The Impact of Electronic Publishing on the Academic Community

Session 1: The present situation and the likely future

Introduction

Ian Butterworth

The Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2AZ, U.K.

It was David Magnusson who suggested that the academia should have some workshops on the impact of information technology (IT) on society in general, and it seemed sensible that we should first think about the impact on the academic community itself. We therefore decided to focus on electronic publishing.

Most of the 80 or so people at the workshop have an involvement in electronic publishing at some level or other. All of us, as academics, are concerned with publishing in general, for scholarly publishing has created the modern academic world. It has not only been the means for transmitting and assembling knowledge, but has also been the means by which academic status has been determined. On a personal level, whether you got onto the faculty of a university or you were elected to bodies like the Academia Europaea depended on your publications. The same is true at the institutional level. In Britain recently, the University Funding Council announced the results of its second research assessment exercise which ranks university departments. A department ranked grade 5 gets 20% more funds than one rated grade 4, and that grading is primarily based on the publication record of its faculty. University departments are being closed as a result of low grading. So, not surprisingly, changes in publication methods are of great importance to us.

And so there was a hidden agenda for this workshop. The main purpose was, of course, as in most workshops, to inform one another and to bounce ideas off of one another. But a second purpose was that in reporting our deliberations, we can hope to influence the thinking of the members of bodies like the Academia Europaea. Like all learned societies, such as the National Academy of Sciences, the Royal Society of London, or the Royal Swedish Academy of Sciences and so on, its members have considerable influence. They network, in the human sense, with the right people; their views are sought by national governments and by the European Commission. But there is a problem. By and large the members get their distinction with age, and the exciting,

new things are done by the young. That is true in any subject, but it is particularly so in the case of the computing and IT areas.

I am sure that IT is changing practices at a speed with which most academics are completely unaware. The World Wide Web was first announced by CERN (the European Organization for Nuclear Research) in 1991, and 2--3 years ago most people had not heard of it. By the end of last year there were 300\000 Web sites. In my own particle physics research group at Imperial College the young PhD students never, but never, look at a printed journal. They get all their information from the display screen. If they don't know how to do something they mainly contact a colleague, if possible in the same room or next door, but if not they e-mail colleagues in Bologna, Hamburg or Amsterdam. These young people are in the happy position of being in one of the largest physics departments in a subject involving strong international collaboration, and they can all access powerful workstations. This is the way the future will soon be in all subject areas. And it is important that workshops, such as the present one, inform the thinking of influential bodies within the academic world and indeed alert them to the problems and dangers.

In order to set the scene for our discussions, the workshop started with a review of current activities and well-defined plans for electronic publishing in various subject areas. A general overview of electronic publishing was given by Arnoud de Kemp, who chairs the Innovations Committee of the Association of Scientific, Technical and Medical Publishers, and was followed by reports of what was happening in physics (Dixon), astronomy (Boyce), molecular biology (Sander; the speaker provided no text for this publication), computing science (Sandewall) and social studies (Harnad). Conscious of a particular and different set of issues involved in clinical studies, we decided at this workshop not to cover electronic publication of medical research.

Enthusiasm for electronic publishing is naturally strongest in subjects where handling of large sets of data, or involvement in international collaborative work, has made researchers familiar with networking, and it is the learned societies in these areas which tend to be leading the way. Thus, Anne Dixon told us that the Institute of Physics, the first publisher to put all its journals, 30 in number, on the Web, gives the on-line version free to anyone subscribing to the printed version, an approach now adopted by the American Physical Society and the American Institute of Physics. Peter Boyce reported that nearly half of the world's peer-review literature on astronomy is now available in electronic form, and that after two years of experience the interlinked resources around the *Astrophysical Journal* are creating a working digital library.

Somewhat surprisingly, the computer science community has been moving rather slowly in the use of electronic publishing, apparently because communication in the discipline is dominated by the use of conferences. However, Erik Sandewall described a new, purely electronic journal in the area of artificial intelligence.

Much of our discussion was, of course, influenced by the success of the Los Alamos electronic archive set up by Paul Ginsparg as a virtually automatic and cost-free 'publishing' operation, originally for theoretical particle physics, but increasingly covering other areas of physics. Some scientific communities, and particularly those in physics and astronomy, have for many years widely distributed copies of research papers at the time when they were 'submitted for publication', but these so-called paper preprints have virtually vanished, replaced by the

electronic preprint or e-print which is submitted via the Internet to the Los Alamos archive. Authors are given simple instructions on what electronic forms are acceptable. Processing, which includes giving the paper a reference number, and its subsequent appearence on the Web is automatic so that the papers immediately become universally available. Some five or so colleagues around the world allow their computers to mirror the Los Alamos server as a local archive to reduce network delays. Again, this daily copying is done automatically. The process is incredibly cheap.

In theoretical particle physics, the main method for reporting scientific progress is now overwhelmingly through the 'archive', thereby ignoring traditional scientific publishing whether by learned society or commercial publisher. Ginsparg argues that peer review can also be at zero-cost, the impact of a paper simply being democratically assessed as it circulates the Web. The consequences of such an approach are discussed in depth in the report of Session 3.

For most of us at the workshop, enthusiasm for electronic publishing is not particularly because it will be cheaper, but because of the new opportunities it brings. The new electronic journals already offer facilities unavailable to users of paper. As pointed out by Dixon and Boyce, these include: (i) the ability to search for particular content within a paper or a complete journal; (ii) the ability to move easily between the table of contents of a journal and the abstract; (iii) facilities to annotate articles to create a personal virtual filing cabinet; and (iv) the ability to move quickly from the article you are reading to one which it cites in its references, and that includes forward citing, i.e. the ability to go to articles that cited the one you have just been reading but which were published later, etc.

Such electronic linking is in principle easy within a given journal or within a stable of journals from the same publisher, provided, of course, that electronic versions are available. As Boyce reported, the American Astronomical Society has digitized into electronic form all the articles in the last 20 years of its journal which can now be accessed, but most publishers are going back only five years. But the reader really wants to be able to move seamlessly between articles in journals from different publishers; we are far from being able to do that. Collaboration is easiest between learned society publishers, and there are some encouraging, if small, examples of collaboration, but progress is slow.

Already, links to large bibliographical databases are being offered so that when you are reading or have just read an article you can check whether there is related published material, but without collaboration between publishers you will not generally be able to get the full text electronically, only the title and abstract.

More exciting are direct links between the electronic journal and databases, for example: the large NASA archive of actual astronomical images or data; crystallographic databases (Acta Crystallographica); protein databases, etc. Such links can dramatically increase the usefulness of the original article.

As yet, few journals support the possibility of including in an electronic article rotatable three-dimensional images of, for example, a protein molecule, or mathematical equations that you can manipulate to see if you agree with the author's conclusions or to fit data to. But many publishers

are working on such facilities, usually using the JAVA language to embed software elements inside a Web document.

Stevan Harnad points out that the present stance of most publishers to produce both a paper and electronic version of a journal is basically unstable and that there is no half-way solution to electronic publishing. He suggests that the costs of peer-review and editing should be borne by authors, and their research funding authorities, rather than readers in the form of subscription costs.

Since the Web is the main form of electronic publishing, Mike Sendall discussed its future and impact. One important development is PICS, the Platform for Internet Content Selection. Established to permit blocking of unsuitable material for children, it could be used for positive labelling to show that electronic material has been refereed, for example, by a learned society.

Overall, from the session it was clear that there is a movement towards greater use of electronic media in all academic disciplines. However a constant theme of our discussion was that one must recognise that there are important differences between disciplines in the benefits, requirements and applications of electronic publishing. What is right for physics or molecular biology is not necessarily right for the humanities, legal studies or computing science.