

---

## SCHOLARLY PUBLISHING IN THE AGE OF INFORMATION TECHNOLOGY

---

C.K. Yuen  
Department of Information Systems & Computer Science  
National University of Singapore  
Kent Ridge  
Singapore 0511

---

### *Summary*

This article considers the many changes that are likely to occur in the sphere of scholarly publishing with the further development of data communication networks, databases and public information systems.

### *Scholarly Publishing*

For scholars devoted to research and education, publishing has always been an important part of their working life. Many forms of scholarly publishing exist, but they may be separated according to intended functions as well as by the publishing media involved. We may identify the following functions:

1. Proclamatory: The publication announces new discoveries and establishes the author's entitlement to receive credit for these;
2. Archival: The publication adds to a store of available knowledge for reference by others;
3. Disseminary: The publication brings knowledge to the attention of a wider audience.

Although any publication can serve all three functions, the difference between objectives makes some media more suitable than others for each purpose. For example, objective 1 is best satisfied by a dissemination medium with a very quick but selective reach: the important thing is to announce the result to other research workers in the same area quickly. Journals that carry brief research reports with short publication schedules are often used, though preprints and technical reports published by the author's own organization and circulated to colleagues also play a very important part, and so do announcements at conferences. It is of interest to point out that, while newspaper and television announcements about research finds are not unknown, these media are not sufficiently selective, and are consequently only playing a small part of this service.

In contrast, objective 2 is best served by media that provide more detailed descriptions of our knowledge, in a form that is relatively easy to store and retrieve. While collections of preprints can serve this function, they are not as convenient to use as journals that can be bound together in volumes and stored on library shelves. To further facilitate retrieval, reviews and bibliographies within various subject areas are compiled, providing a central reference to sets of publications that have appeared in different places at

---

different times. Reference monographs play a similar function, but less effectively because of the large lead time.

Different yet again, objective 3 requires good production and distribution channels, through which material, written in a way useful for informing a non-expert audience, can be produced at relatively low cost, and supplied conveniently to the intended readership. Textbooks and "low brow" journals and magazines are designed to serve this purpose. Although these too can be stored in libraries and play some part in function 2, bibliographic references to their contents are more difficult to compile and they do not work as well as journals for archival purposes.

### *Information Technology*

As publishing is a form of information dissemination, it is critically dependent on the technology for managing information, such as information representation methods, tools for storing and changing information, and systems for transmitting information. The efficiency and capabilities of information management methods naturally affect the techniques adopted for the purpose of scholarly publishing.

The following stages in the current procedures of scholarly publishing can be identified:

1. Text generation: This starts with authors generating their source information, which is prepared in the form of typed text by themselves or by typists. A text validation process is carried out to correct errors and revise the content until the text is satisfactory. The final manuscript is then passed to the publisher's editorial representatives.
2. Scholarly reviewing: The material is reviewed to determine its suitability for publication, in accordance with some editorial criteria. Thus, the main criteria for proclamatory publications are newness and brevity, while those for disseminatory publications are comprehensiveness, relevance and comprehensibility. This process may be long and detailed for books and journals, but short and self-imposed (i.e., by the author himself or by him asking nearby colleagues) in the case of preprints.
3. Print preparation: After material passes reviewing, it is set into type to produce print masters. Sample pages (page proofs) or single column text (galley proof) are then printed from the masters and used in proofreading. Although in some cases proofreading is performed by professional staff under the publisher, the normal practice in scholarly publishing is for the author to perform the proofreading, and sometimes, make final changes.
4. Printing: This may be done in the author's organization using simply photocopiers or low cost printers, or professional equipment at a commercial publisher. In scholarly publishing the print runs are generally small. Even large learned societies would seldom publish journals with circulations exceeding 5-10 thousand. Similarly small press runs are normal for textbooks, even best sellers. Research monographs and journals published by commercial companies generally have even smaller print runs.

- 
5. Distribution: Learned societies generally distribute journals to all members (sometimes stratified into interest groups), while commercially published journals, usually more costly, are subscribed mainly by libraries. Textbooks are marketed via established commercial channels, while general interest magazines may be both subscribed and sold on a retail basis. Generally, libraries take up a large portion of scholarly publications. Preprints, on the other hand, are sent directly to selected recipients on the mailing list of the author's organization.
  6. Archiving: Publications are classified and references to them established by listing them in annual indexes of journals and in area-specific abstracting and search systems. The publications themselves are placed into holding collections but may be retrieved upon request, either by physically retracting the publications themselves or by making copies of them. In the case of books and classifications and index establishment are performed by library staff, whereas currently most journals require the authors to classify and index their own work, with the publishers and abstracting services merely performing the clerical side of the operation.
  7. Citation analysis: As scholarly publications refer to earlier work containing information that has been used to produce the new work, the frequency of a paper's being cited is an indication of its impact. This is a unique derivative of scholarly publishing, and is achieved by an exhaustive count of all the bibliographic references in each published work and compiling the information to-

gether for each area.

It is apparent that publishing is a costly process, since a great deal of labour is required to produce the results. It also seems to be an inefficient process, since the same information is passed through many stages which make no more than minor changes or merely convert the same information from one form to another. It also seems highly inefficient to distribute books and journals containing heavy and bulky sheets of paper, when the information itself is contained only in the *ink* on the paper; the paper itself has no information value whatever.

### *Electronic Publishing*

The digital computer provides us with new methods for handling information, electronically rather than in ink and paper. Since electronically represented information is far easier to change and transport over distance, the technology offers many possibilities to improve efficiency in publishing, but particularly for scholarly publishing.

Computers already play very significant parts in the publishing process. For example, most people today prepare their manuscripts on word processors, storing their text on disks so that corrections and modifications are performed without complete retyping of pages. The modern typesetting systems contain a high element of computerization, with sizes, text and page formatting on computer screens, and print masters made by computer plotters and laser printers. Printing machines may also be computer controlled to regulate print pressure, ink supply, print run, etc. Virtually all publishers and learned societies use computers to maintain

---

mailing lists, and computer based archiving and library search systems are extensively used.

More significantly for the future, the use of computers and computer-based information systems offers the possibility of eliminating many of the intermediary steps in publishing and minimizing the repeated conversions of the same information into various different forms.

For example, suppose everyone involved in scholarly publishing agrees on a standard format for text preparation using a standard computer on a standard diskette, then authors will write their manuscripts on diskettes and send these to editors, who will forward copies to reviewers. The reviewers, who possess the same model computer as the author, can then view the manuscript on their own machines and provide comments on them. The commented diskettes are then returned to the editor, who consolidates all the comments on one disk for returning to the author. The author can then prepare his final manuscript, and the finished diskette is sent to the publisher. The diskette can be immediately used by the publisher's computer to produce print masters, and proofreading is unnecessary since the output will contain exactly the same information as what the author specified.

Further, instead of sending printed journals and books to libraries, indexing journals and citation journals, the publisher can simply send disks or tapes containing the complete contents of publications. Using the information on these magnetic media, computer programs can immediately compile indexing and citation lists without requiring human reading and counting of journal and book contents, while libraries can selectively reproduce any part of a book or journal for

the disk or tape to supply to readers, who can take home a new paper either in printed form (produced by the library computer), on microfilm (if the library computer has a microfilm printing unit) or on diskette to be viewed on his own computer.

To consider a more revolutionary scenario, all the parties involved in a particular subject may be connected together in a computer network, but each network user has only capabilities to use the network in accordance with his status. An author can enter information into the network, but this will be sent to readers only if the editor/publisher, who controls the network information transmission procedures, approves the material after undergoing a reviewing processes. When a new manuscript is received in the network, either a computer program or the editor makes the decision to send it to selected terminals on the network where reviewers for the subject can be found. Comments are electronically returned and revisions performed. Approved material is selectively transmitted to readers who have specified an interest in the particular topic. Material can also be immediately entered into databases connected to the network, and computer programs are executed to link each item to appropriate indexes and count the citations in the item. All unnecessary information conversions are eliminated.

It should be clear that the diskette-based scenario can be achieved with less effort and cost from today's conditions, and is already being implemented though in a piecemeal and unstandardized fashion. On the other hand, the network-based scenario offers certain new advantages, such as high speed. It would not be surprising if research organizations in certain

---

particularly active fields decide to form world-wide computer networks for such purposes, and indeed many are already linked to informal data exchange networks, though formal publishing networks remain to be organized.

### *Paperless Publishing*

Just as computerization in organizations has produced predictions of “paperless offices”, computerization in publishing suggests the possibility of “paperless publishing”, since all information can be stored electronically and may be reproduced by the computer upon request. However, whether electronic information is better than paper and ink information depends on circumstances.

To reach the reader, electronically stored information requires a user interface, whether a computer terminal, a printer, a microfilm maker, or some other device. For network based publishing, the device must be connected to the network, and portability is low, requiring at least dial-up lines through telephone sets. For diskette-based publishing the portability is greater, though it is still difficult to imagine a student going to classes with a box of diskettes and referring to textbooks on terminals. We have not yet reached the point when we can find a standard computer wherever we go, and for an individual, carrying a book around is still convenient, even though all the information is in the ink and all the weight is in the paper.

Thus, for the purpose of knowledge dissemination among a large readership, books still provide a convenient mechanism. However, the handling of bulky and heavy volumes can be reduced within the publishing process by defer-

ring the production of books to the last stages of distribution. It is already common for publishers to have a number of printing centres to which diskettes for the same book or journal are sent to allow each centre to print the necessary copies for distribution in the local region. This, however, still requires a reasonably large run at each point, but technologically it is now possible to use a laser printer to print a complete book by itself immediately upon request. Thus, a bookseller or library might have an instant printing machine together with a stock of diskettes or a line to a publishing network. Any book or paper asked for by a customer is immediately printed on the premises, and book transportation costs are eliminated. However, such commercial issues as copyright and royalties will then need to be solved.

For the purpose of proclamatory and archival publication, an electronic network eliminating the transport of paper seems to be a more logical objective. However, the network will need some mechanism of authentication and certification, such that a record of someone having “published” something at a certain time can be established. It is insufficient to send someone’s paper to all readers interested and enter it into the database. The network has to be able to certify, in an indisputable fashion, that a particular person produced the work at a particular time.

In short, while technologically paperless publishing is only a short step away, many administrative issues need to be resolved. These may prove to be as challenging as the technology itself, and will require the participation of publishers, librarians and information technologists. Having glimpsed the new possibilities of scholarly publishing in the information technology age, we await to see its final shape with interest and fascination.