBNs (international standard book numbers) for books, ISSNs (international standard serial numbers) for serials and ISMNs (international standard music numbers) for printed works of music. Such identifiers are critical for the identification of original works and for the success of electronic commerce on the Internet and other networks.

Publishers are not at all homogeneous and the industry is not standardized, although many common rules exist. The traditional publishing process is complex: it is a complicated system with many different specialists. In the view of an increasing number of scientists as well as information specialists, and in the perspective of changing information technologies and available network capabilities, publishing is costly, slow, inefficient and even a hindrance to the free flow of information. They feel that it needs to be replaced and electronic publishing could provide all the solutions.

One of the most basic activities of publishing seems to be almost unknown or simply ignored. Publishers, whether they are privately owned companies, enterprises, societies, or university presses or departments, finance and administer this publication process. They sell books, journal subscriptions, translation rights and other rights and licenses, and acquire advertising and sponsorships. Society publishers, controlled by their membership, may use part of their dues to offer publications as a service to their members at a more favourable rate. However, all of them have to make that difficult decision whether to publish a work or not. The decision not to publish sometimes seems to be the more difficult. The costs of handling unsolicited and rejected material can be substantial.

The basic roles of publishers are concentrated in the following areas; selecting, reviewing and filtering information. This information may be written on invitation, at the request of a colleague, editor or publisher, as a work for hire or it can arrive unsolicited. Journal articles fall into the category of unsolicited material, but the number of articles and/or reviews that are written on invitation should not be underestimated.

Here we see a first bottleneck. It takes time to review and discuss a new intellectual and/or artistic work until the decision to invest in it and publish it, with an estimation of all the additional work that it entails, has been made. Although electronic mail may speed up the communication process, the intellectual work done by anonymous peer reviewers and in-house editors does not change at all. Some journals may have rejection rates far exceeding 50%. The rejection is always based on quality (relevance), the scope of a journal and its size. Not, as I heard recently from a prominent scientist, because of a lack of paper.

There is some debate about how many times rejected articles are resubmitted for publication, but there are no empirical studies available. From publication experience with some major scientific handbooks in the area of chemistry and physics, I know that the redundancy in primary journal publishing is high and seems to be growing. There is no system available for filtering redundancies at the time of submission and review.

Since the reviewing process is not transparent and nobody likes being rejected, there is a broad discussion of the obvious advantages of unlimited publishing on the Internet. Indeed, this social aspect of free, global and instant communication is one of the most important impacts of the Internet. But we should not forget that the Internet was developed for improved communication (and for secure information in wartime, since it started as the ARPANET network) and does not simulate or even replace the publication process. I often show the (now famous) cartoon of two dogs in front of a computer, "On the Internet nobody knows that you are a dog!" (*The New Yorker*). Most of the information on the Internet at the moment belongs in my opinion to communication, marketing and promotion. From the perspective of official publishing, there is too much unplugged and unbundled information, which is increasingly difficult to retrieve.

Serious attempts to build databases with critically reviewed and officially published full text (written for the sake of the public record) have only been started in recent years. Most of this 'new' information is still based on print publications. Publishers select quality by screening. They add quality in the publication process, both in the printed world and in the digital age. Their buyers and subscribers expect high quality

at a critical cost level. The name of a publisher and the title of a publication can work as a quality stamp.

Other criteria, especially among librarians and information specialists, are continuity and availability. For electronic information, new criteria include accessibility and retrievability.

Electronic publishing is not a new activity. Publishers and authors already concentrated on text editing and processing (input) in the 1970s and 1980s. STM published an interesting survey in 1995 and listed up to 120 different systems, with Word and WordPerfect being the most popular. The last ten years have given us new means of output on diskettes and CD-ROMs. The near future will concentrate on the mode of delivery: print, CD-ROM or online.

Authors are not typesetters and it is a dream to expect that all future manuscripts will be supplied in one global format with perfect SGML coding. The successful integration of text, images, formulae, video and sound requires new skills and teamwork right from the beginning. The modes of delivery should be discussed as early as possible and all information should be collected, engineered and stored in digital archives that allow publishers to be more flexible. But it is still a long way from DTP (desktop publishing) to the complicated world of DTDs (document-type definitions) that are essential for SGML.

Journals, ever since the first publications in 1665, have been and will continue to be very efficient vehicles for the dissemination of selected content. In the future, publishers should carefully indicate what the material is and this requires a new typology of publications and manifestations, which has consequences for targeting and presentation activities.

Springer-Verlag has made a first attempt with its online LINK service. All bibliographic information and abstracts have been pulled out and made freely available as meta-information. The journal (and book) content is defined as: E for electronic publications (electronic *sui generis*); P for print-based publications; and S for multimedia supplements. A combination of E and P indicates an electronic version of a print publication.

In this way all those concerned with information management and archiving know that there is still a paper archive available in libraries. (Interestingly enough, Springer-Verlag makes annual archive editions of its electronic journals as a book with a CD-ROM for the database, including multimedia components).

## **Targeting and presenting information**

A publisher selects and collects relevant new information for defined targets or user groups. They may be the members of a learned society, a community of scientists (that may have started their 'own' journal as a community organizer), the subscribers of a journal or a specified group of scientists, professionals, students or that large group of interested laypeople. The content is then prepared for better reading and understanding within that group. A book for freshmen in medicine requires a different presentation to a reference work with mathematical formulae for physicists.

Apart from vocational requirements in the editorial process, there are also quality requirements in the production process, e.g. better paper for medical illustrations, a larger format, different type sizes and different fonts. It is in this area that the Internet still has to learn from print. In the online environment we are still heavily handicapped by the limitations of ASCII, the development and implementation stage of SGML, the Internet browsers, helpers and plug-ins. We are only at the beginning of typography. It makes little sense to present full text as a scanned image or a PDF (portable document format) file. Computer screens originally were designed for viewing, not for reading. For this reason we still seem to print on local printers. If there is one major paradigm shift worth mentioning, it is the transition from central printing to local or remote printing.

There is one other symbolic aspect of the Internet that is not often discussed: the hypertext link! This demonstrates an important difference in presentation and a dramatic difference in cross-referring between the paper and an electronic version. I am referring to those hypertext links, usually highlighted in light blue, which are clickable. I am in favour of hypertext links as navigation tools (but they do not bring me back!) as long as they are used critically. However, they might also prevent us from reading an entire article and distract us, like zapping on cable television.

In this context I should remark that a sheer transition from a linear paper-based text to an online information product is not the state of the art anymore.

A parallel CD-ROM edition makes sense for archiving purposes and personal documentation. Online information should serve different purposes such as full text searching, linking and rich indexing, and offer far more functionality, beginning with interactivity and enhanced by simulation, visualization, video and sound, etc. Here we see major improvements over print. Multimedia components and supplements, three-dimensional presentations, annotations, moderated forums and active or dynamic documents are all welcome additions to the typology of publications. Electronic products and electronic versions of print products greatly support accessibility and retrievability, important barriers to the use of information in the printed world.

**Making the work known and available: marketing, promotion and distribution**

Making the (unsolicited) work known so that it will be bought, subscribed to and cited in other works is another important task for publishers of all kinds. This is not an easy task as all those who have started a journal themselves will know. Marketing, sales and distribution require considerable resources and often special skills. Here we hope to see the return on our investment within a reasonable period. Classical marketing and promotion with brochures, advertising, sample copies and exhibitions are now greatly facilitated by means of home pages, online catalogues, news and informative texts, online ordering and useful addresses, etc. The availability of sample pages and cover designs has supported book sales in at least one major project (MEDOC) that I know of.

Since the Internet opened to commercial communication, publishers have used it for exactly this purpose. They still lag behind academics and professionals, who started using the Internet for the dissemination of information and research much earlier.

In general, there is still a long way to go from classic World Wide Web to full text and multimedia database management systems. All such work may be limited in the case of unsolicited articles that are published in a journal pipeline, but copyediting, typesetting or converting to another formats or fonts, etc., proofreading, printing on paper, binding, warehousing, packing and shipping still need to be done. Additional costs include those for marketing, promotion, exhibitions, sample copies and review copies. Hidden costs include trade discounts for booksellers, subscription agents, wholesalers and other middlemen who assist in transporting the physical information product within the information chain.

### **Invisible and infinite information**

The biggest problem with electronic information is not that it still requires so many different tools or helpers: i.e. browsers, viewers, players, display software and sound blasters with too many different official and industry standards. The real problem is still its digital rather than physical character. In our cultures we still refer to hard copies or photocopies. The whole process of refereeing and citing is based on the public record. A book or journal is a hard copy of a master. But electronic documents are something else and they can be transmitted everywhere as a clone of something that has existed or still exists somewhere else. We cannot know when and where, since no references are made. Nobody can check its originality, integrity or authenticity anymore, unless we encrypt the information. (Fraud and misuse already take place even in the printed world).

These developments will have a major impact on the information chain as a whole, up to and including archiving, preservation and conservation. Lending between libraries or borrowing then become definitions from the past. National copyright and

intellectual property laws need to be changed and become global. How do we handle interactive books and dynamic articles which are living documents? We not only need new typologies, but also new rules in our global information society. This goes beyond information technologies and the networking of information and must be taken up as a general task for society.

### **Discussion following presentation by de Kemp**

Lindquist asked how much of Springer's publishing business was now in electronic form. De Kemp replied that electronic business accounted for 10% of the company's current investment but less than 1% of turnover. Neuhold considered that charging systems for meta-information would have to change as the material was expensive to supply. De Kemp commented that Springer had, to date, adopted a deliberate policy to supply this information free of charge to the user. Celis asked whether publishers were collaborating in the development of compatible electronic media. De Kemp referred to a number of mechanisms enabling the publishers to talk to each other, but it had to be recognised that different publishers were working at very different scales. In particular the requirements for the electronic publication of books and of journals were very different. In Springer, 120 journals were currently available in electronic form; this would rise to 200 by the end of 1997 and 400 in two years time.

---

Reproduced with permission from de Kemp, A. (1997) *Astrophysics and Space Science* **247(1/2)**, 11--16, Kluwer Academic Publishers, Dordrecht. ©Kluwer Academic Publishers, 1997.

[Copyright Information](#)