

# Information Technology: Catalyst for Change in Scientific Communication

Julie Hurd

*University of Illinois at Chicago*

---

Julie Hurd, "Information Technology: Catalyst for Change in Scientific Communication." *Proceedings of the IATUL Conferences*. Paper 27.

<http://docs.lib.purdue.edu/iatul/1996/papers/27>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact [epubs@purdue.edu](mailto:epubs@purdue.edu) for additional information.

# **Information Technology: Catalyst for Change in Scientific Communication**

*Hurd, Julie M.  
University of Illinois at Chicago, USA*

---

## **Computer-supported Communication as a Change Agent**

The personal computer, teamed with the modem, and low-cost telecommunications networks, first provided new communication channels to permit computer-to-computer information transfer. Wireless technologies including satellite transmitters and cellular telephones serving as receivers now further broaden the communication base and allow emerging nations access to communication channels without necessarily wiring a nationwide communication network. Classrooms, laboratories, libraries, hospitals, offices and homes around the globe can be connected for purposes of transmitting words, data, images, and sounds. It is not surprising that, in the fall of 1995, politicians as far apart in philosophy as Vice President Albert Gore and Speaker of the House Newt Gingrich agreed on the importance of information technologies to bring about sweeping social changes.

Overshadowing most other technological development in already-realized and potential impact is the global network of networks known as the Internet. This system of interconnected computer networks has its origins in 1969 in the Advanced Projects Research Administration Network (ARPANET), a network that was created and funded by the U.S. Department of Defense to link four university-based supercomputers with other key research sites at universities and government facilities. This original network functioned as a high-speed backbone to which other university and government computer networks could link to route messages with great efficiency. The communications protocols and technical standards that continue in use today were developed by a loosely organized and dispersed community of scientist-users motivated by shared interests in scientific research. The linked networks have evolved into the Internet, also referred to in the popular media at this time as the Information Superhighway, cyberspace, and formerly as the National Information Infrastructure (NII) and the National Research and Education Network (NREN). The governance of the Internet continues to be loose; the economic base is increasingly complex as commercial organizations have become active and government funding support has dwindled.

The Internet has been adopted with general enthusiasm by almost every scholarly and scientific discipline and has, as well, encouraged the growth of a vast profit-sector array of organizations and services that market access to virtually any interested computer user. Although the Internet originally was created to facilitate sharing of research data among a very elite community of scientists, it has now broadened its user base to include every segment of society including children in grade schools and average citizens who may use a computer for strictly recreational pursuits. At this time estimates of the number of people connected to the Internet range from 16 million to 40 million, with an exact number of users probably unknowable. Charles

Arthur, writing in New Scientist, suggests that arguing about the precise number of Internet users is the contemporary equivalent of earlier religious debates o how many angels could dance on the head of a pin! [[Arthur, 1995](#)] It is likely sufficient to realize that the number of users of the Internet continues to increase and this network, with all its capabilities, represents a fundamental transformation in human communication.

### **The Garvey/Griffith Model of Scientific Communication**

To provide a context for consideration of new communication models I find it useful to refer to a model of the traditional paper-based system. Over thirty years ago William Garvey, Belver Griffith, and co-workers developed this model of the scientific communication system based on their observations of psychologists. [[Garvey and Griffith, 1972](#); [Garvey, 1979](#) and references cited therein] Garvey asserted that “communication is the essence of science” and that scientific communication as a social process would lend itself to the methodology of social psychology. The Garvey/Griffith model was subsequently demonstrated to be generally applicable across both the physical and social sciences. It outlines the process by which research is communicated and provides details of the various stages within a time frame encompassing from initial concept to integration of the research as an accepted component of scientific knowledge. Although the time scale varies from one discipline to another, the essential elements of the model appear to be universal.

The Garvey/Griffith model was postulated based on the communication channels then operational. These were both informal and formal and included personal (oral) communications to individuals and groups as well as publication in paper-based journals and books. Figure 1 is a general representation of Garvey and Griffith’s model adapted from illustrations in their publications. It outlines the communication of research findings in a typical scientific discipline. The refereed scientific article is the key element in the system; formal and informal communication lead to journal publication as the expected outcome of scientific research. The system serves additional functions including support for a reward structure that had long been the norm for scientists working in higher education. It also has fostered the development of the present array of scientific publishers that includes both non-for-profit associations and commercial organizations who produce both primary journals and the secondary services that facilitate access to the contents of journals by providing indexing and abstracting.

Since Garvey and Griffith developed this model, emerging information technologies have dramatically altered and enhanced options for communicating. The applications of computers to publishing have resulted in online bibliographic databases and large amounts of machine-readable text created to support the publication of books and journals, as well as totally electronic journals., Visionaries such as F.W. Lancaster [[1978](#)] could foresee in these developments a “paperless” future. Although that future has not become reality yet, the technological foundations are in place; the economic, social and political barriers have not been overcome.

## **New Models of Scientific Communication**

Recent organizational and societal changes have created a dynamic environment with pressures for new ways to examine issues and new approaches to scientific problem-solving. Big Science has become “Bigger Science” with increased emphasis on collaborative and team research directed to address problems of global importance whose solutions require expensive facilities and equipment and create enormous data sets. Concurrently, computer-based information technologies have emerged that are beginning to change the ways scientists use, produce and disseminate information.

As we speculate on how computer-based information technologies can both enhance and alter the scientific communication system, we can extend the traditional model just described to a digital environment and can also consider how some current developments suggest very different models. In envisioning future communication systems I have found it helpful to employ terminology borrowed from Clifford Lynch [1993]. He draws a distinction between modernization and transformation of scientific communication. Modernization is defined as the use of new technology to continue doing the same thing, but presumably in a more cost-effective and/or efficient way. Transformation is the use of a new technology to change processes in a fundamental way. At the present time there are more examples of modernization of the communication system than transformation, but that is likely to be attributable to the early stages of the transition. The three models that I will next describe represent both types of responses to technological innovation.

### **Modernized Garvey/Griffith Model**

The original Garvey/Griffith model was developed during a “print-on-paper” era and described scientific communication broadly across many disciplines. With the development of information technologies that provide for production of a manuscript using a word processor and transmission of text and data across computer networks, electronic equivalents have emerged to many paper forms of communication. Figure 2 depicts a modernized Garvey/Griffith model in which electronic media replace paper formats. Computer-based communication offers alternatives to seminar discussions, conferences and other informal means of sharing research findings. The basic unit of distribution continues to be the scientific journal, although now published electronically. This is Lancaster’s “paperless information system” and at this time it co-exists, in many scientific specializations, with the traditional system. Aspects of this model were described over fifteen years ago by Lancaster [1978] in his book on “paperless information systems” even before the Internet was fully developed as far-reaching as we know it today. By recognizing that every element of the traditional model has been affected by information technology, this modernized model can be outlined. Although electronic-based, it retains the key feature of the well-established system by building on the peer-reviewed scientific journal as the basic unit of distribution.

The modernized Garvey/Griffith model retains the basic elements of a traditional paper-based system by building on peer-reviewed journals as the unit of distribution for research. By moving from a paper to electronic medium the communication

process is accelerated and research findings are disseminated more rapidly at all stages in the stream of communication. Perhaps as significant is another feature of networked communication, its potential for opening the process to individuals previously excluded. If listserves are open to all (although it must be recognized that some are closed), scientists can participate in discussions whatever their institutional affiliation or geographic location. Similarly, electronic conferences ignore geographic and financial constraints, although participants must have suitably equipment and connectivity. Pre-Print databases provide access to research as yet unpublished in journals to far wider groups than were privy to such information when paper copies were mailed to lists of colleagues. This model represents a modernized system with enhanced opportunities for faster and wider communication through networked technology; it is not a transformed paradigm.

Next I shall describe two transformed models that are based on developments in high energy physics and human genome research where technology has been appropriated to devise new systems and structures to support communication.

### **High Energy Physics Model**

High energy physics is an example of an environment that could be expected to be particularly receptive to applications of emerging information technologies that hold promise for improving access to recent research findings. It is not surprising, therefore, that an electronic preprint distribution system developed for the high energy physics community by Los Alamos physicist Paul Ginsparg has been readily adopted by users. The electronic pre-print archive maintained by Ginsparg now serves specializations in physics, computational linguistics, economics and other disciplines and is potentially a model for a transformed communication system.

Paul Ginsparg [1994] described the history of the “e-print archives” that he developed and continues to operate from a Los Alamos Laboratory computer and acknowledged the existence of a “pre-print culture” in high energy physics that predated electronic communication networks that certainly contributed to a positive response to his efforts. He identified concurrent developments in computer software and hardware that were essential to the construction of electronic pre-print databases. First, acceptance during the mid-1980s by the physics community of TeX as their scientific word processor standardized the manuscript creation process. Distance collaboration is thus freed from the constraints imposed by the need to mail drafts of manuscripts; instead manuscripts in process could be available in real time to collaborating authors on computers connected to a network. The second essential development then was the great increase in computer connectivity that gave rise to the Internet. Finally, high-powered workstations with high-capacity storage media were needed.

Ginsparg saw potential in these emerging technologies for development of an electronic pre-print archive and distribution network. he wrote supporting software to automate the processes that would allow users to submit and replace papers, search and obtain pre-prints, and receive online assistance in using the system. He incorporated a current awareness function that allows users to subscribe and receive a daily listing of titles and abstracts of new papers added to the database. he intended that minimal computer literacy would be required and designed his system around an e-mail interface.

The acceptance of e-print archives by physicists has already begun to impact their libraries. The Special Libraries Association's Physics/Astronomy/Mathematics Division sponsors an active electronic discussion list (SLA-PAM@listserver.lib.muohio.edu) that has explored the issues related to electronic versus printed pre-prints. During a series of discussions on the list during a 1995 a number of major physics libraries reported that they had discontinued their paper pre-print collection and organization efforts and were relying instead on the e-print archives; other libraries continued to maintain paper files but had noted decreased usage and were monitoring developments closely. [STSL-1995] A significant difficulty cited by some librarians resulted from the desire of users to print copies of papers of interest from the archive. Not all physicists presently have equipment on their desktops that can handle graphics and mathematical notation; compressed graphics files that required UNIX machines for decompressing appeared to be a particular problem. However, those who had access to suitable equipment, whether through their library or elsewhere in the institution, were pleased with the quality of prints they could obtain. It seems likely that, as more physicists upgrade their workstations, these early difficulties will diminish.

At the present time the e-prints are still "pre-prints" and virtually all are submitted to refereed journals for traditional publication. The e-print archive software permits authors to insert a citation to a published article at any time, although not all necessarily do this. The e-print archive that Paul Ginsparg developed as an experiment can be extended by incorporating a reviewing function as Ginsparg has proposed. Figure 3 represents a model of such a system. This model is a "journal-less" paradigm with the e-print as the basic unit of distribution. Just as specialized journals collected articles for their readers, e-print archives in specialized fields would play the same role of partitioning a large body of literature. This model blurs the distinction between informal communication as represented by pre-prints and formal communication as in published journal articles. It also shifts roles of authors and publishers as authors become "publishers" by the act of transmitting an e-print to the database. It offers potential for a more broadly-based type of peer review as any reader of an e-print may comment on a submission. Whether this electronic-based distribution scheme will eventually displace traditional refereed journals is yet to be determined, but the high use that has been measured for the high energy physics prototype provides evidence of the value physicists place on the service.

### **Human Genome Project Model**

Furthest removed from the original Garvey/Griffith model, and representing a genuinely transformed communication system, is a model that evolves from the notion of the scientific collaboratory. The term "collaboratory" was coined in a National Research Council report and melds the notion of collaboration with laboratory to convey an image of a worldwide network of computers supporting a global research community. [Wulf, 1993] Scientists in a collaboratory exchange data, share computer power, and consult digital library resources, interacting across great distances as easily as if they were sharing a laboratory facility. The collaboratory concept is particularly applicable to those projects of "Bigger Science" requiring large

scale instrumentation such as observatories and space satellites or enormous shared databanks such as the one under construction in the Human Genome Project.

Research projects, such as the Human Genome Project, generate data that is transferred to central depositories and made widely available. This data may initially be unrefereed but later will be validated by others. This data is a type of “publication” and makes information available widely at an earlier stage in the research than would have occurred when publication of a refereed journal article would have been the means of distribution of the data. Posting of gene sequencing data in a shared repository diverges from past practice of including data in a refereed journal article; this may still occur at a later stage in the communication process, or may be replaced by a reference to the depository data. What is noteworthy is that data is the unit of distribution here. Figure 4 is a model of a transformed communication system drawn from the Human Genome project.

## **Conclusions**

I believe that new communication models appear first in those scientific specialization's that are experiencing rapid changes and have adapted technologies in innovative ways. The reasons that scientists in these fields can be expected to shift to new modes of communication are likely complex. A number of variables seem to be influential:

- high level of research front activity which fosters a need for very rapid communication of research
- reliance on informal communication which leads to a high level of invisible college interaction
- large-scale collaborative projects at scattered locations
- multi-institutional research projects at scattered locations
- large datasets that can be shared and used for multiple experiments

Each of these change agents can stimulate shifts in communication behaviors that, in turn, lead to new models such those presented here.

All of the innovations that have led to the models presented here - listserves, pre- print archives, and shared depositories of data - have changed communication in significant ways:

- invisible colleges may be more accessible and more freely joined,
- research results are available sooner and in unrefereed forms,
- articles and data are the units of distribution rather than journal issues,
- authors are becoming direct publishers through Web pages, and
- boundaries are blurring between informal and formal communication.

The above changes are beginning to impact journal publishers who are responding with their own innovations, primarily in modes of distribution. The well-established system of peer review may also adapt to an electronic environment in response to

these changes. These and other aspects of scientific publishing might easily be the focus a paper in themselves.

### **The Shift from Print: How Soon?**

Today I have presented a case for the ultimate transformation of scientific communication from a print-based system dominated by the refereed scientific journal to an electronic system in which the basic units of distribution of information may well be individual articles and data. It seems clear that the numerous advantages offered by a networked-based structure facilitate both informal and formal communication among scientists and will eventually transform the present system. The evolution will be gradual with some early changes being merely electronic versions of paper-based communication; this can be seen in some of the electronic journals that endeavor to look exactly as does the printed page, even employing the same typeface and page layout. Ultimately, new formats will emerge, that unlike their paper predecessors, could be dynamic, interactive, multi-media, non-linear, and more. Network-based "publication" may be initiated by authors and could look very much like documents that are found in pre-print databases with peer review adapting to an electronic environment and taking on a more open form by involving more scientists whose commentaries could be read and evaluated by others. An e-print of the future will possess features of both informal and formal formats presently in use.

Print and electronic information sources are likely to co-exist for some time to come although there will likely be major changes in the scientific publishing sector. Libraries and librarians will have new opportunities for participation in the evolving communication system that will result in transformed roles while at the same time preserving some of the traditional functions of libraries. Karen Drabentstott has described a "library of the future" that is increasingly a digital library.

For a while, library users will rely on the paper collections that libraries

"have amassed over the years. As information is increasingly produced in digital artifact form ...Paper collections will slowly fall into disuse and large portions of such collections will be warehoused at remote locations....libraries will continue to be associated with buildings. Although physical collections of books, journals and other materials will no longer consume valuable space in these buildings, we can envision the need for workspaces where users consult state-of the art computer workstations; study spaces where users demand quiet for contemplation and reflection..."

[\[Drabentstott, 1994, 168\]](#)

### **References**

1. Arthur, Charles. (1995, May 13) "And the net total is..." New Scientist, Number 1977, 29-31.
2. Drabentstott, Karen M., (1994) Analytical Review of the Library of the Future. Wash., DC: Council on Library Resources.
3. Garvey, W.D. (1979) Communication: The Essence of Science. Elmsford, NY: Pergamon Press.

4. Garvey, W.D. & Griffith, B.C. (1972) "Communication and information processing within scientific disciplines: Empirical findings for psychology," *Information Storage and Retrieval*, 8, 123-126.
5. Ginsparg, Paul, (1994) "First steps towards electronic research communication," *Computers in Physics*, 8 (no.4), 390-396.
6. Hurd, Julie M., Weller, Ann C. and Crawford, Susan (1996) *From Print to Electronic: The Transformation of Scientific Communication*. American Society for Information Science Monograph. Medford, NJ: Information Today.
7. Lancaster, F.W. (1978) *Toward Paperless Information Systems*. London: Academic Press.
8. Lynch, Clifford. (1993) "The transformation of scholarly communication and the role of the library in the age of networked communication," *Serials Librarian*, 23 (no. 3), 5-20.
9. STS-L summary of SLA-PAM discussion. (1995, April 22). Available via: [sts-l@utkvm1.utk.edu](mailto:sts-l@utkvm1.utk.edu)
10. Wulf, William A. (1993, August 13) "The collaboratory opportunity," *Science*, 261: 854-855.