

The Impact of Networks on Publishers

The potential for publishing electronically rather than printing on paper has been discussed and researched for many years. Professor Senders, reporting at the European Information Management Conference in Brussels in 1980 on the research that he had been doing on an electronic journal, having introduced his paper, placed the work in perspective, outlined what they had hoped to achieve then said that now famous line "Ladies and Gentleman, I have to tell you that we have visited the future and it doesn't work!" The future that does work is already here and although there may be a little disagreement about that, there is still plenty of room for conjecture about timing and level of impact on scholar, librarian, agent and publisher.

Introduction

Let me remind you (as if UKSG members needed reminding) of some of the facts:

1. library budgets are being cut annually in real terms and the trend is for that to continue (reports of up to 20% cuts in renewals for 1992 have been made by one subscription agent);
2. prices of journals and books increase in real terms as publishers endeavour to protect their income from a steadily shrinking customer base;
3. as a result of the applications of IT, it is possible for researchers to do more work, adding significantly to their productivity and to the number of potentially publishable papers;
4. the cost of hardware and software continue to fall dramatically, whilst their sophistication increases;

Richard Coleman
Richard Coleman Associates,
London

5. the telecommunications network is being improved continuously*;
6. the need for scholar to scholar communication is as great as ever.

Method

Every scholar owns or has access to a word processor with communications links. More and more authors are preparing their articles in machine readable form. It is becoming routine to send and receive electronically the text of articles between authors, editors, referees, publishers, typesetters and printers. It is at this point that the traditional technologies take over, i.e. large numbers of copies of each article are printed, bound and mailed to subscribers world-wide (heavily energy-dependent activities).

So why bother to print on paper and distribute? If the article sits in a central database after acceptance and the potential reader is alerted to it through the increasingly sophisticated secondary services, it only remains to communicate directly with the database to have the article delivered straight into a personal computer. If the article has been formatted with codes to drive a desk top production system, then the article that comes off the reader's laser printer would be no different from the article as printed in the learned journal. (Time saved up to three years).

A major stumbling block could be communications technology. However the amount of bandwidth available for data transmission is increasing rapidly (optical fibre, ISDN, microwave links, satellite etc).

* Although a word of caution here let us hope that the attitudes of the telecommunications companies have changed over the last twelve years. The paper after John Senders was given by Neville Moray and I quote "... But the British Post Office reneged on an agreement given a year earlier to let our experiment take place, 10 days before it was due to start. They were completely, utterly and totally obstructive and unwilling to discuss any way we might be allowed to run the system, and took refuge in a blind bureaucratic assertion that 'we and Western Union have a complete monopoly on message transmission across the Atlantic and there is nothing we can do about it'".

A big change will be the impact of Document Image Processing and optical storage making it possible for the new media to become much more powerful facsimiles than previously. I heard recently that in 1994 the National Science Foundation network in the US will be able to carry 1gbyte per sec (550 mbytes to a CD-ROM, less than 1 second to send all the data on it, 200,000 pages). Even so to send the amount of published data via telecommunications would make BT's mouth water.

Enter the research librarian. How many people read any article in a learned journal? There is some dispute here. Some studies show one or two, others ten or more (large numbers scan the title and abstract, but this can be done using the secondary services).

Take a learned journal published monthly, with a thousand subscribers, containing ten articles per issue. It has on this basis, during the course of one volume, at 10 readers per article of the 120 published, only 1200 readers of 1 article from 120,000 sent out (1%). Imagine the cost and energy savings to be made in print, paper and distribution if it were possible to target those 1200 (let alone the time). The time factor and probably the cost too have now created a system whereby the authors send copies of their papers to the people working in the field who they know really need to know, thus bypassing the whole structure. Eventually the article is published but by then the 'in' crowd have already seen the original and probably the revised version as well. Publication in this case is possibly just for the record (the citation record). With this system commercial publishers would have to change their structure, but not dramatically; printers and typesetters would be in even more trouble than they are now with the recession.

But why use the commercial publisher? Traditionally because of the heavy financial investment required to support editorial staff, typesetting, printing, distribution and promotion. These roles could now be changed. The secondary service could fill the role of alerting the potential reader. The

impact of good secondary services on CD-ROM and eventually semi-conductor memory will have a massive effect on usage. Another word of warning here from Trend Monitor Spring 1989 (3 years ago) on electronic media and indexing systems "From the information supplier's point of view, the problem will be making sure that their piece of information, whether video, text, image or sound, will be given a prominent position in the index. The terms of selection will become such an important issue that Trend Monitor predicts that it will become necessary to regulate to ensure minimum fair representation requirements." Librarians say that wherever CD-ROM is installed users fall over each other to use it. The paper product lies untouched on the shelves, as it always did, the expensive online option remains under-used and the user-friendly electronic product gets 'hammered'. This is already having the effect of increasing the demand for the original articles and – since librarians have diminishing budgets – Interlibrary Loans will increase and the photocopiers at Boston Spa will get even hotter.

This then could be the new scenario. The original text could be prepared in the author's institution on a word processor (or DTP system). The editorial boards and referees could operate as they do at present but will need a focus. This role is custom-built for the University Press and the Learned Society publisher. The control of the output of articles would rest with the author's institution or the Learned Society. This has another benefit, universities and other institutions give away their intellectual capital by virtually giving the fruits of their labours to commercial publishers for them to exploit. What I am suggesting is that there will be growth in University Press-Learned Society publishing houses and a consequent migration of publishing staff from commercial houses to institutional ones. I would go further and urge Universities, Learned Societies, Institutions etc to set up publishing houses even if only to begin to control the copyright of the output of their

own scholars, members etc. They can then always license another publisher rights to the journal article or whatever. But the first step is to gain a control over what, by rights, belongs to them.

An Online Journal

The AAAS and OCLC in the US are launching a new online journal called "The Online Journal of Current Clinical Trials" on 1 July 1992. It has the following attributes:

1. A very prestigious publisher;
2. A very prestigious editorial board;
3. Typeset quality on screen including: text, charts, tables and graphs made possible by the increasing sophistication of Microsoft Windows;
4. Quick and easy search and retrieval;
5. Direct hypertext link to Medline, which allows you to view abstracts of articles cited in the reference lists without leaving the article you are reading;
6. Publication 24 hours after acceptance, saving up to 3 years;
7. On-screen alerts to all subsequent correspondence concerning an article;
8. Fax or mail alerts to articles of interest, once you have given them a set of keywords defining your area of interest;
9. Instant access to all back articles, i.e. your subscription covers back issues;
10. Worldwide simultaneous access;
11. A publishing medium for very specialised research whose market would normally be too small to justify publication;
12. Ability to download and print or request typeset quality reprints to be delivered by fax or mail;
13. No host connect time charges. You pay only for your local phone call to your carrier node plus whatever transmission time is charged by the network you are using. (OCLC, Compuserve, PSS or Internet.)
14. All this for \$110.

It would seem therefore that the future is nearly here, and I wonder what will happen this time. They already have over 1000 subscribers. They chose this particular field because they felt that rapid access to this type of data (clinical trials) would be a key factor.

Costs

To finance this sort of activity obviously requires investment initially. For example, studies of document usage may need to be undertaken in order to evaluate the amount of bandwidth actually required for transmission. International standards will have to be agreed. Although I suspect that the amount of bandwidth required for transmitting the scientific and technical literature will be minute compared with the demands of other users (electronic banking, shopping, teleconferencing, e-mail etc.) and also that, by the time international standards have been agreed, there will be slick bits of interface software that will render the current barriers invisible. Once the systems are in place the savings will be substantial. In the case of the University, greatly reduced expenditure on expensive primary journals.

University Presses and Learned Societies, however still need to generate income from their print on paper publishing. Since their costs will be much less, i.e. no typesetting, printing or distribution costs, they will require a much smaller gross income in order to maintain their margins. Taking the example of the traditional journal model outlined earlier, i.e. typeset, printed and distributed, with a reasonable amount of complex setting (maths, tables, figures etc) it would cost with overheads about £100,00 pa. With 1000 subscriptions it could gross say £150,000. Representing a surplus of £50,000.

To access 1200 articles from that database would cost a proportion of the running costs and telecommunications access. A report from the ARL in the US puts this figure at between 30 and 50% of the paper copy costs. To give a gross margin of £50,000 then the annual subscription could be reduced by

between one third and a half of the currently charged.

Table 1

Conventional journal

Costs	£100,000
Number of Subscribers	1,000
Subscription price	£150
Revenue generated	£150,000
Surplus	£50,000

Electronic journal

Target surplus	£50,000
Costs	£30-50,000
Revenue to be generated	£80-100,000
Number of subscribers	1,000
Subscription price	£80-100

This could be collected by:

- (a) increasing the individual member subscription in the case of a professional institution,
- (b) charging the purchasing institution a "corporate" membership fee or
- (c) direct charging against usage.

The latter is appealing because it would give a measure of usage of a document. (Data which the various funding bodies would find useful, i.e. how much interest is this article creating, where from, how much money is it making for us!) It also means that power returns to the user, i.e. it is the end user who makes the decision about whether or not to have a document delivered. Since there would no longer be the delays incurred in typesetting, printing etc., distribution would become highly targetted and instantaneous. The funding for these options would come from the library budget or, since the usage is now so specific, it could come from the budget of an individual researcher or group. By the same token credits could accrue to the individual researcher or group for use made of the documents that each have put on the network.

Potential

And the emerging technologies haven't finished yet. Software is now available which can handle books electronically (including diagrams). You load the disc version of the text of the book into the software and there on screen is the electronic version of the book fully indexed (automatically). This will mean that you can extract the chapters or parts of chapters you require from the books you have identified, add in the articles that you have downloaded, load up any data you require from directory sources, add your own results and conclusions and there is the raw but highly manipulable source for articles, monographs, course manuals etc. Add in new stuff as it becomes available delete the out-of-date, correct errors, incorporate suggestions, rework it, retile it and republish it as often as you can get it past the referees.

This model is then:

fast (no more printing and distribution delays),

cheap (ditto plus paper) and

green (savings in paper, its manufacture and its shipment around the world).

Editing the text

In order to get the most from an electronic retrieval system there will be a need to write in a far more structured way. This will continue to be an important role for editorial staff wherever they are located. Scientific authors are not trained apparently to write in a clear, concise and structured way and more often than not English will be their second language. It will be even more important to spell correctly if you are doing full text searches of electronic text (remember what Trend Monitor said about indexing). On a course recently I came across 'sock control' instead of 'stock control'. The speaker blamed the computer, the fault was of course in his expectation. The spellchecker passed 'sock' because that is how you spell 'sock'.

Until it is possible for the software to check context then this will always be the case.

Nor will it check for missed negatives, a previous boss of mine sent out a memo to the union reps in the organisation saying that “In principle I do agree with . . . , however” . He meant to say “In principle I do not agree . . .”.

Impact on emergent nations

Now to the problems of the Third World and Eastern Europe. With the model outlined above they could now be in a position to benefit from the research carried out in the developed world. The information systems in this model are now being run by Universities and Learned Societies who have a vested interest (i.e. prestige) in making their research available. The price of the electronic product will be fixed to give them a surplus from their traditional marketplace (the West). Additional access could be heavily discounted according to the ability of the various sectors to pay. Third World and Eastern European countries could negotiate rates with the suppliers. Their investment will be heavy in the first instance, computers and infrastructure (telecommunications etc.) but if they are to develop they are doing to need these anyway. It is possible that, since it is a once-only investment, the aid agencies might be more inclined to fund the installation of a satellite receiver dish than to fund annual subscriptions to primary journals. There will no longer be a need for trying to invest in the prohibitively expensive primary journal. There will no longer be intellectual restrictions because of the difficulty of accessing the primary literature. Quality and productivity would both increase significantly. The network would provide, because of the lower all round costs, much cheaper alternatives for industry to invest in research. From the Eighteenth century to the Twenty-first in one bound. There would be much closer international co-operation

between Universities working in similar areas. National learned societies would become international etc.

Impact on the roles of the players

The role of the publisher is unchanged, to package the information and provide the communication pathway; the role of the scholar is unchanged, to research; the role of the librarian is unchanged, to operate as an intelligent intermediary; the role of the agent is unchanged, to pick their way through the minefield of different document supply systems on behalf of customer and supplier. The culture shock on all will be substantial and it will probably be this, above all else, that delays implementation. Technology will advance, readers will become used to the electronic medium, and new technologies will dwarf what today seem like stupendous advances. The new British Library in St Pancras which has a large exhibition area (previously the area set aside for the catalogue cards – now automated) may find, by 1996, that the vast mechanised storage and mechanised delivery systems will have a limited life span.

This article has only looked at traditional journal and book publishing, there are a number of other potential networks which together could create the necessary demand that would make the installation of the necessary structures cheaper and sooner, e.g. local authority papers, government, corporations. The network system would be an ideal medium for the “grey literature” message. Networks are already very heavily used, e.g. bulletin boards, e-mail, etc and increasingly for sending around prepublication drafts of articles.

Time factors

So, when will all this happen? There is no question that it will. It was predicted by Senders in his paper in 1980 that if the baud rate were doubled then the electronic

journal would be a viable product. The baud rate over normal telephone lines in 1980 was 300, today it is up to 8 times that and the first electronic journals are just appearing. I suspect that a major barrier was, and will be for some time to come, the computer. So many of us for so many years had (and still have) so many bad experiences that the sight of screen and keyboard can bring on panic attacks in the most stalwart. How many times when you were in the middle of something did you get (do you get) “everybody off the system in 60 seconds!”, or “the system’s down for the rest of the morning” etc. We encourage each other to think negatively about the computer “Oh it’s the bloody computer again”. I think it has to do with what I will call the computeriness of computers. Let me give you an example of what I mean. Ever since you can remember (and before that even) everyone here started to learn about water. If I say to you ‘water’ you can fill your mind with images of it. When we were little we stood at the sink and played with it, we poured it, sprayed it, spilled it, got soaked by it, discovered it could be very hot and very cold, could taste and smell in an infinite number of ways, that it could be found in a variety of ways tiny drops and great oceans. And we learned more about it each day until today we only think about it when it inconveniences us, like falling out of the sky on us, or coming out of the shower cold. My theory is that since we live in what a speaker at a workshop I was at last year, described as ‘the paleoelectronic age’ it will only be when the generations who have played with computers since before they can remember, learned to summon images to the screen, and interpret those images and the accompanying signs, and be completely at home with mouse, keyboard and joystick, that the true impact of the computer will begin to be felt. It will also be with this generation that teleworking, telecommuting and teleconferencing will really advance. CD-ROM usage is particularly noticeable in the

student population – the first generation to begin growing up with the computer and the generation that has begun to use the keyboard and screen like their parents and grandparents learned to ride bikes.

At another meeting of a network publishing workshop last year someone said “that in general one tends to over estimate what is possible in one year and under estimate what is possible in ten”. At the moment my money is on more than ten years but less than twenty before the full impact is felt.

One action might spur the advance of the global electronic village. Governments could penalize with additional taxes technologies which damage the environment and then with those taxes subsidise environment-friendly technologies. This would advance the impact of network publishing and teleconferencing. Although we would all need quite sophisticated virtual reality equipment in order for us to experience the full spectrum of stimulus available at a UKSG meeting!

Conclusion

So there we are, all of the elements for a period of exciting change are now in place. The challenges are becoming more clearly defined . . . and as far as I can see it is going to be totally chaotic!

Always remember that publishers, agents and librarians are the ‘middlemen’ between the scholars, researchers etc. If we in the middle impede rather than support the flow of information, i.e. do not respond to the requirements of the end-user, then our clients will find another route. (There is already evidence for this.) The current giants of the scholarly communication process will find themselves to be the dinosaurs of the late 20th-early 21st century and some, as yet unrecognised, furry creature running about in the undergrowth will turn out to be the forerunner of the new giants of the next millenium.