An Agenda for Digital Journals: The Socio-Technical Infrastructure of Knowledge Dissemination

Brian R. Gaines
Knowledge Science Institute
University of Calgary
Alberta, Canada T2N 1N4

Abstract

The problems of information overload from the growth of scholarly literature, and the need to use information technology to manage them, were identified by major writers and scientists over fifty years ago. Yet the main form of scholarly communication, the journal, is still circulated in paper form as it has been for over three hundred years. The economic arguments for using computer and communication technology to overcome these problems through a new form of scientific communication, the electronic or digital journal, were vigorously presented in the 1970s. Experimental trials of digital journals with the technologies of the 1970s and 1980s have not been successful. In the 1990s, the continuing value of current journal systems is again being questioned in terms of soaring library costs, the burden of the current refereeing system and the diminishing returns of journal publication brought about by information overload. This paper presents a fundamental examination of the prerequisites for the introduction of digital journals, at one level in terms of the role of journals in the social and economic processes of human knowledge production, and at another in terms of the state of the art in the relevant technologies. Models of the processes underlying the growth of knowledge in the literature on the philosophy, history and psychology of science are first used to analyze the structure and role of the social infrastructure of journals, including the editorial and refereeing systems and the role of commercial publishers and libraries. The motivation for digital journals and past experience is then surveyed, then the learning curves and current costs and performances of the enabling hardware, software, communications and interface, technologies. Examples of the current impact of computer and communications technology on scholarly discourse are given to enable probable changes to be predicted in the structure of journals when they are transferred to digital form. Finally, the social and technological analyses are used to outline some architectures for a first generation of digital journals emulating the current medium, and for the evolution of later generations diverging in characteristics to take advantage of the new medium.

1 Introduction

Some three hundred and twenty five years ago the first two scholarly journals came into being: the *Journal des Sçavans* in January 1665 in France, and the *Philosophical Transactions* of the Royal Society in March 1665 in England. Now, as we prepare to enter the next millennium, the two have grown to some fifty thousand, and what was the blessing of improved scholarly communication has become the curse of information overload. The growth of knowledge that has its seeds in the seventeenth century enlightenment has become a prolific jungle, the understanding and control of which through the information sciences and technologies is now an important knowledge objective in its own right.

The origins of this information explosion, although three centuries in the past, have much relevance to our understanding of the problems of today. The second enlightenment was seen as an emancipation of human thinking from the weight of authority of the church which had taken the fruits of the Greek enlightenment and turned them into a ritual and static form. As Kant phrased it, "Enlightenment is man's release from his inability to make use of his understanding without direction from another." (Kant, *What is Enlightenment?*) This freedom of the mind was inextricably related to the sociology and politics of the times in paradoxical ways. It provided the environment for the French revolution with its overthrow of the authoritarian *ancien régime*, but it also provided that for the industrial revolution based on scientific knowledge that enslaved humanity in new ways (Toffler, 1980).

The progress of the enlightenment was also inextricably related to the technology of printing which enabled new knowledge to be disseminated widely at a reasonable cost. More than 30,000 titles were published in France between 1723 and 1789 and literacy became widespread. A technology that was first used by Gutenberg to make the bible more widely available became the vehicle for disseminating Rousseau's model of the intrinsic rights of every citizen and the need for his or her involvement in communal affairs. As many writers have documented (Innis, 1951; Ong, 1977; McLuhan and Rogers, 1989), innovative technologies that support new media for human discourse are both a response to social needs and major factors in social and cognitive change. Our understanding of those needs at the time, our intentions in design and our expectations of use, are all likely to prove incorrect from later historical perspectives. However, it is difficult to use such past historic perspectives to enable us to manage socio-technical change in our own time more effectively. One might hope that if we plan and design new scholarly communication systems with a view to flexibility and evolution then the expected surprises in general, although intrinsically unexpected in particular, will be less traumatic in their impact. Even that may be an over-optimistic assessment of our capabilities to understand and manage our own social dynamics.

Thus, as we look to electronic, computer and communications technology to provide a new medium for the dissemination of knowledge, the *digital journal* in which electronic access and digital processing will be combined to harness the information explosion, it is important to remember that technology is only one consideration, albeit a very important one. The social framework within which scholarly activities take place, and of which they are an essential part, is that which provides the rationale for the production and dissemination of knowledge, and also for the development of information technologies. Bacon noted that "Human knowledge and human power meet in one" (Bacon, Aphorism III, Book I of The New Organon, 1878), and power, politics and economics are as integral a part of the dynamics of the life world as is the creation and dissemination of knowledge. The role and nature of power structures in the application of new technology will be as significant as it has been in the old, and whatever structures for publication that we design will evolve to become part of the post-modern socio-economic infrastructure, whatever that may become (Habermas, 1985; Heller, 1990).

This paper brings together the social and technical aspects of the development of digital journals, analyzing those of existing, printed journals, extrapolating this to a first generation of digital journals that will emulate the functionality of current journals, and finally attempting to transcend the past by projecting how new forms of knowledge dissemination may evolve. It commences with the motivation for digital journals, analyzes the social role of scholarly

communication, examines the state of the enabling technologies, uses these to outline possible architectures for digital journals, and puts this in a setting of trends in scholarly communication using information technology other than journals to speculate about future developments.

2 Motivation For Digital Journals

This sections examines motivation for the digital journal in terms of identification of information overload as an impediment to scholarship in the late 1930s, failure to ameliorate the growing problem during the next fifty years, experiments with unsuccessful digital journals in the late 1970s and early 1980s, and lessons learnt about issues relevant to the assessment of further developments.

2.1 Early Concerns with Information Overload

When we examine the growth of science and technology in general and that of information technology in particular, we see two exponential processes in competition with one another. As knowledge has grown, one of its byproducts has been the technology to manage the impact of the growth of that knowledge, but for a very long period it seems that the growth of raw knowledge has surpassed the growth of the knowledge of how to manage it.

The problems of matching these growth curves have long been recognized. For example, in 1937, just prior to the advent of computer technology, Wells was promoting the concept of a "world brain" based on a "permanent world encyclopaedia" as a social good through giving universal access to all of human knowledge, and he remarks:

> *"our contemporary encyclopaedias are still in the coach-and-horses phase of development, rather than in the phase of the automobile and the aeroplane. Encyclopaedic enterprise has not kept pace with material progress. These observers realize that the modern facilities of transport, radio, photographic reproduction and so forth are rendering practicable a much more fully succinct and accessible assembly of facts and ideas than was ever possible before." (Wells, 1938)*

Wells' world brain concepts and objectives have continued for over fifty years to be an active theme in the information systems community (Goodman, 1987).

In 1939, Bernal echoed the same issues from the viewpoint of the scientist:

> *"In the old ideal of science, communications were the only link between scientists. Now the very quantity of scientific information has made its diffusion an enormous problem, with which the existing machinery has utterly failed to cope. The present mode of scientific publication is predominantly through the 33,000 odd scientific journals. It is, as we have already shown, incredibly cumbersome and wasteful and is in danger of breaking down on account of expense." (Bernal, 1939)*

Some six years later Bush was re-echoing these sentiments in his famous article in *Atlantic Monthly* which is often cited as the first expression of the need for hypertext:

> *"Science has provided the swiftest communication between individuals; it has provided a record of ideas and enabled man to manipulate and to make extracts from that record so that knowledge evolves and endures throughout the life of a race rather than that of an individual. There is a growing mountain of research. But there is increased evidence that we are being bogged down today as specialization extends. The investigator is staggered*

by the findings and conclusions of thousands of other workers—conclusions which he cannot find time to grasp, much less to remember, as they appear. Yet specialization becomes increasingly necessary for progress, and the effort to bridge between disciplines is correspondingly superficial. Professionally, our methods of transmitting and reviewing the results of research are generations old and by now are totally inadequate for their purpose...The difficulty seems to be not so much that we publish unduly in view of the extent and variety of present-day interests, but rather that publication has been extended far beyond our present ability to make real use of the record." (Bush, 1945)

We are acutely aware of scientific and technological progress in the last fifty years, and marvel at the continuing exponential trend lines in information technology that result in surprising new capabilities year by year, but surely these statements give the lie to such feelings. If H.G.Wells, John Bernal and Vannevar Bush presented these statements at a session of the American Society for Information Society today, would not they appear timely, significant and utterly to the point? Their worlds of science and technology, and the problems of knowing and navigating them, appear simple compared with our own. The volume of scientific publication has soared since these remarks were made but the means of disseminating the information in those journals has remained unchanged, that of physically transmitting print on paper. There have been major advances in the technology of printing during this period, notably the development of low cost photolithographic printing, phototypesetting and computer typesetting, but the end product remains basically the same as it was in 1665 when the *Philosophical Transactions* were first published.

2.2 The Failure to Cope with Information Overload

Information technology has been applied to help manage information overload, but the the impact has been palliative and not addressed the root problem. Computer-based information systems have been developed to maintain indices of the scientific literature and to search these indices in increasingly complex ways, but the literature itself is still in paper form. Moreover, most of the indices and abstracts have to be entered manually, possibly supported by optical character recognition, and this keeps the cost of information services based on them at a much higher level than if the text were already available in digital form.

Some journal publishers have established text databases as a by-product of computer typesetting but this is still rare, and trends in this direction are offset by other trends towards authors supplying 'camera-ready' copy that can be photographed directly to printing plate with no intervening setting. The anarchical mix of processes used in the production of scientific journals is becoming more and more absurd year by year as authors increasingly come to use computer-based word processors to develop their manuscripts. Publishers who still undertake typesetting are increasingly coming to accept authors discs as a way of bypassing the initial stages of keyboarding text, but they rarely accept figures in digital form, and in general are only using the digital material from authors as part of the printing process, not as entries to a database.

The impact of lack of progress in capturing scientific knowledge in digital form and taking full advantage of information technology is most apparent in terms of the economic problems faced by librarians in providing both physical storage and funds for books and journals (Simpson, 1989). This may be quantified through analysis of library budgets which have grown at rates far exceeding inflation (Baumol and Marcus, 1973). Soaring costs are now a major impediment to

the dissemination of knowledge in that few institutions can provide immediate access to other than a small proportion of the published literature (American Council of Learned Societies, 1979).

These financial problems are often blamed on the commercial publisher standing between the community of researchers who freely contribute publications and the same community of researchers who are forced to pay high prices for access to these publications. As one librarian has remarked:

> *"Scholarship today has become victimized by a predatory commerce that is preying on its weakness. The knowledge chain in which libraries are the penultimate link is under particular pressure beyond the volume and complexity of communication. It is traumatically stressed by the aggressive efforts of those who would live off the avails of scholarship; those who I've referred to elsewhere as the 'pimps and pushers of scholarly publishing.' In the past decade or more, commercial publishers have markedly increased their incursion into the scholarly publishing process. We are now faced with an unwelcome economic concentration focussed particularly in the areas of science, technology and medicine. The noisy bleating and moaning of your librarians as they crumble under the weight of cost increases is merely symptomatic of these wolves at the door." (MacDonald, 1990)*

MacDonald also criticizes, as do others (Carrigan, 1991), academic promotion procedures for encouraging excessive publication and and counter-productive forms of publication. He argues that digital publishing can overcome the economic problems, but only if is taken out of the hands of commercial publishers so that the cost of access is not established at a high level compared with the cost of the medium.

There is scope for new modes of publication not involving any commercial enterprise through the use of existing digital networks, and this low-cost competition should act to counter the increasing costs of current journals. However, it takes several years to build the reputation and status of a new journal, and, while new journals taking advantage of the innovative capabilities of the digital medium will be very significant in the long term, it is probably the parallel publication of existing journals that will do most to establish the value of the new medium in the short term. Some commercial publishers are already experimenting with digital publication through databases offering access to the full text of certain journals (Lerner, Metaxas and Scott, 1983; Tenopir, 1984; Lynch and Preston, 1990), and are in a position to play a leadership role in developing digital journals. In particular, professional societies operating major journals and operating as commercial publishers, but responsible to their membership rather than to shareholders, have the status and resources to establish effective digital journals in parallel with their existing publications. A competitive environment in which the new contends with the old would provide a very healthy scenario for new developments and for the creation of pressures to minimize the cost of existing publications.

The only restriction on competition might come from improper use of copyright laws to enforce monopolistic control of material. Most publishers require authors to assign them the copyright in papers for the purpose of publication, but assign back to authors the right to use the material freely for their own purposes and in their own publications. Thus authors are already able to issue papers in digital form for parallel publication and there is no reason why this cannot be

institutionalized through non-commercial archives and indexes. It would be a very negative development if there were attempts to restrict this by use of the copyright transfer authorizing paper publication to prevent parallel digital archiving. It is not only commercial publishers who might be tempted to misuse copyright procedures. The litigation between the Dialog Corporation and the American Chemical Society (O'Leary, 1991) indicates the possibility for use of copyright protection by professional societies in a way that is against the interests of their members. The proposal by the Canadian Government to ensure a monopoly for Canadian companies on the distribution of books within the country by amendment of copyright laws shows that national institutions can be responsible for initiating abusive practice.

Knowledge is a valuable resource and one should expect to see significant economic influences at work in any change in the method of its distribution. It is reasonable to expect commercial publishers to attempt to protect their investments in establishing and marketing existing journals. It is is reasonable to expect governments to attempt to regulate knowledge flows in such a way as to advantage their nations—the proposal to tax the 'brain drain' of skilled workers from lesser developed countries (Bhagwati and Partington, 1976) is a pre-digital example, and a reasoned argument to legitimize what has been the practice in many countries. Existing regulations on cross-border information flows are nominally designed to prevent misuse of personal data, but, as a side-effect, give a commercial advantage to local companies offering services to process that information. Given government regulatory power over communications (Samarajiva and Mukherjee, 1990), there will always be a danger that attempts may be made to restrict knowledge flows to national advantage.

Factoring out these politico-economic considerations, however, much of the commercial and academic distortion of the publication process can be viewed as the natural outcome of the poor economics of continuing to use an obsolescent technology when substitution of a new one could achieve the same ends more effectively at lower cost. This raises the question as to whether current information technology is as yet adequate to substitute for printing and paper technology. Computing costs have fallen steadily as library costs have risen, but low cost personal computers, graphic workstations, CD-ROMs, laser printers and international digital networks, are products of the 1980s only recently coming into widespread use. They show every promise of providing a new medium to support scientific knowledge dissemination, but are they yet ready to do so?

2.3 Early Experiments with Digital Journals

It is salutary to remind ourselves that the arguments for digital journals have been apparent for a long time, that experiments commenced over a decade ago, and that surveys in the late 1980s indicate that none had succeeded in establishing a viable publication (Freedman, 1987). Some digital publications apparently providing a useful service during this period, such as *Clinical Notes On-Line*, have now been discontinued (Dixon, 1988).

In 1978 Lancaster documented the basis of increasing library costs in great detail, and saw "paperless information systems" as providing a solution to the problem (Lancaster, 1978). Following an economic analysis (Senders, 1977), action was taken at that time to investigate the possibility of operating a digital journal, and an experiment took place with NSF funding in which a psychological journal on mental workload was developed and operated on the EIES

computer-based conferencing system (Hiltz, 1984). The rationale and expectations for the new technology were made very clear in the final report on that experiment:

> *"The traditional journal system has among its strengths a defined subject-matter, a developed body of subscribers, an experienced editorial staff, a group of referees chosen for competence in the subject matter, and a form that is convenient in numerous ways, including, for example, portability. Authors submit articles to the editors; the editors choose referees to review the material; the referees submit their judgements to the editor; and the editor makes a decision whether to accept or reject material. If the decision is to accept the article, it is edited and put into the production pipeline, eventually being included in a volume that is printed, bound, and mailed to the subscribers. On the other hand, the drawbacks of the traditional journal include such problems as the fact that each volume contains material that many subscribers are not interested in; that delays between acceptance of an article and its publication are often long, as much as a year to two years; that library storage is expensive and space consuming; that finding specific material can be difficult and time-consuming; and that the limitations on the size of a volume drive editors and authors toward excessive brevity and limits the publication of raw data." (Sheridan, Senders, Moray, Stoklosa, Guillaume and Makepeace, 1981)*

The report remarks that "the results fell short of initial aspirations" and "many of the participants, not themselves experienced with computers, found EIES inconvenient or difficult to use" (p.1). This inspired John Senders infamous aphorism, "I have seen the future and it doesn't work" (Senders, 1981). However, dissatisfaction with the technology is not surprising at a time before the advent of UUCP and USENET, when access to the system was at 110 or 300 baud, largely through teleprinters, and with no possibility of transmitting anything other than textual material in a discipline that traditionally makes heavy use of graphs and figures. Experiments with the use of interactive time-shared computers for elementary computer-assisted learning a decade before had resulted in equally bitter remarks by students and their parents (Suppes, Jerman and Brian, 1968)—over-expectations leading to premature use of a new technology can be a disheartening experience.

A small-scale experiment in the early 1980s involving an electronic journal, *Selected Papers from Social Sciences and Humanities,* again resulted in criticism of the technology—"a major obstacle in the path to a wide acceptance of electronic journals is the inability to provide an environment conducive to convenient and efficient text reviewing" (Coward and Standera, 1985). Technical limitations were also apparent in the larger British Library BLEND experiment in the UK from 1981 through 1983 (Shackel, 1982b; Shackel, 1982a; Oakeshott, 1985). This was undertaken in major part because the telecommunications authority in that country, the British Post Office, had vetoed participation in the NSF experiment because the digital communication involved would infringe its monopoly. However, it also had severe and arbitrary limitations on participation that made it meaningless as a realistic experiment. It was not only the technology that was not ready for the digital journal. Political institutions also were unready for the information age and prepared to abuse their monopolistic powers to impede its development. In examining the potential for digital journals today, it is important to remember that even a decade ago the technology and culture were clearly inadequate to support them effectively, although this was not very apparent to those of us involved at the time.

Experimental publication of digital journals continues, and there have been positive reports on some activities in the 1990s such as the *Public-Access Computer Systems Review* on BITNET (Bailey, 1991), although it will be many years before there it is possible to evaluate the staying power of these later developments. The most important of recent developments has been the publication by the American Association for the Advancement of Science in the Spring of 1992 of a new journal, *The Online Journal of Current Clinical Trials*, as an online, peer-reviewed journal (Palca, 1991) making full use of new electronic documentation technologies (Walter, 1992). There are also many innovative uses of the network to support scholarship that replace or complement some of the functions of journals and these are reviewed in Section 5.

2.4 Issues in Developing Future Digital Journals

One major issue in assessing the feasibility of digital journals today is the current and coming states of the relevant technologies. However, it is also important to take into account that our views of digital journals are heavily influenced by the structure and operation of existing journals which in turn are biased by the print and paper medium that we currently take for granted after three hundred years of use. It is reasonable to assume that we are blinkered by our experience of current media, and that much of our thinking about the use of new media will seem naive once they have evolved into new patterns of use. It is important to go back to fundamentals and base our thinking as much as possible on the basic rationale of journals, their function to capture and disseminate knowledge, rather than their existing structures and modes of operation. Such a fundamental analysis is necessary in assessing the potential impact on scholarly communication of new representational structures, such as those of hypermedia, rather than an analysis based only on what is technologically feasible.

Prigonine adopted a foundational perspective in a recent keynote address to a conference on library automation and networking, taking up Attali's analysis of human knowledge as made up of "nomadic entities," similar to other products of our civilization, but having no fixed place (Attali, 1990), and remarked:

> *"la bibliotèque n'est plus une armoire inerte, mais un objet quasi-vivant, qui évolue selon ses contraintes propres, en même temps qu'en réponse aux pressions des usagers: elle se construit ainsi comme site privilégié, à la recontre de l'homme, de sa civilisation nomades que cette civilisation produit." (Prigogine, 1991)*

It is this capability of our repositories of knowledge to become living, growing, evolving organisms integrated with our own growth and evolving in response to it that we are currently stunting through dependence on an obsolescent medium. However, much is happening already in the use of computers and computer networks for the support of scholarly communication that does seem to have a 'life of its own' and to be generating new modes of creating and communicating knowledge. It is important to take this into account in designing and operating digital journals.

In addition, as already noted, the dissemination of knowledge, is not an end in its own right but rather part of a social process. As described in Sheridan's remarks above, it involves an authority structure of editors and referees mediating between authors and readers. It is not sufficient to assume that either this structure will continue to exist in its current form, or that it may simply be obsoleted by changes in the technology. Its role has to be analyzed not in regard to the medium

and its operation, but more fundamentally in terms of the social processes underlying knowledge generation and dissemination.

3 Social Processes in Scholarship

Journals are the major medium for discourse in the scholarly community and, as such, are intrinsically part of the social processes in that community. This section reviews the role of social processes in knowledge production, first from studies in the philosophy and sociology of science, and then more generally in terms of the function of feedback processes in efficient management of a society of cognitive agents. It summarizes these review in terms of a general classification of the interactions between scholars and their knowledge environments, and emphasizes the significance of journals in ascribing priorities to intellectual innovations as part of a social reward system encouraging knowledge production.

3.1 Social Processes in the Production of Knowledge

Because the material in journals has become detached from the activities, mental processes and existence of its originators, it may be seen as a record of those activities and external expression of those processes, completely available for the critical assessment of others independent of the originator. This gives rise to the 'objective knowledge' perspective on scholarly material, as products of human activity that are ultimately, in some sense independent of it, and autonomous in their own right (Popper, 1972). Some of these products are of major interest outside the scholarly communities generating them since they contribute to activities in society as a whole, and provide an economic return for investment in scholarship.

Hence one role of the journal is to act as a repository of knowledge and to make this widely available. This role involves the nature of knowledge, particularly as it is perceived by the client community who are not so much concerned with critical development of scholarship but with use of 'knowledge' as something that is reliable in application and whose sources can be trusted. The philosophical definitions of knowledge as "justified, true belief" provides a generative principle for the quality control that is applied to scholarly publication. The objective is *truth* but that this arises from the expression of the *beliefs* of authors through arguments which *justify* those beliefs. The refereeing processes of current journals have been developed to apply standards of 'truth' and 'justification' to the material submitted so that certain minimal levels can be relied on as applying to all material in those journals.

Justifications in the literature are generally not independent of one another but cross-reference other justifications to form a network of interdependent material. At a coarse level much of this structure is apparent through the citations between publications, and citation analysis provides a useful overview of the structures of scientific disciplines and their interrelations (Garfield, 1979; Bayer, Smart and McLaughlin, 1990; Braam, Moed and Raan, 1991a; Braam, Moed and Raan, 1991b). The formal structure of science as objective knowledge may also be modeled as a network of linked theories (Sneed, 1977; Balzer, Moulines and Sneed, 1986), whose dynamics of change may be modeled in terms of the underlying structural dependencies (Stegmüller, 1976) which are again reflected in the citation patterns in the literature.

This 'objective knowledge' perspective on journals, emphasizes the significance of the product, its quality, and the technical linkages between items. There are complementary perspectives that see knowledge production and dissemination as a human social process, and emphasize the

producers, their quality, and social linkages between them. Like all human discourse, the communications of scholars through journals serves many functions and carries many messages other than those of what is apparently overtly communicated. The review process that ensures material is of adequate quality has the side effect that the successful publication of material in a journal reflects well upon the authors—their statuses are enhanced by that of the journal in which their work is represented. In disciplines where scholarly research is dependent on access to limited resources, whose allocation is itself dependent on estimates of the status of the researchers, the journal publication process becomes part of the economic dynamics of the research program itself. This is an obvious phenomenon in disciplines requiring expensive equipment but, to a greater or lesser extent, it is ubiquitous in that the time of a scholar is itself a resource and access to employment that makes that time available for scholarship is also dependent on status often evaluated in major part through publications. More subtly, access to other scholars and their recent unpublished work through meetings, workshops and conferences, is an important resource that is again dependent on perceived status and hence to some extent on journal publication. The link between perceived capabilities and access to research resources significant in enhancing those capabilities is a positive feedback loop that gives rise to what has been termed the "Matthew effect" in science (Merton, 1968), that "unto every one who hath shall be given."

This analysis of the role of journals within the social processes of scholarship, and the interaction between the social processes and the production of knowledge, may be developed in detail through use of the literature on the development of knowledge, the history and philosophy of the scientific process, and on the psychology, sociology, politics and economics of science and scholarship in general. Feyerabend focuses on the sheer anarchy of the origins of knowledge, that there are *no universal methods* underlying scholarship and science (Feyerabend, 1975). Gellner emphasizes that our notions of 'justified' and 'true' derive from processes for the *legitimation of belief*, rather than *vice versa* (Gellner, 1974). This lack of absolutes in both the activities and value systems of scholarship may be seen as necessitating the establishment of the working *paradigms* which provide criteria for rationality in particular disciplines over particular periods, and whose change Kuhn has identified with what are perceived as *scientific revolutions* (Kuhn, 1962). The definition of these paradigms is rarely overt, and they become identified with the social norms of cultures corresponding to sub-disciplines of scholarship that Crane has termed *invisible colleges* (Crane, 1972). The editorial and refereeing processes of journals associated with these invisible colleges support these cultures at the differing levels Hall has identified (Hall, 1959): at the *informal* level by providing examples that may be mimicked; at the *formal* level by accepting or rejecting material for publication; and at the *technical* level by publishing meta-level descriptions of the aims, objectives and methodologies of the sub-discipline.

The 'rules of the game' whereby a community of scholars accepts or rejects changes in knowledge have been described by various writers from many different perspectives ranging from the purely empirical to the purely prescriptive. A major set of these rules where there seems to be some degree of consensus has been investigated by a number of historians of science, each taking specific documented phenomena of scientific development and analyzing them for evidence for, and against, each of the conjectured rules (Donovan, Laudan and Laudan, 1988). If we take advantage of the capabilities of digital journals to structure and index material in terms of the type of contribution it makes, these, and similar studies, are important in supplying a framework for classifying contributions according to their role in a process of scholarship.

The intricate involvement of journals in the processes and value systems of scholarship, and particularly the direct impact of journal publication on the allocation of resources to individual scholars, imply that the development of digital journals is not solely a technical matter. They have to be designed to play an effective role within the social infrastructure of scholarship, and any changes and extensions to the nature of the publication process have to be examined in relation to both that social structure and the underlying dynamics of knowledge itself. This is a major requirement going beyond making knowledge overt to making the metaknowledge concerned with how it comes into being overt. The following subsection gives a brief account of some of the processes involved and the role of publications in these processes.

3.2 Social Processes in the Formation and Dissemination of Expertise

In examining the social structure of science it is useful to commence from a perspective that sees it as functional, as part of a process through which our species allocates human resources appropriately for it survival (Gaines, 1989). We may view the species as consisting of a relatively undifferentiated set of cognitive agents, each of which has potential to perform useful functions, and where too much duplication would be wasteful yet detailed planning in the use of agents is not feasible because of the uncertainties of the environment. In these circumstances the positive feedback processes of the Matthew effect are functional in supporting the development of highly differentiated individual expertise in an initially uniform population through a distributed mechanism involving no central planning (Gaines, 1988). That is, the allocation of resources enabling further learning to agents with existing evidence of competence is a systemic basis for developing a diversity of capabilities in a society.

The structure and dynamics of this process of focusing the development of specialized knowledge within individuals can be analyzed in terms of a model which Hawkins has abstracted from industrial experience in developing mineral exploration expert systems and proposed as a model of human expertise relevant to expert systems (Hawkins, 1983). The model is summarized in the central diagram of five linked circles in Figure 1: the expert elicits data about the problem from the client; develops a minimal model that accounts for the data provided; generates advice based on the model and feeds this back to the client; the client may accept the advice, or query it and, possibly, the model; the queries lead to further data elicitation, and repeat of the modeling/advice/query cycle. Thus, in this model, the client plays an active role in further developing the model by providing more data until he or she is satisfied with the model and consequent advice. Expert advice giving and taking is part of a cycle of negotiation around a process of model formation.

The positive feedback processes already discussed may be analyzed in terms of this model by noting that the process of negotiation is also a basis for learning by the expert and client, and particularly so on the part of the expert who has exposure to many client situations and clients. As shown in the lower part of Figure 1, the client community provides access to practical experience through a variety of problems, particularly novel ones that go beyond existing expertise, and it also manages the growth of expertise through systems of reward, criticism and access to resources. The role of publications can also be analyzed by noting that learning from experience is a slow and error-prone process, and socially significant areas of expertise become associated with professional communities that attempt to expedite learning and reduce errors through the sharing of experience. The upper part of Figure 1 shows these other processes, such as apprenticeship, instruction, training, education, workshops, conferences, journals and books.

Professional communities usually also play a major role in refining and directing the client community's reward, criticism and resource allocation systems.

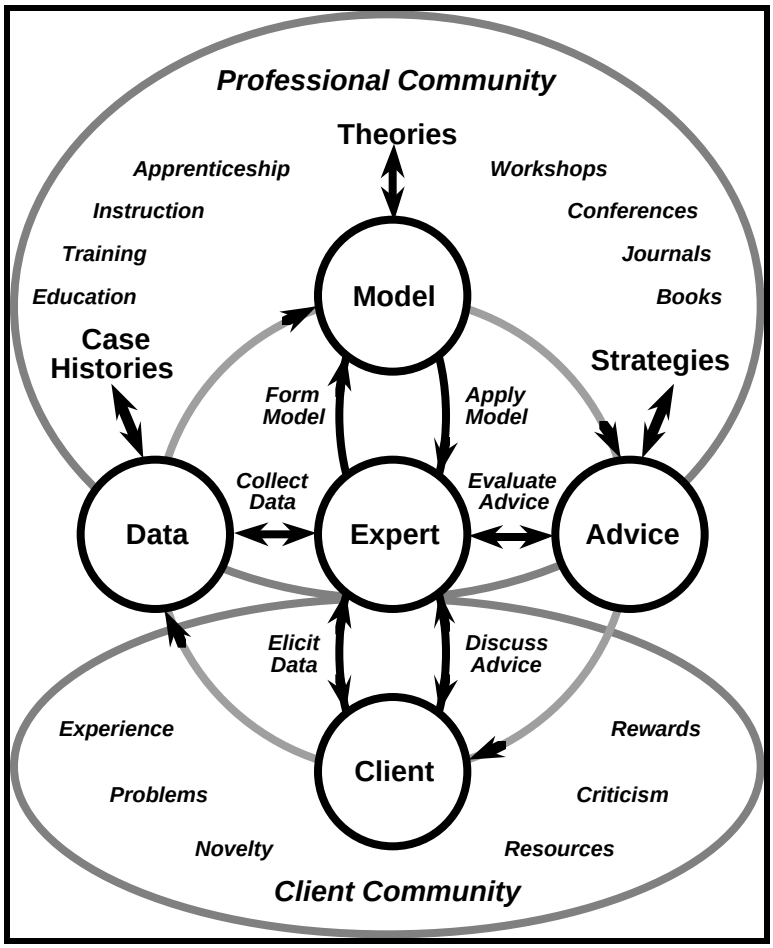

Figure 1 Processes in the formation and dissemination of expertise

In terms of overall system dynamics we may see the positive feedback processes of the Matthew effect as acting to concentrate knowledge in particular individuals and to allocate the individuals themselves as resources to different areas of knowledge generation. However, individuals have finite lifes after which direct access to their knowledge is lost, and also they cannot generate all the knowledge that they require through direct experience, so a second system is necessary that encourages the sharing of individual knowledge. This involves rewards for the social dissemination of knowledge that complement those encouraging its individual concentration. The most significant award seems to be that of allocation of priority in an act of scholarship, typically a scientific discovery, but generally any innovation perceived of value by the community (Merton, 1957). From an objective knowledge perspective it is irrelevant except as a fact of history as to who first developed an innovative idea. The society of scholars and human society at large, however, pays homage to that individual to whom an innovation is attributed, and this creates social pressures both to innovate and to provide clear evidence of having done so in a way that is widely recognized.

From a more abstract perspective the creation of an artificial reward system through ascription of priorities to innovators of ideas may be seen as precisely the kind of cultural mechanism that

Coleman suggests is necessary to bias the instability of social behavior in relation to public goods from away from 'freeriding' (Coleman, 1990). Our society treats knowledge in large part as a public good that should be made freely available through universal access to education and libraries. Patent and copyright legislation gives some potential for exploitive ownership, but are relevant to only relatively small domains of directly applicable knowledge that can result in commercial products. The vast majority of knowledge is not of this nature and it is society as a whole that benefits from its production and dissemination so that some cultural norms are necessary to encourage individuals to become participants in its production and dissemination.

The scholarly journal was perceived from its inception to be the primary mechanism whereby priorities in intellectual innovation might be registered. The first editor of what became the *Transactions of the Royal Society* wrote to the physicist Robert Boyle in encouraging him to publish his material:

> *"the Society being very carefull of registring as well the person and time of any new matter, imparted to ym, as the matter itselfe; whereby the honor of ye invention will be inviolably preserved to all posterity." (Oldenburg, 1966)*

Some two hundred years later when Ira Remsen was joining the faculty of John Hopkin's University in 1876 he asked permission to publish his investigations in a form that became the *American Chemical Journal*:

> *"1st, that we may be recognized as soon as possible as belonging to the working chemists of the country; 2nd, that the results of our labors may be insured to us, or, in other words, to establish our priority." (Carrigan, 1991)*

This social reward system of attributing priorities requires that the publication process be seen to have integrity in attributing date received and authorship, publishes in a timely fashion, and disseminates the material as widely as possible. This generates ethical requirements for the editor and publisher, and also adds another quality control dimension to refereeing apart from truth and justification, that of correct attribution. In terms of new technology, a digital journal that focuses only on knowledge capture and dissemination and does not fully support the registration and attribution of priority in innovation cannot be expected to play a significant role in scholarship.

Figure 1 also provides a microcosm in which to identify the type of information that requires capture and dissemination through journals. The "data", "model" and "advice" shown generalize to be all forms of relevant data and experience, all forms of theory, and all bases for applications and technology. One problem with existing paper-based media is that they are primarily targeted on the dissemination of theory and applications, and that the relevant data is often only available on a very erratic basis. This leads to problems of overstatement and fraud in reporting data analysis, and to problems of replication in comparing data. Information technology is already being used to alleviate these problems by establishing data banks and encouraging publicly available archiving of data. One can expect digital journals to accelerate these trends by providing links from published analyses to both the raw data and the data analysis techniques used. Indeed with the advent of low cost CD-ROM publication, there is little excuse for researchers not to give full access to a wealth of material relevant to their summarized results, including photographs and videos of laboratories and procedures, field studies, and son on.

As CD-ROM technology comes into widespread use, we may expect the culture to change in such a way that report work is suspect if not fully supported by the provision of such background

material. Claerbout's recent publication of a book on geology in parallel paper and CD-ROM form is an early example of this changing culture. The paper book is conventional (Claerbout, 1991) but the CD-ROM publication is a major innovation in publishing, containing the full contents of the book, a related PhD thesis, technical reports, data and programs. On the disk the data and programs are linked to the plots in the book as "interactive figures" such that the results plotted may be re-analyzed as the reader wishes (Claerbout and Dellinger, 1991). Claerbout is explicit about the rationale for this dual publication:

> *"The scandal of Nobelist Dr. David Baltimore being accused of fraud and ultimately stepping down as President of Rockefeller University carries many interesting lessons for laboratory scientists. There are lessons too for us whose research is heavily computer dependent. From a journal article alone, the task of reproducing the computations and data analysis leading to computer-generated illustrations is often formidable. Typically, the authors themselves, if stripped of their files, starting over again from only their data and their published paper, might require weeks or many months to reproduce the illustrations on which their conclusions are based. Referees and readers rarely attempt to reproduce anything because the effort could be comparable to the first two years of work on a dissertation. I thought this reproducibility problem hopelessly mired in human nature when, much to my amazement, I solved it by accident using paper-plus-electronic publication technology."*

3.3 Interactions between Scholars and their Knowledge Environments

Figure 2 summarizes the general relations between a scholar and his or her intellectual environment in the creation and dissemination of knowledge. On the left the scholar interacts with some research domain which is of the nature of Popper's (1968) *World 1*, sometimes the physical world, but more generally a circumscribed domain defined by a set of natural or artificial constraints. On the right the scholar interacts with some research community which is of the nature of Popper's *World 2*, the mental and social processes that combine to form the culture of a human community concerned with the domain. In the center the scholar interacts with some research literature which is of the nature of Popper's *World 3*, the overt knowledge of that community expressed in a form intended to be largely independent of individual mental processes.

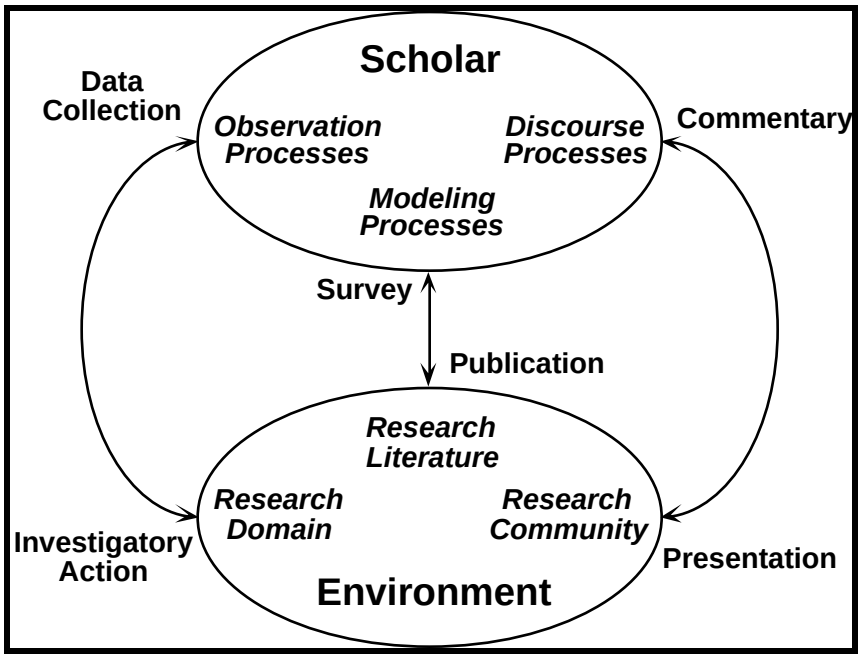

Figure 2 Interactions between scholars and sources of knowledge

The journal, in paper or digital form, has a clear role to play in the central interaction. However, one must remember that it is not the only medium at work. In particular, books play a significant role in the research literature which overlaps with, but is in major part distinct from the functions of journals. Books generally bypass the intellectual gatekeepers of the refereeing process, often being reviewed more pragmatically for their potential sales rather than for truth and justification. This is a healthy counterbalance to over-evaluation in the heavily refereed and condensed literature of journals whose editorial requirements may impede publication of more speculative material that often also needs greater length in which to present an exposition of nonstandard assumptions. At the other extreme textbooks require extended length to give a tutorial exposition of generally accepted positions.

Alternative forms of publication to journals, such as conference proceedings, are also associated with the less formal community processes on the right of Figure 2. It is common in many disciplines for presentations and discussions and the circulation of manuscript drafts to precede the publication of papers in journal form. Thus, in general there is a migration of knowledge from working papers, through workshop proceedings to journals and then to books. This is very compartmentalized currently and there appears to be no technical reason why much the same information technology should not serve all these purposes in digital form. In considering such a universal medium, however, it is important to take into account the quality control issues discussed in relation to journal publication, both in the maintenance of standards of justification in particular disciplines and in the ascription of status and priority in innovation in those disciplines.

3.4 The Impact of the Matthew Effect on Publication Opportunities

There have been a number of detailed studies of the operation of the editorial and refereeing systems of current journals, documenting the procedures that are followed, attention to fairness, standards of quality, acceptance rates in different journals and different disciplines, and the fate

of papers rejected by a publication (Zuckerman and Merton, 1971; Abelson, 1990; Crawford and Stucki, 1990; Garfield, 1990). There are also many accounts of the refereeing process (Matkin and Riggar, 1991) both from the point of view of the author (Graham and Stablein, 1985; Schoorman, 1985; Leslie, 1989) and from that of the reviewer (Daft, 1985; Pondy, 1985; Schwab, 1985). The literature gives a picture of a typical humanly managed process in which many of the rules of the game are implicit, and there is much variation in process, but which is seen overall as performing reasonably well in promoting quality without being able to give any guarantees of 'truth' or solid defense against fraud.

One concern of newcomers to a field is the difficulty of getting past the gatekeepers operating the editorial and refereeing process. Indeed the general model of positive feedback processes in science outlined above may raise concerns that the refereeing system could be too rigid in institutionalizing an obsolete paradigm and stultifying innovation. This could indicate that one role for digital publication would be to allow individuals more freedom to publish without the restrictions imposed by the limited resource of the availability of space in paper publications, such as control by an over-rigid refereeing system. However, empirical studies suggest that quite the contrary is true, that the Matthew effect is very small in relation both to opportunities for publication and in the evaluation of publications (Cole and Cole, 1973). Papers that are rejected for one journal are accepted by others, and it is relatively easy to launch a new journal supporting an alternative paradigm. The prestige of a tightly refereed and accepted journal may be missing when an innovative paper is first published, but that paper becomes available for further assessment apart from the refereeing system and, if it becomes recognized and widely cited, both it and the journal in which it is published gain in prestige. The history of scholarship is full of innovations that were rejected initially only to be recognized as major advances at a later stage, and the search for good ideas in obscure sources is a major activity in scholarship. The main impact of the Matthew effect in all its aspects is to influence the career trajectories of individual scholars in ways that affects the research they actually carry out (Nowakowska, 1975), rather than prevent the publication of work already undertaken..

The problems of the current system of journal publication are not ones of difficulty in publication but rather ones of volume creating information overload, one impact of which is on the refereeing system itself. This does not seem to have been documented, but as a journal editor, I can report anecdotally that finding scholars with the time and capability to undertake reviewing is a major problem and one that seems to be worsening with the increasing volume of publication. The problem can be seen in quantitative terms if one considers a single-authored paper in a journal that obtains three reviews for each paper submitted and has an acceptance rate of fifty per cent. Since the communities of successful authors and suitable referees are the same, the author of an accepted paper in these circumstances 'owes' the journal six reviews of other papers if the 'accounts' of papers and reviews are to balance. Most authors seem to regard that as an excessive demand. The anonymity of refereeing in particular means that there is no overt reward for what can be a major contribution to the quality of a paper, and quite often a significant original contribution to its content. Information technology is already being used to support editors in managing the selection of referees (Sanford, 1991), but it does not currently support the reviewers in their task or give any incentive to undertake that task.

Our current system of anonymous refereeing is set up as a public good with incentive to be a freerider, and it may be one of the areas most suited to major change as digital journals develop.

For example, publications might be made available without refereeing and then their quality and status assessed through open peer commentary linked back to the original paper. In this way the contributions of those improving a publication would be recognized, and the assessment of a paper would evolve over time. This is not markedly different from current processes of peer commentary, but digital publication makes it very much more effective. An erratum note, or a critical note pointing out an error in a paper, published in a subsequent issue of a journal is easily missed. It is important to do as much quality control as possible in advance. However, a digital publication can evolve organically as a linked network of versions, annotations and contributions. It will always be to the advantage of an author to get as much right as possible initially, and slipshod papers will be far less likely to attract commentary. Quality control will still be very important but it will be less of an all-or-nothing activity than it is now.

4 Technological Aspects of Digital Journals

Since the availability of adequate technology is a critical enabling factor to the implementation of digital journals, this section reviews whether information technology in the 1990s has advanced in such a way as to overcome the problems that undermined electronic journal experiments in the 1970s and 1980s. Overall trends in information technologies are first reviewed, then the capabilities and costs of those essentially required for digital journals, the problems of the diversity of standards for document interchange, and the digital support of access control, authentication, registration and management of replicated archives.

4. 1 Learning Curves in the Infrastructure of Information Technology

A foundational perspective on trends in information technology may be gained by examining the *learning curves* that characterize innovation and diffusion in all technologies and determine the dynamics of technological substitution (Ayres, 1968; Marchetti, 1980). Logistic curves have been found to be a useful model of the introduction of new knowledge, technology or product in which growth takes off slowly, begins to climb rapidly and then slows down as whatever was introduced has been assimilated. Such curves arise in many different disciplines such as education, ecology, economics, marketing and technological forecasting (Dujin, 1983; Stoneman, 1983). From the learning curves for information technology it should be possible to determine the state of the technology in relation to digital journal requirements during past experiments and the present and coming generations of the technology.

It has also been noted in many disciplines that the qualitative phenomena during the growth of the logistic curve vary from stage to stage (Crane, 1972; De Mey, 1982; Gaines and Shaw, 1986). The era before the learning curve takes off, when too little is known for planned progress, is that of the inventor having very little chance of success. When an inventor makes a *breakthrough*, very rapidly his or her work is *replicated* at research institutions world-wide. The experience gained in this way leads to *empirical* design rules with very little foundation except previous successes and failures. However, as enough empirical experience is gained it becomes possible to inductively model the basis of success and failure and develop *theories*. This transition from empiricism to theory corresponds to the maximum slope of the logistic learning curve. The theoretical models make it possible to *automate* the scientific data gathering and analysis and associated manufacturing processes. Once automation has been put in place effort can focus on cost reduction and quality improvements in what has become a *mature* technology.

The fast, sustained, learning curve for electronic devices, and the scope for positive feedback in the information sciences, together result in a tiered infrastructure for the information sciences and technologies which is fundamental to their nature. It involves a succession of learning curves as rapid advances in one level of technology trigger off invention in others as shown in Figure 3 (Gaines, 1990a; Gaines, 1991b). Advances in digital *electronic device technology* in the 1930s allowed the *virtual machine architecture* that detached computing from electronics to be developed in the late 1940s. This in turn triggered off developments in *problem orientated languages* in the 1950s, and the increasing reliability of computers enabled *interactive activity systems* to be developed in the 1960s. It is this last development that supports digital communications, interactive systems and networking and is most critical to digital journals.

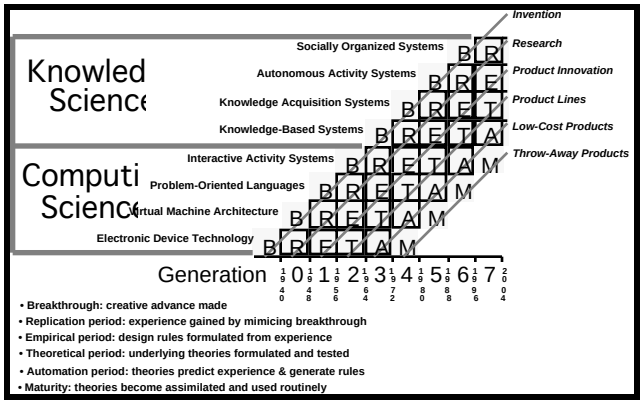

Figure 3 Learning curves in the infrastructure of the information sciences

A major transition from information technology to knowledge technology took place in the 1970s with the advent of *knowledge-based systems*, and this in its turn has stimulated research in *knowledge acquisition systems*, *autonomous activity systems*, and *socially organized systems*. This research on the computer processing of knowledge, as contrasted with its digitization, communication and delivery, is not critical to the digital journal as a substitute for the paper journal. However, it may be expected to have a major impact on the evolving role of computers in scholarly activities. Developments in artificial intelligence and computational linguistics, such as semantic networks and text modeling, may add such significant functionality to digital journals that they radically change their nature. Such developments belong to the new era of knowledge science that focuses more on the content of computer-based media than on the underlying technologies of hardware, software, communications and human-computer interfaces.

One important factor in the viability of digital journals that is apparent in Figure 3 is that there is a significant time interval between inventions and product innovation based on them. This seems to be remarkably constant at about seventeen years across diverse areas of technology (Mensch, 1975). In terms of the learning curves this interval corresponds to the *replication/empirical* periods when the technology is first being investigated. There is a corresponding interval during the *theory/automation* periods before products become mature, and automated mass-production at low cost is feasible. The resultant trajectories of *invention*, *research*, *product innovation*, *product lines*, *low-cost products* and *throw-away products* where replacement is cheaper than maintenance are shown superimposed on Figure 3. It becomes reasonable for pioneering applications to experiment with the new technology during the product innovation era, but it is not until low-cost products become available that widespread use will become common, and it is

the mature technology of throw-away products that becomes used routinely as the expected mode of operation.

The learning curves for the first four technologies are all critical to digital journals, and the limiting one is the last to develop, that for interactive activity systems. From the time scale of development along this curve in Figure 3, one would expect:

- experimental use in the fourth generation era from 1972 through 1980, the era of time-sharing systems and the advent of the Apple II personal computer;
- routine use in selected professional areas where it was cost-effective during the fifth generation era from 1980 through 1988, the era of the IBM personal computer, Sun workstation, and the growth of USENET;
- widespread use during the current sixth generation era from 1988 through 1996, the era of workstations at the price of personal computers, ready international availability of Internet, CD-ROM storage at virtually zero cost, and low-cost, high-quality laser printing.

It will probably not be until the seventh generation era from 1996 through 2004 that cost reduction in the required technology will make electronic publication the preferred technology to paper publication, and we make expect the transition to be complex because when substitution occurs in major technologies there is substantial repositioning of products based on existing technology by commercial vendors. However, at least the learning curves model provides a framework within which to analyze digital journal technology within general trends in information technology, and also a basis for future projections as to how the nature of digital journals may change from a technological perspective.

4.2 Specific Technologies Critical to the Digital Journal

The following sub-sections provide some basic figures on the current state and trends in the technologies relevant to digital journals. Figure 4 puts these technologies in context by illustrating the typical working environment of a modern scholar in relation to information technologies. It should be emphasized for all ensuing discussion that this environment is not specific to computer scientists, or the physical sciences or the professions, but is becoming increasingly common across all scholarship including the arts and humanities. Even the ‘laboratory’ shown is not specific to what are classically regarded as laboratory sciences—the ‘computer laboratory’ brings instrumentation to what have been purely intellectual activities involving no technology in the past. For example, as the use of electronic mail, computer conferencing and computer-supported collaborative research grows across all disciplines there are opportunities for the study of human discourse and knowledge processes that offer new research possibilities in the humanities.

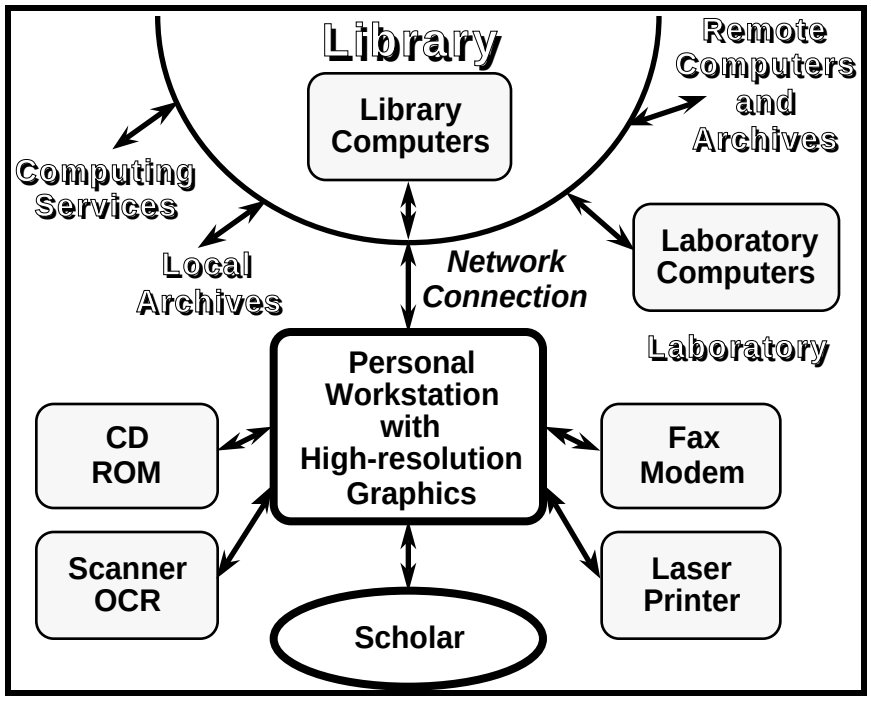

Figure 4 The working environment of the modern scholar

The technologies specifically related to digital journals in Figure 4 are:

- The high resolution graphic workstation allowing typographic quality text and illustrations to be both browsed and created (Earnshaw, 1987).
- The laser printer allowing such material to be selectively printed locally.
- The fax modem allowing material to be rapidly transmitted to, and received from, locations not having such computer facilities.
- The scanner with optical character recognition allowing printed pictures and text to be entered into the system.
- The CD-ROM disk reader allowing high volumes of material to be accessed from compact storage that may be mailed at low cost (Lambert and Ropiequet, 1986; Hendley, 1987).
- The network connection through cable or telephone providing access to (Quarterman, 1990):
  - — Laboratory computers allowing direct capture of data for publication.
  - — Computing services allowing direct capture of analyzed data and simulations.
  - — Local archives of documents put up for local and remote access.
  - — Library catalogs and archives of published material.
  - — Remote computers and archives giving communication with collaborators and access to research materials at other sites.

4.3 Cost and Performance of Workstations

The cost of the workstation and peripherals shown is now under US$7,000. This may be put into an perspective as being equivalent to about one month’s salary for the scholar, a comparison that will become increasingly favorable as computer costs continue to decline at about 30% a year

while salaries increase at about 4%. Twelve years ago the workstation configuration shown would have cost a quarter of a million dollars and been equivalent to some eight researcher-years. These cost comparisons illustrate how sharp transitions in working practice arise from the deceptively smooth learning curves of the underlying technologies. Major changes become appropriate as cost thresholds are passed.

In terms of accessing the medium, the technical capabilities of the system shown are such that two full pages of text may be shown on the computer screen with adequate resolution for normal typography and line-art diagrams, and documents may be paged through at rates similar to that of manipulating a journal. These comparisons assume a screen resolution of about one million pixels, and buffer storage of at least one Mbyte. They are insensitive to disk transfer rates and seek times, provided the bus controller, file organization and software can cope with large block transfers. The high-speed display of typographic text and images requires a powerful processor and highly optimized software, but both are now routinely available with personal computers.

Monochrome photographs may be viewed at a newspaper level of resolution, but high quality color reproduction requires a more expensive system and also slows down paging due to the transfers of large data structures required. Thus journals requiring numbers of high-quality color plates may still be problematic in digital form, but on current technology trends they will not be so within the decade. The accurate color rendering and image processing capabilities of graphical workstations will greatly increase the value of access to photographic material, and make digitized continuous tone material the preferred medium in the long term. Apart from storage requirements, there are no significant additional problems in providing access to video material, and this again will increase the capabilities of the medium.

Similar considerations apply to sound digitization, and it is to be expected that a transition to multimedia communication will be one of the first major changes in the nature of scholarly communication brought about by the development of digital journals. It is important to take into account this possibility when assessing storage and communication requirements for digital journals, and when planning for parallel publication in paper form. For the remainder of this section, calculations will continue to be based on the replication of existing journal functionality rather than its extension to multimedia.

4.4 Cost and Performance of Storage

In terms of storage and communication, current CD-ROMS can hold some 750 Mbytes of data which is adequate for about 500 books or 10,000 papers that are mainly text, about half that number if diagrams predominate, and about one tenth if high quality color pictures are common. The basic cost of the medium is such that CD-ROMs may be replicated at about $1 each in quantities of a few thousand. The capital cost of the manufacturing technology is high but there are now bureaus creating smaller quantities at somewhat higher costs that may be used by institutions and individuals.

The cost of archives held on magnetic storage is now under $1,000 for a Gbyte of storage. It again puts this in perspective if one envisions a library archive holding the 500,000,000 books estimated to have been published up to 1975 (Gore, 1976). The capital cost of on-line storage would be under one million dollars and dropping at 30% a year. Archives on CD-ROM could cost substantially less, but CD-ROM drives are unsuited for on-line access as they cost about the same as magnetic storage which is much higher in speed of access. In assessing these figures one

should note that they are raw technology costs that do not take into account the cost of capturing material in digital form or royalty payments for access to the material. They indicate, however, that the costs of large-scale archives for digital journals will not be a limiting factor in their development.

4.5 Cost and Performance of Communications

Remote communication costs are important to estimate because they affect the economics of distribution of material, whether to transmit physically through CD-ROM or electronically through the network, and whether the communications cost of access to remote archives is cost-competitive relative to the storage costs of local archives. However, network costs are difficult to estimate because they involve a ramified distribution system in which use of telecommunication facilities is optimized under a complex tariff system, and the cost of expensive resources such as international trunk links is distributed somewhat *ad hoc.* Individual users in many institutions do not contribute directly to the costs of network access, but rough estimates from institutional budgets would suggest that allocated costs are a few hundred dollars for each scholar.

Speed of network communication is also extremely variable and varies with local conditions and patterns of access. A typical figure on long-haul access over Internet might be about 1,000 bytes/second allowing a twenty page scientific paper to be transmitted in about a minute. As usage of the existing network increases problems of overload may occur, but the network bandwidth is being continually upgraded to keep pace with demand, and technology cost reductions are keeping the cost of access comparatively stable, so it is reasonable to project that the current capabilities of the network will improve rather than worsen. In comparing the technical merits of network and CD-ROM distribution it is important to remember that ease of network access is still largely restricted to the developed nations, and scholars in third world countries in particular may not have effective access to networks (Samarajiva, 1988; Matta and Boutros, 1989)—in this respect, as in others, the technologies are complementary rather than competitive.

The current network system provides transportation facilities for digitally encoded material but leaves much to be desired in supporting ease of access to the network. This is being addressed by many new software subsystems that organize user access to the network more effectively. Wide Area Information Services (WAIS) are one important development that provides a protocol for access to a large number of information services world-wide (Stein, 1991). The WAIS protocol is a modified form of the ANSI-NISO standard for library information retrieval, Z39.50-1988 (also proceeding within ISO as DIS 10162 and DIS 10163). Server software for this protocol has been placed in the public domain, as have a number of client systems for workstations and personal computers. The WAIS protocol is itself a vehicle for many other services that specify linkage structures within the documents being retrieved so that, for example, one may retrieve those documents referenced in a paper by a simple procedure involving the reference. World-Wide Web is such a system, again in the public domain, that provides access to a wide variety of information services through simple, context-dependent procedures designed to be easy to use. Internet Gopher is another protocol originally designed to support campus-wide information systems but being essentially independent of the extent of the network. World-Wide Web links into WAIS, Gopher and other such services internationally providing users with fairly seamless to an ever-growing range of information resources. Information about these systems is difficult to cite in a paper publication because it is largely available over Internet through the services

themselves. The USENET news groups *alt.gopher, alt.hypertext, alt.wais* and *comp.mail.multi-media* are the primary source of communications about these activities.

4.6 Standards for Document Interchange

Thus, the raw technology for digital journals exists at a reasonable cost and is already in use for related applications. However, there is a major impediment to the use of that technology to support effective journals and that is the lack of standardization of the typographic, image and page layout information that constitute the digital representation of journal material. The basic character codes for computer programs, non typographic text, and numeric data, have long been standardized but beyond that there is currently chaos. Computer typesetting systems have used proprietary coding schemes designed to lock users into particular vendors rather than facilitate data interchange. Word processing and page layout programs on personal computers have similar proprietary formats, although Microsoft's Rich Text Format (RTF) (Walden, 1987) has become a *de facto* partial interchange standard as a result of commercial needs to offer users ways of migrating from one program to another.

In the academic world TEX (Knuth, 1986) has become widely used by mathematicians and computer scientists since its adoption by the American Mathematical Society, and is treated as a standard in related publications. The device independent intermediate output format of TEX, DVI format, has become widely used for archives of papers in computing and mathematics since it is reasonably compact and readily converted to output on laser printers. The lower level image format of PostScript (Adobe Systems Incorporated, 1986) which is the 'machine language' of many laser printers is also used as a common archiving format since it is non-proprietary and can be exported from many word processing and page layout programs.

The difficulty with going to as low-level a format as Postscript is that it is not compact, not universally usable as an output standard, and not indexable or editable as a document. The difficulty with any higher level standards currently is that they are substantially less universal, and the more proprietary ones are also not indexable or editable except in the related application software. There are two solutions to this problem on the horizon. One is that the growing need for document interchange between word processor users has encouraged the development of a substantial secondary industry offering data conversion programs. Many such programs are now available at low cost for personal computers and it is becoming reasonably safe to assume that one will be able to convert a document developed on one mainstream document processing system to be accessible on another mainstream system.

The second solution to the lack of formatting standards is the development of graphic and typographic structuring languages that are designed for document interchange. The Standard Generalized Markup Language SGML (Bryan, 1988; Goldfarb, 1990; Herwijnen, 1990) is the major contender having been adopted by the International Organization for Standardization as ISO 8879, with parallel adoption in the US as FIPS-PUB-152, the UK as BS 6868 and Europe as European Standard 28879. SGML is not a text layout or page description language as are those discussed above. It is rather a way of embedding a knowledge structure in a document that specifies the roles, within a predefined structure of potential roles, of each component of the text, for example, title, author, first-level heading, and so on. In word processor terminology, SGML supports the capability to label text components from what can be a very elaborate 'style sheet.' The primary objective of SGML is to mark up a text with the knowledge that a copy editor might

use in making typographic and layout decisions, rather than with the typographic and layout information itself.

SGML is a very flexible standard, essentially providing a meta-language for defining markup conventions through formal Document Type Definitions (DTDs), and this enables it to be applied to the representation of complex document structures such as those for mathematics, music, chemical formulae, and hypertextual links. The SGML standard may also be used to define document structures that are not so much concerned with typography and layout as with information retrieval. The protocol used in World-Wide Web documents is HTML (hypertext markup language), a particular format within the SGML rules. SGML does not address issues of image interchange formats but it supports the markup of images as being files defined in other ISO, or industry, standards.

The major advantage of this from a conventional publishing perspective is that different 'house styles' can be applied to the same SGML document. An associated disadvantage to an author is that he or she loses the control of layout that comes from supplying 'camera-ready' copy. However, the author in this situation can also use an SGML publishing system on their personal computer to apply their own 'house style' to the SGML text. The only loss then relative to a word processing or page make up system is that they do not have a 'what you see is what you get' user interface for the editing process. It is not clear yet how significant these issues will be in practice, and they can, in any event, be addressed by developments in word processor technology to combine SGML with continuous page layout. Brailsford and Beach (1989) provide an interesting discussion of these issues in the context of their planning for the journal *Electronic Publishing* which commenced in 1988 to itself become an electronic publication at a later date, and Alexander and Water (1990) provide a critical commentary on SGML from a publisher's point of view.

The more profound advantages of SGML come from its supporting knowledge structures for free text databases. The information provided in the markup language is exactly that necessary for automatic intelligent indexing and the re-presentation of the material in documents in a variety of forms, such as table of contents, outliner form, and with various forms of hypertext linkage. The preparation of high-volume document databases on CD-ROM has been one of the major areas of successful application of SGML. The problem of converting from word processor formats to SGML is already been tackled by service bureaus who have developed the necessary translation software and, since it is similar to that of the conversions required between word processors, we may expect such conversion to become routinely available on personal computers as the need for it arises.

The Office Document Architecture (ODA) is another standard (ISO 8879) that uses SGML to encode presentational information also such as document layout and graphic formats. It has been used in the EXPRES Project (Rosenberg, Sherman, Marks and Akkerhuis, 1991) at Carnegie Mellon University to provide a comprehensive environment for multimedia document preparation, communication and presentation under X-Windows. Various personal computer manufacturers are committed to providing ODA support as part of their operating systems, primarily to provide an interchange and presentation technology for complex multimedia document architectures. However, this also means it is likely to become widely available for standard documents in word processors and page layout programs on personal computers. There is also a standard under development for a Document Style Semantics and Specification

Language (DSSSL) that complements SGML, and it is not yet clear how this will relate to ODA (Walter, 1989).

This variety of document formats, industry 'standards,' and ISO current and developing standards, is a major impediment to digital journal dissemination. However, it is not peculiar to journals but a problem for electronic data interchange in many significant applications in various industries. Hence, it is being solved in a variety of ways, notably by conversion utilities that are becoming increasingly effective and are already low in cost. There are many options available to the developer of an effective electronic journal system, such as accepting material in a variety of formats and converting to a standard format which is both widely acceptable and where readers can be provided with a low-cost viewing software. There are many document archives already in existence that adopt some effective approach to this problem, and it is not so much a block to digital journal production as yet another technological problem that has to be managed taking into account changing technologies and evolving standards.

The proposed Internet MIME protocol for multipurpose Internet mail extensions (Borenstein and Freed, 1993) is not based on any of these standards but rather provides a meta-standard within which plain text, RTF, PostScript, SGML, ODA, and other document formats can be used as required. MIME will enable a wide range of text, typographic text, fully made up documents, pictures, video and sounds, to be transmitted over Internet through a well-defined protocol. It is expected that many convertors and filters will be developed for MIME so that material send in any format can be viewed in systems, or on equipment, that does not fully support that format. MIME will probably be adopted also for other media such as CD-ROMs since similar considerations apply to the converting and viewing of material.

4.7 Impacts of Digital Publication on Citation Practice

Whereas the emulation of existing practice will be important for the acceptance of digital journals, at least in the early stages, there are some important aspects of existing practice which are peculiar to paper publication and where emulation would be a highly inappropriate restriction. Page numbers are the outstanding example. Paper publications are frozen into an immutable format that is related to physical access, and page numbers are a valid form of citation useful for access. However, a major feature of digital documents is flexibility in presentation, and pagination should be seen as under control of the user rather than the originator. This is not only a question of image size but also of the variation of type faces and sizes for optimum legibility on the users' systems.

The main impact of allowing flexibility in pagination is upon citation practice where the volume and page numbers of a paper have become the standard basis for reference. This may be seen as a historic accident since the accession number of a paper within a volume, a year, or within the lifetime as the journal, would be an equally valid basis for citation. It is simple to resolve this problem: a digital journal may use a sequential accession number to reference a particular paper; one paralleling a paper journal may continue to use the volume and page numbers, but now as a code only to reference the paper not as an indicator of actual 'pages.' What is important is that the paper-based notion of a fixed page is not emulated as a restrictive anachronism in digital journals.

Where more specific citations are desired, such as those for specific quotations, these can be done on the basis of context, such as section heading and paragraph. Searches on contextual

information are simple in digital publications, and one may expect changes in citation style to take advantage of this capability. This will become even more significant if there is a significant trend towards publications which do not follow a linear form but use various modes of digital information retrieval to allow users to tailor their access to specific needs.

4.8 Access Control

The problem of copyright in electronic documents is a significant one, as is access control to document archives if charges are intended to be made to provide commercial revenues or contributions to support costs. Access to documents in digital form offers possibilities for reuse without consent of the copyright holder and for plagiarism of scholarly material. Normal methods of access control and monitoring can be bypassed as soon as one legitimate accessor has a copy of a document in digital form that can be passed on to others. Controlling these problems through computer techniques without removing many of the benefits of users having full access to digital documents is virtually impossible. Legitimate and illegitimate reprocessing are not distinguishable through the information available to the computer system. They are differentiated by legal and social norms that reference the relevant situation and the intentions of the human agents involved. A similar situation applies to copying, converting and reusing paper publications, and it has been argued that current copyright law is adequate to cover electronic publication (Garrett, 1991). In particular, plagiarism of digital documents within digital documents is very easy to demonstrate through computer analysis and is already being used in legal cases (Stone, 1991).

There are, unfortunately, already software vendors attempting to attract publishers to elaborate copy protection schemes for CD-ROMs that go further than many of those used in the early days of software sold on magnetic disks. While such schemes may be appropriate for some highly commercial information distributed digitally, they so undermine the benefits of digital publication that they will hopefully never come into use. The commercial failure of similar exercises for magnetic disks should provide the information that protection schemes have a negative influence on sales and customer relations, and are not effective in the long term.

It is also possible that digital publication may radically change attitudes towards copyright. Brent has developed a very thought-provoking analysis of the possible future of electronic knowledge based on the historical development of the oral tradition and its replacement by printed media. He notes:

> *“I believe that computer mediated communication provides a totally different metaphorical message, one that can take theories of collaborative knowledge out of the realm of language philosophy and stamp them indelibly in the consciousness of the entire society. Let us begin by looking at what is now the most mundane aspect of computer-mediated communication, word processing. Remember that one of the most important psychological effects of writing in general and the printing press in particular is the fossilization of text as an exteriorized object. However, composing on a word processor divides the production of the text into two distinct stages. Ultimately the text issues in a final stage of more or less complete closure, once a "final" draft is published in a hard codex. But the word processor greatly extends the fluid stage of text, abolishing the sense of discrete drafts and smaller divisible units (pages) and turning the text into a long continuous document, a scroll examined through a twenty-five line sliding window...A key*

aspect of this form of text is that it can easily be recombined with other texts. Skilled writers who use word processors are well aware of how often they cannibalize their own older texts for quotations, well-turned paragraphs, ideas cut out of drafts and saved for future works in which they might be more appropriate. But this effect does not become truly significant until the writer's own text begins to interact with other sources of text available on-line. The word processor is often seen as a preliminary stage of conferencing, for posted text is often prepared initially on some kind of word processor. However, this metaphor can be reversed: the word processor is coming to be fed by on-line information as much as the reverse. As other sources of text become available in machine-readable format--texts received through electronic conferences and on-line publications, texts downloaded from databases, et cetera--the awareness of intertextuality that LeFevre speaks of becomes increasingly objectified, its implications increasingly unmistakable." (Brent, 1991)

Brent speculates that because of these phenomena the notion of ownership will become diminished in digital publications and issues of copyright will tend to become less significant as the network of collaborative knowledge becomes more and more apparent.

4.9 Authentication

A rather more serious problem where technological solutions are essential is in the *authentication* of digital documents, in ensuring that the document the reader has is precisely that intended by the author and not an edited copy. This is not only a problem of error and fraud, but also one of document dynamics in that the author may have issued a series of variants of the 'same' document, for example, through the refereeing process, in response to criticism and to the detection of errors. Readers need to be able to determine what version of the document they have, be assured that it has not been edited, and determine what is the latest version of the document available.

These requirements apply to any form of electronic mail (Sherblom, 1988; Robertson, 1991), not just digital journals, and are being resolved by the use of encryption based on public and private key systems (Tanaka and Okamoto, 1991). The basic principle is that a document encrypted by you using your private key can be decrypted by anyone knowing your public key. Indeed this can be taken further in that a document encrypted with your public key can only be decrypted with your private key, so that double-decryption supports both authentication and access control. However, for digital journals authentication will generally be all that is required. Like document standard support, authentication is generally important to network communication and standards and services in this area are being supported by all the major computer manufacturers.

4.10 Registration

Technology for authentication can also be used to support another function of current journals and that is the *registration* of the date of receipt of a paper from an author to record priorities of ideas and inventions. This is an important side-effect of the journal publication process whose social function is analyzed in the following section. Registering a paper with an archive, that serves the role of a journal editorial office in date-stamping a paper and encrypting it using the archive's private key, allows a paper to be issued such that no-one, including the author, can change the contents or the date-stamp. The prevention of fraud through tampering with the archive server then becomes significant if there are significant incentives for such fraud to take

place. However, it will be readily detectable if a number of reputable users or sub-archives have taken copies of a document at the time of its original accession, in the same way that changing a received date on a published paper is ineffective once it has been disseminated.

4.11 Replicated Archives

It is to be expected that, as trade-offs between communication costs and local storage costs are optimized, documents will become replicated in a variety of archives on an anarchistic basis. Fortunately, the essence of a published work is that it can not be changed without issuing a new version, so the requirement is for *immutable* storage making the problems of maintaining a networked, distributed database fairly straightforward compared to more general requirements. The main problems with replicated archives is to automate the effective use of communications, for example, transferring bulk materials on the network during low-usage periods, and to ensure that all sites have access to status information about the sequence of versions, particularly the latest one. Protocols for the operation of replicated repositories have already been developed that can deal effectively with these problems (Prusker and Wobber, 1991).

4.12 Information Retrieval

The emphasis in the technology evaluation above has been on the digital journal as a parallel to, or substitute for, the paper journal, largely in terms of cost, quality and convenience. A major advantage of digital publication is, however, the computer access to the content of publications allowing them to be processed in various ways including both general and specialized indexes. There are already available low cost software packages for microcomputers that maintain indexes to large bodies of text automatically and offer excellent information retrieval capabilities. There is a large body of research literature on the computer analysis of text for purposes of scholarship (Tenopir, 1984; Callon, Law and Rip, 1986; Deerwester, Dumais, Furnas, Landauer and Harshman, 1990; Marshall, 1991), and the availability of digital publication will facilitate the routine use of such tools.

There is also scope for integration of existing information retrieval services with digital publication document archives to allow very much more effective access to the literature. Information retrieval is a heavily researched and studied subject in its own right and there is a wealth of services and ideas that can be incorporated as part of the infrastructure of digital journals (Oddy, Robertson, Rijsbergen and Williams, 1981; Sievert, 1990; Kuhlthau, 1991). In particular, it is highly desirable that citations in digital publications become actively linked to information retrieval services so that readers can access abstracts and contents of references simply and rapidly while reading a document online.

4.13 Hypertext

It is difficult to determine to what extent the linear presentation of material in paper journals is an artefact of the medium and to what extent is a natural result of our cognitive processes in relation to argument forms. Some proponents of hypertext have assumed that the freedom to generate branching presentations is itself a major advantage of digital publication, but empirical studies show problems of cognitive disorientation in users of hypertextual documents (McKnight, Dillon and Richardson, 1991), and it is not clear whether these are intrinsic, the result of poor technology, the result of poor use of the medium, or the result of cognitive bias due to long experience of linear media.

However, there is a growing body of literature on the effective use of hypertext in scholarship (Barrett, 1988; Barrett, 1989; Delany and Landow, 1990), and the availability of contextual links into, and out of, an otherwise conventionally structured document is an important feature of digital publications even if they are not written as branching documents. In particular, it allows annotation, criticism and commentary to be incrementally added to a document and made readily accessible to a reader, and this supports many scholarly activities. The capability for hypertext linkage is technically extremely simple to provide, including the retrofitting of documents initially published without it, and it is reasonable to assume that it will become a routinely available feature of digital publication.

4.14 Multimedia

As noted in the section on document standards, many existing and proposed standards include the encoding and presentation of material such as video and sound that go beyond those currently supported by paper publication. The support of interactive multimedia material has been a major thrust in the commercial development of personal computers (Ambron and Hooper, 1988), and a number of standard word processors now support the incorporation of video and sound material. There are also standards developing for the transmission of such material on Internet, and it is reasonable to assume that multimedia facilities may be a requirement attracting some potential users of digital journals. This has significant implications for the planning and feasibility of parallel paper publications. The hardware, software and application of multimedia is a major area of research currently and on should expect a rapid pace of change in all aspects of the technology (Huston, 1990; Blair and Davies, 1991; Gibbs, 1991; Hoepner, 1991; Meghini, Rabitti and Thanos, 1991; O'Docherty and Daskalakis, 1991; Philips, 1991; Rosenberg et al., 1991).

4.15 Knowledge Representation and Cognitive Maps

The dynamic processing and graphic capabilities of personal workstations make it possible to support facilities for knowledge structures ranging from concept maps (Novak and Gowin, 1984; Lambiotte, Dansereau, Cross and Reynolds, 1989) to visual languages for formal knowledge representation and deduction (Gaines, 1991a), and to link these to documents and information retrieval systems. Interactive graphic programs for eliciting conceptual structures about topics and comparing and contrasting them between individuals and across groups have already been linked with email services (Shaw and Gaines, 1991). We shall increasingly come to view digital publications as part of generalized *knowledge support systems* (Shaw and Gaines, 1987) in which information technologies developed in library and information science, artificial intelligence, tutoring systems, and so on, are combined to support human knowledge processes (Gaines, 1990b).

Word processing and page makeup programs that support active knowledge structures are already in use. For example, an "electronic paper" on a knowledge-based system that was itself a working example of that system was recently published in parallel paper and electronic forms. The abstract of that paper explains the technology:

> *"This paper is written in a document production tool that appears to a user as a word processor but also acts as an expert system shell with frame and rule representations supporting deductive inference. The electronic version of the document is active, providing typographic text and page layout facilities, versioning, hypermedia sound and movies, hypertext links, and knowledge structures represented in a visual language. It*

can be read as a hypermedia document and also interrogated as a knowledge-based system for problem-solving. The paper version of the document, which you are now reading, is produced by printing the electronic version. It loses its active functionality but continues to act as a record of the knowledge in the document. The overall technology has been developed as an alternative approach to the dissemination of knowledge bases. It also provides a different interface to knowledge-based systems that emulates document interfaces with which many users are already familiar." (Gaines and Shaw, 1992)

4.16 Computer-Supported Collaborative Work

Digital publication offers opportunities to support the social processes of scholarship more overtly, in particular, collaborative research between scholars. Research on computer supported collaborative work in general has grown rapidly in recent years and there is now a rich literature of ideas and experience (Greif, 1988; CSCW, 1990; Bannon, Robinson and Schmidt, 1991; Nunamaker, Dennis, Valacich, Vogel and George, 1991). In the USA the Federal High Performance Computing Programming (Bromley, 1989) has as one component a very high speed National Research and Educational Network (NREN) designed to support scholarly collaboration (Doty, Bishop and McClure, 1990) and whose functionality is being developed through research and development programs such as the National Science Foundation's National Collaboratory initiative (Lederberg and Uncapher, 1989). In Sweden the MultiG project is a national effort concerned with multimedia support of collaborative work on multi-gigabit networks based on optical fibres (Pehrson, Gunningberg and Pink, 1992).

There are already commercial word processors designed to support writers collaborating over a network, and research studies of a variety of user interfaces and protocols for such collaborative editors (Seliger, 1985; Dalton, 1987; Delisle and Schwartz, 1987; Fish, Kraut and Leland, 1988; Leland, Fish and Kraut, 1988; Lewis and Hodges, 1988; Catlin, Bush and Yankelovich, 1989; Neuwirth, Kaufer, Chandhok and Morris, 1990; Malcolm and Gaines, 1991). Most of these systems currently operate through local area networks but they can be extended readily to wide area networks. The development of word processors that support the versioning and annotation required for collaborative work also provides an important subset of those facilities needed to support digital journals in general.

4.17 Reading Publications on Digital Displays

There is no intrinsic reason to assume that digital journals will necessarily or generally be read directly from the screen of a computer. Low cost, high quality laser printers make it feasible for those receiving digital material to print it out rapidly and economically, and produce a publication comparable in quality to a paper journal. However, material will usually be browsed on the screen before printing, and many users certainly intend to use the screen as much as possible. In particular, digital storage is much more compact than paper storage and hence it is attractive to avoid unnecessary use of printed document in both private and institutional libraries. However, the resolution of computer screens is typically about 80 pixels/inch compared with 300/inch for standard laser printers and 1000/inch for phototypesetters. There are also limitations on the speed at which pages may be displayed and scrolled compared with manually browsing a paper publication, although these are no longer significant with proper use of hardware and design of software. There are also less obvious differences such as the flicker of video output that may have a negative effect. There have been many studies of differences in reading text

from paper and from the screen, both in terms of task performance and subjective assessment, and these have been surveyed and the data collated (Mills and Weldon, 1987). The overall results indicate slight, but consistent reductions in task performance using screens, but slight and consistent preferences for screens. However, most studies have been with obsolescent technology and data processing populations so these generalizations are suspect.

In relation to digital journals, it is important that printed output as well as browsing on the screen are both properly supported, and that these are presented as equally available options to potential users. There is no intrinsic commitment to 'paperless' publication. It is important that users be given flexibility to adopt whatever work practices are most effective to them, and that they do not see the technology as circumscribing this freedom. In the longer term it will be interesting to examine whether cognitive differences emerge in scholarship using material on the screen, as have been found for the generation of material using word processors (Heim, 1987).

5 Existing Innovations in Using Networks for Scholarly Discourse

The failure of experiments with digital journals to date that was documented in the first part of this paper obscures the fact that the growth of availability and use of digital networks is already having a major affect on processes of knowledge dissemination. This section reviews the relevant services, gives examples of their use by communities of scholars, and uses existing experience to indicate both the potential for new forms of digital publication and some of the problems that have to be overcome

5.1 New Forms of Publication through Network Services

The growth of the Unix-based networks offering email through UUCP and news through USENET from 1979 onwards, and their linkage with other networks to provide the world-wide Internet of today (Quarterman, 1990) has allowed many communities to develop networked interactions at very low cost that have become a routine part of scholarly activities in particular disciplines. The ubiquity of, and ease of access to, the network means that it is very difficult to track these activities in general. There is a minimum of bureaucracy associated with establishing a news group, and virtually none associated with establishing a mail or archive server, so that activities can commence, grow and die out with very little impact on anyone other than those directly involved. This anarchy has the advantage of allowing a natural evolution of services, but it also has disadvantages in that those with relevant interests may have difficulty in determining that an activity exists, even though they may have no difficulty in joining it once they know of its existence. Hence, there are meta-services developing that support the registration and indexing of the basic services, using the same technology of archive servers to provide information about what is happening on the network.

There is a wide variety of types of service based on the network already but, in relation to digital journals, there are four major dimensions along which they may all be classified:

• *Private—Public*: whether material is available to only restricted participants or generally available. Email is essentially private and most news groups are public, but many forms of semi-restricted access along this dimension are possible. For public newsgroups an interesting subdimension is whether users are registered so that one can tell who has expressed interest, and a subdimension of such registration is whether one tell which users are active recipients. That is,

there are possibilities for instrumenting communities and measuring knowledge flows in ways which are much more difficult to achieve with paper-based media.

• *Discursive—Archival*: whether the service essentially supports discourse or the storage of materials, or both. Email and news are both generally discursive, whereas data repositories are archival. However, again many variations are possible along this dimension since some news groups are also archived, and archives with sequencing, for example through accession dates, can be used to support the type of "discourse" that is typical in journals where one paper specifically comments on a previous one. An important subdimension here is the period of time over which material remains accessible. News is generally accessible for short periods, typically two weeks, unless archived, and archives generally maintain material for some years, but there is often no guarantee of this, and rarely a definite policy for the updating of versions of the "same" material.

• *Moderated—Unmoderated*: whether material has to pass some checking process before being disseminated or is accepted without question. Email is unmoderated as are many news groups, but checking is also common to support groups that wish to communicate information satisfying certain quality controls such as relevancy. There are many significant subdimensions here concerned with how, by whom and on what basis, the checking is done.

• *Requested—Sent Automatically*: whether material is sent only on request or transmitted automatically to an individual or group of users. Email is sent automatically but news, even though it is sent automatically to sites, is accessed only at the request of users. Archives generally only respond to requests, but mailing list servers send material through email to everyone registered with them. An important subdimension here is the visibility of material not sent automatically. News has a high visibility through specialist access programs designed to support browsing but remote archives may be totally invisible to someone who does not know of their existence. Some archives provide automatic notification of updates and good browsing facilities, while others provide minimal directories to compressed files with non-mnemonic names.

In terms of these dimensions, current scientific journals are public, archival and moderated. They are both sent automatically to individual subscribers and available on request to others through library services. There are two forms of activity common on the net that have some of these characteristics. *Mailing list servers* are used to support special interest groups who wish to receive material from one another on a routine basis. One generally joins by sending a subscribe message to the list server located at some Internet address. One sends material to a related address where the server automatically resends it to all subscribers, possibly after human moderation. Some or all the material send to the server may be archived for access on request. *Archive servers* are also used to support such groups, but their most significant use is to provide a publication outlet for individuals and institutions such as research divisions in companies or academic departments. Material is put up on such a server at a specific Internet address, generally on an unmoderated basis, although it is often technical reports that have undergone some local checking or papers that are being refereed for conferences or journals The servers operate in two modes: responding to requests sent by email through email, including requests for an index of archived material; and making the whole archive available for open remote access through the Unix anonymous file transfer protocol.

The ways in which the network is already being used for processes of scholarship is very significant to the design of digital journals, both because it suggests trends in their evolution, and also because it provides case histories of problems with some existing uses that digital journals have to overcome. Much of the remainder of this section will be anecdotal because there have been very few anthropological studies as yet of scholars in their new digital habitat. Such studies should be encouraged because we are dealing with profound changes in cultures which are not in themselves very well documented or understood. The design of new technologies for scientific discourse requires information and insights that can only come from studies of what is actually happening in a very open and anarchical environment, but one that is much more easy to instrument than most social arenas.

5.2 A Mailing List Server Supporting Discourse in the Conceptual Graphs Community

The conceptual graphs list server provides an interesting example of the intensive support of a small and well-defined scientific community. Conceptual graphs are a form of knowledge representation developed by John Sowa and defined primarily in a book written by him and published in 1984 (Sowa, 1984). Conceptual graphs have been used by many different research groups in artificial intelligence for applications ranging from natural language analysis to process control. However, there is no journal dominated by the relevant research, and not even one in which this work routinely appears. There is an annual meeting of conceptual graph researchers which publishes an extensive proceedings of some forty papers, many of which are of journal quality so that it would not be unreasonable for a journal to have developed around this work.

The list server was set up some three years ago and now sends out several messages a day. These are typically technical notes several pages long, generally rather longer and with greater attention to scientific arguments than in most news groups. The server is unmoderated but otherwise might well be compared with a ‘letters’ journal publishing short research notes. Where it most markedly differs is in the rapidity of critical comment that technical proposals attract. It is rare for any communication not to receive some commentary, typically support and extension or criticism and counter examples, within a day. Sowa uses the group to critique short expository papers in which he develops the foundations of conceptual graphs, and many applied researchers use it to request solutions to specific problems. The conceptual graphs mailing list server is a very interesting model of the way in which information technology can support scholarly communication and very open collaboration at a pre-journal stage in which the objectives are primarily intellectual problem solving rather than refereed status and polished archival results.

In the past year conceptual graphs have come into prominence as providing a strong contender for an official standard for the highest conceptual level of database repository systems. This has led to interest from others groups concerned with standards and with knowledge representation in general, and the conceptual graphs list server has been cross-coupled to list servers supporting these other communities. This has led to intensive cross-disciplinary discussions of technical and social issues that are a common concern of groups which previously has little direct interaction. It is of interest that some of the policy issues involved have been the subject of journal papers (Neches, Fikes, Finin, Gruber, Patil, Senator and Swartout, 1991), and have also been targeted at the annual conference of one of the groups and supposedly ‘resolved’ by a panel discussion and open forum. It is clear from the network discussion that this resolution was not widely known or accepted, and that the issues need exposure to a much wider community over a longer period of time before any basis for consensus might emerge. The list servers, and their connection across

communities, are proving very effective in supporting a very open, and yet highly focussed discussion of major technical issues in knowledge representation and associated standards.

5.3 Problems in Formalizing Network Communication

The success of mailing list servers in providing a mechanism for intense technical discussions in a scholarly community has made it attractive for many groups to propose more formal structures for such communication. There are a number of newsletter-style publications disseminated through news groups, list servers or archives on the net. There have been many more proposals for such publications modeled on journals with material submitted from a community, and generally these have failed to develop. The newsletters that operate effectively all seem to be written by one dedicated individual or by such a person supported by a few colleagues. There seem to be insufficient incentives for a critical mass of general contributors to develop, and this seems true also of other types of archive such as those for particular forms of software or bibliographies of material in a certain discipline. These are generally started by an energetic individual with great enthusiasm who gathers a certain amount of material initially but then finds further contributions do not occur.

One cannot currently do more than speculate on the reasons for reluctance to contribute to more formal network archives. Lack of community recognition of the activity may be a major problem. Those putting up archive servers for particular institutions usually report the statistics of access with some pride, but this does not propagate back to individuals who contribute to more general archives and newsletters. “League tables” of how often a contributor’s material has been accessed that are themselves given some prominence may not appear to have a direct relevance to the operation of an archive, but they may provide both a significant reward mechanism for contributing and also feedback on what activities the community finds particularly worthwhile.

There is some evidence for this speculation in the major public services performed by some individuals through the network which have established very favorable reputations for them in the communities served. For example, this has been the case with John Norstad and Chris Johnson, the author of the public domain anti-virus programs, *Disinfectant* and *Gatekeeper*, respectively. In many newsgroups some individuals have established themselves as expert commentators and information providers, and they have a high status and reputation within large communities that have never met them individually and are generally unaware of their backgrounds except for what has been generated on the network. This is not professional recognition in the normal sense and cannot generally be ‘cashed-in’ for academic tenure and promotion or professional society fellowships, yet it seems in practice a significant reward for those involved.

5.4 A Case History of Reactions to Proposals for Formalization

The proposals in news groups for more structured use of the network, somewhere between the current informal usage and a formal refereed journal, have at times been extremely detailed. The critical reactions to such proposals are of interest because they give a direct indication of the current culture of usage of the network for scholarly communication. In December 1991 Stodolsky posted to *comp.groupware* a 4,000 word article proposing a new form of network publication, a “consensus journal,” in which:

“Readers submit reviews that evaluate articles on agreed dimensions. A statistical procedure is used to identify the most knowledgeable representative of each consensus position and these persons are invited to submit articles that justify the review judgments they have submitted. A major advantage of this approach is the ability to develop reputation without article publication.” (Stodolsky, 1990)

The proposal is of interest in its own right as a contribution to the discussion of digital journals, and would seem a wholly appropriate contribution to a news group having some 35,000 readers interested in collaborative information technology. However, most of the ensuing discussion was not so much about the detailed proposal as about the appropriateness of both the proposal and the length of the posting. The sample following is highly representative and gives an indication of very common attitudes to the use of USENET:

> *“You have suggested a very elaborate format for the generation and electronic publishing of high quality scholarly papers. Alternatively, you have provided a way of channeling the flow of discussion in an environment like usenet such that discussion is kept more focused and would tend to carry more content-per-line relative to initial articles, etc.*
>
> *But Albert (IMHO) is right, this is not the real issue in the usenet, and ultimately, I suppose, because it’s users don’t want it. Most people are not here to participate in scholarly style debate. Communication for its own sake has become the primary reason for the volume of the usenet. It is *entertaining*. If it happens also to be enlightening, so much the better, but that is no longer the great motivator.”*
>
> *“Yes but the question is, would THIS message have been accepted under David’s guidelines? If not, why should anyone else have to live under rules David himself finds inconvenient. Or if the guidelines have room for such "informal" chatter, then I submit that this is the kind of traffic which will in fact predominate, just as it does now.*
>
> *Offered an essentially free-form environment, Usenetters have created what you see today. Attempting to legislate something else instead is precisely the kind of bureaucratic oppression Usenet was built to get around.*
>
> *There are rigorous academic environments where something fancier may be needed. You are welcome to make available tools of your own design for this purpose, and piggyback them on Usenet if you can convince people to carry the groups thus formed. It ain’t gonna change what we already have.”*
>
> *“True, the fitness of the IDEAS IN David Stodolsky’s proposal cannot legitimately be commented upon unless the contents are understood. But the ‘main point’ (IMHO) might be that a 500-line-plus article, carefully worded and thus requiring careful reading, is itself NOT APPROPRIATE FOR USENET!*
>
> *At least, I myself rarely read postings more than 40 or 50 lines. I think that Usenet is mainly for quick reactions to short, to-the-point, often-provocative postings. If others feel as I do, then your carefully-worded proposal is a self-defeating (rather than self-fulfilling) comment.”*

"Yes, I tend to find that even in the news group I'm most interested in (comp.sys,acorn), I much prefer short postings - I have a habit of skipping longer ones, or saving them for 'later'.

However, I feel there is still a need for handling longer articles. If usenet is not appropriate then what/where? Can't honestly see a better alternative at the moment. Do we need to wait for better hardware/software to make it easier to read long postings comfortably - cetainly to progress we MUST (?) move to the situation where we can routinely incorporate graphics and more sophisticated methods of presentation. BUT how can this be achieved? Cost and standardisation are some issues - everybody buying brand XYZ - is not the answer. Other main problem is the transition phase, don't want to lock people out that only have text or limited groaphics capabilities, yet we don't want to be locked into today's technology!"

This type of discussion is common across all news groups where there are generally negative reactions to the external imposition of disciplines on communications and to very long postings. The news groups are seen as having a very different function from journals, and to be playing an important role of their own in providing a forum for public discussion, often of complex technical issues but generally in the form of short commentaries rather than long papers. In the more academic groups citations are often given to published works for those wishing to follow up on background material. One very important function of these groups is to draw attention to relevant papers and to new work that may have been missed for some reason, such as its publication in a journal that is not mainstream for the field. This is a particularly significant function in relation to the dissemination of knowledge that is available, in a journal or on an archive server, where awareness of existence is as important as capability to access.

5.5 Problems of Social Norms in Digital Communications

The example of interchanges in *comp.groupware* shows the informality of network communications which are often composed directly at the terminal and not edited for content or presentation. Because of this spontaneity of response, they may also become very personal in ways that might be acceptable in individual interaction but can appear unreasonable when broadcast across the network. One of the documented problems of unmoderated news groups is the reduced social constraints on expression that have been noted to be a problem in email (Brotz, 1983). Even in highly technical newsgroups it is common for someone to originate a "flame" in which complaints are made about someone else's communication in very blunt and unpleasantly personal terms. These would be usually considered inappropriate to scholarly communication, although they are by no means universally absent from it—authors of reviews, peer commentaries and books in particular sometimes take the opportunity of reduced or absent editorial control to make highly personal attacks.

One problem with flames is that they often generate a massive flow of communication and metacommunication, in defense of the person attacked or about the propriety of making such comments on the network, and this can in itself be highly disruptive of the normal activities of the news group. Moderation can overcome this in principle but flames are often directed at moderators by those whose material is excluded or censored in some way, and the moderator may then feel ethically bound to convey the criticism. It is also appropriate to consider that unmoderated communication may serve a purpose and that, while a digital journal might exclude

it from its mainstream publication, it might well support it through a general commentary facility.

5.6 Problems of Volatility in Digital Archives

Some interesting problems arise when archive servers are used to make papers available under the control of individuals or groups. While writing this paper I had occasion to check an archive server operated personally by a colleague in another country who makes all his papers available in this way. I noted that two papers I had previously copied from the archive were no longer present. They had both been submitted to conferences, rejected, and removed. However, those of us who had taken copies might quite reasonably have cited them as working papers available on the archive, and anybody attempting to follow up these citations would no longer have been able to obtain them. A variant problem was that another paper had greatly increased in length without change in title or version number. It had been rejected for a conference and been greatly extended and submitted to a journal. Again someone following up a citation to the earlier paper could have been very much confused by retrieving the changed work. These are problems which one tolerates for the value of having access to significant research in its early stages, but they are ones that could cause major problems for scholarly practice if they become common on a large scale.

In parallel with this I had some interesting experiences with an archive server operated by an industry research group in another country. In this case I did not have an archive address but was able to obtain it from *Archie*, an archive server that maintains an index of what material is available on other archive servers. The system I finally accessed services requests by email and when I communicated with it in January 1992 the index showed that two reports that had been available previously had been withdrawn so that they might be edited to be consistent with software that was “to be issued in September 1991” (sic). Furthermore, as a result of my access, in February 1992 I received postal mail from the research group informing me that I would be sent the reports in September 1991. As already noted, maintaining archives is a tedious task that may easily be neglected once the initial enthusiasm has waned. For this type of situation, since the archive server has access to the current date, a better software system could at least prevent the issue of inappropriate messages. However, the root problem is one that we have to accept as part of the price of more rapid but less formal publication systems. Technically, some of the problems noted could be overcome by having more formal archives at which one could register an authenticated version of a paper. However, it is probable in both the cases mentioned that the authors or institutions would not wish to give up their control over the informally ‘published’ material, for example, to leave a rejected or outdated version of material permanently available.

As a user of both these archives I would also prefer to put up with problems caused by the volatility of the material rather than lose access to it. In both cases I retrieved a number of valuable documents that I was able to print out with full typography and embedded figures at a quality that was as good as any other source and much better than a photocopy. I was accessing material at least a year before its availability through paper publication and was not involving the authors in any additional effort or costs. They were in a position to monitor my access to determine the level of interest in their material and obtain some indication of who might be using it. As a mechanism for knowledge dissemination networked access to digital archives is very effective indeed. The volatility is acceptable in many circumstances, and could be overcome by use of authentication and registration as already discussed. Indexing, awareness of indexes, and

knowledge of how to use them, are important and sources of problems, but they parallel similar functions in conventional libraries and can be supported through services and training as they are in libraries.

5.7 Concerns about the Growing Use of Digital Communication

Serious concerns have been expressed about the growing use of digital networks for discourse between scholars. A survey undertaken by Meadows in Britain showed many correspondents to be concerned about the development of electronic cliques of senior scholars that were invisible to those not invited to join (Anderson, 1991). Concern was also expressed about the circulation of data and theories that has not been adequately assessed and yet were being used by other scholars as if they were published results. Doubt was expressed about the capability of electronic publishing to take over the functions of a conventional journal unless a standard refereeing process was used. It was also noted that the availability of a well-defined base of publications acts to create coherence in a scholarly community, and that the publication of an item of scholarly knowledge freezes it into a form subject to criticism which continuous updating would undermine.

Similar doubts have been expressed by Kassirer in the *New England Journal of Medicine* about electronic medical journals such as AAAS *Online Journal of Clinical Trials*. He raises a large number of issues relevant to the effectiveness of digital journals that serve as an excellent check list for any designer of the technology. One of his most telling statements is:

> *"We prize our editorial peer-review system, the current structure of our scientific papers, and the regular publication of journals because they increase the chance that data are valid, deter the inappropriate interpretation of data, and freeze both data and interpretations in time. Without editorial peer review, how can we be confident of the validity of available data? If information floats and merges, will healthy scientific criticism be lost? And, finally, who owns the data? If the outcome of research consists only of data shared on a computer network, how will investigators receive credit for their work? Such attribution is essential in making decisions not only about promotion and tenure but also about grant support for worthy work. When medical investigators share their ideas informally on a computer network, as computer scientists have done for many years, they jeopardize attribution and ownership. In addition, because such shared information, like hearsay, suffers from distortion as it changes hands, the quality of the information deteriorates." (Kassirer, 1992)*

These are very clearly stated and relevant comments that have to be addressed by any digital journal design.

In more informal terms, there has been extensive discussion of the need for "virtual journals" in the the USENET group *sci.virtual-worlds*, and one of the participants made a very clear statement of the issues relating to the status of digital publications that Kassirer raises:

> *"There is one MAJOR problem: Recognition of effort during the transition. When I go to the effort to put together a publishable paper, I want the recognition for it in my university's tenure-and-promotion scheme. My dean of engineering (may he live forever) probably will not be too impressed with a paper in some new form of publication that he has never heard of, much less that he doesn't know how to read. (This is not intended to*

be derogatory to my dean, my provost, your boss, or any other similar person. It is simply a reflection of the fact that the powers-that-be are not familiar enough with the technology to accurately assess its power and appreciate its advantages.)

Until the electronic journal is recognized as a respectable avenue for publication of real research, no one will submit the kind of research and write-ups that will earn that respect. We are caught in a vicious cycle:

No respect -> No good papers submitted -> No respect -> No good papers....

One possible answer is to take it one step at a time: Start out with an electronic journal which has some major players backing it. We need to get somebody like ACM to sponsor the first one, or failing that, an editorial board with some real heavy-weights. Since this is a VR list, some of the obvious candidates would be those who have already published heavily in the field, or who have the credentials in related fields and the interest in this one: Fred Brooks comes to mind, along with similar people such as Tom Furness. With the right support and backing (prestige is really what we need in the beginning more than funding, etc.) this could have a chance to break the ice.

If I could go to my department head and tell him that my paper was accepted for publication in a new journal upon the personal recommendation of Fred Brooks (wouldn't THAT be nice!!), that might carry enough weight to do the job. That doesn't necessarily mean that the big names would have to do all the work, but they would have to be committed enough to the endeavor to do the equivalent of 'Editorial Board' work, and be willing to take a few hard knocks if the quality dropped (which should guarantee that it wouldn't).

In the beginning, it would have to be refereed at least (read 'more') stringently than the current paper journals, in order to build respectability. Perhaps (again, in the beginning) it would be permitted for papers from the electronic journal to be republished later, in paper, so the author could get credit with his/her work with deans, etc. (This may run into opposition from the conventional paper journals, since they may rightly perceive such a movement as the beginning of the end of their monopoly.)

The key to developing this kind of new publication medium is to build a level of respectability which would allow me and other academics like me to take advantage of this."

Thus, not only are the issues perceived but they are also seen as obstacles that can be overcome through social means using existing resources.

5.8 Recent Developments in Digital Publications

The quotation from Brent in Section 4.8 is taken from the third issue of a digital journal that commenced publication in 1991 through a mailing list server at the University of Albany. *EJournal* satisfies the criteria suggested in the quotation at the end of the preceeding section by having an editorial board of senior researchers with a strong track record of research publications and journal editing and well-defined refereeing policies. Its statement of policy is:

"_EJournal_ is an all-electronic, Bitnet/Internet distributed, peer-reviewed, academic periodical. We are particularly interested in theory and practice surrounding the creation, transmission,

storage, interpretation, alteration and replication of electronic text. We are also interested in the broader social, psychological, literary, economic and pedagogical implications of computer-mediated networks.

> *The journal's essays will be available free to Bitnet/Internet addresses. Recipients may make paper copies; _EJournal_ will provide authenticated paper copy from our read-only archive for use by academic deans or others. Individual essays, reviews, stories--texts --sent to us will be disseminated to subscribers as soon as they have been through the editorial process, which will also be "paperless." We expect to offer access through libraries to our electronic Contents, Abstracts, and Keywords, and to be indexed and abstracted in appropriate places."*

Its copyright notice is particularly interesting:

> *"This electronic publication and its contents are (c) copyright 1991 by _EJournal_. Permission is hereby granted to give away the journal and its contents, but no one may "own" it. Any and all financial interest is hereby assigned to the acknowledged authors of individual texts. This notification must accompany all distribution of _EJournal_."*

EJournal has an electronic mailing list of some 1,000 subscribers, and another journal based on the same system at the same site but specializing in post-modern issues largely in the arts has a list of some 2,000 subscribers. These are high subscription rates for journals that have been operating only for a short period.

The EJournal approach to publication and copyright is very different from that of another recent development, AMIX, which is an online service that allows anyone to upload documents for sale or purchase documents that have been published. Anyone capable of selling documents may also be contacted for consulting services, with negotiations taking place and payments made entirely through AMIX. There is a variety of markets in which documents may be bought and sold. AMIX is being run as a commercial operation through a subsidiary of Autodesk which takes a proportion of the payments made for AMIX services in return for providing the computational and commercial infrastructure. It provides a medium for digital publication and it would be possible to run a journal on a commercial basis through this infrastructure.

The two extremes of the free, refereed EJournal aiming at academia, and the commercial, market-driven AMIX aiming at industry, coupled with the examples cited earlier of the AAAS journal, *The Online Journal of Current Clinical Trials*, Claerbout's parallel publication of book and CD-ROM, the conceptual graphs list server, and many other diverse activities, all show that digital publication is an active area of experiment with many possibilities. These are all phenomena of the 1990s and it will be several years before it is possible to judge them in critical terms—notably in terms of survival, growth, effectiveness, respect and impact. However, they already provide a diversity of role models for further experiments.

6 An Agenda for Digital Journals

This section stands alone as a statement of objectives for the development of digital journals, based on the discussion of the preceding sections, but summarizing it in terms of action points rather than derivations.

6.1 Statement of Aims and Objectives

A1. To enhance scholarship by systematically improving the creation, dissemination and utilization of knowledge.

The difficult concept here is "systematic improvement" which may seem impossible given the anarchy of knowledge creation processes, and undesirable given that much of that anarchy may be essential to innovation. However, it is the very fact that we recognize these issues that makes it reasonable to attempt to support knowledge processes systematically, including the freedom to innovate without constraints. We can use the capabilities of information technology to allow the media supporting scholarship to evolve from being passive repositories to becoming the life worlds of our nomadic knowledge products, and this means supporting the continuing evolution, and co-evolution with ourselves, of the knowledge worlds that we create.

A2. To improve the productivity of individuals and groups generating and using knowledge.

We should not assume that improvements in the processes of recording and disseminating knowledge will automatically result in improvements in scholarship. It is important to establish the requirement for advances in the technology in terms of support for the knowledge processes of individual scholars and communities of scholarship. In particular, these processes should be studied in depth, and the impact of new technology on them should be monitored with a view to continuous improvement and enhancement of services.

A3. To reduce the adverse impact of the growth of knowledge by improving access to knowledge sources.

The major problem currently is not the lack of opportunities to publish but the information overload created by the growth of knowledge and the freedom to publish. It is the capabilities to organize, index, search and use knowledge in digital form that are most important in the design of digital journals. For example, it would be absurd to design digital publications in which the user did not have complete freedom to analyze and restructure the published material. Facsimile-style replication of paper publications might speed access to the material but it would be a short-term expedient that does nothing to alleviate the fundamental problem of information overload.

A4. To increase the speed of knowledge dissemination.

The processing lag between a publication being available in draft form and being published in journal form is several years in many disciplines, particularly for the most highly regarded journals. Scholars attempt to overcome this by informal dissemination of manuscripts and this process should be supported through digital archives.

A5. To increase awareness of relevant material.

All forms of dissemination, and particularly the informal availability of material prior to formal publication, are ineffective unless potential users are aware of the existence and availability of the material. The indexing of archives should be supported through the evolution of increasingly effective information retrieval systems taking advantage of access to the full text of the material.

A6. To maintain openness of access to material.

Concern has been expressed about the development of electronic cliques in which senior scholars maintain a discourse which is not open to others, either because they are unaware of it or because

they are not allowed to join. There will always presumably be discussions which are appropriately private or confidential for good reasons. However, it is important to support and encourage the open access to knowledge which is already part of the culture of scholarship and of our society's attitude to knowledge. Digital communications have the advantage of supporting the critical discussions commonly associated with workshops without the physical restrictions on attendance that limit participation in conventional meetings. Experience with current mailing list servers shows how much can be learnt from access to the ongoing discussions of major scholars about critical issues, and it is important that this type of peer commentary be supported and encouraged in the operation of digital journals.

A7. To improve access to existing knowledge.

A critical mass of material available in digital form will be most rapidly achieved through incorporation of conventionally published material as soon as possible. In particular, pre-circulation and parallel publication of material to be published in paper journals should be encouraged by establishing public archives in which scholars may deposit their publications, including already published material.

A8. To increase cross-disciplinary access to knowledge.

Much of the compartmentalization of scholarship can be attributed to the need to manage the growth of knowledge and avoid information overload. New developments should not take current disciplinary boundaries for granted, but should question their roles and placement. In particular, there will be scope once large amounts of material are available digitally for new approaches to the structuring of knowledge, and these should not be impeded.

A9. To support the development of overt conceptual structures for knowledge.

Past media have encouraged the linear presentation of material and the implicit embedding of argument structures within the text. Digital media give scope for major innovations in presentation, such as labelled hypertextual links, and the shadowing of informal arguments by formal and operational ones. It is probable that few of the experiments in hypertext and computer-based knowledge representation to date will prove to be effective in the long term, and that the major innovations are yet to come involving approaches that not yet been developed. Innovation and flexibility in the use of digital media should be encouraged, and explicitly supported, not impeded.

A10. To use modern information technology to support the achievement of these objectives.

This is perhaps a rather obvious presupposition to the preceding objectives, but it is proper to place it late in the sequence as the servant to the other objectives, not their master. We will always have much activity that is 'technology-driven' where those who see the elegance and potential of new technologies attempt to deliver it in useful form. However, it is important to balance the technology thrust with thoughtful planning that is 'market-led' where those who see the essential needs of scholarship attempt to mobilize appropriate technology.

A11. To support existing innovations in scholarly communication.

Ease of access to existing network services had allowed many scholarly communities to develop new modes of operation on an informal basis. This makes available valuable empirical data that is important to the development of effective digital journals. In particular a perspective that sees

the digital journal as a formalization of this network discourse, in the same way that paper journals formalized postal and verbal discourse, is a useful counterbalance to one that sees digital journals as the emulation of existing paper journals.

A12. To minimize the disruptive aspects of the introduction of new technology.

Much of the 'human factors' of the introduction of digital journals is subsumed in this statement. Designs that are based on existing usage of word processors on personal computers and existing access to networks are far more likely to succeed than those that require major changes in existing technologies or work practice.

A13. To maintain flexibility allowing enhancements in technology and changing requirements to be incorporated with the minimum of disruption.

Digital systems can be designed for change with automatic conversion between standards and continuous enhancement of capabilities without disruption of services. Often, however, they are not and flexibility should be stated as an explicit requirement.

A14. To encourage cooperation between those operating existing media and services and those developing new approaches.

Digital journals not only have close relationships to existing journals, their editors, refereeing systems, authors, readers and publishers, but also to a wide range of associated activities such as abstracting, indexing and information retrieval services, and to many aspects of library services and librarianship. It would be of greatest service to scholarship to mobilize the knowledge and resources available through these existing institutions to develop digital journals as expeditiously and effectively as possible. It is possible that some of those potentially involved may see the new medium as a threat and not be prepared to cooperate, but it would be counterproductive to presume lack of cooperation in advance. There is most to be gained in the short term by sharing information, resources and opportunities and attempting to enhance scholarship through open collaboration on a professional basis.

A15. To prevent any abuses of monopolistic control or copyright legislation that restrict developments in scholarly discourse.

The community of scholarship at large should be made aware of the new possibilities for the dissemination of knowledge and mobilized to protect them from abusive practice if necessary. Ultimately, scholars are the major producers and consumers of knowledge and in a position to regulate the market place to the best advantage of the community at large.

6.2 Dimensions of Digital Publication

There are many different ways of attempting to realize the above objectives using information technology. One way of coming to terms with alternative approaches, their relative merits and roles in scholarship, is to classify them along major dimensions of variation. Four of these have already been discussed in Section 5.1:

D1. Private—Public: whether material is available to only restricted participants or generally available.

D2. Discursive—Archival: whether the service essentially supports discourse or the storage of materials, or both.

D3. Moderated—Unmoderated: whether material has to pass some checking process before being disseminated or is accepted without question.

D4. Requested—Sent Automatically: whether material is sent only on request or transmitted automatically to an individual or group of users.

There are a number of additional dimensions relevant to the operation of digital journals:

D5. Standard Format—Nonstandard Format: whether the material is available in one of a number of well-defined and interchangeable formats. As already discussed, it is unreasonable to require a single standard format, but it is possible to operate with a range of formats that can be converted into a locally usable form.

D6. Accessible Content—Inaccessible Content: whether the full content of the material is available for search, indexing or reuse, or whether only an image description is available.

D7. Multi-Media—Textual: whether the documentation representation supports fully typographic text, diagrams, pictures, sounds, video, and so on, or only text.

D8. Transient—Permanent: whether the material is available only for a period or indefinitely. Permanent availability is important to the 'freezing' of a particular statement of knowledge to make it a well-defined subject of critical commentary.

D9. Mutable—Immutable: whether the material is fixed so that the apparent publication is not subject to change, or whether it can be edited without this being apparent except through content. This is again important to scholars having access to, and citing, precisely the same statements. Immutability does not imply that later versions of a document cannot be issued, only that these are treated as later descendants, not replacements.

D10. Authenticated—Unauthenticated: whether the material can be checked to be an authentic, unedited copy of the version issued. This ensures that immutability can be propagated to replicas of documents.

D11. Registered—Unregistered: whether the material has been registered with some independent authority to establish date of publication. To be useful this registration has to be of an authenticatable and hence immutable document.

D12. Indexed—Not Indexed: whether the material is indexed in well-known, publicly accessible archives. This is important in supporting awareness of relevant material. The utility of the index may vary dependent on the amount of contextual and classificatory material entered, and on the quality of associated services, but some degree of minimal indexing is essential to the functions of a journal.

D13. Annotatable—Unitary: whether the material is structured in a way supporting precisely defined internal citation. Page and line numbers have never been very satisfactory but, at least, they were functional for paper documents. We need to establish new conventions for digital publications that can be supported by software. Uniform standards cannot be assumed, and the conventions will need to be fairly pragmatic for each form in which publications might be issued. This is very important to support peer commentary, conceptual and argument form annotation, and hypertext linkages.

D14. Attributed—Anonymous: whether the author of material is identified. This is primarily relevant to refereeing where referees are usually anonymous. There are a few publications that do not supply the identities of authors to referees, but this is not usually effective since they are often obvious and also the publication should be refereed from the perspective of a reader who will have the full information and can interpret statements in terms of knowledge of the author or authors. There can be circumstances in which pseudonyms are appropriate also, for example, because an author does not wish knowledge of his or her identity to be used in interpreting a statement, or wishes to establish a separate identity. However, there have already been serious social problems generated by the establishment of false identities on CompuServe (Gelder, 1991), and it seems likely that present conventions of attribution and accountability on Internet should be continued.

In terms of these dimensions and the discussion in this paper, publications in an effective digital journal should be public, archival, moderated, available to be sent automatically, in a standard format, with content fully accessible, supporting multi-media, permanent, immutable, authenticated, registered, indexed, annotatable, and attributed. As already noted, none of the current applications of Internet satisfy these requirements, and no past or currently established digital journals satisfy them either. However, none of them are technically difficult to satisfy.

6.3 Action Plans

The action points in this section are intended as examples to focus attention on a number of simple initiatives open to the scholarly community that could develop digital publication services by mobilizing existing resources at a minimum cost and with a high chance of success.

P1. A Digital Journal Consortium

Establish a digital journal consortium as a loose confederation of interested parties communicating through Internet with a view to sharing ideas and technology, encouraging necessary developments, and supporting experiments with digital journals.

This is a typical Internet activity and has the usual advantages of mobilizing a diverse community without requiring a central power base, and of completely transcending national and disciplinary boundaries. This consortium can have fruitful collaboration with those concerned with document standards, multi-media communication, groupware and so on. It will serve to facilitate change and the evolution of *de facto* standards through experience. The actual operation of particular journals will be through editors, anonymous reviewers, commentators, and so on, as it is now, and it is important that these disciplinary communities have access to reliable, fully functional technology which they can use without responsibility for development and maintenance. It is probable that the first communities attracted to digital publication will be those with specific requirements better served by digital, rather than paper, publication, and it is important that such requirements are identified and addressed. It is also probable that a printed form of most digital journals will be required for the foreseeable future, and that support of printed output of a parallel paper publication should be treated as a major initial requirement.

P2. Publication Archives

Establish public publication archives in which scholars can place digital versions of works published in peer-reviewed paper journals.

These archives should support authentication so that the copies obtained from them can be distributed with the means to check that they have not been edited. Generic archives for all disciplines and publications are suggested so that they can be put in place rapidly without dependence on action in particular disciplines. Restriction to published works is suggested so that the archives do not become overloaded with material, or have to establish their own moderation procedures. This suggestion relies on scholars having the right to make whatever use of their material they see fit, and is an extension of the existing practice of distributing reprints. It seems unlikely that even publishers who do not explicitly return this right will object to the archives. Most scholars would assume that they have this right anyway and would not continue to publish in journals whose publishers attempted to claim otherwise. Note that the works would not be placed in the public domain and that the scholars would retain their copyright. The use of material without due acknowledgement is very easy to detect in digital publishing and offenses ranging from professional discourtesy to outright plagiarism would be much more visible than they are now.

P3. Public Indexes to Personal Archives

Establish public index archives of personal or institutional archives of material that has not been peer reviewed.

This is intended to support the current practice of placing submitted manuscripts, draft documents, reports and data in public archives under local control, but to make the contents more accessible to the community. It balances the restriction of the publication archives above to papers published in peer-reviewed journals by supporting any amount of publication based on the resources of the author or his or her institution. Permanence, immutability and authentication should be encouraged in these archives, but these can also be achieved by setting up sub-archives that transfer selected material from the personal archives, probably operated by some disciplinary sub-community.

These three action plans are complementary and together, including also current network activities, they would give an adequate basis for a major acceleration of the progress towards digital publication. They form a basis for the development of extended services based on content searches, hypertext linkage of peer commentary, knowledge structures for disciplines, and so on. However, before these new developments occur we need a critical mass of material available on the net and a significant community of scholars drawing upon it on a routine basis. That has to be our first practical objective, and is that addressed most directly by these three action plans.

7 Conclusions

This paper has presented a fundamental examination of the prerequisites for the introduction of digital journals, at one level in terms of the role of journals in the social and economic processes of human knowledge production, and at another in terms of the state of the art in the relevant technologies. Models of the processes underlying the growth of knowledge in the literature on the philosophy, history and psychology of science have been used to analyze the structure and role of the social infrastructure of journals, including the editorial and refereeing systems and the role of commercial publishers and libraries. The motivation for digital journals and past experience have been surveyed, together with the learning curves and current costs and performances of the enabling hardware, software, communications and interface, technologies.

Examples of the current impact of computer and communications technology on scholarly discourse have been given to enable probable changes to be predicted in the structure of journals when they are transferred to digital form. Finally, the social and technological analyses have been used to outline some architectures for a first generation of digital journals emulating the current medium, and for the evolution of later generations diverging in characteristics to take advantage of the new medium.

If there is one lesson to be learnt from the successful growth of UUCP and USENET facilities it is that a grass-roots, incremental development of services that are of widespread importance to the scholarly community can be a very effective of harnessing information technology. These services were based on software developed by a few skilled programmers, placed in the public domain, and 'marketed' as a communication mechanism enabling mutual support of unix systems. The integration of these facilities with those of other, more formally developed, networks to form the current Internet shows that the use of informally developed and supported services is not an ideological issue but rather a pragmatic approach to the evolution of services in response to user needs. The many successful commercial network services operating in parallel with Internet and serving strongly overlapping communities also demonstrates that the provision of very low cost services, apparently free to many users, does not prevent the development of commercial services having some attractive added value.

This is an excellent model for the development of digital publications: that a few technological groups be encouraged to develop suitable software to be placed in the public domain as the basis for a new generation of knowledge sharing services supported the scholarly community; that the use of these services be 'marketed' through their utility in improving scholarship, and, in particular, in improving the existing uses of the net by scholarly communities; and that commercial services freely compete with Internet services by providing added value such as improved search procedures or better knowledge processing tools. There is scope for commercial ventures in the development of extensions to existing word processors that interface well to networked digital publication procedures. There is scope for publishers to use their existing networks of contacts with the scholarly community, and their expertise in the management of the publication process, to develop new products. In particular, the existing base of published knowledge is invaluable and its reissue in a variety of digital forms will be significant to scholarship and a source of revenue.

How the publication of knowledge will become restructured in the long term is extremely difficult to forecast, but the time horizons involved are long in terms of most organizations' planning cycles and give enough time for major changes to be assimilated without disruption. However, this will only be the case if the inevitability of change is recognized now and the planning process is taken seriously. Publishers who do not respond to the possibilities of the new medium will begin to see the flow of material through their publications decline as other, more attractive outlets become available. Scholars who do not begin to take advantage of the new medium will start to lose contact with the intellectual discourse in their discipline. Institutions that do not support, and encourage the use of, the new services will begin to see their status diminish as other institutions use the new capabilities to systematically enhance their scholarly productivity.

These are perhaps foolish remarks because, as this paper has documented, they have been made with as much fervor and conviction many times before over the past fifty years—and nothing has

happened. It is proper that someone be foolish enough to cry wolf every decade so that at least we reexamine the issues periodically. The problems of the growth of knowledge will be with us forever, and will only worsen with time if we do not address them. I have argued that the technology is now available to make major advances in our capability to deal with them. It needs to be harnessed to the needs of scholarship based on a responsiveness to those needs and changes in them as the new medium of digital publication comes into widespread use.

Acknowledgements

Financial assistance for this work has been made available by the Natural Sciences and Engineering Research Council of Canada.

References

Abelson, P. (1990). Mechanisms for evaluating scientific information and the role of peer review. **Journal American Society Information Science 41**(3) 216-222.

Adobe Systems Incorporated (1986). **PostScript Language Reference Manual**. Reading, Massachusetts, Addison-Wesley.

Alexander, G. and Walter, M. (1990). A fresh look at SGML: the conventional wisdom changes. **Seybold Report on Publishing Systems 20**(7) 3-16.

Ambron, S. and Hooper, K., Ed. (1988). **Interactive Multimedia**. Redmond, Washington, Microsoft Press.

American Council of Learned Societies (1979). **Scholarly Communication: The Report of the National Enquiry**. Baltimore, Maryland, John Hopkins University Press.

Anderson, A. (1991). Networks for thinking in cliques? **Science 253**(5019) 506.

Attali, J. (1990). **Lignes d'Horizon**. Paris, Fayard.

Ayres, R.U. (1968). Envelope curve forecasting. **Technological Forecasting for Industry and Government: Methods and Applications**. pp.77-94. Englewood-Cliffs, New Jersey, Prentice-Hall.

Bailey, C.W. (1991). Electronic (online) publishing in action: the *Public-Access Computer Systems Review* and other electronic serials. **ONLINE 15**(1) 28-25.

Balzer, W., Moulines, c.-U. and Sneed, J.D. (1986). The structure of empirical science: local and global. Marcus, R.B., Dorn, G.J.W. and Weingartner, P., Ed. **Logic, Methodology and Philosphy of Science VII**. pp.291-306. Amsterdam, North-Holland.

Bannon, L., Robinson, M. and Schmidt, K., Ed. (1991). **ECSCW'91: Proceedings of the Second european Conference on Computer-Supported Collaborative Work**. Dordrecht, Holland, Kluwer.

Barrett, E., Ed. (1988). **Text, Context and Hypertext: Writing with and for the Computer**. Cambridge, Massachusetts, MIT Press.

Barrett, E., Ed. (1989). **The Society of Text: Hypertext, Hypermedia and the Social Construction of Information**. Cambridge, Massachusetts, MIT Press.

Baumol, W.J. and Marcus, M. (1973). **Economics of Academic Libraries**. Washington, D.C., American Council on Education.

Bayer, A.E., Smart, J.C. and McLaughlin, G.W. (1990). Mapping intellectual structure of a scientific subfield through author cocitations. **Journal American Society Information Science 41**(6) 444-452.

Bernal, J.D. (1939). **The Social Function of Science**. London, Routledge.

Bhagwati, J.N. and Partington, M., Ed. (1976). **Taxing the Brain drain I: a Proposal**. Amsterdam, North-Holland.

Blair, G.S. and Davies, M. (1991). Incorporating multimedia in distributed object-oriented systems: the importance of flexible management. **Proceedings International Workshop on Object Orientation in Operating Systems**. pp.36-41. Los Alamitos, California,

Borenstein, N. and Freed, N. (1993). MIME (Multipurpose Internet Mail Extensions) Part One: Mechanisms for specifying and describing the format of internet message bodies. Internet. RFC 1521.

Braam, R.R., Moed, H.F. and Raan, A.F.J. van (1991a). Mapping of science by combined co-citation and word analysis. I. Structural aspects. **Journal American Society Information Science 42**(4) 233-251.

Braam, R.R., Moed, H.F. and Raan, A.F.J. van (1991b). Mapping of science by combined co-citation and word analysis. II. Dynamical aspects. **Journal American Society Information Science 42**(4) 252-266.

Brailsford, D.F. and Beach, R.J. (1989). *Electronic Publishing*—a journal and its production. **Computer Journal 32**(6) 482-493.

Brent, D (1991). Oral knowledge, typographic knowledge, electronic knowledge: speculations on the history of ownership. **EJournal 1**(3)

Bromley, D.A. (1989). The Federal High Performance Computing Program. Executive Office of the President, Office of Science and Technology Policy.

Brotz, D.K. (1983). Messages system mores; etiquette in Laurel. **ACM Transactions Office Information Systems 1**(2) 179-192.

Bryan, M. (1988). **SGML: An Athor's Guide to the Standard Generalized Markup Language**. Wokingham, UK, Addison-Wesley.

Bush, V. (1945). As we may think. **Atlantic Monthly 176** 101-108.

Callon, M., Law, J. and Rip, A., Ed. (1986). **Mapping the Dynamics of Science and Technology**. Basingstoke, UK, MacMillan.

Carrigan, D.P. (1991). Publish or perish: the troubled state of scholarly publication. **Scholarly Publishing 22**(3) 131-142.

Catlin, T., Bush, P. and Yankelovich, N. (1989). InterNote: Extending a hypermedia framework to support annotative collaboration. **Hypertext'89 Proceedings**. pp.365-378. New York, Association for Computing Machinery.

Claerbout, J.F. (1991). **Earth Soundings Analysis: Processing versus Inversion**. Cambridge, Massachusetts, Blackwell.

Claerbout, J.F. and Dellinger, J.A. (1991). **Stanford Exploration Project: SEP-CD-1: (1) Earth Soundings Analysis—Processing versus Inversion; (2) Anisotropic Seismic Wave Propagation; (3) Progress Reports SEP-72 and SEP-70**. Palo Alto, California, Stanford University Press.

Cole, J.R. and Cole, S. (1973). **Social Stratification in Science**. Chicago, University of Chicago Press.

Coleman, J.S. (1990). **Foundations of Social Theory**. Cambridge, Massachusetts, Harvard University Press.

Coward, H. and Standera, O. (1985). Refereeing and editorial problems in electronic journal publication. **Computer Compacts 3**(2) 48-51.

Crane, D. (1972). **Invisible Colleges: Diffusion of Knowledge in Scientific Communities**. Chicago, University of Chicago Press.

Crawford, S. and Stucki, L. (1990). Peer review and the changing research record. **Journal American Society Information Science 41**(3) 223-228.

CSCW (1990). **Proceedings of theThird Conference on Computer-Supported Cooperative Work**. New York, Association for Computing Machinery.

Daft, R.L. (1985). Why I recommended that your manuscript be rejected and what you can do about it. Cummings, L.L. and Frost, P.J., Ed. **Publishing in the Organizational Sciences**. pp.193-209. Homewood, Illinois, Irwin.

Dalton, R. (1987). Group-writing tools: Four that connect. **Information Week** 62-65.

De Mey, M. (1982). **The Cognitive Paradigm**. Dordrecht, Holland, Reidel.

Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K. and Harshman, R. (1990). Indexing by latent semantic analysis. **Journal American Society Information Science 41**(6) 391-407.

Delany, P. and Landow, G.P., Ed. (1990). **Hypermedia and Literary Studies**. Cambridge, Massachusetts, MIT Press.

Delisle, N.M. and Schwartz, M.D. (1987). Contexts—A partitioning concept for hypertext. **ACM Transactions on Office Information Systems 5**(2) 168-186.

Dixon, B. (1988). Science and the information society. **Scholarly Publishing 20**(1) 3-12.

Donovan, A., Laudan, L. and Laudan, R., Ed. (1988). **Scrutinizing Science: Empirical Studies of Scientific Change**. Dordrecht, Holland, Kluwer.

Doty, P., Bishop, A.P. and McClure, C.R. (1990). The National Research and Educational Network (NREN): An empirical study of social and behavioural issues. **ASIS'90: Proceedings 53rd Annual Meeting American Society Information Science**. pp.284-299. Medford, New Jersey, Learned Information.

Dujin, J.J. Van (1983). **The Long Wave in Economic Life**. London, George Allen & Unwin.

Earnshaw, R.A., Ed. (1987). **Workstations and Publication Systems**. New York, Springer.

Feyerabend, P., Ed. (1975). **Against Method**. London, NLB.

Fish, R.S., Kraut, R.E. and Leland, M.D.P. (1988). Quilt: A collaborative tool for cooperative writing. **Proceedings of the Conference on Office Information Systems**. pp.30-37. New York, Association for Computing Machinery.

Freedman, D.T. (1987). The false start of the electrical journal. **ASIS'87: Proceedings 50th Annual Meeting American Society Information Science**. pp.79-82. Medford, New Jersey, Learned Information.

Gaines, B.R. (1988). Positive feedback processes underlying the formation of expertise. **IEEE Transactions on Systems, Man & Cybernetics SMC-18**(6) 1016-1020.

Gaines, B.R. (1989). Social and cognitive processes in knowledge acquisition. **Knowledge Acquisition 1**(1) 251-280.

Gaines, B.R. (1990a). Intelligent systems as a stage in the evolution of information technology. Ras, Z.W. and Zemankova, M., Ed. **Intelligent Systems: State of the Art and Future Directions**. Chichester, UK, Ellis Horwood.

Gaines, B.R. (1990b). Knowledge support systems. **Knowledge-Based Systems 3**(3) 192-203.

Gaines, B.R. (1991a). An interactive visual language for term subsumption visual languages. **IJCAI'91: Proceedings of the Twelfth International Joint Conference on Artificial Intelligence**. pp.817-823. San Mateo, California, Morgan Kaufmann.

Gaines, B.R. (1991b). Modeling and forecasting the information sciences. **Information Sciences 57-58** 3-22.

Gaines, B.R. and Shaw, M.L.G. (1986). A learning model for forecasting the future of information technology. **Future Computing Systems 1**(1) 31-69.

Gaines, B.R. and Shaw, M.L.G. (1992). Documents as expert systems. Bramer, M.A. and Milne, R.W., Ed. **Research and Development in Expert Systems IX. Proceedings of British Computer Society Expert Systems Conference**. pp.331-349. Cambridge, UK, Cambridge University Press.

Garfield, E. (1979). **Citation Indexing: Its Theory and Application in Science, Technology and Humanities**. Philadelphia, ISI Press.

Garfield, E. (1990). Response to the panel on the evaluation of scientific information and the impact of new information technology. **Journal American Society Information Science 41**(3) 229-230.

Garrett, J.R. (1991). Text to screen revisited: copyright in the electronic age. **ONLINE 15**(2) 22-24.

Gelder, L. Van (1991). The strange case of the electronic lover. Dunlop, C. and Kling, R., Ed. **Computerization and Controversy: Value Conflicts and Social Choices**. pp.364-375. Boston, Academic Press.

Gellner, E., Ed. (1974). **Legitimation of Belief**. Cambridge, UK, Cambridge University Press.

Gibbs, S. (1991). Composite multimedia and active objects. Paepcke, A., Ed. **OOPSLA'91: Conference on Object-Oriented Programming Systems, Languages and Applications**. pp.97-112. New York, ACM Press.

Goldfarb, C.F. (1990). **The SGML Handbook**. Oxford, Clarendon Press.

Goodman, H.J.A. (1987). The "world brain/world encyclopaedia" concept: its historical roots and the contributions of H.J.A. Goodman to the ongoing evolution and implementation of the concept. **ASIS'87: Proceedings 50th Annual Meeting American Society Information Science**. pp.91-98. Medford, New Jersey, Learned Information.

Gore, D. (1976). Farewell to Alexandria: The theory of the no-growth, high-performance library. Gore, D., Ed. **Farewell to Alexandria**. pp.170-171. Westport, Connecticut, Greenwood Press.

Graham, J.W. and Stablein, R.E. (1985). A funny thing happened on the way to publication: newcomer's perspectives on publishing in the organizational sciences. Cummings, L.L. and Frost, P.J., Ed. **Publishing in the Organizational Sciences**. pp.138-154. Homewood, Illinois, Irwin.

Greif, I., Ed. (1988). **Computer-Supported Cooperative Work: A Book of Readings**. San mateo, California, Morgan Kaufmann.

Habermas, J. (1985). Modernity—an incomplete project. Foster, H., Ed. **Postmodern Culture**. pp.3-15. London, Pluto Press.

Hall, E.T. (1959). **The Silent Language**. New York, Doubleday.

Hawkins, D. (1983). An analysis of expert thinking. **International Journal of Man-Machine Studies 18**(1) 1-47.

Heim, M. (1987). **Electric Language: A Philosophical Study of Word Processing**. New Haven, Yale University Press.

Heller, A. (1990). **Can Modernity Survive?** Berkeley, University of California Press.

Hendley, T. (1987). **CD-ROM and Optical Publishing Systems**. Westport, Connecticut, Meckler.

Herwijnen, E. van (1990). **Practical SGML**. Dordrecht, Holland, Kluwer.

Hiltz, S.R. (1984). **Online Communities: A Case Study of the Office of the Future**. Norwood, New Jersey, Ablex.

Hoepner, P. (1991). Synchronizing the presentation of multimedia objects—ODA extensions. **ACM SIGOIS Bulletin 12**(1) 19-32.

Huston, M.M. (1990). New media, new messages: innovation through the adoption of hypertext and hypermedia technologies. **Electronic Library 8**(5) 336-343.

Innis, H.A. (1951). **The Bias of Communication**. Toronto, University of Toronto Press.

Kassirer, J.P. (1992). Journals in bits and bytes: electronic medical journals. **New England Journal Medicine 326**(3) 195-197.

Knuth, D.E. (1986). **The TEXbook**. Reading, Massachusetts, Adison-Wesley.

Kuhlthau, C.C. (1991). Inside the search process: information seeking from the user's perspective. **Journal American Society Information Science 42**(5) 361-371.

Kuhn, T.S. (1962). **The Structure of Scientific Revolutions**. Chicago, University of Chicago Press.

Lambert, S. and Ropiequet, S., Ed. (1986). **CD ROM: The New Papyrus**. Redmond, Washington, Microsoft Press.

Lambiotte, J.G., Dansereau, D.F., Cross, D.R. and Reynolds, S.B. (1989). Multirelational semantic maps. **Educational Psychology Review 1**(4) 331-367.

Lancaster, E.W. (1978). **Toward Paperless Information Systems**. New York, Academic Press.

Lederberg, J. and Uncapher, K. (1989). Towards a National Collaboratory. **Report of an Invitational Workshop at The Rockefeller University**.

Leland, M.D.P., Fish, R.S. and Kraut, R.E. (1988). Collaborative document preparation using Quilt. **Proceedings of the Conference on Computer-Supported Cooperative Work**. pp.206-215. New York, Association for Computing Machinery.

Lerner, R.G., Metaxas, T. and Scott, J.T. (1983). Primary publication systems and scientific text processing. **Annual Review of Information Science and Technology 18** 127-149.

Leslie, L.Z. (1989). Manuscript review: a view from below. **Scholarly Publishing 20**(2) 123-128.

Lewis, B.T. and Hodges, J.D. (1988). Shared books: collaborative publication management for an office information system. **Proceedings of the Conference on Office Information Systems**. pp.197-204. New York, Association for Computing Machinery.

Lynch, C.A. and Preston, C.M. (1990). Internet access to information resources. **Annual Review of Information Science and Technology 25** 263-312.

MacDonald, A.H. (1990). Publish and perish: a voyage into the black hole of scholarly communications in the 90's. **National Meeting of Vice-Presidents (Academic)**, Calgary, University of Calgary. 1-20.

Malcolm, N. and Gaines, B.R. (1991). A minimalist approach to the development of a word processor supporting group writing activities. **COCS'91: Proceedings of Conference on Organizational Computing Systems**. pp.147-152. ACM Press.

Marchetti, C. (1980). Society as a learning system: discovery, invention and innovation cycles revisited. **Technological Forecasting & Social Change 18** 267-282.

Marshall, R. (1991). Manipulating full-text scientific databases: a logic-based semantico-pragmatic approach. **Computer Journal 34**(3) 245-253.

Matkin, R.E. and Riggar, T.F. (1991). **Persist and Publish**. Niwot, Colorado, University Press of Colorado.

Matta, K.F. and Boutros, N.E. (1989). Barriers to electronic mail systems in developing countries. **Information Society 6**(1/2) 59-68.

McKnight, C., Dillon, A. and Richardson, J. (1991). **Hypertext in Context**. Cambridge, UK, Cambridge University Press.

McLuhan, M. and Rogers, B.R. (1989). **The Global Village: Transformations in World Life and Media in the 21st Century**. New York, Oxford University Press.

Meghini, C., Rabitti, F. and Thanos, C. (1991). Conceptual modeling of multimedia documents. **IEEE Computer 24**(10) 23-30.

Mensch, G. (1975). **Stalemate in Technology: Innovations Overcome the Depression**. Cambridge, Massachusetts, Ballinger.

Merton, R.K. (1957). Priorities in scientific discovery. **American Sosciological Review 22**(6) 635-659.

Merton, R.K. (1968). The Matthew effect in science. **Science 159**(3810) 56-63.

Mills, C.B. and Weldon, L.J. (1987). Reading text from computer screens. **ACM Computing Surveys 19**(4) 329-358.

Neches, R., Fikes, R., Finin, T., Gruber, T., Patil, R., Senator, T. and Swartout, W.R. (1991). Enabling technology for knowledge sharing. **AI Magazine 12**(3) 36-56.

Neuwirth, C.M., Kaufer, D.S., Chandhok, R. and Morris, J.J. (1990). Issues in the design of computer support for co-authoring and commenting. **Proceedings of theThird Conference on Computer-Supported Cooperative Work**. pp.183-195. New York, Association for Computing Machinery.

Novak, J.D. and Gowin, D.B. (1984). **Learning How To Learn**. New York, Cambridge University Press.

Nowakowska, M. (1975). Measurable aspects of the concept of scientific career. Knorr, K.D., Strasser, H. and Zilian, H.G., Ed. **Determinants and Controls of Scientific Development**. pp.295-322. Dordrecht, Holland, Reidel.

Nunamaker, J.F., Dennis, A.R., Valacich, J.S., Vogel, D.R. and George, J.F. (1991). Electronic meetings to support group work. **Communications of the ACM 34**(7) 40-61.

O'Docherty, M.H. and Daskalakis, C.N. (1991). Multimedia information systems—the management and semantic retrieval of all electronic data types. **Computer Journal 34**(3) 225-238.

O'Leary, M. (1991). Dialog and the American Chemical Society play a high stakes game. **ONLINE 15**(1) 15-20.

Oakeshott, P. (1985). The 'BLEND' experiment in electronic publishing. **Scholarly Publishing 17**(1) 25-36.

Oddy, R.N., Robertson, S.E., Rijsbergen, C.J. van and Williams, P.W., Ed. (1981). **Information Retrieval Research**. London, Butterworths.

Oldenburg, H. (1966). **Correspondence of Henry Oldenburg**. Madison, University of Wisconsin Press.

Ong, W.J. (1977). **Interfaces of the Word: Studies in the Evolution of Consciousness and Culture**. Ithaca, New York, Cornell University Press.

Palca, J. (1991). New journal will publish without paper. **Science 253**(5027) 1480.

Pehrson, B., Gunningberg, P. and Pink, S. (1992). Distributed multimedia applications on gigabit networks. **IEEE Network Magazine 6**(1) 26-35.

Philips, R.L. (1991). MediaView: a general multimedia digital publication system. **Communications of the ACM 34**(7) 75-83.

Pondy, L.R. (1985). The reviewer as defense attorney. Cummings, L.L. and Frost, P.J., Ed. **Publishing in the Organizational Sciences**. pp.210-219. Homewood, Illinois, Irwin.

Popper, K.R. (1968). Epistemology without a knowing subject. Rootselaar, B. Van, Ed. **Logic, Methodology and Philosophy of Science III**. pp.333-373. Amsterdam, North-Holland.

Popper, K.R. (1972). **Objective Knowledge: an Evolutionary Approach**. Oxford, Clarendon Press.

Prigogine, I. (1991). Les objets nomades et la bibliotèque. Liebaers, H. and Walkiers, M., Ed. **Library Automation and Networking: New Tools for a New Identity**. Munich, Saur.

Prusker, F.J. and Wobber, E.P. (1991). The siphon: managing distant replicated repositories. Paris Research Laboratory, Digital Equipment Corporation. PRL-RR-11.

Quarterman, J.S. (1990). **The Matrix: Computer Networks and Conferencing Systems Worldwide**. Reading, Massachusetts, Digital Press.

Robertson, B. (1991). Electronic mail—is it safe to use? **Computers & Security 10**(1) 17-19.

Rosenberg, J., Sherman, M., Marks, A. and Akkerhuis, J. (1991). **Multi-media Document Translation: ODA and the EXPRES Project**. New York, Springer.

Samarajiva, R. (1988). Appropriate technology for secondary information delivery to small third world countries: print, microform, online or CD-ROM? **ASIS'88: Proceedings 51st Annual Meeting American Society Information Science**. pp.17-22. Medford, New Jersey, Learned Information.

Samarajiva, R. and Mukherjee, R. (1990). Information services and the intelligent network: policy implications from US and Canadian experience. **ASIS'90: Proceedings 53rd Annual Meeting American Society Information Science**. pp.326-333. Medford, New Jersey, Learned Information.

Sanford, C.C. (1991). Augmenting the gatekeeper's role: a decision support system for a journal editor. **IEEE TransactionsProfessional Communications 34**(3) 140-146.

Schoorman, F.D. (1985). Publishing in the organizational sciences: the dilemma of values. Cummings, L.L. and Frost, P.J., Ed. **Publishing in the Organizational Sciences**. pp.155-170. Homewood, Illinois, Irwin.

Schwab, D.P. (1985). Reviewing empirically based manuscripts: perspectives on process. Cummings, L.L. and Frost, P.J., Ed. **Publishing in the Organizational Sciences**. pp.171-192. Homewood, Illinois, Irwin.

Seliger, R. (1985). Design and implementation of a distributed program for collaborative editing. Master's Thesis. Massachusetts Institute of Technology.

Senders, J.W. (1977). An on-line scientific journal. **The Information Scientist 11** 3-9.

Senders, J.W. (1981). I have seen the future and it doesn't work: The electronic journal experiment. **Proceedings of the Society for Scholarly Publishing 2nd Annual Meeting**. Washington, Society for Scholarly Publishing.

Shackel, B. (1982a). The BLEND system: programme for the study of some 'electronic journals'. **Computer Journal 25**(2) 161-168.

Shackel, B. (1982b). Plans and initial progress with BLEND—an electronic network communication experiment. **International Journal Man-Machine Studies 17**(2) 225-233.

Shaw, M.L.G. and Gaines, B.R. (1987). KITTEN: Knowledge initiation and transfer tools for experts and novices. **International Journal of Man-Machine Studies 27**(3) 251-280.

Shaw, M.L.G. and Gaines, B.R. (1991). Extending electronic mail with conceptual modeling to provide group decision support. **COCS'91: Proceedings of Conference on Organizational Computing Systems**. pp.153-158. New York, ACM Press.

Sherblom, J. (1988). Direction, function and signature in electronic mail. **Journal of Business Communication 25**(4) 39-54.

Sheridan, T., Senders, J., Moray, N., Stoklosa, J., Guillaume, J. and Makepeace, D. (1981). Experimentation with a multi-disciplinary teleconference and electronic journal on mental workload. Massachusetts Institute of Technology.

Sievert, M.C. (1990). Ten years of the literature on online searching: an analysis of *Online* and *Online Review*. **Journal American Society Information Science 41**(8) 560-568.

Simpson, D.B. (1989). Libraries and the changing scholarly process. Milwaukee School of Library and Information Science, University of Wisconsin, Milwaukee, Wisconsin. TR-FD-90-08.

Sneed, J.D. (1977). **The Logical Structure of Mathematical Physics**. Dordrect, Holland, Reidel.

Sowa, J.F. (1984). **Conceptual Structures: Information Processing in Mind and Machine**. Reading, Massachusetts, Adison-Wesley.

Stegmüller, W. (1976). **The Structure and Dynamics of Theories**. New York, Springer.

Stein, R.M. (1991). Browsing through terabytes: wide-area information services open a new frontier in personal and corporate information services. **Byte 16**(5) 157-164.

Stodolsky, D.S. (1990). Consensus Journals: Invitational journals based upon peer consensus. Roskilde University Centre, Institute of Geography, Socioeconomic Analysis, and Computer Science. Datalogiske Skrifter No. 29/1990.

Stone, R. (1991). Court test for a plagiarism detector? **Science 254**(5037) 1448.

Stoneman, P. (1983). **The Economic Analysis of Technological Change**. Oxford, Oxford University Press.

Suppes, P., Jerman, M. and Brian, D. (1968). **Computer-Assisted Instruction: Stanford's 1965-66 Arithmetic Program**. New York, Academic Press.

Tanaka, K. and Okamoto, E. (1991). Key distribution for mail systems using ID-related information directory. **Computers & Security 10**(1) 25-33.

Tenopir, C. (1984). Full-text databases. **Annual Review of Information Science and Technology 19** 215-246.

Toffler, A. (1980). **The Third Wave**. New York, William Morrow.

Walden, J. (1987). **More File Formats for Popular PC Software**. New York, Wiley.

Walter, M. (1989). ODA (office document architecture): What is it? What is it good for? **Seybold Report on Publishing Systems 19**(7) 3-20.

Walter, M. (1992). Scholarly journals: case studies of SGML for electronic publishing. **Seybold Report on Publishing Systems 21**(18) 3-20.

Wells, H.G. (1938). **World Brain**. New York, Doubleday.

Zuckerman, H. and Merton, R.K. (1971). Institutionalized patterns of evaluation in science. **Minerva 9**(1) 66-100.