providing individualized assistance to students.[21]

The driving force behind the advance of information technology has been the development of faster, smaller, and cheaper electronic devices, which can be used to produce machines with greater capabilities for manipulating and processing information. These machines have in turn inspired the production of more powerful and imaginative programs and solution techniques (computational methods or algorithms) for solving problems that would be intractable without this new computational power. The availability of increased computational power, in turn, has enabled the design of new computer hardware and software, producing a snowball effect in which each new generation of system facilitates the design of its successor. This process can be expected to continue until designers reach the fundamental limitations of physics and exhaust all technological alternatives, which does not appear imminent. An improvement in computational power of six orders of magnitude (a factor of a million) over the past two decades can be attributed to roughly equal improvements (three orders of magnitude each) in hardware and software.[22] It is not unreasonable to expect a comparable improvement to occur over the next two or three decades. As a result, in the next few decades an unimaginable amount of computational power will be available to scholars. This capacity compels the archival profession to determine the implications of the use of information technology by scholars for conventional archival practices.

Although the future evolution of information technology is fairly predictable in broad outline, predicting precise details of how the technology will evolve is more difficult. For our purposes, however, it is the broad outline of these trends that is most important. Our discussion of technology, therefore, avoids mentioning specific devices, techniques, or research results. Instead, the next section examines trends of information technology that are likely to have the greatest impact on scholarly communication—and, by implication, on archives management. The focus here is on broad descriptions and projections most relevant to the future of scholarly research. Later in this paper we examine how scholars are actually using information technology in their current work.

OVERVIEW OF INFORMATION TECHNOLOGY TRENDS

The two most obvious—and for the purpose of this paper, the most important—information technology trends that pertain to scholarly communication are *end-user computing* and *connectivity.* These trends are distinct and separable, and each is discussed in detail below. Ultimately, however, it is the integration of the two that will have the greatest impact on scholarly communication. End-user computing enhances the autonomy of the researcher, i.e., the researcher's ability to use the power of computation to conceptualize and execute research without sacrificing intellectual control by delegating computational tasks to specialists. Connectivity enhances the researcher's abilities to access data, collaborate, seek input and feedback, and disseminate ideas and results. The confluence of these trends produces a rich interplay of synergistic effects, which are explored below.

A number of more specific technology

---

[21] See Miall, *Humanities and the Computer,* 4; and Jean-Claude Gardin, "The Future Influence of Computers on the Interplay Between Research and Teaching in the Humanities," *Humanities Communication Newsletter* 9 (1987): 17–18.

[22] *Grand Challenges: High Performance Computing and Communications, The FY 1992 U.S. Research and Development Program, A Report by the Committee on Physical, Mathematical, and Engineering Sciences, Federal Coordinating Council for Science, Engineering and Technology, Office of Science and Technology Policy* (1991), 14–15.

trends are also likely to affect scholarly communication. Most of these are examples of end-user computing or connectivity (or the integration of the two), but each warrants attention in its own right. The most relevant of these appear to be artificial intelligence, end-user publication and distribution, hypermedia, and visualization and virtual reality.

End-User Computing

In the current context, *end-user computing* refers to the direct use of computers by researchers.[23] The general trend toward the increased use of computers is understandable. Computers continue to become better, cheaper, more accessible, and more usable. Software continues to become more application-oriented, and user interfaces continue to improve. Databases continue to become larger and more relevant. As the use of computers becomes more common, users continue to increase in number and sophistication, generating greater and greater demand for computation while driving prices even lower by expanding the size of the market. But the increasingly *direct* use of computers by their end-users is a more recent and more interesting trend, and its implications for research are profound.

The term *end-user* refers to someone who physically uses a computer—the person who touches the keyboard and reads the screen. The end-user may or may not initiate or consume the results of the computation. It is useful to distinguish the end-user from the "ultimate user" of a computer: someone who initiates and consumes the results of a computation, without necessarily touching or seeing the machine. The ultimate user is the person who causes a computation to be performed and who uses the results of the computation, i.e., the person whose work involves computation, whether or not it involves using a computer directly.

End-user computing occurs when the end-user and the ultimate user are the same. The crux of end-user computing is that the end-user is able to initiate computations and get results without going through an intermediary. To some extent, this is a detail: What difference does it make if a computation is performed by a researcher or a programmer? But the distinction is an important one, since it bears on how central the computation is to the researcher's thought process. If a researcher is the ultimate user of a database, for example, but is not the end-user, then some intermediary (librarian, data archivist, programmer, secretary, or assistant) is interposed between the researcher and the database, limiting the researcher's ability to interact directly with the data, to browse through it, to explore its idiosyncrasies, and to become intimate with it. Similarly, if a researcher asks someone else to write a program to compute summary statistics, the researcher will be unaware of the decisions embedded in that program or the problems encountered in writing it.[24] This kind of insulation from the computational process may free the researcher from menial tasks, but it also limits his or her ability to define the computation

---

[23] For most users, the trend toward direct access began with personal computers (PCs), but it actually began soon after the advent of the modern computer. The very first computers of the early 1950s were essentially single-user machines and, since users had to be very aware of their machines' foibles (and typically had to be present while running their programs in order to deal with problems), they necessarily became intimate end-users. Later, more reliable mainframe computers often ran jobs in batch mode (batches of work were run together instead of individually) to improve their utilization, which tended to distance users from their machines. In the early 1960s, however, timesharing reintroduced direct access by allowing multiple users to share a mainframe machine remotely from their terminals.

[24] Although writing a program does not guarantee that one will become—let alone remain—aware of its implications and limitations, using a program written by someone else virtually guarantees that the user will *not* be aware of them.

correctly, use it appropriately, and understand the implications of its results.

From a practical standpoint, end-user computing is attractive because of its convenience. An end-user need not find a programmer or data processing specialist (and an available machine) to get an answer to a computational problem. This reduces the threshold of effort required to perform computation, allowing users to consider it a more integral part of their work style.

The ramifications of end-user computing in the research process are deeper and more subtle than they may first appear. Only by becoming intimate with the computational process can a researcher fully realize the potential of computation in performing research. Only when the researcher is an end-user does computing become familiar enough and convenient enough to be a natural part of the research process. This is not an end in itself, but it is important because it allows the researcher to conceive of new kinds of research that become possible only when computation becomes an integral part of research. End-user computing is an important trend because the activity of computation allows researchers to reconceive the nature of research itself, i.e., the kinds of questions posed, the methodologies used, the type and extent of sources analyzed, and the form of presentation of the findings. (Examples are discussed in a later section.)

To summarize: End-user computing means direct access to computational capability; the key implication of this in the current context is that it allows computation to become an integral part of a researcher's thought process—and therefore of the research itself.

Ubiquitous computing. One trend that is still relatively new is the advent of portable computing, using laptop, notebook, or even pocket-sized ("palmtop") computers. This portability means more than just being able to carry a computer from one location to another. It implies the ability to carry a part of one's working context (database, text, notes, and correspondence) in a machine that can be used on location, in meetings, or while traveling. This context may be "downloaded" to a portable machine from a researcher's home machine and used for on-site research or during interactions with other researchers to modify data, record notes, work on evolving documents, and many other tasks. The results of this work can then be "uploaded" to the researcher's home machine, by a telecommunications link from the remote location or by a direct transfer of data after the researcher returns home.

In addition to portable machines themselves, cellular modems (modulator/demodulators) allow computers to communicate over cellular telephone links. This allows the user to link computers while traveling anywhere that cellular telephone coverage is provided; it is already possible to connect to a remote computer or database from a portable computer while riding in a taxi in any major city in the United States. Whether this kind of remote computing will ultimately become a common activity depends on tradeoffs between the size, cost, and capacity of portable versus remote computers and the attendant telecommunications costs.

The important point is not the size and capability of portable machines, but rather the freedom they give the user to perform computations and to access data from any location. For example, another way of achieving the same result would be to provide computer terminals in public places; this would be analogous to the use of standard (noncellular) telephones, which are ubiquitously available anywhere in the developed world. The French government has implemented just such an approach to computing in its Minitel system, which is available in homes and post offices throughout France.[25] Because of these alternatives, it

---

[25] David L. Margulius, "C'est la France, C'est Min-

is useful to think of this as a trend toward "ubiquitous computing" rather than "portable computing." This is discussed further under Connectivity below.

End-user interfaces. The design of software for end-users has also had a tremendous impact on the growth of end-user computing. For end-users who are not computer specialists, "access" to computation means more than simply having a computer or communicating with one. To use a computer effectively, such users need software that allows them to work in ways that are natural to them, without having to learn the intricacies of an arcane computer system. Software for end-user computing must have two key attributes: It must provide functionality that is of use to the end-user, and it must present an interface that is usable by an end-user.

Appropriate functionality requires that software be either generically useful (such as word processors, electronic mail, databases, spreadsheets, and mathematical programs) or designed for some specific task that the user performs. Task-specific programs (or *applications*) tend to be written for users in a given industry or type of work.[26] But if its interface makes it difficult to use, neither generic nor task-specific software is of much value to any but the most dedicated and tenacious of end-users.

The trend toward improving end-user interfaces began in the early 1960s.[27] Many of the principles of current user interfaces were developed by Engelbart's group at the Stanford Research Institute (SRI) in the 1960s and early 1970s.[28] This led to the development of a number of systems at Xerox Corporation's Palo Alto Research Center (PARC) in the late 1970s, culminating in the introduction of the Star in 1981.[29] The Xerox Star pioneered the point-and-click, window- and menu-based "desktop metaphor" that is currently in vogue. This trend toward better user interfaces gained momentum with the development of personal computers, and it has now reached a point where many systems can be learned and used effectively by most users without any formal computer training. Although the term *user friendly* has become such an advertising cliché that it is now all but meaningless, its overuse is a measure of the extent to which the computer industry recognizes the importance of user interface design for end-user computing.

The "online transition." One of the key factors that facilitates end-user computing is an "online transition"[30] in which com-

---

itel," *PC Computing* 2 (January 1989): 194; Ellis Booker, 'Vive le Minitel," *Telephony* 215 (8 August 1988): 24; and S. Nora and A. Minc, *The Computerization of Society: A Report to the President of France* (Cambridge, Mass.: MIT Press, 1980).

[26] Both general-purpose and task-specific programs become more useful when they can be tailored to the needs of a particular end-user. Examples of this are word processors that allow users to define their own document formats, function keys, "macros," etc. The ultimate general-purpose program is a programming system (or language) that allows end-users to define new computations at will (i.e., to write programs); end-users may become programmers to a limited extent by tailoring software to their own needs.

[27] For example, Cliff Shaw's JOSS system is widely regarded as one of the earliest successful timeshared systems designed for direct access by researchers. See J. C. Shaw, *JOSS: Conversations with the Johnniac Open-Shop System* (Santa Barbara, Calif.: RAND Corporation, P-3146, 1965); J. C. Shaw, "JOSS: A Designer's View of an Experimental On-Line Computing System," in *American Federation of Information Processing Societies Conference Proceedings* (Fall Joint Computer Conference), Vol. 26 (Baltimore, Md.: Spartan Books, 1964): 455-64.

[28] In addition to inventing the mouse, this visionary group developed many of the concepts that form the foundation of modern user interface design, as well as producing one of the first hypertext systems. For an early description of this work, see D. C. Engelbart and W. K. English, "A Research Center for Augmenting Human Intellect," *American Federation of Information Processing Societies Conference Proceedings* (Fall Joint Computer Conference) vol. 33. (May 1974), 395-410.

[29] J. Johnson, T. L. Roberts, W. Verplank, D. C. Smith, C. H. Irby, M. Beard, and K. Mackey, "The Xerox Star: A Retrospective," *IEEE Computer* 22 (September 1989): 11-26.

[30] The term *online* originated in the electric power industry. Generating equipment is said to be "online"

puting becomes more useful the more it is used. If a user is still bound to the telephone, paper mail, paper documents, paper files, and paper memos, then computation remains an infrequently used tool that does not integrate with the rest of the work environment. When electronic mail (e-mail) begins to replace telephone and paper messages and when machine-readable electronic documents and files begin to replace paper, the user's working context is integrated in new ways.

The online transition produces a new phenomenon: Many previously separate forms of communication become integrated by being stored in electronic form. For example, if telephone messages and telephone directories are both electronic, users can forward information from a phone message in e-mail and can use telephone numbers or other information from a phone message to search their phone directories for information about callers. Many messages that traditionally have come by telephone will in the future be sent by e-mail instead, since e-mail is asynchronous (the recipient does not have to be present to receive an e-mail message) and provides a more legible and reliable medium for messages containing text or data. Similarly, users can easily copy text from letters, memos, and informal messages into new documents and search their contents electronically, rather than visually scanning voluminous printed material.

In the early stages of the online transition, computation does not fully realize its potential because it is not yet integrated into the user's work style. This creates a chicken-and-egg problem. Users are not motivated to use computation until its benefits outweigh the cost of learning to use it (and changing one's work style to make use of it); but its benefits are realized only after it becomes an integral part of one's work style. This problem produces a learning curve in which progress initially is slow, but it accelerates as the online transition proceeds. This curve rises steeply above a certain point, when a critical mass of the user's context becomes integrated online.

Summary. The exact ways in which computation will be delivered to end-users in the future will be determined by factors that involve trade-offs among the costs of computers, various kinds of memory and communication, and issues of privacy, convenience, and control. The form in which computation is delivered will continue to evolve as the relative costs and benefits of various alternatives change. Ultimately, the end-user may not even know—and should not care—whether the response to a request is generated locally by the machine sitting on the user's desk, remotely by a special-purpose processor, or by some combination of the two. The importance of the trend toward end-user computing for researchers lies not in the details of its implementation but rather in its potential to transform scholarly communication by making computation an integral part of the researcher's thought process and work style.

Connectivity

The trend toward end-user computing is intimately related to the equally important trend of *connectivity.* This term describes the researcher's ability to access data, processing capabilities, and other researchers electronically in ways that facilitate the research process. Connectivity is a broader

---

when it is connected to a power distribution grid. The term is used in information science to refer to information and other resources being electronically accessible to users by means of computers and communication devices. Similarly, it refers to users being able to access their work resources electronically, i.e., having terminals, communication facilities, computer accounts, etc., as needed to work in this way. (Information that is not accessible in this way, or users who do not have access to their work in this way are referred to as being "offline.") The term *online* as used in the database and library domains is derivative and analogous but considerably narrower. It is used here in its more general sense.

concept than communication. Like communication, connectivity includes the ability of computers to talk to each other and to access remote databases, but it also includes the ability of researchers to work together in useful ways, to solicit feedback from each other, to disseminate their ideas and results, and to integrate their research sources and products. Connectivity requires communication, but it further assumes that information is in a usable form that facilitates interchange and integration.

Many aspects of end-user computing rely on connectivity. The online transition requires that a sufficient critical mass of the user's context be available online. That is, the various categories of data that comprise this context (such as telephone messages, e-mail, memos, and documents) must all be accessible electronically and must be stored in a common, interchangeable form, so that data can be shared and exchanged among these different categories. Conventional wisdom recognizes that a critical mass of users must be online before they will truly benefit from their connectivity, but it is at least as important that a critical mass of information and tools be online if users are to reap the benefits of connectivity. Furthermore, convenient and effective interchange must be available across this critical mass of information and tools before a user can profitably make the online transition.

Access to databases also requires connectivity, especially if the user needs to see the most up-to-date version of dynamic data. Access to dynamic data is particularly important for research, where the most recent additions to a database (representing new publications, ideas, data, or research) are often the most valuable, even though they may change only a small fraction of the overall database. If a database is static (i.e., does not change very often), it can be copied onto local systems, either by physically sending disks to different sites or by downloading data over a network (which again requires connectivity). However, if a remote database is dynamic, a user can see the most up-to-date version of the data only by either viewing the updated database over a network (relying on connectivity) or by updating a local copy of the database on demand (again, over a network) and viewing the copy. Access to dynamic data therefore depends on connectivity.

An infrastructure of connectivity allows computation to be performed and data to be stored wherever it is most cost-effective, given that the relative costs of memory, computation, and communication are continually changing. Connectivity allows computation and data to be reallocated from local to remote resources (computers, disks, etc.) as these costs change. This reallocation has traditionally required physical changes to system configurations (such as moving disk drives or rewiring buildings with cables), but in principle this can be done without physical intervention, responding automatically to changing costs or shifting demands. Connectivity therefore facilitates end-user computing by allowing it to take advantage of evolving cost factors.

The trend toward ubiquitous computing—whether provided by portable computers, publicly available terminals, or other alternatives—relies on a similar form of connectivity to link users to their working "office" contexts by remote or portable access. Ultimately, it will become irrelevant whether a user's working context exists in a single place or is distributed over a number of sites and machines. Connectivity will allow users to access their computational and informational contexts wherever and whenever they need them.

Access to computational and human resources. Although access to data and one's working context is the most obvious aspect of connectivity, it has other implications as well. In general, connectivity allows users to access resources. These may be data resources, but they may also be specialized computational or human resources. Two related initiatives intended to encourage such interactions by providing widely available, high-capacity networking are the National Research and Education Network (NREN) and the High Performance Computing and Communications (HPCC) efforts. The capacity of a network is measured by its bandwidth, which is the number of bits of information it can transmit per second.[31] The NREN and HPCC efforts are targeted to produce gigabit (billion-bit per second) transmission capacities during the next decade.[32] In addition to providing high-capacity "backbone" communications, related initiatives include efforts aimed at integrating the communication of text, images, voice, video, and other media. The NREN is intended to support the transmission of other media as well as text, although it should be noted that nontextual media require much greater transmission capacity. When fully implemented, NREN should greatly facilitate collaboration and resource sharing among researchers.

Efforts such as NREN also are important because, despite the evolution toward cheaper computers, there may always be state-of-the-art computing facilities that remain too costly for individual researchers to own. For example, large parallel computers may allow searching through huge databases for complex patterns, but the most powerful of such machines may always be too expensive for any one researcher or even any one research facility to justify their purchase. Connectivity will allow researchers to share such facilities through remote access.

Beyond access to machines, connectivity allows researchers to communicate and collaborate with each other and with specialists in other fields. The vast web of interconnected networks (sometimes referred to informally as "WorldNet") already allows researchers to broadcast or direct queries and requests by e-mail to a large proportion of the researchers in a given field, regardless of their nationality or location. This process is not always directly controlled by the initiator of a request: Queries may be forwarded by their initial recipients across networks and gateways between networks to individuals, electronic mailing lists, and electronic bulletin boards,[33] eliciting responses from distant and unlikely places. Integrated networking is greatly facilitated by an open systems approach, allowing multivendor software and hardware to communicate using standard protocols. The International Standards Organization's Open Systems Interconnection (OSI) reference model serves as a standard for interconnection of this kind.[34] These developments are producing a truly global communication capability, which is expanding rapidly and spontaneously.

The communication aspect of connectivity goes beyond the use of e-mail for asking questions or broadcasting general information. It is causing a major shift in the way many researchers collaborate and interact.[35] The use of e-mail allows arbitrary

---

[31] An average page of text consists of approximately 20,000 bits, although this volume can be reduced (compressed) for transmittal.

[32] *Grand Challenges*, 17–19, 54.

[33] Electronic bulletin boards are analogues of their physical counterparts. They allow online users to remotely view notices posted electronically by other users.

[34] The OSI reference model is discussed in detail in A. S. Tanenbaum, *Computer Networks*, 2d ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1988), 14–34.

[35] We are unaware of any research on e-mail use among scholars, but for recent studies on the use of e-mail and other collaborative electronic media in international organizations, see T. K. Bikson and S. A. Law, *Electronic Mail Use at the Bank: A Survey and Recommendations* (Washington, D.C.: Information, Technology, and Facilities Department, World Bank, September 1991); and Tora K. Bikson and Sally Ann Law, "Electronic Information Media and Records Management Methods: A Survey of Practices in United Nations Organizations," *ACCIS Electronic Information Media and Records Management Survey Report*, A RAND Note (N-3453-RC) (Santa Monica, Calif.: RAND Corporation, 1991).

text and data files to be transmitted in simple, linear text formats, without concern for machine compatibility or knowledge of remote file systems. Researchers can generally transform any relevant information into text and send it as the body of a message. Transforming formatted information (such as structured documents or page layouts) into linear text so that it can be exchanged in this way requires that the sender and recipient have software capable of performing the appropriate transformations. Standards for transforming such information into linear text are evolving in response to this need. For example, the Standard Generalized Markup Language offers a standard textual representation for structured documents, whereas PostScript[36] offers a widely used *de facto* standard textual representation for formatted page images. Such standards already allow users to send textually encoded documents, pictures, or formatted page layouts by e-mail instead of on paper. The e-mail recipient can view or print the transmitted information after transforming it back to its original form. This capability will continue to improve as standards for graphics and other media evolve.

Connectivity also promises to "erase the geography" that separates students from teachers, classes, or other resources of interest. The educational notion of "distance learning" has evolved from the correspondence course to the use of televised instruction, but networking allows a much richer form of educational interaction. Particularly in upper-level scholarly subjects, it is now possible to envision geographically distributed seminars that bring together interested scholars and students without regard to their physical locations.

The use of e-mail, teleconferencing, and remote windowing is producing a new phenomenon: computer-supported cooperative work (CSCW).[37] Through CSCW, groups of researchers can work together, sharing their context and coordinating their work, regardless of their locations, schedules, and work styles. Connectivity allows cooperation in all phases of research, including concept formation, literature and background search, analysis, publication, peer review, and dissemination. This trend has the potential to both reduce the time required to perform and publish research and improve its quality through earlier and wider review. CSCW also facilitates interdisciplinary research through online discussion forums that are open to all interested parties, not just credentialed members of a particular discipline. This openness makes it easier for researchers from different fields and institutions to collaborate, which may broaden the perspective of scholarly communication. Finally, the trend toward sharing the research process may well change the conception of the research product itself into something more multidimensional than a traditional document, allowing it to reflect multiple views and opinions. (See the section on hypertext and hypermedia later in this paper.) Note that the implications explored here are not derived from technological determinism: The technology itself does not produce such changes. Rather, the changes result from the trend toward sharing and collaborating, which the technology facilitates.

The trend toward interchange standards. True connectivity involves the ability to interchange information, which requires that information be represented in a standard form. The relative youth of information science as a field and the rapid evolution of computers and communication

---

[36] Adobe Systems, Inc., *PostScript Language Reference Manual* (Reading, Mass.: Addison-Wesley, 1990).

[37] For an excellent annotated bibliography of current work in CSCW, see Saul Greenberg, "An Annotated Bibliography of Computer Supported Cooperative Work," *SIGCHI Bulletin* 23 (July 1991): 29–62.

technology have produced chaotic alternatives for representing and communicating information. This may be unavoidable in a field in which technology and paradigms are still evolving. By their very nature, novel ideas do not always fit into previous patterns. Similarly, new computational capabilities often produce new information structures that do not easily translate into existing standard forms. Furthermore, the development of new standards is a slow process because it requires compromise and consensus. The development of standards is therefore a difficult undertaking, and they tend to lag behind the latest technological advances. Nevertheless, the growing emphasis on interchange standards is a vital and worthy trend, without which the promise of connectivity cannot be realized.

Standards are beginning to evolve for text (as discussed in the section on Computer-Assisted Analysis Achieved Through Conversion), and ultimately they will extend to graphics, voice, three-dimensional modeling, animation, video, and other media as well. In the early stages of this process, the goal is to develop usable initial standards quickly, without precluding their extension and modification in the future. This trend toward extensible standards is motivated by a recognition of the inevitable lag between standards and technological advance. Developing such extensible standards is a major technical challenge, involving a significant effort to translate among different standards and different versions of evolving standards. Ideally, such translation will minimize the need for the user to be aware of the underlying standards, and inexpensive computation will provide transparent translation among standards without user intervention.

In addition to interchange standards, a trend is developing toward defining standards and policies for privacy and authorization of access. As collaboration becomes more common, it will become increasingly important for researchers to be able to protect their data, analysis, and results. Plagiarism, theft, tampering, and sabotage will undermine the advances of connectivity if technical, administrative, and legal solutions to these problems are not implemented. Even the computation and collaboration processes themselves must be protected from unauthorized auditing and analysis. Various agencies or individuals could easily misuse or abuse knowledge of the kinds of questions a researcher asks and the thought processes involved in formulating research. The trend toward increasing interest in privacy and security issues is evidenced in a number of recent conferences and publications.[38]

A false dichotomy: distributed versus centralized control. One of the most intriguing implications of the trend toward connectivity is its potential to redefine the meaning of control over intellectual artifacts. In particular, the traditional dichotomy between distributed and centralized control may no longer be appropriate. This dichotomy is based on the natural but outdated assumption that control is a function of location in the physical world. Traditionally, a resource has been considered to be under centralized control if it exists in only one physical location and is maintained by agents residing at that location. Conversely, a resource is considered to be under distributed (decentralized) control if it consists of multiple copies or parts that are dispersed among multiple locations and

---

[38] Computers, Freedom and Privacy Conference, sponsored by Computer Professionals for Social Responsibility, San Francisco Marriott, Burlingame, Calif. 25–28 March 1991; The National Conference on Computing and Values (NCCV), held at Research Center on Computing and Society, Southern Connecticut State University, New Haven, Conn. 12–16 August 1991; and the seventh Annual Computer Security Applications Conference, sponsored by Aerospace Computer Security Associates and American Society for Industrial Security, and the Association for Computer Machinery and the Institute of Electrical and Electronics Engineers, St. Anthony's Hotel, San Antonio, Tex., 2–6 December 1991.

maintained by agents dispersed among those locations. This dichotomy applies reasonably well to physical resources, but it fails to work for resources created by electronic connectivity.

The physical location of a resource has little meaning in the electronic domain. Connectivity allows resources to be replicated and distributed among numerous physical locations while behaving as though they existed in only one location (and vice versa). The key to this phenomenon is the separation between an electronic resource's physical location and its availability: A database may reside on a storage device in one location while being viewed or modified via a terminal in another location. Similarly, a database that appears to exist in only one location may actually consist of pieces distributed and replicated among numerous locations and may be viewed or modified by numerous agents via computers at different locations. This characteristic is the definition of connectivity: Access becomes independent of location. The notions of centralized and decentralized (distributed) control simply do not apply in this context. New forms of control—and policies for when to employ them—are likely to evolve as connectivity replaces physical access to resources.

Summary. End-user computing and connectivity have been discussed separately here for expository reasons, but their full impact lies in their mutual synergy. Connectivity elevates end-user computing above simple word processing or calculation by allowing end-users to access remote databases, share information in many different media and forms, connect to their working contexts wherever they are, communicate with their peers, and collaborate in all phases of research. End-user computing in turn provides one of the main motivations for improving connectivity: Networks do not connect machines, they connect people. The combined trends of end-user computing and increasing connectivity will shape the evolution of research (along with many other endeavors) well into the next century.

Specific Technology Trends Affecting Scholarly Communication

The major trends of end-user computing and connectivity will manifest themselves in many ways. This section identifies a number of specific technology trends that will superimpose themselves over this background. Each subsection discusses an area of technology that is expected to have a particular impact on research. Although not exhaustive, this examination includes some of the technology that are likely to have the greatest influence over the next decade, i.e., artificial intelligence, end-user publication and distribution, hypermedia, visualization, and virtual reality.

Artificial intelligence. Current trends in artificial intelligence (AI) have the potential to affect scholarly research in a number of ways. AI may provide intelligent aids for analyzing and interpreting sources; automated "agents" that can help researchers stay abreast of new findings; and tools to help formulate research concepts. AI may also enable researchers to model their subject areas to test hypotheses. Finally, AI has the capacity to produce intelligent tutors that may help researchers leverage their teaching skills.

The recent commercial success of expert systems (and more generally, knowledge-based systems) has brought AI out of the ivory tower where it had evolved since the early days of computing. A number of general-purpose programming languages and environments (expert system shells) for building expert systems have appeared on the market, allowing users with little or no formal training in AI to take advantage of some of the most common AI techniques. Yet AI encompasses much more than just expert or knowledge-based systems. As one of the frontiers of computing, it attempts to find ways of using computers to solve problems they cannot now solve. AI is driven by dual motivations that sometimes conflict with and sometimes enhance each other. The first of these, which can be thought of as a "modeling" motivation, seeks to use computers to model and understand intelligence. The second, which can be thought of as an "engineering" motivation, simply seeks to solve difficult problems, by whatever means. AI efforts that are motivated by modeling tend to focus on defining intelligence, understanding cognitive processes, and addressing problems whose solutions are acknowledged to require intelligence. AI efforts motivated by engineering simply try to solve difficult, worthwhile problems, using any available techniques, regardless of whether the techniques simulate human intelligence.

Because of these dual motivations and because AI is a frontier (and therefore necessarily dynamic and evolving), it tends to include many disparate activities and technology, ranging from the automation of formal mathematical logic to the design of artificial neural networks. Several themes run through AI, such as representing knowledge, language, and meaning and finding relevant patterns or solutions among large, complex sets of alternatives. The primary influences of AI on scholarly communication are likely to be its ability to analyze linguistic and pictorial information, its ability to find patterns, its ability to create automated "agents" that act on a user's behalf, and its ability to model reality and formulate concepts.

The bulk of scholarly data is currently in textual form, and text will undoubtedly continue to be the major target of scholarly research for some time. Other forms of data, such as visual imagery (including drawings, paintings, photographs of sites or artifacts, holograms, and film and video), spoken language, sounds, and music may, however, play greater roles as the technology for their encoding and analysis improves. AI software's growing ability to understand the semantics (and eventually the pragmatics) of language and to analyze relationships and identify patterns will make it an increasingly attractive tool for performing scholarly analysis. In addition, AI has developed a number of techniques for dealing with beliefs and uncertain, contradictory, or hypothetical information, which may help researchers who must often generate hypotheses and rely on contradictory or uncertain conclusions and beliefs in order to find patterns and relationships. Coupled with growing databases of encoded text and fast processing, these techniques will enable researchers to look for new, unexpected patterns across a wide range of subject areas. Similar capabilities eventually will extend to visual imagery and sound, allowing integrated analyses of text, speech, music, and pictorial data. Although it will probably be some time before AI will be capable of truly understanding literary text[39]—and even longer before it will be capable of understanding spoken language or visual imagery—it is already capable of filtering large bodies of text to find literary aspects or relationships that are of particular interest to a researcher. In this role, AI will not replace the analytic insight of the researcher, but it will enhance the researcher's ability to scan large collections of information and find patterns worthy of analysis.

One of the major emphases of AI research has been to develop intelligent agents that can behave autonomously on behalf of their users. Robots (which are still largely experimental) are the most dramatic examples of such agents, but another class of agents is more relevant to scholarly research. These are informational agents, such as literature-search or SDI (selective dis-

---

[39] See Nancy M. Ide and Jean Veronis, "Artificial Intelligence and the Study of Literary Narrative," *Poetics* 19 (1990): 37–63.

semination of information) agents, which can search for information of interest to a researcher, using criteria specified in a form similar to a database query. Such agents ultimately may perform a number of services, such as translating a researcher's query into the form required by particular databases; periodically repeating a query or search; monitoring activity on a network or in a database and alerting the user when "interesting" events occur; soliciting, collecting, and filtering information from many sources; responding to routine requests from other researchers for information or to other correspondence; and coordinating the schedules and activities of a collection of researchers engaged in collaborative effort. Such agents will eventually take over many of the traditional activities of a secretary: They will make up for their relative lack of initiative and creativity by being tireless, dedicated, and inexpensive.[40]

In addition to its role in the analytic phase of research, AI may have an impact on the concept formation that leads to research. In this earliest conceptualization phase, researchers often generate informal hypotheses about a subject area, in an attempt to define interesting research thrusts. A number of tools currently emerging from "knowledge acquisition" efforts in AI have the potential to help identify viable hypotheses and useful concepts. These concept-formation tools help the user form concepts by asking questions that can discriminate between examples and counterexamples of an evolving concept, based on attributes that the user declares as defining the concept. For example, a researcher might attempt to define a concept such as "adolescent imagery" in a body of text in terms of attributes such as age, immaturity, and sexual embarrassment. A concept formation tool might attempt to find examples of such images, asking the user to rate each candidate passage according to each attribute. Based on these ratings, the tool might then show which of these passages appear to be examples of the concept and which ones appear to be counterexamples, thereby helping the user form a consistent and useful definition of the desired concept.

Much of AI research focuses on modeling. In order to act intelligently or solve complex problems, AI systems often create models of reality about which they can reason or which they can manipulate in order to decide how to act in the real world. Traditional simulation and mathematical modeling techniques are severely limited in the types of questions they can answer. Simulation users, for example, typically specify the initial state of a simulated world and then run the simulation to see what happens. This "toy duck" view of modeling ("wind it up and see where it goes") corresponds to asking questions of the form "what if . . . ?" (i.e., what would happen if the world were to proceed from this given initial state?). This ability to ask "what if . . . ?" questions is often touted as the ultimate analytic capability, but many other kinds of questions are at least as important in many situations.[41] These include such questions as: Why did some agent take a particular action? Why did a given event happen? Can a particular event ever happen? Under what conditions will a given event happen? Which events might lead to a particular event? How can a desired result be achieved? Ongoing AI research in this

---

[40] For research on intelligent agents, see Robert E. Kahn and Vinton G. Cerf, *An Open Architecture for a Digital Library System and A Plan for its Development, The Digital Library Project, Volume 1: The World of Knowbots* (Washington, D.C.: Corporation for National Research Initiatives, March 1988).

[41] M. Davis, S. Rosenschein, and N. Shapiro, *Prospects and Problems for a General Modeling Methodology* (Santa Monica, Calif.: The RAND Corporation, N-1801-RC, June 1982); and J. Rothenberg, "The Nature of Modeling," in *Artificial Intelligence, Simulation, and Modeling,* edited by L. Widman, K. Loparo, and N. Nielsen, 75–92 (New York: John Wiley & Sons, August 1989).

area is producing powerful new techniques for modeling intentions, causality, goals, beliefs, and other phenomena to allow answering questions that go beyond "what if . . . ?"[42]

This trend toward model-based systems will provide researchers with techniques for conducting experiments, evaluating hypotheses, and exploring alternative interpretations of reality with minimal cost and risk (since they are carried out within a computer). As a simple example, sociological or cultural models could be built to explore alternative hypotheses about an ancient civilization, using the model to make predictions that can be compared with historical evidence. AI techniques such as these may help researchers conceptualize research as well as perform analyses.

The modeling capabilities of AI are also the key to its use in education. Intelligent tutors are an outgrowth of joint research in education and AI; typically, they involve a model of the subject matter to be taught (a domain model) and a model of the student. The domain model elevates an intelligent tutor above the level of simple programmed instruction because it enables the tutor to answer unanticipated questions about the subject matter. Students can therefore ask a much wider range of questions and pursue many alternative paths of instruction. Similarly, the student model helps the tutor determine which concepts the student is having trouble understanding. This helps the tutor address the student's underlying problem rather than simply repeating new material or backing up blindly to review previous material. Although intelligent tutors are still largely experimental, they appear to hold great promise for improving the educational process, particularly for students who are self-motivated and self-paced. Ultimately, this should allow scholars to leverage their teaching skills by developing tutors that embody their expertise.

In summary, current trends in artificial intelligence may affect scholarly research by

- providing analysis aids that can help find and interpret relevant source data, text, and other media.
- creating informational agents that can perform some of the routine tasks of keeping abreast of new findings, acting as tireless monitors of developments in a field.
- providing tools to help researchers explore, formulate, and refine research concepts and hypotheses.
- enabling researchers to model their subject areas to try out hypotheses and predict where to find confirming (or falsifying) evidence.
- facilitating the development of intelligent tutors that can help researchers disseminate their knowledge and teaching skills to wider audiences.

Since AI is one of the frontiers of information science, it is also not unlikely that additional developments in this field will have unforeseen consequences for the evolution of scholarly research.

End-user publication and distribution. An equally important though less exotic computing trend is the growing ability of end-users to publish and distribute their own work. This is already creating alternatives to traditional publication in scholarly journals, not only reducing the time it takes to publish research but, more importantly, changing the channels of distribution, redefining the review process, and transforming dissemination by means of electronic connectivity.

The most prosaic form of end-user publication is the production of camera-ready printed documents, suitable for publication or reproduction and dissemination without

---

[42] See J. Rothenberg, "Using Causality as the Basis for Dynamic Models," in *Proceedings of the Third International Working Conference on Dynamic Modelling of Information Systems* (DYNMOD-3) (Delft, The Netherlands: Delft University of Technology, 1992), 277–92.

further typesetting or layout work (sometimes referred to as "desktop publishing"). Even this simple modernization of the traditional publication process has profound implications. As with all forms of end-user computing, end-user publication involves the author of a document much more directly in its production. Because of this availability of layout and production tools during the draft phase, a document approaches its final form at an earlier stage of development. For example, figures, footnotes, and final formatting can be incorporated into early drafts, giving reviewers a more readable product and helping to eliminate errors and, in general, to improve the product. Ideally, the author's control over questions of typography, graphics, and layout means that the final document represents a more accurate and integrated reflection of the author's overall intent. The corresponding disadvantage is that authors must learn new publication skills, for which they may have little inclination, patience, or talent. Of course, end-user publication does not preclude the use of secretaries, graphic artists, or publication specialists to reintroduce traditional expertise in the publication process, but this intervention tends to subvert the advantages of end-user publication by slowing the process and reducing the author's control.

Beyond modernizing the traditional publication process, end-user publication allows authors to publish their work electronically, bypassing the production and distribution of paper documents entirely. Electronic documents can easily reproduce most of the desirable attributes of paper, and they provide increased flexibility for correction, revision, access, and dissemination. During the production phase of a document, these features facilitate remote collaboration and early review and they greatly simplify the revision process. End-user publication also facilitates a radically different view of the research process, in which ideas are disseminated for review and feedback in the earliest stages of research, i.e., prior to documenting or even performing the research. (Examples of this are discussed later in this paper.)

Electronic dissemination makes use of increasing connectivity to bypass traditional distribution channels, reduce the cost of reproduction and mailing, and enable recipients of a document to redistribute it by forwarding it in electronic form.[43] The copyright and other legal implications of electronic dissemination are only beginning to be explored. Similarly, direct, online access to the source of a document makes it easier than ever to plagiarize ideas, text, and even complex graphics without leaving any trace. These problems must be addressed by technical, legal, administrative, and, ultimately, cultural policies. Such policies are likely to evolve more slowly than the technology they seek to civilize, leaving a gap between practice and policy for at least the next decade or two; this gap is part of the cost of the technological revolution of scholarly research.

Hypertext and hypermedia. All research studies must explicitly or implicitly address a number of questions that represent different dimensions of inquiry, such as What is the problem? What assumptions were made about the problem? What related research exists? What is original about the study? What methodologies were considered? What approach or method was chosen, and why? What sources and data were used? What analysis was performed? What were the results? How should the results be interpreted? What other interpre-

---

[43] Computers and networks are being used in the commercial sector as well, both to help automate the process of publishing traditional books and journals and to develop novel electronic products. This electronic publishing industry has so far had little impact on end-user publication, but it may be too soon to tell whether this industry will ultimately attempt (or manage) to appropriate and commercialize the new channels of distribution and dissemination that end-users are currently developing for themselves.

tations are worth considering? How do the results and interpretations depend on the researcher's assumptions? What are the implications of the research? It is difficult to answer all such questions without inundating and confusing the reader.

Similarly, presenting complex subject matter to students requires answering analogous questions about context, background history, alternative approaches or formulations, and relationships to other disciplines. Traditional textbooks and other instructional materials seldom address these issues adequately.

Such questions are inherently interrelated and multidimensional. Answering them in a strictly sequential, linear fashion is often constraining and unrevealing. Yet written documents necessarily present their arguments linearly. In addition, an expository sequence that provides insight to one reader or audience may not be enlightening to another. Cross-references, references to other documents, repetition, overviews, and summaries can ameliorate these problems, but only at the cost of redundancy and added work for the reader (flipping pages to find cross-references or consulting other documents). Furthermore, documents, which are inherently static, are hard-pressed to portray processes or other dynamic phenomena. The effectiveness of graphics is similarly limited by the static nature of the printed image. Oral presentations can be less linear than documents, can be tailored to specific audiences, and are better suited to presenting dynamic phenomena, but they are ephemeral and cannot provide the depth of the printed word.

Electronic information technology promises to transcend these limitations by delivering research results in an interactive, electronic form that is nonlinear and multidimensional and that integrates written, spoken, and graphic media in a permanent, dynamic, customizable presentation. The terms hypertext and hypermedia suggest the novel characteristics of this new approach:

1. It provides rich, dynamic linkages among the elements of a presentation. For example, using electronic retrieval and display, a reference from one item of text to another (whether a cross-reference, a bibliographic entry, or a citation in another work) can be viewed instantly in a window without the user's having to turn pages or find another document. Such links can be used to present different dimensions of analysis, alternative sequences of exposition, optional degrees of elaboration or depth, supporting evidence, references, data, or contextual background. The multidimensional nature of such structures is denoted by the prefix "hyper." Authors can use this linking to present different kinds of information or to define alternative paths that generate different presentations or variants from a single master document.
2. Hypermedia combines several media that currently can be presented electronically, such as text, color graphics, and sound (including voice). These can all be linked together as easily as text, producing presentations that combine the features of documents and oral presentations.
3. These media can be presented dynamically. This allows animating graphics, synchronizing voice with animation to describe processes, and controlling the pace of a presentation, as in an oral briefing.
4. This approach is interactive, allowing the reader to control the sequence, speed, depth, and focus of the presentation, within limits set by the author.

The concept of a nonlinear document[44]

---

[44] Although hypertext and hypermedia products are very different from traditional documents, they are generally referred to as "documents" for want of a better word.

can be traced back at least as far as the seminal paper "As We May Think," by Vannevar Bush, in 1945.[45] The electronic implementation of this concept is beginning to transform the traditional notion of a document into a multimedia, nonlinear form of presentation. The publication of research results in hypermedia form may make them more accessible and more captivating, thereby greatly increasing the impact and influence of research, particularly outside the traditional scholarly community. The result may be greater public recognition of policy issues identified by research—such as the need to preserve historic sites or artifacts—in much the same way that popular televised documentaries have increased public awareness of myriad scientific, cultural, and environmental issues. Furthermore, the use of hypermedia may transform the research process itself by providing a natural way to represent and keep track of interrelated facts, references, hypotheses, and arguments, as well as reactions, revisions, and annotations to support collaboration. Finally, hypermedia may transform educational material into a new, multidimensional experience that will capitalize on the exploratory tendencies of scholarly students.

Visualization and virtual reality. Recent trends in visualization and virtual reality have the potential to transform the way scholarly researchers interact with their data and perform their analyses. The world of scientific computing has begun to develop techniques that allow scientists to *visualize* the results of complex computations. Graphic techniques and animation are being used to display complex data in ways that attempt to make significant patterns leap out at the user. Abstract relationships are often easier to grasp if they are translated into graphical presentations, such as false-color maps, cluster plots, or adjacency graphs. These techniques apply equally to any field in which complex data, patterns, and relationships must be understood. Many areas of scholarly communication may profit from this technology by visualizing quantitative or qualitative data to gain insight into its meaning or to present complex results in a perspicuous form.

Though it is typically viewed as a very different trend, the technology of virtual reality is closely related to visualization. A virtual reality is a simulated world created in a computer, using traditional simulation or AI modeling techniques such as those discussed above. The user "enters" a virtual reality by wearing a display helmet or goggles to create the visual illusion of being in the simulated world (e.g., showing different views as the user's head turns). The user interacts with the virtual reality by wearing devices such as instrumented gloves or suits that sense the user's hand or body position, thereby allowing the simulated world to react. The result is something like an intensified video game, in which the user perceives the virtual reality and interacts with it for some purpose.

The power of virtual reality is that it harnesses the user's full sensory and motor capabilities in exploring an abstract world, rather than relying on more limited faculties such as reading and typing. Coupled with modeling and visualization, this has the potential to allow a researcher to interact intimately with a virtual world created out of data or analytic results and to explore this world in a much more direct, experiential way than would be possible by reading numbers or even by viewing a graphical display. In addition to its potential for transforming certain aspects of the analytic process, virtual reality technology might also be of use during concept formation (allowing researchers to explore abstract spaces of concepts, represented as visual worlds) or for bringing the education of scholarly subjects to life (allowing students to ex-

---

[45] Vannevar Bush, "As We May Think," *Atlantic Monthly* 176 (July 1945): 101-08.

perience subject matter as a virtual world). Virtual reality may also be viewed as a logical extension of hypermedia, in which research results may be presented as a virtual world to be explored, rather than as a document to be seen or heard.

Caveats. The trends described herein are not without their dangers. The legal issues surrounding electronic dissemination and connectivity have been pointed out above, as have some of the possible violations of privacy that result from working in an open, networked environment. Every technological advance has its own risk for misuse, whether this risk is legal, ethical, or merely a matter of lost productivity and quality. For example, the indiscriminate use of end-user publication and distribution may bypass carefully established mechanisms for editorial and peer review, leading to a proliferation of low quality, unprofessional publications. Similarly, the use of hypermedia by authors who are not trained in graphic design or media presentation may produce a flood of incoherent research products whose complexity makes them inaccessible to their intended audiences. The naive use of modeling tools, visualization techniques, and virtual reality may seduce researchers into believing results that seem compelling despite the fact that they have not been validated. Researchers and audiences alike may tend to accept conclusions based on state-of-the-art computations, such as AI, with less than the required skepticism, especially if these computations exhibit a veneer of intelligence.

These dangers are real and may well plague scholarly researchers for decades to come, as they adopt new methods empowered by technology. Nevertheless, these trends appear inevitable and are likely to change the form and substance of scholarly communication in fundamental ways. Whether this change will ultimately improve the quality of that research is a verdict that only the future can deliver.

Summary

The availability of quantitative data and numerical techniques for analyzing them have had a marked effect on scholarly communication over the past several decades. The technology trends discussed here, as well as others that may prove to be important, are likely to have an even more profound impact. This impact will do more than simply change the work styles of scholarly researchers: It will affect their thought processes as well, suggesting new kinds of research questions and new kinds of answers. It will change the way researchers collaborate and interact with their peers and the way they produce their results. It will change the form of these results, the way they are distributed and disseminated, their audiences, and the impact they have on the research community and the public. These changes, already under way, will have profound implications for the information services, libraries, and archives that serve the research process.

SCHOLARLY COMMUNICATION AND THE USE OF CURRENT INFORMATION TECHNOLOGY

The previous section explored key trends in information technology most relevant to scholarly communication. This section considers the use of currently available information technology by social science and humanities scholars to advance scholarship and intellectual productivity. The use of technology across the full spectrum of scholarly communication is considered by examining how researchers rely on technology to: (1) identify sources, (2) communicate with colleagues, (3) interpret and analyze data, (4) disseminate research findings, and (5) develop curriculums and aid instruction. Case examples of scholarly practices illustrate broader tendencies within