Research Article

Scholarly Communication and Information Technology: Exploring the Impact of Changes in the Research Process on Archives

AVRA MICHELSON AND JEFF ROTHENBERG

About the authors: *Avra Michelson is a member of the Technology Research Staff at the National Archives and Records Administration. Jeff Rothenberg is a Senior Computer Scientist in the Social Policy Department of The RAND Corporation.*

Avra Michelson would like to express her gratitude to William Holmes and Charles Dollar for the opportunity to do this research. Jeff Rothenberg would like to thank Tora Bikson for establishing the connection with NARA that led to this research and Bob Anderson for insightful comments on the technology issues discussed in the report. The authors also would like to extend a special thanks to Paul Conway for serving as a wise sounding board throughout the research and writing of the report; Anne Kenney, for taking the time to explain and consider strategies for preservation microfilming and digital imaging; Richard Cox for his instructive readings of many drafts; Tora Bikson for her thoughtful review of the final version; and Lida Churchville for extraordinary assistance with database searching. The recommendations that accompany the report were written as a product of the authors' participation in the 1991 Research Fellowship Program for the Study of Modern Archives administered by the Bentley Historical Library, University of Michigan, and funded by the Andrew W. Mellon Foundation and the University of Michigan. The authors are deeply grateful to David Bearman, Richard Cox, Charles Dollar, Anne Kenney, Paul Evan Peters, Michael Sperberg-McQueen, Joan Warnow-Blewett, and Jerome Yavarkovsky for their review of drafts of the report, including an early version circulated prior to the Bentley symposium, and for their assistance in developing policy recommendations for the report. Also, the authors would like to acknowledge the valuable comments received from many people mentioned in the report who aided the research or were kind enough to review sections.

Contents

EXECUTIVE SUMMARY

The emergence and use of information technology is this century's most significant development affecting archival practice. In response to this development, members of the archival profession have explored both the ways in which new technology can improve the management of archives, and the most appropriate methods for managing electronic records that result from the use of automated systems. But these two issues only partially address the impact of technology on archives. A third but indirect influence also deserves examination: technology's impact on scholarly research methods, which has consequences for the use and management of archives. This article considers the policy implications for archives of trends resulting from the infusion of information technology into the scholarly research process.

The article considers the interaction of two distinct kinds of trends: trends in information technology and trends in research practices, particularly among social scientists and humanists. Although much of the rapid growth and evolution of information technology may be unrelated to scholarly research, and aspects of scholarly research may be evolving in ways that have little connection with information technology, there is nevertheless a strong and important interaction occurring between these two evolutions. The possibilities created by new technology are prompting transformations in scholarly practice, and these transformations are in turn stimulating new needs among researchers and are inspiring further technological breakthroughs. Understanding the nature of this interaction is necessary for forecasting the most likely ways in which new scholarly methods will demand innovative services and responses from the archival community.

This article explores two fundamental trends in information technology affecting scholarship: end-user computing and connectivity. Several other technologies of relevance to scholarship are also considered, including artificial intelligence, end-user publication and distribution, hypertext and hypermedia, and visualization and virtual reality. Changes in the research process resulting from scholarly use of information technology are considered within the broad framework of scholarly communication. The scholar's use of currently available technology to search for sources, communicate with colleagues, interpret and analyze source materials, disseminate research findings, and prepare curriculums and instructional applications is examined. Our key finding is the exploding use among researchers of information technology on research and education networks to advance scholarship. Far from being visionary, this future is already present: It is currently being experienced by significant and increasing numbers of scholars from many disciplines. The library profession is responding to the emergence of network-mediated scholarship by promoting global connectivity, performing conversions of print sources to machine-readable form, undertaking the software engineering of full-text delivery systems for online materials, and collaborating with technologists in the use of computing and communication technology to meet specialized researcher needs.

The report recommends that the archival profession:

1. establish a presence on the Internet/NREN.
2. make source materials available for research use over the Internet.
3. create documentation strategies to document network-mediated scholarship and the development of research and education networks.
4. develop archival methods suitable for operation with NREN.
5. take user methods and future computational capacity into account in establishing policies on the management of software-dependent records.
6. recognize and reward initiatives that (a) advance archival management of electronic records, (b) respond to scholarly use of information technology, or (c) promote a network-mediated archival practice.

This article is the result of nearly two years of collaboration between Avra Michelson and Jeff Rothenberg. Earlier versions or derivative presentations of the article were reported at annual meetings of the Society of American Archivists, National Association of Government Archivists and Records Administrators (NAGARA), National Net '92, and the Library of Congress Workshop on Electronic Texts. The article is available electronically on the file server operated by the Coalition for Networked Information. (Contact craig@chi.org for instructions.)

INTRODUCTION

The emergence and use of information technology is this century's most significant development affecting archival practice.[1] In response to this development,

[1]The term *archives* refers broadly to historic sources of enduring value that document the activities of governments, organizations, or individuals; it also refers to the repositories responsible for preserving and making available the historic record.

members of the archival profession have explored both the ways in which new technology can improve the management of archives[2] and the most appropriate methods for managing electronic records that result from the use of automated systems.[3] But these two issues address only a portion of the impact of technology on archives. A third though indirect influence that deserves examination is technology's impact on scholarly research methods, which has consequences for the use and management of archives. This article considers the policy implications for archives of trends resulting from the infusion of information technology into the scholarly research process.

The term *information technology* refers to the computing and communications technology used to obtain, store, organize, manipulate, and exchange information. The definition includes computer hardware and software, as well as the telecommunications devices and computer-based networks that connect them.[4] The influence of information technology on the research process, already evident, promises to deeply penetrate scholarly practice as we enter the twenty-first century. This technology is enabling academics to change significantly the way they communicate and collaborate, identify and analyze sources, store and retrieve data, and disseminate the products of their research. Although technology affects the research process across a spectrum of disciplines and professions, this article focuses on changes in the social sciences

[2]See Marion Matters, ed., *Automated Records and Techniques in Archives: A Resource Directory* (Chicago: Society of American Archivists, 1990), 12–37, for a bibliography on the topic. A selection of the seminal literature includes: Thomas H. Hickerson, *Archives and Manuscripts: An Introduction to Automated Access,* SAA Basic Manual Series (Chicago: Society of American Archivists, 1981); Richard H. Lytle, "An Analysis of the Work of the National Information Systems Task Force," *American Archivist* 47 (Fall 1984): 357–65 (see also other articles in this issue of *AA*, especially Thomas E. Brown, "The Society of American Archivists Confronts the Computer," pp. 366–82); David Bearman, *Towards National Information Systems for Archives and Manuscript Repositories: The NISTF Papers* (Chicago: Society of American Archivists, 1987), as well as Bearman's *Archives and Museum Informatics* technical reports and quarterly newsletter; and two special issues of the *American Archivist* devoted to "Standards for Archival Description" (Fall 1989 and Winter 1990). More recently, archivists have begun to explore the use of specific technologies to support archival functions. See, for instance: *Optical Digital Image Storage System: Project Report* (Washington, D.C.: National Archives and Records Administration, Archival Research and Evaluation Staff, March 1991); Avra Michelson, *Expert Systems Technology and Its Implications for Archives,* National Archives Technical Information Paper no. 9 (Washington, D.C.: National Archives and Records Administration, Archival Research and Evaluation Staff, March 1991); and Anne R. Kenney and Lynne K. Personius, "The Future of Digital Preservation," *Advances in Preservation and Access*, vol. 1 (Westport, Conn.: Meckler Press, forthcoming).

[3]Charles Dollar identifies salient literature on this topic in his work, *The Impact of Information Technologies on Archival Principles and Methods* (Macerata, Italy: University of Macerata Press, 1992). A selection of seminal publications includes: Charles Dollar, "Appraising Machine-Readable Records," in *A Modern Archives Reader: Basic Readings on Archival Theory and Practice*, edited by Maygene F. Daniels and Timothy Walch (Washington, D.C.: National Archives and Records Service, 1984); Margaret L. Hedstrom, *Archives and Manuscripts: Machine-Readable Records*, SAA Basic Manual Series (Chicago: Society of American Archivists, 1984); Harold Naugler, *The Archival Appraisal of Machine-Readable Records: A RAMP Study with Guidelines* (Paris: General Information Programme and UNISIST, UNESCO, 1984); United Nations, Administrative Committee for the Coordination of Information Systems, Technical Panel on Records Management, *Electronic Records Guidelines: A Manual for Policy Development* (New York: United Nations, 1989); and *Research Issues in Electronic Records.* (St. Paul, Minn.: Published for the National Historical Publications and Records Commission, Washington, D.C., by the Minnesota Historical Society, 1991). See also Tom Ruller, "Managing and Appraising GIS Data: Issues and Strategies," unpublished paper presented at the 1991 annual meeting of the Society of American Archivists, Philadelphia.

[4]John R. B. Clement, "Increasing Research Productivity Through Information Technology: A User-Centered Viewpoint," unpublished paper, 19 October 1989, p. 3.

and humanities because scholarly patrons of archives tend to be drawn most heavily from these fields.[5]

Undertaking the research for this report was motivated, in part, by efforts in the archival profession to provide answers to questions related to the use of source materials, such as the following: Who are the (potential) users of primary sources? What are the characteristics of the modern research process? How do patrons frame research questions?[6] In the past few years, several empirical studies have been conducted on patterns of research use within or across repositories or specific disciplines.[7] Although these studies provide valuable insights on users and patterns of use for the period of study, they typically fail to consider their findings within the context of a broader analysis of scholarly research trends.

Archivists need more than snapshots as a basis for policy formulation. An accurate depiction of current research practices is necessary, but archival strategic planning must also involve an analysis of significant trends. This article addresses the interaction of two distinct sets of trends. Electronic information technology as a phenomenon is experiencing rapid growth and evolution, much of which may be unrelated to scholarly research. At the same time, aspects of scholarly research may be evolving in ways that have little connection with information technology. Nevertheless, a strong and important interaction is occurring between these two movements. The possibilities created by new technology are prompting transformations in scholarly practice, and these transformations are in turn stimulating new needs among researchers and further inspiring technological breakthroughs. Understanding the nature of this interaction is necessary for forecasting the most likely ways in which new scholarly methods will demand innovative services and responses from the archival community.

Trends analysis is inherently somewhat circular, since technological changes "drive" changes in scholarly practice only to the extent that the new technology provides capabilities that scholarly researchers can use in meaningful and productive ways. It involves more than the description of arbitrary technological trends: Their relevance must be derived from the perspective of scholarly research. It also involves more than the description of current trends in scholarship: To the extent that scholarship uses information technology, it is necessarily constrained by what is currently possible. Only by considering the joint evolution of technology and scholarly methods can a

[5]The terms *scholar* and *researcher* generally are used throughout the paper to refer to social scientists and humanists unless specified otherwise. Nevertheless, we believe that the research trends identified in this report apply to a broader range of the research community.

[6]A selection of key literature that has advanced the archival profession's conceptual framework includes: Mary Jo Pugh, "The Illusion of Omniscience: Subject Access and the Reference Archivist," *American Archivist* 45 (Winter 1982): 33–44; Elsie T. Freeman, "In the Eye of the Beholder: Archives Administration from the User's Point of View," *American Archivist* 47 (Spring 1984): 111–23; Paul Conway, "Facts and Frameworks: An Approach to Studying the Users of Archives," *American Archivist* 49 (Fall 1986): 393–407; and Lawrence Dowler, "Availability and Use of Records: A Research Agenda," *American Archivist* 51 (Winter/Spring 1988): 74–86.

[7]A selection of the key studies includes: *Major Findings, Conclusions and Recommendations of the Researcher and Public Service Component Evaluation Study* (Ottawa: Public Archives of Canada, 1985); Paul Conway, "Research in Presidential Libraries: A User Study," *Midwestern Archivist* 11 (1986): 35–56; William J. Maher, "The Use of User Studies," *Midwestern Archivist* 11 (1986): 15–26; David Bearman, "User Presentation Language in Archives, *Archives and Museum Informatics* 3 (Winter 1989–90): 3–7; Paul Conway, *Partners in Research: Towards Enhanced Access to the Nation's Archives* (Washington, D.C.: National Archives and Records Administration, forthcoming); and Ann D. Gordon, *Using the Nation's Documentary Heritage: The Report of the Historical Documents Study*, supported by the National Historical Publications and Records Commission in cooperation with the American Council of Learned Societies (Washington, D.C.: National Historical Publications and Records Commission, 1992).

convincing picture of the future be constructed. The remainder of this article attempts to create such a picture in order to examine its implications for archives during this decade and beyond the turn of the millennium.[8]

This article presents a conceptual framework for understanding long-term trends relevant to the scholarly research process. The topic is introduced by a discussion of scholarly communication and the early use of computers among academics. An analysis of information technology trends most pertinent to the conduct of research follows. The third section explores, through case examples, trends in the use of currently available information technology by social scientists and humanists. The fourth section considers model efforts by those in the library profession to respond to changes in the research process. The article concludes with policy recommendations that address key changes needed in archival practices and methods to respond to transformations in scholarly research methods, and the growing prominence of a new electronic communication medium—research and education networks.

BACKGROUND

Scholarly inquiry represents a timeless human quest to understand the world around us. Although this quest for understanding is a sustaining element of human culture, the techniques of the scholar have changed over time. No longer characterized by oral tradition and forum dialogues, the modern research process is commonly understood to entail five processes: (1) identification of sources, (2) communication with colleagues, (3) interpretation and analysis of data, (4) dissemination of research findings, and (5) curriculum development and instruction for preparing the next generation of scholars. Refinement of the scholar's original idea or hypothesis occurs throughout these more tangible processes. The impact of information technology on these processes is resulting in unprecedented transformations in scholarly communication.

Scholarly communication is the term used to refer to the interrelationship of the five processes of modern scholarship.[9] The term implies both a dynamic exchange of information and ideas and an interdependence among publishers, librarians and others in the support of scholarship and the advancement of knowledge. Scholarly communication is generally understood to involve the social exchange of intellectual and creative activity from one scholar to another.[10] As a concept, it denotes a recognition of the mutual reliance of researchers, publishers, professional associations, and libraries and archives in fostering intellectual pur-

[8]Because this paper examines the interaction of two distinct trends, differing frameworks are used to organize the key sections (Overview of Information Technology Trends and Scholarly Communication and the Use of Current Information Technology). The former uses information technology trends as the organizing framework, whereas the latter uses the elements of scholarly communication as a structuring framework. The relationship between technology and scholarship is both dynamic and complex, and our understanding of it continues to evolve. Although it was suggested to us that the framework used to explore information technology trends should be used as the organizing framework for the section on current scholarly practices as well (e.g., a more technological determinist approach), we consider the dual frameworks, and the analysis of the relationships between them, one of the paper's key virtues.

[9]The American Council of Learned Societies popularized the term *scholarly communication* among academics as a result of their mid-1980s survey on the experience of more than five thousand humanists as authors using scholarly publications, libraries, and computers. The findings of the report appear in Herbert C. Morton and Anne J. Price, *The ACLS Survey of Scholars: Final Report of Views on Publications, Computers, and Libraries* (Washington, D.C.: Office of Scholarly Communication and Technology, American Council of Learned Societies, 1989).

[10]Thomas W. Shaughnessy, "Scholarly Communication: The Need for an Agenda for Action—A Symposium," *Journal of Academic Librarianship* 15 (May 1989): 69.

suits. This interdependence implies that a change in the practice of any one of these agents is capable of inspiring changes in the entire paradigm. In transforming the way in which academics learn of primary source materials, search and gather data, interpret and analyze sources, and report findings to the scholarly community, information technology is influencing significant aspects of scholarly communication. Consequently, changes in scholarly research patterns have ramifications for archives and libraries.[11]

The influence of modern technology on scholarly communication began with the birth of computers. More than forty years ago, the scientific community was the first of the academic disciplines to introduce computers into the research process. As computing power expanded, geographically dispersed scientists began collaborating on research questions requiring computers. In 1969, in response to the needs of this community, the U.S. Defense Department's Advanced Research Projects Agency (ARPA) developed the ARPANET, a telecommunications network designed to allow the sharing of expensive computer resources among government and academic research laboratories.[12] Scientific computing has evolved to include the use of electronic networks for electronic mail (e-mail) and for access to supercomputing processing power and to software that facilitates group work.[13]

Since the 1970s, a large and complex array of networks has emerged to support collaborative scientific research. As the scientific need for connectivity increased, network infrastructures at institutions, organizations, commercial enterprises and regions expanded. Today, more than three thousand regional, federal, commercial, and organizational networks connect an estimated 5 million scholars in seventy countries.[14] The Internet, the existing network of research and education networks, comprises thousands of trunk lines that currently carry from 1.5 to 45 million bits per second.[15] The National Research and Education Network (NREN), authorized in 1991 and due to be operational by 1995, will be capable of transmitting 1 billion bits of data—the equivalent of fifty thousand typewritten pages—every second.[16]

In recent years, the global expansion of electronic networks has allowed for worldwide collaboration among scientists. Further, the connectivity provided by greater bandwidth lets scientists process previously unimaginable amounts of data. Expanding the volume of data able to travel across networks permits scientists to explore new types of questions because greater amounts of data are available with less time required for analysis. Equally important, the prom-

[11]For a historical consideration of the relationship between scholarly communication and libraries, see Phyllis Dain and John C. Cole, eds., *Libraries and Scholarly Communication in the United States: The Historical Dimension,* Beta Phi Mu Monograph, no. 2 (New York: Greenwood Press, 1990).

[12]Clifford A. Lynch and Cecilia M. Preston, "Evolution of Networked Information Resources," *Proceedings of the Twelfth National Online Meeting May 7–9, 1991 N.Y., N.Y.,* Martha E. Williams, ed. (Medford, N.J.: Learned Information, 1991): 221–30.

[13]See, for instance, a recent book of readings, edited by Irene Greif, *Computer-Supported Cooperative Work* (San Mateo, Calif.: Morgan Kaufmann, 1988); for an assessment of the information needs of the scientific scholar, see *Communications in Support of Science and Engineering: A Report to the National Science Foundation from the Council on Library Resources* (Washington, D.C.: The Council, August 1990); for a discussion of state-of-the-art collaboration-oriented software, see Daniel Williams, "New Technologies for Coordinating Work," *Datamation* 36 (15 May 1990): 92–96.

[14]Clifford Lynch, "Telecommunications and Networking: A Tutorial," presentation made at the American Society for Information Science 54th Annual Meeting, Washington, D.C. (29 October 1991).

[15]Lynch and Preston, "Evolution of Networked Information Resources."

[16]From a presentation by Paul Peters, executive director of the Coalition for Networked Information, to the National Archives and Records Administration on 7 May 1991; see also Ralph Alberico, "The Development of an 'Information Superhighway'," *Computers in Libraries* 10 (January 1990): 34.

ise of increased computing power and advances in telecommunications will allow scientists to expand the graphical display of research results, alleviating many problems associated with interpreting very large data sets.[17] The trends characteristic of modern scientific inquiry—greater collaboration, increased use of computer-assisted analysis of machine-readable sources, and expanded use of global research and education networks—increasingly represent trends in the social sciences and humanities as well.

In the humanities, scholars initially used computers simply to store and retrieve data. In what is commonly believed to be the earliest project of its kind, Father Roberto Busa in 1949 began his effort to compile an index and concordance to the work of St. Thomas Aquinas.[18] But apart from the hard sciences, the field of political science is typically regarded as the discipline most responsible for transforming computer processing into an accepted scholarly method. What began as a simple use of computers by political scientists for processing survey data and analyzing national opinion polls became a standard social science methodology: quantitative analysis. During the past four decades, following the lead of survey researchers, a range of scholars within academic disciplines began to use computer technology to process large sets of numeric data.[19]

The advance of information technology over the past several decades has astonished even the most visionary technologists. Although certain predictions have proved too optimistic, the overall rate of advance has matched or surpassed the prophesies of most experts, and it shows every sign of continuing unabated during the next few decades. Indeed, from 1980 to 1985, the period that marked the birth of personal computers, their use among scholars soared from nonexistent to more than 50 percent.[20] Today, the scholarly use of personal computers extends beyond storage and retrieval of data and includes text editing, formatting, and text analysis. Increasingly scholars are turning to technology to do statistical analysis, create databases, produce spreadsheets, and compile graphical images of data. Many scholars consider technology an essential instructional tool for generating simulations, capturing data, and

[17]Clement, "Increasing Research Productivity," 3; for a discussion of the role of imagery in human understanding, see Mary Alice White, "Imagery in Multimedia," *Multimedia Review* (Fall 1990): 5–8.

[18]See David S. Miall, ed., *Humanities and the Computer: New Directions* (Oxford: Clarendon Press, 1990), 2.

[19]As their numbers grew, quantitative scholars successfully campaigned for the establishment of data archives, special repositories designed to preserve and provide access to machine-readable collections of survey, census, polling, and legislative data. See Kathleen M. Heim, "Social Scientific Needs for Numeric Data: The Evolution of the International Data Archive Infrastructure," *Collection Management* 9 (Spring 1987): 1–53.

[20]Morton and Price, *ACLS Survey of Scholars*, 33. The ACLS study represents the only currently available direct survey of scholars on their use of computers. But the survey polled only scholars who are members of professional associations. For the past few years, EDUCOM and the University of Southern California have conducted an annual survey of academic computing directors on campus planning, policies, and procedures affecting the use of desktop computers. According to reports by academic computing centers, 39.5% of faculty at two year public and four year public and private colleges and universities have access to or own computers. This figure, however, is considered unreliable, as it is based on estimates by academic computing staff, rather than on direct counts. Furthermore, no one believes that actual usage has dropped from 1985 to 1991, as implied by the discrepancy between the ACLS and EDUCOM/USC figures. According to Kenneth C. Green, the EDUCOM/USC survey developer and author of the report on the findings, "our limited knowledge about student and faculty access to and use of technology is appalling." Green argues that a direct survey of scholars is needed to identify actual computer usage. See USC Center for Scholarly Technology *Newsletter*, "Despite Budget Cuts, Campuses Attempt to Maintain Computing Services," (October 1991); Kenneth C. Green, "A Technology Agenda for the 1990s," *Change* 23 (January/February 1991): 6–7; and Kenneth C. Green and Skip Eastman, *Campus Computing 1990* (Los Angeles: University of Southern California, Center for Scholarly Technology, 1990).

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.

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.

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 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 the field.[46] For analytical purposes, the implications of these practices for archival administration are discussed in the final section (Conclusion and Recommendations) of this report. Although the discussion focuses primarily on practices in the social sciences and humanities, the emerging patterns exhibited in these professions mirror those found in a broad range of disciplines and occupations.[47]

The old assumptions commonly shared by archivists and librarians about the research process characterize a decreasing segment of the scholarly community.[48] Instead, a paradigm shift is occurring in the research styles of social scientists and humanists, as in the scientific community, where: electronic communication is gaining prominence; direct online searching is replacing intermediary searching; research collaborations are becoming more common; electronic sources available in homes and offices are becoming an alternative to reading room visits; source materials originally created in print are being converted to machine-readable form; standard scholarly research practices are extending to the use of artificial intelligence to interpret and analyze materials; and electronic publishing and nonlinear technology, such as hypermedia, are prompting the development of new forms of scholarly research products. The following explores how scholarly communication practices among social scientists and humanists are changing as a result of the use of currently available information technology.

Identification of Sources

According to the professional literature, the key way scholars learn about relevant research materials is through their colleagues. But in the last few years, word of mouth has been supplemented by new forms of electronic searching through online public access catalogs (OPACs). For instance, most campuses provide academics with direct access via personal computers to the institution's online library catalog.[49] Instead of visiting the library, researchers can now explore descriptions of the library's holdings from their offices. Furthermore, if the institution's catalog proves insufficient, scholars can access more than two hundred major American library catalogs, including those of the universities of California, Michigan, Pennsylvania, and Wisconsin, via the Internet.[50] For a comprehensive search of this nation's library holdings, the Research Libraries In-

[46]The authors do not endorse particular techniques, uses of technology, or the validity of results reported in the case examples. Because empirical data on scholarly use of information technology does not exist, this section relies on case examples intended to be illustrative of broader trends within the social sciences and humanities. Academic computing officers, however, are beginning to recognize the need for such data, and some have expressed interest in conducting campuswide or intercampus surveys on scholarly use of technology.

[47]For a discussion of the impact of information technology on the research process of intelligence analysts see Michael R. Leavitt, *The Analyst and Technology—2000*, prepared for the U.S. Intelligence Research and Development Council (January 1991).

[48]Including such assumptions as: patrons discover source materials essentially through word of mouth and through supplemental assistance by intermediaries; humanities and social science scholars conduct research basically as individuals; primary sources, by nature, require viewing in reading rooms fortified with professional assistance; primary sources are best stored and viewed in their original form or on microfilm; the qualitative methods used in the analysis of sources typically preclude computation; and the standard scholarly products (e.g., publications) are linear documents distributed in print form.

[49]See Clifford A. Lynch, "Library Automation and the National Research Network," *EDUCOM Review* 24 (Fall 1989): 22; and *Communications in Support of Science and Engineering,* 1–7.

[50]Conversation between Avra Michelson and Paul Peters, executive director of the Coalition for Networked Information, 7 May 1991. Clifford Lynch, director of library automation for the University of California, reports that as many as 30 percent of the log-ons to the universitywide MELVYL library catalog are from remote sites.

formation Network (RLIN) is available over the Internet, and plans are under way to make OCLC available on the Internet as well.[51] Humanist scholars report that, by providing a comprehensive means to browse through libraries in their homes and offices when convenient, direct access to bibliographic databases represents a source of intellectual empowerment.[52] The use of online catalogs probably represents the most widespread example of scholarly practices in the social sciences and humanities that involve end-user computing and connectivity.

Communication with Colleagues

The search for sources and the need to refine intellectual ideas motivate academics to communicate with their colleagues. Indeed, communication of this sort is fundamental to the advancement of scholarship. Beyond the most common methods of communication (such as face-to-face discussions, telephone conversations, written correspondence, or public presentations) scholars are using e-mail and a variety of new electronic communication formats derived from it for academic interchange. Scholars naturally still talk to one another, but many information exchanges occur through network communications rather than through oral discourse.[53] E-mail exchanges are growing at an astonishing rate, and currently constitute approximately half the traffic on research and education networks.[54] The global spread of e-mail has been rapid, and it is now possible for American scholars to communicate via e-mail with colleagues in close to 140 other countries. The popularity of e-mail among scholars emphasizes the increasing importance of network connectivity in the daily life of academics.[55]

As an outcome of e-mail, scholars are creating new formats for substantive exchange to supplement conventional communication. For example, nearly thirty thousand public-access electronic bulletin boards are currently available through research and education networks. This is up from fourteen thousand such applications counted one year earlier.[56] BITNET, a network developed during the mid-1980s to provide rapid communication among researchers, educational institutions, and funding agencies, reports more than two thousand listservs. Listservs are discussion groups that allow people with common interests to communicate with one another by sending to a special network address mail that is automatically distributed to each person who has subscribed to a particular list.[57]

[51]See Clifford A. Lynch, "The Growth of Computer Networks: A Status Report," *Bulletin of the American Society for Information Science* 16 (June/July 1990): 10; and Robert Weber, "Libraries Without Walls?" *Publishers Weekly* 237 (8 June 1990): S20–S22.

[52]Stephen Lehmann and Patricia Renfro, "Humanists and Electronic Information Services: Acceptance and Resistance," *College and Research Libraries* 52 (September 1991): 411.

[53]Many argue that although less interactive, for brief exchanges e-mail is a far deeper medium for communication than oral discourse. Unfortunately, there are currently no studies on the nature and extent of the use of e-mail among scholars, but its significance as a new communication medium is indisputable.

[54]Presentation by Paul Peters to a joint meeting of the National Association of Government Archives and Records Administrators Committee on Information Technology, and the SAA Committee on Automated Records and Techniques, Washington, D.C., 22 April 1992.

[55]The EDUCOM/USC Survey of Desktop Computing in Higher Education estimates that 25 percent (extrapolated figure) of faculty at four year public and private universities and colleges use e-mail. See Green and Eastman, *Campus Computing* 1990, 23. See the two studies by Tora K. Bikson and Sally Ann Law cited in the previous section for studies on the use of e-mail within an office environment.

[56]Christopher Lindquist, "Ferret Lovers Unite and Download," *Computerworld* 25 (12 August 1991): 1.

[57]Theodore J. Hull, *NNXA Reference Report,* Center for Electronic Records, National Archives and Records Administration (June 1991 draft), 2; and Eric Thomas, *Revised List Processor* (Listserv@frecp11), Release 1.5d, Ecole Centrale de Paris, from Lis-

Of the thousands of electronic discussion groups, or conferences, operating on the Internet, close to 600 are devoted to scholarly topics in the social sciences and humanities.[58] The rate of growth of these scholarly electronic conferences is astonishing. From 1990 to 1991, 200 new conferences were identified on the Internet. For the eight months prior to March 1992, an additional 150 conferences in the social sciences and humanities were added to the existing directory of listings.[59] Scholars have established conferences in virtually every field within every discipline. For example, there are conferences on topics such as Hellenic culture, folklore, modern British and Irish literature, the Vietnam War, and eighteenth-century world history. There are conferences devoted to the study of countries or regions, such as Peru, Iberia, Latin America, and the Baltic states. There are conferences on the works of single authors, such as James Joyce, John Milton, Thomas Pynchon, and Hegel, and there are conferences devoted to concepts such as libertarianism, intuition in decision making, ethics, and fraud in science.[60]

The *Humanist*, an electronic conference established several years ago serves as a focal point for discussions of humanities computing techniques and research methods. It also broadcasts announcements and includes a column for ongoing queries and responses that cover a broad range of issues of interest to humanist scholars. The *Humanist* is transmitted daily to about two thousand readers, including subscribers in Europe and the Near East.[61] A British counterpart, *Humanities Online Bulletin*, operates as a forum for humanists to exchange experience, solicit advice and information, notify one another of projects, review publications, and make announcements. The almost thirteen hundred registered readers are mostly members of humanities departments in British universities.[62] These electronic discussion groups serve a unique role in scholarly communication in that they permit the rapid interchange of current information, ideas, and perspectives. No other medium has permitted scholars to communicate with an international group of peers quickly and effortlessly at the front end of the research process. (The scholarly implications of the new exchange mediums are examined further in the section Dissemination of Research Findings later in this report.)

Interpretation and Analysis of Sources

The use of information technology to assist in interpreting and analyzing data represents one of the most important paradigm shifts toward end-user computing in scholarly research practices. Scholars are both converting primary textual sources to machine-readable form to allow for conventional computational processing and using artificial intelligence to do new types of machine-assisted interpretation and analy-

tserv@indycms.iupui.edu (13 June 1991 17:55:18), 1.

[58]Diane Kovacs, *Directory of Scholarly Electronic Conferences,* 3rd ed. (Kent State, Ohio: Kent State University Libraries, August 1991), [available on Bitnet/Internet at Listserv@kentvm or FTP from ksuvxa.kent.edu. The directory, an indispensable, growing resource, is also available in print as *Directory of Electronic Journals, Newsletters and Academic Discussion Lists* by Michael Strangelove and Diane Kovacs, edited by Ann Okerson, 2nd ed. (Washington, D.C.: Association of Research Libraries, March 1992). The conference figures cited reflect updated information that Diane Kovacs was kind enough to share with Avra Michelson.

[59]These figures may underrepresent actual scholarly activity, as Kovacs warns that the directory's coverage of Usenet is less than comprehensive.

[60]For a description of these conferences, see the Kovacs directory cited earlier.

[61]Elaine Brennan and Allen Renear are the current co-editors of the *Humanist*. Information on the *Humanist* from a telephone conversation between Avra Michelson and Allen Renear, 17 December 1990, and from the description of the conference that appears in the Kovacs directory.

[62]Brendan Loughridge, "Information Technology, the Humanities and the Library," *Journal of Information Science* 15 (July–September 1989): 280.

sis. The use of computing to perform interpretation and analysis is a developmental trend with broad implications for scholarship. These practices suggest fundamental changes in scholarly methods, and each will be examined in depth.

Computer-assisted analysis achieved through conversion. Social scientists and humanities scholars use both quantitative and qualitative methods to analyze and interpret sources. Typically, the search for and evaluation of evidence involves both types of methods. At one end of the continuum, quantitative analysis involves the use of mathematical processes such as a count of frequencies and distributions of occurrences, or higher level statistical techniques. At the other end of the continuum, qualitative analysis typically involves nonmathematical processes oriented toward language, interpretation, or the building of theory.[63]

Scholarly analysis often involves the processing of large and sometimes massive amounts of textual sources.[64] But researchers have discovered that many of the methods of interpretation and analysis related to both quantitative and qualitative methods are processes that can be performed by computers. For example, computers can count (e.g., they can count words, births, deaths, marriages, commercial activity, and even brush strokes used in a Rembrandt painting). Computers can perform regression analysis to suggest cause and effect relationships. Through the use of advanced technology, computers can perform pattern recognition, do semantic analysis, analyze text, and model concepts. And computers can perform these processes faster, over more sources, and with greater precision than scholars who must rely on manual interpretation of data.

But if computers are to be used for these purposes, source materials must be in machine-readable form. For this reason, many scholars, once they have identified the key sources for their research, are converting them to machine-readable form so that they are in a form amenable to computer-assisted analysis.[65]

Scholarly conversion of sources to machine-readable form has been occurring for at least forty years. At first the practice was generally limited to numeric data. But in more recent years, the scholarly appetite for machine-readable data has extended to text as well. Textual conversion projects undertaken by individual scholars or under the auspices of academic institutions are far more prevalent than one might expect, especially in the fields of linguistics, classics, religion, and even history. The Center for Electronic Text in the Humanities estimates that there are currently eight thousand series of converted electronic text.[66] The conversion efforts among scholars are an example of the manifestation of end-user computing, in an effort to store, retrieve, manipulate, and analyze large amounts of sources in electronic form. The availability

[63]Nigel G. Fielding and Raymond M. Lee, eds., *Using Computers in Qualitative Research* (London: Sage Publications, 1991), 4.

[64]The use of nontextual sources of evidence, such as photographs, film footage, artifacts, and sound recordings is significant as well.

[65]See Avra Michelson, "Forecasting the Use of NREN by Humanities Scholars," paper presented at the panel "New Constituencies for the NREN," 27 March 1992, National NET '92, Washington, D.C. Available electronically on the Coalition for Networked Information fileserver.

Contact craig@cni.org for transfer information.

[66]Conversations between Avra Michelson and Marianne Gaunt, Center for Electronic Texts in the Humanities, Rutgers University, on 30 October 1990, and 14 May 1991. Rutgers and Princeton universities recently announced the creation of the jointly sponsored Center for Electronic Texts in the Humanities to respond to the information needs of a new generation of scholars. The center will develop an international inventory of machine-readable textual source materials, provide catalog entries through the Research Libraries Information Network (RLIN), and ultimately make electronic textual source materials available to researchers on research and education networks.

of an electronic corpus of sources on a topic encourages new types of questions to be asked and hypotheses to be explored.

The conversion of paper-based textual source materials to machine-readable form occurs worldwide. The earliest American conversion project, the Thesaurus Linguae Graecae (TLG), was founded in 1972 by Theodore F. Brunner at the University of California at Irvine to create an electronic data bank of extant ancient Greek texts from the period of Homer (ca. 750 B.C.) through about A.D. 600. The massive electronic file is used by researchers in Greek language and literature, linguistics, ancient history, philosophy, and religious studies to access Greek texts and related documents in full text. In conjunction with the American Philological Association, many members of the classicist profession participate in the ongoing compilation. Today the TLG is an immense, growing database of more than eight thousand works of classical Greek literature stored on CD-ROM, copies of which are available at two hundred locations in this country and abroad.[67]

Another conversion effort, the American and French Research on the Treasury of the French Language (ARTFL) draws on the work of the French government since 1957 to create a new dictionary of the French language. In conjunction with the development of the dictionary, the French developed an electronic database of approximately 150 million words derived from major literary and philosophical works and scientific and technical texts. For instance, the auxiliary database contains the novels of prominent and popular authors, correspondence, literary criticism, an extensive collection of poetry and theater, travelogues, biographies, historical works, political documents, biblical commentary, philosophical and economic essays, and writings on biology.

In 1979, the National Endowment for the Humanities (NEH) granted funds to the University of Chicago to conduct a survey of North American French literary scholars and historians whose work focused on the eighteenth to twentieth centuries. The purpose of the survey was to evaluate the potential usefulness of the ARTFL database to their work. Based on the scholars' endorsement, France deposited the corpus of fifteen hundred machine-readable texts at the University of Chicago in 1982. After the database was restructured to allow for text analysis, the electronic materials were made available to researchers. As an ongoing project at the University of Chicago, scholars continue to augment the database with, for example, a collection of troubadour poetry estimated to include 65 percent of the genre's extant poems; a collection of texts from the 1848 revolution, including radical newspaper articles, pamphlets, posters, speeches, and manifestos by proletarian leaders; and a collection of seventeenth-century French theater pieces.

A variety of scholars use the ARTFL, including Keith Baker, a University of Chicago historian of ideas. Baker's research concerns the attempt to redefine traditional terms during the Enlightenment to conform with the new political and social order. According to Baker, the "advantage of the ARTFL Project is that it provides a broad basis for systematic analysis of . . . key terms,"[68] such as the occurrence of important political phrases like "opinion publique" in eighteenth-century texts.[69] Other

[67]For information on TLG, see Theodore F. Brunner, "Data Banks for the Humanities: Learning from Thesaurus Linguae Graecae," *Scholarly Communication* 7 (Winter 1987): 1, 6–9; and David S. Miall's "Introduction," in *Humanities and the Computer: New Directions* (Oxford: Clarendon Press, 1990), 5.

[68]Alice Musick McLean, Robert Morrissey, and Donald A. Ziff, "ARTFL: A New Tool for French Studies," *Scholarly Communication* 8 (Spring 1987): 8.

[69]For another example of research devoted to the historical analysis of language, see Mark Olsen and

instructors have used ARTFL to teach French nationalism or to study the literary myth of Charlemagne as recorded from the middle ages to the nineteenth century. ARTFL is available online to scholars and students at institutions that participate in an inter-university fee-based consortium.[70]

A third large file, the Medieval and Modern Data Bank (MEMDB) was founded in 1982 at Rutgers University by Rudolph M. Bell and Martha C. Howell to establish an electronic library for medieval and early modern historians. The data bank consists of text descriptions of currency exchange rates, including a master data set of tabular works concerning medieval and early modern history. More than thirteen thousand medieval currency exchange quotations from the mid-twelfth century to 1500 A.D. are available, covering Europe, Byzantium, the Levant, and North Africa. MEMDB is a growing database; plans for expansion include adding taxation records, wills and inventories, parish records, vital statistics, company records, import/export records, household/estate accounts, palaeopathology studies, and such reference aids as glossaries of weights and measures, gazetteers of Latin and vernacular place names, and calendars of dates. The Research Libraries Group (RLG) is preparing a CD-ROM version of MEMDB for release.[71]

The TLG, ARTFL, and MEMDB represent discipline-specific electronic compilations, but many smaller and often more diverse humanities conversion projects also exist.[72] For instance, under the direction of Robert Hollander at Princeton University, the Dante Project converted to electronic form the complete text of sixty commentaries on Dante in Italian, Latin, and English. Before the Dante conversion project, many of these works were unavailable in the United States.[73] The purpose of Victoria Kirkham's Penn Boccaccio Project at the University of Pennsylvania is to develop an electronic archives that establishes links between the author's writings and the seven thousand illustrations of his work that were created contemporary to his lifetime in the fourteenth century through the sixteenth century.[74]

Several archival conversion projects are under way in England. For example, the Brotherton Library is compiling a complete database of its seventeenth and eighteenth century manuscript verse. The University of York History Department initiated a joint effort with the York Archaeological Trust both to develop a computerized database of the town's title deeds and to create a reconstruction of the region's topographical evolution between the twelfth and sixteenth centuries. At the University of Southampton, scholars are developing an online database of the papers of the first Duke of Wellington.[75]

At Bar-Ilan University in Israel, Yaacov Choueka is constructing a Jewish culture

Louis-Georges Harvey, "Computers in Intellectual History: Lexical Statistics and the Analysis of Political Discourse," *Journal of Interdisciplinary History* 18 (Winter 1988): 449–64.

[70]McLean, et.al., "ARTFL," 1, 6–9.

[71]Information reported in a phone conversation by Marianne Gaunt, the Center for Electronic Texts in the Humanities, Rutgers University to Avra Michelson 14 August 1991. See also Loughridge, "Information Technology," 281; and Rudolph M. Bell (Rutgers University), "User Perspectives and Requirements: Creator of Non-bibliographic Databases Has to Share with Others," unpublished paper presented to the Library of Congress Network Advisory Committee Meeting, 29–31 March 1989, Washington, D.C.; also information reported by Rudolph Bell to Avra Michelson in a phone conversation 13 November 1991.

[72]Georgetown University's Center for Text and Technology has compiled a database of descriptions of more than three hundred conversion projects, many of which comprise hundreds of series.

[73]Constance Gould, *Information Needs in the Humanities: An Assessment,* Prepared for the Program for Research Information Management of the Research Libraries Group, Inc., Stanford, Calif.: 1988, 27.

[74]Ibid., 27.

[75]Loughridge, "Information Technology," 281.

database, the Global Jewish Database/Responsa Project. This database includes about fifty thousand rabbinical answers to questions about Jewish life and culture, the Babylonian Talmud, Midrash literature, medieval commentaries, Maimonides Code, and the full text of the Hebrew Bible. When completed, the database will contain full text of nearly all written Hebrew works up to the tenth century as well as about one thousand major sources on Jewish culture.[76] At the University of Pennsylvania, Robert Kraft and John Abercrombie, working in conjunction with the Packard Humanities Institute, issued a CD-ROM containing at least ten versions of the Bible as well as a dictionary of New Testament Greek, classical Latin texts, Greek inscriptions, and various texts including Sanskrit sources.[77]

England's Oxford Text Archive (OTA) is a large repository of machine-readable text and includes text bases in more than twenty-five languages. Recently it served as a key source for a dissertation on Jane Austen's novels.[78] In Pisa, Italy, the Istituto di Linguistica Computazionale, one of the oldest and largest repositories of machine-readable classical and modern texts, has converted an extensive variety of materials, including Italian newspapers and periodicals, modern novels and poetry, and works of nonfiction.[79] Similarly, the British Domesday project assembles a variety of textual and visual information on contemporary Great Britain.[80]

Other conversion efforts involve domains such as Italian Renaissance music and lyric poetry, Spanish texts, medieval medicine-related drawings and illustrations, the papers of Charles Sanders Peirce,[81] and the works of literary greats such as Shakespeare, Shelley, Faulkner, and Milton.[82] Besides these institution-based conversion efforts, hundreds of smaller projects are addressing the needs of particular teams of researchers. Humanities scholars predict that the millions of words of text already available in machine-readable form represent only a minute fraction of source materials to be converted in the next ten to fifteen years.[83] Scholars contend that the reuse of textual databases by those other than the original converters will soon—if it does not already—constitute the predominant use.

In an effort to compile a massive electronic text corpus that will serve as a comprehensive research resource, language scholars have initiated the Data Collection Initiative (DCI). Sponsored by the Association for Computational Linguistics, the

[76]Gould, *Information Needs in the Humanities*, 39, and Loughridge,"Information Technology," 280.

[77]Gould, *Information Needs in the Humanities,* 38.

[78]The Center for Electronic Texts is cataloging the records of the OTA. The catalog is made possible through an NEH grant, and the center is describing the approximately eight hundred records that comprise the OTA and making the descriptions available through RLIN. For information on the OTA, see Miall, ed., "Introduction," in *Humanities and the Computer*, 5, and Gould, *Information Needs in the Humanities*, 27.

[79]Gould, *Information Needs in the Humanities*, 27.

[80]Miall, "Introduction," in *Humanities and the Computer: New Directions,* 5.

[81]This new effort involves a consortium that includes the current documentary editing project on Peirce centered at Indiana University-Purdue University, working with the philosophy departments at Harvard and Texas Tech universities, Georgetown University's Center for Text and Technology, Brown University's Computing and Information Services, and George Washington University's Department of Communication. The consortium plans to convert Peirce's large print manuscript collection housed at the Houghton Library, along with secondary commentaries on Peirce's work, to machine-readable form. The database would also include provisions for electronic scholarly communication on the vastly interdisciplinary work of Peirce.

[82]References appear on the database of electronic texts compiled by Georgetown University's Center for Text and Technology.

[83]Association for Computers and the Humanities, the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing, "Proposal for Funding for an Initiative to Formulate Guidelines for the Encoding and Interchange of Machine-Readable Text," unpublished proposal prepared for the National Endowment for the Humanities, 1988, 12.

DCI is the most extensive international collaboration of its kind. The ultimate goal of the project is to develop a global electronic library of text available for online research, primarily to serve the needs of computational linguists. Coordinated by Mark Liberman at the University of Pennsylvania's Department of Linguistics, the DCI includes a broad sample of materials, such as the archives of the *Challenger* investigation commission, which constitutes about 2.5 million words of deposition and hearings transcripts; portions of the Library of America volumes; 200,000 U.S. Department of Energy scientific abstracts; U.S. Department of Agriculture Extension Service fact sheets; the Federalist Papers; the King James Bible; computing journals; and sample correspondence and dictionaries.[84]

Besides acquiring a large corpus of electronic text, scholars are developing encoding standards for documents, to ensure that converted files can be read on a variety of computers and software. The Text Encoding Initiative (TEI) is a collaboration among the Association for Computers and the Humanities, the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing, which received funding from the National Endowment for the Humanities (NEH), the European Economic Community, and the Mellon Foundation to determine the elements and the methods for encoding machine-readable text for electronic exchange.[85] The first phase of funding is devoted to the needs of literary, linguistic, and text-oriented historical research.[86]

The TEI encoding standards closely follow the International Standards Organization's standard ISO 8879, the Standard Generalized Markup Language (SGML). This interchange format specifies how to encode (or mark up) texts so that they can be shared in a machine- and software-independent form by different research projects for different purposes. The TEI encoding standards use delimiters and tags to distinguish markup from text and to express specific information about the format of a document.[87] A draft version of the TEI standards is circulating to scholars and industry for review.[88]

The extraordinary projects under way by scholars to convert source materials to machine-readable form, assemble an electronic corpus of textual data, and establish data format standards for the interchange of text are in essence efforts aimed at facilitating end-user computer-assisted analysis of sources within the social sciences and humanities.

Computer-assisted analysis with artificial intelligence. Some scholars are con-

[84]Information on the DCI from phone conversations between Avra Michelson and Don Walker, Bellcore (14 May 1990), and Mark Liberman, AT&T Bell Laboratories (5 June 1990); see also Mark Liberman, "Report to the ACL Executive Committee on the ACL/DCI," (5 June 1990).

[85]Some scholars consider questions of what to encode as serious a concern as how to encode. For instance, should encoding indicate the physical condition of a document by marking the presence of ink spots, water stains, brittleness of paper, etc.?

[86]Association for Computers and the Humanities et al., *Proposal for Funding,* 59.

[87]SGML is a standard set of instructions for composing machine-readable tag sets and grammars. SGML applications, such as the TEI guidelines, establish tags and delimiters for the interchange of all types of text, including rules for encoding many types of document structures and data elements. The encoding allows computers, using appropriate software, to "read" the structure of a document (e.g., to know that an anthology of poems contains individual poems and that each possesses a title, stanzas, and lines), and to present it as such to the user; for further explanation of SGML, see C. M. Sperberg-McQueen and Lou Burnard, eds. *Guidelines for the Encoding and Interchange of Machine-Readable Texts* (Chicago, Oxford: Text Encoding Initiative Version 1.1, October 1990).

[88]Scholars in Europe have formed the History Working Party, a subgroup of the Text Encoding Initiative to ensure that TEI encoding guidelines address the needs of historians (e-mail via Internet from Donald A. Spaeth to Avra Michelson, 9 August 1991).

verting records to machine-readable form so that artificial intelligence (AI) can be used to assist in data interpretation and analysis. The use of AI in scholarly research signals a new phase in social science and humanities end-user computing.[89] As early as 1986, a panel of specialists brought together by the National Science Foundation reported that AI methods held great promise for research in the social sciences, especially in relation to the analysis and interpretation of complex situations, research design, and theory formation.[90] Within the humanities, scholars contend that the ability to process incomplete and inconsistent data with software that supports uncertainty and changes in beliefs makes AI uniquely suitable for many research efforts.[91]

In the area of AI, political scientists currently are the most sophisticated experimenters outside the hard sciences. Their prominence with AI calls to mind their earlier role as the pioneer users of computational processing with electronic numeric data. They are using artificial intelligence, especially in the area of international relations, to model decision making for the study of "deterrence, escalation control and war termination."[92] The applications involve the choices defense programs and military operations confront during peace, as well as methods for evaluating choices during a conflict. They explicitly address complications that decision makers face, such as conflicting principles and objectives, ill-defined alternatives, the complexity of problems, and the pervasive uncertainty of assumptions. As a result of intensive work, some existing prototypes are evolving into more advanced applications.[93]

Besides the "conflict-oriented" projects, other examples include AI prototypes that interpret Sino-Soviet negotiating sessions,[94] recognize patterns over a large, complex data set of historical events for purposes of prediction,[95] and generate hypotheses by exploring data to induce rules. One application of this last type analyzes the factors that influence different satisfaction levels of state legislators with legislative outcomes. The developers contend that the existing application can be adapted for use with similar research questions.[96]

The discipline's innovators argue that AI techniques should be considered "standard components in every political scientist's tool kit."[97] In making their case, they argue that many foreign policy questions represent suitable AI applications, such as the degree to which the Soviet economy declined under Brezhnev or the impact of the development of a navy on China's foreign policy.[98] One of the discipline's journals,

[89]See, for instance, Miall, *Humanities and the Computer*, 2; or for an earlier discussion, E. Casetti et. al., "Regarding the Feasibility and Desirability of Conferences on 'The Methodological Research Frontiers and the Social Sciences,' " Final Report to the National Science Foundation (NSF Award No.: OIR 8406230), 10 September 1986, 13.

[90]Casetti, et al., "Regarding the Feasibility and Desirability of Conferences," 13.

[91]Miall, *Humanities and the Computer*, 6.

[92]See Paul K. Davis, "A New Analytic Technique for the Study of Deterrence, Escalation Control, and War Termination," in *Artificial Intelligence and National Security,* edited by Stephen J. Cimbala (Lexington, Mass.: Lexington Books, 1987), 35–60.

[93]Ibid., 35–55.

[94]See William deB. Mills, "Rule-Based Analysis of Sino-Soviet Negotiations," *Social Science Computer Review* c 8 (Summer 1990): 181–95.

[95]Philip A. Schrodt, "Pattern-Matching, Set Prediction, and Foreign Policy Analysis," in *Artificial Intelligence and National Security,* 89–107.

[96]G. David Garson, "The Role of Inductive Expert Systems Generators in the Social Science Research Process," *Social Science Microcomputer Review* 5 (Spring 1987): 11–18.

[97]William deB. Mills, "Rule-Based Analysis," 182; and Paul A. Anderson, "Using Artificial Intelligence to Understand Decision Making in Foreign Affairs: The Problem of Finding An Appropriate Technology," in *Artificial Intelligence and National Security,* 133.

[98]See also, for instance, the ten or so articles in *Artificial Intelligence and National Security.*

Social Science Computing Review, keeps readers current on AI software with regular reviews of expert systems shells.

Unlike political scientists, most historians using AI tend to apply it to a narrower range of research questions. The chief use of AI in historical research is in applications designed to build nominal record linkages to reconstruct the population history of past societies.[99] This technique is usually used with family and community reconstruction, an area of study already quite computer-oriented. Nominal record linkage involves the analysis of parish and census-like records to reconstruct individual identities and relationships among individuals. It is a complex process that requires much interpretation because of the prevalence of homonic names (multiple names, with the same sound and often the same spelling, which refer to different people), name variations, and the need to link evidence related to the same individual from separate records. Historians typically consider an individual's vital dates, residence, profession, filiation, and other available data to decide whether several pieces of evidence refer to the same person.

Nominal record linkages typically involve analysis of a large and diverse set of records. Once the records of an individual have been linked, then a similar process must be performed to link the records of families and, ultimately, of communities. Historians are finding, however, that AI can be used to perform some interpretations associated with the task. For example, in France, historians at the Institut de Recherche et d'Histoire des Textes are using expert systems technology to identify unambiguously individuals, based on thirteenth and fourteenth century parish registers.[100] Similarly, the Cambridge Group for the History of Population and Social Structures has been using artificial intelligence for both nominal records linkage and to disambiguate household relationships. In the Cambridge project, AI performs some of the rudimentary aspects of analysis but still leaves the hard questions of interpretation to the historians. Kevin Schurer, a member of the group, describes it this way:

> The study of history should be driven by theory rather than fact. AI techniques may help historians to examine the relationship between facts more closely, and may add to the understanding upon which interpretations are made, yet they can never act as a substitute. In the examples given, expert systems may help us to determine the degree of household complexity in the past, or the levels of fertility. They may "positively" identify that females married on average at age 24 and had a completed family size of between five and six at the beginning of the 19th century, compared to an average age of 26 and a completed family size of around three at the end of the century, yet it is the task of the historian to theorize why this transition occurred.[101]

Although the primary use of AI among historians has been to reconstruct kinship

[99]See Kevin Schurer, "Artificial Intelligence and the Historian, Prospects and Possibilities" in *Interpretation in the Humanities: Perspectives from Artificial Intelligence*, Library and Information Research Report no. 71, edited by Richard Ennals and Jean-Claude Gardin, 169–95 (Cambridge: Cambridge University Press, 1990); and Joaquim Carvalho, "Expert Systems and Community Reconstruction Studies," *History and Computing II*, edited by Peter Denley et. al. (Manchester: Manchester University Press, 1989), 97–102.

[100]Caroline Bourlet and Jean-Luc Minel, "A Declarative System for Setting Up a Prosopographical Database," in *History and Computing*, edited by Peter Denley and Deian Hopkin (Manchester, England: Manchester University Press, 1987), 190.

[101]Schurer, "Artificial Intelligence and the Historian," 190.

and community relationships, other uses also are being explored. For instance, French social historian Beatrice Henin developed a computer file of leasehold documents created by notaries and property inventories taken at the time of death to study seventeenth-century Marseilles. Toward the end of her research, Henin became interested in the interior decor of houses from different social classes. Her use of artificial intelligence to analyze textual descriptions of pictures on the walls of rooms, largely with religious themes, led her to develop a new model for understanding Protestant and Catholic families in seventeenth-century England.[102]

Another European effort, the RESEDA Project, uses AI to respond to historical questions from a biographical database of French public and private figures during the fourteenth and fifteenth centuries. In addition to biographical information, the database contains abstract data about individuals, such as their beliefs, intentions, opinions, and mental attitudes. Using a hypotheses template, the system sorts the information to discover relevant facts, and infers information from the data to answer questions involving conjecture.[103]

In yet another type of project, a scholar is using AI to extend the findings of Tzvetan Todorov's *The Conquest of America: The Question of the Other* (1985). Todorov's work concerns the use of "signs and communication (and failed communications) within the cultural encounter."[104] Jim Doran at the University of Essex uses AI to add another dimension to Todorov's analysis by analyzing the belief systems and their impact on the behavior of the key persons and cultural groups examined in Todorov's book. Using evidence for beliefs already embedded in Todorov's work, Doran furthers the analysis by systematically examining the relationship between the beliefs and the conquest of America. This effort suggests one way in which scholars are exploring the use of artificial intelligence to extend an existing analysis of source materials. It uses AI to examine the relationship between reasoning and beliefs, to categorize "faulty" belief systems, and to consider metabeliefs—beliefs about beliefs.[105]

In the field of history, the principal investigators using AI in their research tend to be credentialed as historians, not as computer scientists. There is, however, an interesting exception. Kenneth L. Jones is an avid avocational genealogist who works with the Cartographics Application Group at the Jet Propulsion Laboratory (JPL) in Pasadena, California. As a hobby, Jones began using his AI background to unravel his family genealogy. The system he developed was fairly comparable to those already described: It provides records linkages by disambiguating individuals, families, and geopolitical boundaries. But the application's level of sophistication caught the attention of the American intelligence community. In developing the system, Jones produced a form of knowledge representation (the depiction of knowledge as symbols in a form that a computer can manipulate), which he refers to as "knowledge visualization." Knowledge visualization entails the use of graphics to clarify or make more intelligible the relationships among interrelated fragments of knowledge. Conferring with colleagues at the JPL,

[102]Richard Ennals, *Artificial Intelligence: Applications to Logical Reasoning and Historical Research* (Chichester, England: Ellis Horwood Limited, 1985), 125.

[103]Gian Piero Zarri, "Artificial Intelligence and Information Retrieval: A Look at the RESEDA Project," in *The Analysis of Meaning: Informatics 5,* edited by Maxine MacCafferty and Kathleen Gray (London: Queens College Oxford, 1979), 166–72.

[104]Jim Doran, "A Distributed Artificial Intelligence Reading of Todorov's *The Conquest of America: The Question of the Other,* by Tzvetan Todorov, 1985," in Ennals and Gardin, *Interpretation in the Humanities,* 166.

[105]Ibid.

Jones realized that the intricate matrices he was developing to assist in family research could be applied to any problem that involves the conceptualization of complex interrelationships among objects, such as tracking money-laundering or counter-terrorism activities. Jones's work on this system continues, with funding from the U.S. Army's Joint Tactical Fusion Office.[106]

Besides research-oriented applications, serious efforts are under way to use software engineering to develop a scholarly workstation devoted to the needs of historians. Manfred Thaller is a historian and key participant in the Historical Workstation Project sponsored by the Max-Planck-Institut fur Geschichte in Gottingen, Germany, an institute dedicated to fundamental research in the humanities.[107] Since 1978, the institute's research has been designed to improve software for historians. The workstation project focuses on the development of three components: software that can access information from both current and historical sources, databases that are as available and easy to use as books, and knowledge bases that allow the other components to draw upon information in historical reference works. The developers plan to use artificial intelligence to provide transparent interaction between subsystems, to create new rules in the knowledge bases when new facts are inferred, and to guide users to relevant information. Various elements of a production prototype of the workstation are being tested. Some are still under development, and some of the more difficult aspects of context-sensitive interpretation are still in the design phase.

Among sociologists, Edward Brent refers to the current era as "the first hint of what it might be like to have computers that act less like clerks and more like colleagues."[108] His remarks pertain to the early benefits sociologists report in using AI for theory development, especially to differentiate dependent variables from independent variables, to develop theories based on causal models, and to extend sociological theory by transforming theoretical assertions into logical ones. Sociologists are developing applications using artificial intelligence for these purposes.[109] In the field of literature, scholars are using natural language understanding for the rapid disambiguation of words stored in machine-readable dictionaries and, within limited domains, to comprehend the "meaning" of a story. In other literary uses, expert systems have been developed that log and analyze differing interpretations of text among readers.[110]

Scholars are beginning to use artificial intelligence as a tool to assist in the interpretation and analysis of sources in nearly every corner of the social sciences and the humanities.[111] In addition to those mentioned, researchers in the fields of archaeology, linguistics, music, art history, and design are exploring the value of "intelli-

[106]From a presentation made by Kenneth L. Jones at the Eighth Annual Intelligence Community AI/Advanced Computing Symposium, Greenbelt, Maryland, 12 March 1991.

[107]See Manfred Thaller, "The Historical Workstation Project," unpublished paper delivered at the seventeenth International Congress of Historical Sciences, Madrid 29 August 1990.

[108]Edward Brent, "Is There a Role for Artificial Intelligence in Sociological Theorizing?" *American Sociologist* 19 (Summer 1988): 164.

[109]Ibid., 160–64.

[110]See for instance, Nancy M. Ide and Jean Veronis, "Very Large Neural Networks for Word Sense Disambiguation," paper presented at European Conference on Artificial Intelligence, Stockholm, August 1990; Nancy M. Ide and Jean Veronis, "Artificial Intelligence and the Study of Literary Narrative," *Poetics* 19 (1990): 37–63; and David Miall, "An Expert System Approach to the Interpretation of Literary Structure," in Ennals and Gardin, *Interpretation in the Humanities,* 196–214.

[111]The Foundation for Intelligent Systems in the Social Sciences, Arts and Humanities is a new organization that publishes a quarterly newsletter, *Intelligent Systems,* devoted to applications in these disciplines. For further information, contact the foundation's director, Stephen Toney, at 2205 Gabriel Drive, Las Vegas, Nevada 89119.

gent" tools, such as expert systems shells and specialized software, capable of performing functions attractive to a variety of disciplines.[112] During this decade, as primary sources become more available in machine-readable form and as commercial AI software becomes more sophisticated and prevalent, it is likely that scholars will turn increasingly to AI for research assistance.

Dissemination of Research Findings

The scholarly obligation to report research findings is typically fulfilled through the publication of articles in peer-reviewed print journals or monographs. Until recently, the defining feature of a publication was its linear and printed format. But the emergence of electronic publishing and hypermedia are challenging this definition of a document. The scholarly use of electronic publishing and hypermedia is a result of the dual trends toward end-user computing and greater connectivity. Considered together, these new dissemination and presentation formats are beginning to transform the manner in which findings are shared in the scholarly community.

Electronic publishing. Introduced less than a decade ago, electronic publishing already represents a $6.5 billion business according to current estimates.[113] The most viable commercial electronic publishing efforts involve indexing and abstracting texts and electronic versions of full-text print journals. Through electronic publishing, it is increasingly possible for researchers to access on their computers full-text versions of "newspapers and newswires, popular magazines and scholarly journals, financial and directory sources, and reference books."[114] For example, electronic versions of more than forty medical journals are available in full text, as are some of the most important scientific and technical journals.[115] More than three hundred full-text newsletters can be accessed through either NewsNet or Dialog files. Business and industry periodicals enjoy wide coverage in electronic form, as do specialized titles like marketing reports.[116] Unlike bibliographic databases developed primarily for use by information specialists, full-text databases generally are designed for the end-user. Researchers, enthusiastic about the convenience of these databases, also find electronic publishing attractive because it promises to increase the pace of publication and expand opportunities for dialogue among scholars.

An electronic resource directory created by Bibliofile, *Fulltext Sources Online*, identifies more than fifteen hundred full-text and information sources available on-

[112]For further information on shells, see Avra Michelson, *Expert Systems Technology and Its Implications for Archives,* National Archives Technical Information Paper no. 9 (Washington, D.C.: National Archives and Records Administration, March 1991), 9–10. An example of specialized software is Ex-Sample which helps researchers determine an appropriate sample size for a study. Ex-Sample is reviewed in Edwin H. Carpenter and Rick D. Axelson, "Statistical and Graphical Research Methods: State of the Art," In *Social Science Computer Review* 7 (Winter 1989), 508. Another example, IXL's Discovery Machine, performs pattern-matching over large amounts of data that typically would go undetected through manual analysis. For a report on its use, see Karen D. Schwartz, "Agencies Use Software to Dig Up Links Among Data," *Government Computer News* 19 (15 October 1990): 60.

[113]Council on Library Resources, *Communications in Support of Science and Engineering,* Report to the National Science Foundation. Washington, D.C.: Council on Library Resources, August 1990, II–8.

[114]See Ruth A. Pagell, "Primary FTDBs for the End User: New Roles for the Information Professional," *Online Review* 13 (April 1989): 143.

[115]Ibid., 146. The Hunt Library at Carnegie Mellon University is compiling an electronic full-text corpus of extended runs of computer science journals on artificial intelligence. The specific journals and runs are cited in a subsequent section of this paper (see "The Library Profession's Response to New Forms of Scholarship/ Software Engineering" section).

[116]Pagell, "Primary FTDBs for the End User," 143–46.

line.[117] The trend watchers in the industry estimate that by the year 2000 much of scholarly and professional publishing will occur electronically, involving the transmittal of journals and books over high-speed networks by authors to the publishers, and then from publishers to readers.[118]

Further, publishers are discovering that the electronic versions of certain printed products are beginning to turn a profit. Indeed, Harry Boyle of Chemical Abstracts Service (CAS), one of the world's largest indexing and abstracting companies, describes the shift occurring in his company in this way:

> The revenue base for the printed product is shrinking. The revenue base for the electronic product is growing. Fifteen years ago the printed product was paying the bills. In the next five years, the electronic form of the product will be the dominant way that the database is used and the printed will become secondary. We are rapidly approaching the point where the electronic use of the product is in fact generating a lot of the revenue needed to build the database, and the printed product is becoming the secondary concern. I don't think we will stop the printed product. But if you look at the economies inside the company, you'll know that electronic use is paying the bills and it is subsidizing the printed product which is an exact reverse of what we saw fifteen years ago.[119]

On the surface, electronic publishing seems to imply only a change in the form of distributing publications. But scholars in the social sciences and the humanities have begun to use the existing research and education networks to engineer a new form of publication distinct from commercial efforts. These publications are academic-based, scholarly created and controlled, (often) refereed, electronic-only, network-delivered journals. Although scholarly electronic journals were invented only several years ago, already about three dozen have sprung up in an array of disciplines, along with sixty newsletters and the thousands of electronic conferences used for less formal communications.[120]

PSYCOLOQUY is one of the best examples of the innovative genre of electronic journals.[121] The journal's editor, Stevan Harnad, a cognitive psychologist at Princeton University, has edited an influential nonelectronic journal (*Behavioral and Brain Sciences*) for more than fifteen years. Harnad decided to edit a scholarly electronic journal as a result of his experience participating in an early electronic conference. He characterized early users of networks as primarily computer enthusiasts and graduate students. These two audiences possessed enough time and motivation to venture into the new medium of conferencing, a unique form of communication that allows people, dispersed in time and place, to share ideas, ask questions, comment on work, and sustain narrative discussions. As

[117]Richard Van Orden, "Content-Enriched Access to Electronic Information: Summaries of Selected Research," *Library Hi Tech* 31 (1990): 28.

[118]Robert Weber, "The Clouded Future of Electronic Publishing," *Publishers Weekly* 237 (29 June 1990): 76.

[119]Jeffrey K. Pemberton, "*Online* Interviews Harry Boyle on CAS's New License Policy . . . Effects on Searching Prices," *Online* 12 (March 1988): 21.

[120]Michael Strangelove, *Directory of Electronic Journals and Newsletters,* ed. 1, July 1991. (To retrieve electronically, contact the author at <441495@uottawa>; the directory is also available in print through the Association of Research Libraries, Washington, D.C). Ann Okerson, of the Association of Research Libraries, provided updated information on current journal numbers to Avra Michelson in March 1992.

[121]The *PSYCOLOQUY* discussion is from notes on a presentation by Stevan Harnad at the "Refereed Journals" session on 21 March 1991, at the National Net'91 Conference in Washington, D.C.

an early participant in an AI conference, Harnad decided to transmit work in a form more polished than customary, as if he were writing for a peer-reviewed journal. To his great surprise, he found the exchange tremendously helpful to his intellectual work. Instead of waiting several years to receive peer responses, he received instantaneous reactions to his work over the networks. Further, the responses arrived at the beginning of his intellectual process rather than at the end, as happens with conventional publishing. Inspired by his conference experience, Harnad wondered what it would be like to experience with the best minds in his field the same kind of instantaneous dialogue he had established with computer enthusiasts and graduate students. This prompted him to create *PSYCOLOQUY*, a fully refereed, scholarly, electronic-only journal, sponsored by the American Psychological Association.

PSYCOLOQUY is an interdisciplinary journal that publishes articles and reviews concerning psychology, neuroscience, cognitive science, behavioral biology, linguistics, and philosophy. Its editorial board of fifty scholars reflects the range of disciplines published by the journal. Journal submissions, refereeing, editorial work, and distribution are handled entirely electronically. There are currently more than two thousand individual subscribers on Bitnet. A large number of institutional subscribers also receive *PSYCOLOQUY* through Usenet, a network connected to most of the universities and research institutions of the world, allowing all individuals at these sites to access the journal. In 1990, *Library Journal* named *PSYCOLOQUY* one of the year's best journals.

Harnad contends that the most important difference between electronic journals and print publication is not the form of distribution but the medium's potentially revolutionary contribution to the furthering of scholarship and the creation of knowledge. The real contribution of the electronic medium is that it does what no other medium can do. Instead of waiting a year or two for peer feedback (the typical amount of time it takes to publish and then respond in print), and instead of receiving the feedback when already strongly invested in the next research project, scholars enjoy rigorous intellectual dialogue with one another, freed from the constraints of time and place, at the front end of the research process.[122] The instantaneous distribution of ideas among peers permits a new and critically important type of interaction that furthers scholarly inquiry in a way not possible previously. The electronic medium is unique in its capacity to support interactive improvement of scholarship at a speed much more commensurate with the speed of thought.

Other examples of scholarly, electronic-only journals include *Post Modern Culture* (North Carolina State University), an interdisciplinary journal of literary theory, culture, and creative writing; *Artcom*, devoted to the interface of art and communication technology; *Quanta* (Carnegie Mellon University), an electronic journal of science fiction and fantasy; the *Bryn Mawr Classical Review*, a review journal of books on Greek and Latin classics; *Online Journal of Distance Education and Communication* (University of Alaska), devoted to the development and practice of distance education; *Ejournal* (State University of New York at Albany), an interdisciplinary journal on the theory and practice of electronic communication; *New*

[122]Stevan Harnad, "Scholarly Skywriting and the Prepublication Continuum of Scientific Inquiry," *Psychological Science* 1 (November 1990): 342. A similar point is made by Cliff McKnight in his article, "Using the Electronic Journal," in *Scholarly Communication and Serials Prices: Proceedings of a Conference Sponsored by the Standing Conference of Naitonal and University Libraries and the British Library Research and Development Department 11–13 June 1990*, edited by Karen Brookfield (New York: Bowker-Saur, 1991).

Horizons in Adult Education (Syracuse University Kellogg Project), a refereed journal for the field; *Journal of the International Academy of Hospitality* (Virginia Polytechnic Institute and Blacksburg State University), publishing refereed articles on basic and applied research on hospitality and tourism; and the *Public-Access Computer Systems Review*, exploring electronic access to library materials.[123] At a meeting recently convened by the Association of Research Libraries (ARL), electronic journal editors established the Association of Scholarly Journal Editors, a "closed" electronic communications list for discussing common concerns and new publishing efforts.[124]

Aside from scholarly controlled electronic journals, commercial publishers are beginning to explore the profitability of publishing electronic-only academic journals. The American Association for the Advancement of Science (AAAS), in conjunction with the Online Computer Library Center, Inc. (OCLC), announced the publication of its first electronic-only journal in 1992. The publishers expect *The Online Journal of Current Clinical Trials*, a new peer-reviewed medical journal, to distribute the findings of original research several months faster than its print counterparts. The journal represents the first commercial electronic effort to display typeset-quality graphs, tables, and equations. The editors are Edward J. Huth (former chief editor for nineteen years of the *Annals of Internal Medicine*), Curtis Meinert (of the Johns Hopkins Center for Clinical Trials), and Thomas C. Chalmers (associate director of the Technology Assessment Group, Harvard School of Public Health).[125] If the commercial publication of scientific journals proves successful, it is likely that their counterparts will emerge in the social sciences and humanities.

Hypermedia. During the last ten years, hypermedia has developed into a mature tool that supports electronic browsing by allowing users to follow links through text, images, and audio and visual records. The electronic links that characterize the technology also make it possible to compose and deliver research products in new ways.[126] As hypermedia becomes a mainstream technology during this decade, scholars are encountering the prospect of redefining the modern product of research. Should a hypermedia document provide automatic links that take a reader from a footnote to the actual cited work? Should hypermedia documents chronicle through links the intellectual process of discovery? What new types of authoring guidelines are necessary for research products developed in hypermedia? Scholars are beginning to tackle some of the hard questions raised by the availability of a technology that allows for a more complex organization of ideas. The scholarly creation and consumption of hypermedia documents is another example of the trend toward end-user computing, further stimulated in this case by the online transition.

One historian argues that the power of hypertext (hypermedia restricted to text) is that "it produces documents not intended

[123]Information on electronic journals from Strangelove, *Directory of Electronic Journals and Newsletters;* for a discussion on publishing a scholarly electronic journal, see Charles W. Bailey, Jr., "Electronic (Online) Publishing in Action . . . *The Public-Access Computer Systems Review* and Other Electronic Serials," *Online* 15 (January 1991): 28–35.

[124]Informal presentation made by Ann Okerson (Association of Research Libraries) to a session on "Non-Commercial Publishing" at the Spring meeting of the Coalition for Networked Information, 19 March 1991, Washington, D.C.; also, letter from Ann Okerson to Avra Michelson dated 8 July 1991.

[125]*The Online Journal of Current Clinical Trials,* brochure published by the American Association for the Advancement of Science and OCLC, ca. 1991.

[126]For an introduction to hypermedia, see Jeff Conklin, "Hypertext: An Introduction and Survey," in *Computer-Supported Cooperative Work: A Book of Readings,* edited by Irene Greif (San Mateo, Calif.: Morgan Kaufmann Publishers, 1988), 423–75.

to exist in printed form."[127] He describes the contrast between a standard history textbook and a hypertext product through this example:

> Imagine a computerized book of documents. As you open it to the Monroe Doctrine, you see the several paragraphs of the President's address which make up the statement of foreign policy. Gliding a mouse-directed cursor over the words, an icon pops up next to the words, "Russian Imperial Government." By clicking the mouse, you reveal a brief essay on the Russian Czar's interest in Alaska. The word "Czar" in that subtext can bring up the Czar's actual statements on the subject, and "Alaska" can trigger a map of the Pacific Northwest. After folding these asides back into the original document, you reach the phrase, "With the existing colonies or dependencies of any European power we have not interfered and shall not interfere," and clicking the mouse reveals an annotated list of interventions prior to 1823. That screen will activate a map of Central and South America showing the new revolutionary governments and the dates of their independence from Spain.[128]

Scholars already have begun to produce research projects in nonlinear formats. At Stanford University, for example, a hypermedia Shakespeare application created by Larry Friedlander allows users to view on a video monitor filmed versions of Shakespearean plays, while viewing on another screen a synchronized presentation of the play's text and stage blocking material. At any point, users can refer to dictionaries and historical notes to increase their understanding of the performance. The system also allows users to create animated versions of plays, provides interactive tutorial instruction on theater topics, and supports note taking.[129]

Another hypermedia application, designed for use in an undergraduate poetry course, uses software to convey the ideas that poems are related to other poems, that they may be related to other art forms, and that they may be related to both other poems and other forms of art simultaneously.[130] Since poems often refer to lines from other poems, use a painting to develop an analogy, quote a piece of literature, or allude to a music score, a hypermedia document can make poetry truly come alive by using links to demonstrate concretely the cultural attachments among art forms.

Two other projects are representative of efforts to use hypermedia as a new authoring medium. The Faculty of Art and Design at Coventry Polytechnic in England considered the possibilities of using hypermedia as an authoring medium for four years. As a result of their deliberations, the faculty decided to allow students to submit the curriculum's required thesis in hypertext. The thesis is a research product on the historical and theoretical portions of the curriculum. Hypermedia enables the art and design students to incorporate their design and visualization aptitudes into the organization and presentation of a theoretical work. After this experiment with student theses, the faculty will evaluate the effectiveness of hypertext as an authoring me-

[127]James B. M. Schick, *Teaching History with a Computer: A Complete Guide* (Chicago: Lyceum Books, 1990), 63.

[128]Ibid.

[129]Charles W. Bailey, Jr., "Intelligent Multimedia Computer Systems: Emerging Information Resources in the Network Environment," *Library Hi Tech*, 8 (1990): 31.

[130]John M. Slatin, "Text and Hypertext: Reflections on the Role of the Computer in Teaching Modern American Poetry," in *Humanities and the Computer: New Directions*, edited by David S. Miall, (Oxford: Clarendon Press, 1990), 129–31.

dium and will consider the most appropriate contexts for its use.[131]

Finally, in anticipation of the widespread use of hypermedia as an authoring tool, the British Library Research and Development Department is funding Project Quartet, a research effort to develop a standard set of guidelines for creating hypermedia documents. The principals on the project argue that researchers authoring in hypertext need guidelines for establishing nodes and links to provide the necessary hooks for readers. They contend that the skills used for writing in paper media do not adequately serve the needs of scholars authoring in the electronic age. The project hopes to establish global taxonomies for hypertext authoring that can be used across systems.[132]

Curriculum Development and Instruction

The enormous amount of literature on computer-aided instruction makes it appear that faculty in the social sciences and humanities use computer technology to improve their teaching to an even greater extent than for research. This is not surprising, since the fundamental aspect of the technological revolution is that faster, smarter machines affect the ways we think and learn. According to Mary Alice White, director of the Electronic Learning Laboratory at Columbia University's Teachers College, information technologies change "how we represent information, and therefore how we view a problem . . . how we analyze problems, and because they change that view and that analysis, they can change how we make decisions. These are *intellectual* tools, the very stuff and excitement of education."[133] The scholarly use of computers to develop instructional applications is another example of the trend toward end-user computing, while connectivity represents the key trend that allows for new styles of distance education.

Teachers at every educational level are revising curriculums to include computer-supported instruction, such as simulations, cognitive modeling, and individual-oriented learning. The trend for an increasing portion of academia is toward "computer campuses" where students are required to purchase a specific set of computer equipment upon enrolling. Some universities have begun to fund positions devoted exclusively to helping faculty develop instructional software or incorporate information technology into the classroom.[134] Com-

[131]Alan Dyer and Kate Milner, "An Examination of Hypertext as an Authoring Tool in Art and Design Education," in *Humanities and the Computer: New Directions,* 137–48.

[132]Cliff McKnight, John Richardson, and Andrew Dillon, "Hypertext Authoring: Some Basic Issues," *Humanities Communication Newsletter* 11 (1989): 25–29. See also other publications issued by this group, such as *Hypertext in Context,* The Cambridge Series on Electronic Publishing (Cambridge: Cambridge University Press, 1991); "Human Factors of Journal Usage and Design of Electronic Texts," *Interacting with Computers* 1 (1989): 183–89; "The Effects of Display Size and Text Splitting on Reading Lengthy Text From Screen," *Behaviour & Information Technology* 9, no. 3 (1990): 215–27; and Bill Tuck, Cliff McKnight, Marie Hayet and David Archer, *Project Quartet,* Library and Information Research Report no. 76, (Wetherby, England: British Library, 1990). Cornell University is also experimenting with the usability of an online hypermedia presentation of thousands of articles published by the *Journal of the American Chemical Society.* See Michael Alexander, "But Can You Read It Like A Book?" *Computerworld* 24 (19 November 1990): 18.

[133]Mary Alice White, "The Third Learning Revolution," *Electronic Learning* 7 (January 1988): 6.

[134]See Schick, *Teaching History with a Computer: A Complete Guide*, 207–08. Schick cites Drexel, North Carolina/Chapel Hill, North Carolina State, the University of Southern California, and Stanford among others, as campuses that have hired staff to stimulate computer-oriented curriculum. Richard Kesner, chief information officer at Babson College, reports that his campus is hiring staff for this purpose as well. The EDUCOM/USC Survey of Desktop Computing in Higher Education estimates that more than 40 percent of two and four year public and private colleges and universities provide support for faculty developing computer-based instructional courseware. See Kenneth C. Green and Skip Eastman, *Campus Computing 1990* (Los Angeles: University of Southern California, Center for Scholarly Technology, 1990), 15.

puter simulations are especially popular in many disciplines because they submerge students in a different social context, allowing them to consider "what-if" scenarios. The simulations tend to be particularly effective in promoting an understanding of history because the first-person experience of time and circumstance helps students appreciate that the past is shaped by individuals reacting to social events and forces.[135]

An extensive array of simulation software is available for instructional purposes, including hundreds of applications in the field of history alone. For example, the experience of the Constitutional Convention, complete with delegate selection and the ratification process, is available to students using simulation software. Other simulations allow students to experience U.S. congressional committee debates on readmitting southern states to the union after the Civil War, or to participate in the presidential decision on whether to take action in the Pullman Strike of 1894. A National Geographic Society product simulates the construction of the Transcontinental Railroad, challenging students to decide about such issues as construction, labor, and relations with Native Americans. Another simulation focuses on the military tactics used by the Soviet Union with Nazi forces. At Stanford University, a French History professor developed a simulation that establishes a seventeenth century bourgeois context for students to negotiate a strategic marriage, consider proper investments, and manage the family's inheritance to promote their stature. Some simulations employ artificial intelligence techniques to demonstrate more fully the meaning of historical context. For instance, AI-enhanced simulations are available for such events as the Russian Revolution and the development of the European Economic Community.[136]

Although simulations are one of the most popular forms of computer technology found in the classroom, other types of applications are also in use. In an effort to computerize a full discipline's curriculum, Gregory Crane, of Harvard University's Classics Department, established the Perseus Project. The Perseus database attempts to provide an interactive multimedia curriculum on classical Greek civilization. It contains a vast corpus of the discipline's sources, including translations of major Greek texts, introductory materials designed for novice students, Greek language texts for more advanced students, color images and line drawings of archaeological artifacts and maps, essays and themes on key facets of Greek literature, a chronology, and a classical encyclopedia. The hypermedia application provides course materials for such disciplines as art, archaeology, classics, history, law, philosophy, and political science. In early 1992, Yale University Press released version 1.0 of the Perseus database, which runs on MacIntosh computers with the HyperCard program.[137]

Anticipating the availability of large volumes of humanities source materials online, the Stevens Institute of Technology, with support from the Humanities Grant Program of the New Jersey Department of Higher Education, is exploring how access to electronic source materials is apt to restructure humanities education. In particular, they are interested in learning how electronic texts, such as those compiled for

[135]Schick, *Teaching History with a Computer,* 101–02.

[136]For an extensive critical bibliography of current simulations available for the study of history, see Schick, *Teaching History with a Computer,* 122–45, and for a description of the Stanford University simulation see pages 100–01. The *Social Science Computer Review* regularly reviews commercial simulation software designed for educational purposes. The AI simulations are mentioned in Ennals, *Artificial Intelligence*, 125.

[137]See a brief review of the project in *Social Science Computer Review* 7 (Summer 1989): 211; also, Loughridge, "Information Technology," 281.

specific disciplines, can best be integrated into undergraduate course work. As part of their research effort, they plan to evaluate student learning patterns and actual student performance with electronic curriculums.[138]

In a different approach, a group of faculty at Manchester Polytechnic in England is developing "viewbooks" for use in history curriculums. These disk-based books take the form either of annotated historical documents with introductions and conclusions or of texts and tables. Approximately twenty-five different books are available, and they permit information to be retrieved through various techniques from the databases. The developer is currently designing a viewbook shell that will allow instructors to insert the text of their choice into the database. This type of application will benefit from advances in document-conversion scanning technology.[139]

Commercial software specially designed for particular disciplines is becoming available, which will facilitate the use of electronic source materials in classrooms. One recently released package, for example, displays nineteenth-century statistical cross-tabulations and regressions on France, England, and Wales.[140] Faculty also are making use of AI shells; a history instructor found a particular software shell enhanced with artificial intelligence well-suited for an application on the Norman Invasion. This same instructor also chose an expert systems shell to construct a learning tool for the study of the Middle Ages.[141] The use of information technology by the British in history curriculums is so great that the country's academics have established a formal organization to support the exchange of technical resources. Headquartered at the University of Bath, the National Information for Software and Services organization coordinates the sharing of historically oriented software and data files among college professors.[142]

Apart from computer-assisted curriculums, teachers in the United States are using information technology to support a new style of education. Distance learning, in essence an improved successor to correspondence course work, interactively links teachers and students in scattered locations. During the past few years, a majority of states have become active proponents of distance learning. The findings of a recent survey show that thirty-two states "currently have at least one statewide network for distance learning, and nearly half have more than one."[143] Enthusiasm for distance learning seems to emanate as much from advances in storage and retrieval technology as from telecommunication's networks that expand the ability to use information at distant locations. Indeed, a study conducted by the U.S. Office of Technology Assessment found that distance learning no longer serves only isolated rural schools. Rather, it has become the vehicle for bringing advanced, specialized course work and an array of experts to many classrooms. Existing programs make it possible for a high-school student in Mississippi to study Japanese, and for Washington State to provide advanced

138See Edward A. Friedman, James E. McClellan III, and Arthur Shapiro, "Introducing Undergraduate Students to Automated Text Retrieval in Humanities Courses," in *Humanities and the Computer*, 103–12.

139See Richard H. Trainor, "History, Computing and Higher Education," in *History and Computing II*, 38–39; for a discussion of scanning technologies, see Timothy C. Weiskel, "University Libraries, Integrated Scholarly Information Systems (ISIS), and the Changing Character of Academic Research," *Library Hi Tech* 6 (1988): 15.

140See review in *Social Science Computer Review* 7 (Summer 1989): 211.

141See Martyn Wild, "History and New Technology in Schools: Problems, Possibilities and the Way Forward," in *History and Computing II*, 30.

142See Trainor, "History, Computing and Higher Education," in *History and Computing II*, 40.

143Barbara Kurshan and Marcia Harrington, *Statewide Education Networks: Survey Results* (Roanoke, Va.: Educorp Consultants, April 1991), 2.

placement English courses to all who qualify. In Maine, teachers enrolled in a masters' program attend after-hours graduate courses in their classroom via distance learning, instead of undertaking a four- to five-hour commute.[144] The feasibility of using distance learning to maximize university students' control over the time, place, and pace of education is being evaluated through experimental courses. The flexibility of a distance-learning program is apt to be particularly attractive to full-time employed students enrolled in advanced degree programs.[145]

The infusion of technology into educational programs is occurring rapidly. Examining the effectiveness of technology as an educational tool represents a popular area of research, though findings are still somewhat preliminary. One study on the impact of the use of AI tutors in high-school geometry classes found that the individually paced applications fostered a healthy competitiveness among students.[146] In a traditional classroom, the students never had the opportunity either to get ahead of or fall behind one another. With the AI tutor, however, self-paced learning stimulated students to rival one another, as they would call out in class the "page" on the monitor they had advanced to through correct answers.

The study also observed that the majority of students enjoyed the AI tutoring more than conventional classroom instruction and that the enjoyment translated into increased motivation. In addition, the study found that students appreciated the independence from adult control and that with the computer they were free to vent anger and frustration unacceptable with teachers. But probably most important, the research discovered that the students experienced the tutor as a game and thus associated it with play. The electronic games popular among youth, combined with computer-assisted learning, in essence are preparing the next generation for a new era. As a result of changes occurring in education and play, young people are being thoroughly indoctrinated into the computer culture. The use of information technology and electronic communication will be deeply ingrained in the next generation of researchers, who will have been computer veterans since elementary school. The current demands for electronic information available through networks in homes and offices can only escalate and deepen among tomorrow's scholars.

Summary

As the preceding section indicates, the clear trend in the modern research process is toward scholarly identification, use, interpretation, and analysis of sources in electronic form, and the gaining prominence of new forms of computer-assisted communication and instruction. The research process is already changing, and this change is accelerating and spreading across a wide range of disciplines. Because a key factor promoting this change is the availability of new information technology, analyzing how trends in information technology interact with current trends in scholarly practice can help predict the future evolution of the research process.

The analysis of information technology undertaken above points to two major technology trends that are likely to transform scholarly practice: increased end-user computing and increased connectivity. This analysis also implies that a number of more

[144]See U.S. Congress, Office of Technology Assessment, *Linking for Learning: A New Course for Education,* OTA-SET-430 (Washington, D.C.: U.S. Government Printing Office, November 1989), 2–3, 54.

[145]See Gil Rogers, "Teaching a Psychology Course by Electronic Mail," *Social Science Computer Review* 7 (Spring 1989): 60–64.

[146]See Janet Ward Schofield, Debra Evans-Rhodes, and Brad R. Huber, "Artificial Intelligence in the Classroom: The Impact of a Computer-Based Tutor on Teachers and Students," *Social Science Computer Review* 8 (Spring 1990): 24–41.

specific technology, including artificial intelligence, end-user publication and distribution, hypermedia, and visualization and virtual reality, are likely to have a significant impact on the research process. The effects of these trends, along with changes in scholarly practice that are already under way, point to a future in which researchers use computation and electronic communication to help formulate ideas, access sources, perform research, collaborate with colleagues in their own and other disciplines, seek peer review, publish and disseminate results, and engage in many professional and educational activities. Far from being visionary, this future is already present: It is currently being experienced by significant and increasing numbers of researchers from many disciplines.

How should the archival profession respond to these changes in scholarly practice? Are the techniques and functions developed by the archival profession to manage printed media adequate for the needs of researchers who operate in a global electronic networking environment? Should established archives convert printed material to machine-readable form? If so, what selection criteria should be used? What constitutes the "reference function" in the age of research and education networks and electronic communication? These issues first are addressed through case examples drawn from the experience of the library community, and then by a set of recommendations specifically designed for the archival profession.

RESPONSES BY THE LIBRARY PROFESSION TO CHANGING RESEARCH PRACTICES

On several occasions in the recent past, libraries and professional associations have sponsored inquiries into scholarly use of technology. For example, the American Council of Learned Societies conducted a survey in 1985 to 1986 that noted the rapid increase in the use of technology by the scholarly community.[147] In a more recent study sponsored by the Harvard College Library and the American Council of Learned Societies, the Conference on Research Trends and Library Resources brought social science and humanities scholars together to explore new trends in research methods. Scholars spent several days considering the impact of new technology, interdisciplinary research, and the use of innovative formats of materials on their work.[148] In another effort, the American Academy for Arts and Sciences sponsored an exchange between scholars and librarians to develop policy recommendations to improve access to library materials. A key observation shared by these inquiries is that scholars increasingly want online access to electronic source materials available through personal computers in their homes or offices.

Visionary leaders within the library community are beginning to implement pilot projects designed to improve the library's role in advancing scholarship and its response to changing research methods. These projects hold particular interest for archivists as the key distinction between the printed form of archival and library materials is disappearing. Indeed, in an electronic environment, concepts, such as "unique" and "multiple," which have been used to distinguish archival sources from library materials, are less meaningful. It is not surprising that librarians hold differing opinions regarding the most appropriate role for libraries in the electronic environment. Some librarians argue for continuity—the continued commitment to collection devel-

[147]Morton and Price, *The ACLS Survey of Scholars: Final Report of Views on Publications, Computers, and Libraries,* 33.

[148]Lawrence Dowler, "Conference on Research Trends and Library Resources," 22–23, February 1990, unpublished draft report (Cambridge, Mass.: Harvard University, Widener Library, n.d.)

opment. Those who hold this position argue for consolidating library resources in the activities of selection and collection management and for relinquishing a role for libraries in converting source materials to electronic form. In contrast, the proponents of change claim that the continuity approach could mark the end of the era of free access to information because commercial vendors would step in to convert library materials and make them available for a fee in electronic form. The advocates of information-based institutions champion a new vision of the library without walls—an enterprise comprising many electronic libraries (including commercially produced products) that provide network access to patrons. Regardless of their perspective, both sides agree that patron demands for electronic access to library materials will be met by someone.[149] This section examines several leading projects and programs undertaken by the library community to address changes in the research environment, focusing on four new trends in professional activity: (1) promoting high-performance connectivity, (2) conversion of printed materials to machine-readable form, (3) software engineering for next-generation systems, and (4) transformations in professional roles.

Promoting Connectivity

In the last few years, library leaders have forged a new political alliance with academic computing centers and the telecommunications industry to support the development of high-performance computing networks capable of rapidly transmitting huge amounts of data and high-resolution graphics. A high-performance computing network is needed because the several thousand academic, governmental, regional, and private networks that already operate worldwide cannot transmit data and images fast enough or in large enough chunks to keep pace with the needs of scientific research. Furthermore, faster networks with higher bandwidths will expand infrastructure support for scholarly exchange of visually-oriented material (such as that required for medical research), online electronic publishing, and high-speed interchanges of text and graphics in the arts and social sciences.

Recognizing the need for infrastructures (or "highways") to disseminate materials electronically, the Association of Research Libraries (ARL) in 1990 joined with academic and administrative computing centers to form the Coalition for Networked Information (CNI). CNI is a collaboration among three distinct groups—EDUCOM, CAUSE, and the ARL—who have united to "promote the creation of and access to information resources in networked environments in order to enrich scholarship and enhance intellectual productivity."[150] The most immediate focus of the coalition's work is to establish the National Research and Education Network (NREN), a federally supported high-performance computing network. In the interim, NSFNet (a network administered by the National Science Foundation), in conjunction with the thousands of other existing networks, serves as the precursor for the future operational NREN.

The coalition is optimistic about implementing NREN as a gigabit-per-second network. In 1991, Congress passed the High Performance Computing Program that establishes the mandate for NREN. Although the original motivation for NREN emerged from the scientific community's requirements, the broader constituency rep-

[149] For two perspectives on the topic see: Stephen E. Ostrow and Robert Zich, in *Research Collections in the Information Age: The Library of Congress Looks to the Future,* edited by John Y. Cole (Washington, D.C.: Library of Congress, 1990).

[150] From Coalition for Networked Information, *Mission Statement,* March 1990.

resented by CNI envisions a network devoted to kindergarten through high school (K–12) programming, as well as leading-edge research. Indeed, EDUCOM recently designated a full-time staff position for the development of network K–12 programs. CNI's commitment is to the development of a network available to all the nation's teachers, students, and researchers.[151]

When fully implemented, NREN will allow researchers at universities, national laboratories, nonprofit institutions, government research centers, and private industry to exchange sources, communicate in real time, share preliminary findings, and disseminate publications electronically. Indeed, the dramatic changes in the ways research is conducted and information is exchanged are key factors driving the development of NREN. Through remote access hookups, NREN will provide the nation's researchers and students, regardless of the type and size of their college, with the same computing tools, data files, supercomputers, electronic libraries, specialized research facilities, and educational technology.[152] It is anticipated that NREN will support the transmittal of at least 1 billion bits of data every second by 1995.

Recognizing the impact a network with such unprecedented speed and capacity will have on their institutions, librarians have joined with other information professionals to support the development of NREN. As coalition members, librarians are participating in a range of NREN-related activities, including CNI's seven working groups on: (1) encouragement of academic publishing; (2) expansion of commercial electronic publishing; (3) development of network architectures and standards; (4) formation of proposals for legislative codes, policies, and practices; (5) organization of directories and resource information services; (6) creation of teaching and learning programs; and (7) improvement of network management and user education.

Through the activity of building a high-performance network, a new vision of the library is emerging. No longer simply a place to visit, libraries are becoming "virtual enterprises" of electronic information.[153]

Conversion

As a concrete step toward the realization of networked electronic libraries, some repositories have begun to convert to machine-readable form records originally created on paper. The American Memory Project at the Library of Congress (LC) represents a leading example of this type of effort.[154] Over the next five years, the Library of Congress, with nearly $1 million per year in congressionally appropriated funds along with private donations, will convert into electronic form large archival collections from their holdings relating to

[151]From Kenneth King, president, EDUCOM, unpublished paper presented at the "NREN Governance and Policy" session at National Net'91 Conference, 22 March 1991, Washington, D.C.

[152]See *NREN: The National Research and Education Network* (Washington, D.C.: Coalition for the National Research and Education Network, 1989).

[153]For additional information on NREN see most recent issues of *EDUCOM Review*; also, Jean Loup, *National Research and Education Network: Overview and Summary* (Washington, D.C.: Association of Research Libraries, July 1990); Charles E. Catlett, "The NSFNet: Beginnings of a National Research Internet," *Academic Computing* 3 (January 1989): 18–21, 59–64; Stephen B. Gould, "Computing and Telecommunications in the Federal Government," *CRS Review* 11 (July/August 1990): 12–15; for information on CNI, see organizational papers available from Paul Peters, CNI, 1527 New Hampshire Avenue, N.W., Washington, D.C. 20036.

[154]Additional programs now under way include the Hunt Library at Carnegie Mellon University and the Image Transmission Program at the National Agricultural Library. Other libraries are creating CD-ROMs on specialized subject areas. The Marine Corps has announced that it is compiling an online version of the Marine Corps University warfighting collection that will allow marines to "fight smart" wherever they are stationed; see Kevin M. Baerson, "Marines Put Library On-Line," *Federal Computer Week,* 5 (2 September 1991): 1, 4.

American culture and history.[155] The purpose of the project is to use advanced technology to make electronic versions of collections available to libraries across the country.

The collections chosen for the initial round of conversion primarily document aspects of turn-of-the-century life in America. They are drawn from a cross-section of original formats, including rare pamphlets, early motion pictures, sound recordings, personal papers, and still photographs. A variety of image, text, and audio types will be linked to catalog information in the standard MARC (MAchine-Readable Cataloging) format.

In fiscal year 1991, the Library of Congress prepared four collections for electronic dissemination, including about 300 broadsides from the Continental Congress and Constitutional Convention; three hours of sound recordings of speeches (sixty examples) of political leaders during World War I and the presidential election of 1920; two dozen short motion pictures of President McKinley at the start of his second term and at the 1901 Pan-American Exhibition in Buffalo, New York; and about 25,000 photographs from a well-known postcard and scenic-view company founded by William Henry Jackson. By the end of 1992, the library will supplement these with collections of Civil War photographs, approximately 350 African-American pamphlets (11,000 printed pages written between 1820 and 1910), local history books from California, early films of New York City, and life histories from the Federal Writers' Project.

The library's selection process attempts to strike a balance between popular, readily available collections and unprocessed collections that comprise a backlog arrearage. Selecting an arrearage collection provides an impetus for processing it. As selections are made, the planners consult both with Library of Congress curators and with outside scholars. The first set of American Memory collections is being evaluated in forty school, university, public, and special libraries to assess patterns of use. The results of this evaluation will provide further guidance.

Compared with all the holdings of the Library of Congress, American Memory will convert only a relatively small amount during the first few years. The program's extent reflects the high cost of conversion, the institution's desire to reduce its arrearage, and the typical difficulties encountered in the introduction of a new technology. To maximize the use of what it has prepared, however, the library is placing special emphasis on educational applications. Besides providing the collections proper, American Memory's presentation also will include introductory information in interactive, computerized form and in print.

The ultimate goal of the American Memory project is to make materials available via telecommunications, but this goal will be fully realized only in the later 1990s. Until then, the collections will be disseminated on disks: CD-ROMs for digital information and analog videodiscs for motion picture and some still photographic collections. But whether on disk or in a network, every American Memory working prototype will model what Ricky Erway, an American Memory associate coordinator, describes as a "library without walls." American Memory will be operating as a pilot project through 1995.[156]

[155]The American Memory project has received gifts from the David and Lucille Packard Foundation, the Annenberg Fund, Inc., Armand Hammer's Occidental Petroleum Corporation, and Jones International, Ltd., as well as gifts or loans of equipment from Apple Computer, IBM, and Pioneer. See Library of Congress, "American Memory," LC *Information Bulletin* (26 February 1990): 83–87.

[156]For further information on American Memory, contact the Library of Congress, Special Projects Of-

Software Engineering

Many libraries are considering ways to expand bibliographic access as part of their plans to develop next-generation library systems.[157] But few are taking as ambitious or comprehensive an approach to the process as the staff at Carnegie Mellon's University Libraries. With a $1.2 million grant from the Pew Memorial Trust and several million dollars of donated hardware from Digital Equipment Corporation, the library is developing a system that will provide the university's faculty, students, and administrators with access to bibliographic databases, full-text documents, and network gateways.[158] Library Information System II (LIS II), implemented in 1991, is designed to improve the quality of retrieval and delivery of textual information to users. In a bold departure from the standard approach to library automation, Carnegie Mellon separated its public catalog from other library administrative functions. As such, LIS II is devoted strictly to user-oriented retrieval, whereas OCLC's LS/2000, an automated system with integrated modules, is in use for other aspects of library administration.

The technical goal of LIS II is to produce for networked campuses an affordable library retrieval system that adheres to available standards. During the first phase, the system will run on University Library installed workstations. Since January 1992, LIS II has been available across campus through workstation or VT 100 access. A Macintosh interface is scheduled to be released by the end of 1992. The application goals of the current system are to provide the following:

- Online bibliographic access to all university resources
- Bibliographic access at the article level to journal literature
- Electronic access to external databases
- Online access to a range of campus information
- Online access to textual information[159]

The system's distributed architecture has been designed to support further research and development toward the realization of an electronic library.

Although the system's software supports standard bibliographic retrieval, it also provides enhanced access to select anthologies, plays, edited collections, exhibition catalogs, and conference proceedings. Several thousand bibliographic records for these

fice, Washington, D.C. 20540, (202) 707-6233. Information on the project from discussions by Avra Michelson with Ricky Erway on 28 December 1990 and Erway and Carl Fleischhauer on 4 February 1991 and from documents supplied by the Library of Congress.

[157]For the development of enhanced bibliographic records, see, for instance, Van Orden, "Content-Enriched Access to Electronic Information," 27–32; Flo Wilson, "Article-Level Access in the Online Catalog at Vanderbilt University," *Information Technology and Libraries* 8 (June 1989): 121–31; and Katharina Klemperer, "New Dimensions for the Online Catalog: The Dartmouth College Library Experience," *Information Technology and Libraries* 8 (June 1989): 138–45; the Klemperer article also discusses Dartmouth's approach to the development of an integrated campuswide information system.

[158]Information for this section is from a site visit by Avra Michelson to the University Libraries that included meetings with Thomas Michalak, Tom Dopirak, and Denise Troll on 27 March 1991. See also two reports on the work of the project: Denise A. Troll, *Library Information System II: Progress Report and Technical Plan,* Mercury Technical Report Series, no. 3 (Pittsburgh, Pa.: Carnegie Mellon University, 1990); and Nancy H. Evans et. al., *The Vision of the Electronic Library,* Mercury Technical Report Series, no. 1 (Pittsburgh, Pa.: Carnegie Mellon University, 1989). Also see William Y. Arms and Thomas J. Michalak, "Carnegie Mellon University," in *Campus Strategies for Libraries and Electronic Information,* edited by Caroline Arms (Bedford, Mass.: Digital Equipment Corporation, 1990), 243–73.

[159]The University of California at Berkeley's Office of Information Systems and Technology also is developing a campus networked information system to support bibliographic and nonbibliographic databases, full-text documents, nontextual documents, and hypermedia links.

types of publications have been embellished manually or by establishing system links with nearly one dozen commercial products that include tables of contents, title pages, and book reviews. One-page abstracts are included in the bibliographic records of campus-issued scientific and technical reports. The intent of this type of record enhancement is to improve the relevance of system retrievals.

Besides record enhancement, the staff plans to mount two types of full-text databases, journal articles and campuswide information, on the system. Elsevier, Pergamon, and the Association of Computing Machinery (ACM) have agreed to provide the University Libraries with machine-readable journals and technical reports in the subject field of computer science. ACM will provide extensive runs of four of its publications: *Computing Reviews* (ten years), *Collected Algorithms* (twenty-five years), *Communications* (two years), and *Guide to Computing Literature* (ten years). Carnegie Mellon is also negotiating an agreement to make the publications of the American Association for Artificial Intelligence available in machine-readable form, and it is working with academic research institutions to collect machine-readable computer science technical reports. Concentrating the full-text offerings in an area such as artificial intelligence and computer science will allow the University Libraries to further evaluate scholarly information needs by studying the use of textual information in a single discipline.

The University Libraries also are installing a CD-ROM jukebox system from University Microfilms, Inc. That system includes full-text images of general and business journals linked to bibliographic citations in tape-mounted databases on LIS II. In the final phase of the project, the images will be delivered to workstations across campus.

The full-text, campus-oriented documents require an indexing scheme entirely different from that developed for standard bibliographic data. The new system will provide campus software licensing and availability information, career and placement resources, the Carnegie Mellon Policies and Procedures Manual, the undergraduate catalog, user help files for other campuswide systems, listings of faculty and staff publications (including research profiles), and indexes and full text of campus newspapers. Standard office reference materials, such as phone books, encyclopedias, and dictionaries, are already available.

Development of the system's user interface is based on staff findings on user work habits and information-seeking behaviors. According to the research, patrons rarely refer to documents in isolation from other activity. For this reason, the LIS II architecture has been designed to integrate with a larger work environment, supporting linkages to word processors, databases, e-mail, and parallel applications. Toolkits (special software routines) permit LIS II users to make individual databases available across the network. Other features allow patrons to store searches for reuse, move in one keystroke from a journal article citation to the full text of the article, and improve queries by browsing indexes that reveal how often terms are used. The windowed screen environment can be customized by each user.

The creation of an electronic library linked to other electronic libraries requires sustained effort. LIS II provides in substantial measure an architecture to support full-text electronic delivery of documents in libraries. In creating this system, the developers clarified many issues and resolved other important issues in the areas of distributed storage and retrieval systems, information capture and representation, information retrieval and delivery, and management and economic concerns. Carnegie Mellon plans to make the software developed for LIS II available to other libraries.

Transformations in Professional Roles

Library literature contains many proposals for new roles for library professionals in the electronic age.[160] Among these the programmatic achievements of the Laboratory for Applied Research in Academic Information serves as one of the best operational models for redefining the librarian's role on campus. A division of the William H. Welch Medical Library at The Johns Hopkins University, the laboratory is a collaboration among academic scholars, scientists, and librarians. They share responsibility for the creation, structuring, representation, dissemination, and use of scholarly knowledge through the use of computing and communication technology. Created in 1987 by Nina W. Matheson and Richard E. Lucier, the Laboratory explores strategies for integrating the library more fully into the scholarly communication process.[161] Lucier has developed what he terms the "knowledge management model," which extends the library's traditional storage and retrieval and information transfer functions to include a third function, knowledge management.

In the knowledge management model, librarians are teamed with content specialists, software engineers, and social scientists to identify the specialized information needs of a constituency and then address the needs with the aid of information technology. In this model, the laboratory performs three types of work: (1) knowledge base and software development; (2) research and scientific support through ongoing needs assessments and quality control of data, education and training; and (3) service through the management of the computing and communications infrastructure. The social scientists assess information needs by using standard methodologies, such as participant observation, formal and unstructured interviews, and document analysis.

The laboratory recently received a three-year grant from the Council on Library Resources (CLR) to document the knowledge management model and explore the feasibility of implementing the model in nonmedical environments. The CLR funds also support an invitational symposium on knowledge management. The laboratory's key projects have been the development of the Online Mendelian Inheritance in Man (OMIM) and the Genome Data Base, which are comprehensive scientific sources used by geneticists worldwide for gene mapping, genetic disease diagnosis, and patient care. These online projects allow an international group of scientists to collect, organize, and electronically distribute mapping and disease information on approximately 100,000 genes that regulate human health and development. The constantly evolving Genome Data Base is maintained by more than one hundred scientists around the world. Lucier considers the database to be a form of dynamic, interactive publication that, unlike static print publications, always provides the most current information and analysis by the most respected scientific authorities.

[160]See, for instance, the ideas developed by Eldred Smith in his book *The Librarian, the Scholar, and the Future of the Research Library* (New York: Greenwood Press, 1990), especially 60–63 and 83–84. Articles by Bert B. Boyce and Kathleen M. Heim, "The Education of Library Systems Analysts for the Nineties," and John Corbin, "The Education of Librarians in an Age of Information Technology," in *Computing, Electronic Publishing and Information Technology: Their Impact on Academic Libraries,* edited by Robin Downes (New York: Haworth Press, 1988), 60–63 and 83–84, respectively; and Timothy C. Weiskel, "University Libraries, Integrated Scholarly Information Systems (ISIS), and the Changing Character of Academic Research," *Library Hi Tech* 6 (1988): 7–27.

[161]This section is based on briefings of Avra Michelson by Richard Lucier and Valerie Florence, 7 May 1991; see also Richard Lucier, "Knowledge Management: Refining Roles in Scientific Communication," *EDUCOM Review* 25 (Fall 1990): 21–27. For information on particular projects, see *Welch Library Issues,* vol. 2, nos. 1, 4, and 6.

Through the development of the Genome Data Base, OMIM, and other projects, the laboratory has demonstrated that knowledge management represents a "practical working alternative to existing roles and relationships in the creation and management of scholarly knowledge."[162] Lucier will expand his work in the development of the new Center for Knowledge Management at the University of California at San Francisco.

This section reviewed some of the library community's strategies. The next section recommends actions that the archival profession can take to respond to changing research methods. These actions are an important step toward confronting the transformation of scholarly practice that is as imminent as the new millennium.

CONCLUSION AND RECOMMENDATIONS

The scholarly use of information technology is resulting in dramatic changes in research practices. Essentially two trends are evident: one toward end-user computing and the other toward connectivity. To an increasing extent, social scientists and humanists are performing their own computation in the context of ever greater connectivity. The scholarly use of computers and communication technology for research and information exchange has both short-term and long-term ramifications for archival practice. In the short term, the archival profession needs to address the increasing prominence of network-mediated scholarship. In the long term, the role of the archival profession in the development of next-generation archives that operate in conjunction with global networks needs to be defined. The following recommendations suggest concrete actions the archival profession can take to address both of these issues during the next decade:

- Establish a presence on the Internet/NREN.
- Make source materials available for research use over the Internet.
- Create documentation strategies to document network-mediated scholarship and the development of research and education networks as a new communications medium.
- Develop archival methods suitable for operation with NREN.
- Take user practices and computational capacity into account in establishing policies on the management of software-dependent records.
- Recognize and reward initiatives that advance (a) the archival management of electronic records; (b) the response to scholarly use of information technology; and (c) a network-mediated archival practice.

These recommendations are considered in the three-part discussion below.

Part I: Establishing a Network-Mediated Archival Practice

The archival profession, first and foremost, must respond to the emergence of network-mediated scholarship. New methods of searching for sources, communicating with colleagues, disseminating research findings, and providing instruction suggest that scholarly communication is increasingly mediated through electronic networks. The existing Internet and the future NREN represent the new meeting ground where scholars turn for bibliographic information, scholarly dialogues and feedback, the most current publications in their fields, and high-level educational offerings. Increasingly, full-text versions of journals, magazines, newsletters, and even

[162]"Knowledge Management: A Collaboration of Academic Scholars, Scientists and Librarians," unpublished statement on the three-year project sponsored by the Council on Library Resources, The William H. Welch Medical Library, Laboratory for Applied Research in Academic Information (15 July 1990).

primary sources are available through networks. In response to this new phenomenon, the archival profession needs to establish a presence on research and education networks and to evaluate the implications of new forms of scholarly communication for standard archival practice.

But before attempting to introduce policy or collaborative action, the archival profession must start using the networks. Indeed, the use of networks is the chief action archivists can take in response to changing patterns of scholarly communication. A presence on the Internet is essential if archivists are to establish credibility as legitimate network collaborators.

Establishing an archival presence on the networks is affordable. Telecommunications hookups involve a modem, communication software, and an e-mail address provided through a link to an already existing network connection. For archivists who do not already possess a modem and who choose not to use public domain communication software, the cost entails a one-time expenditure of, at most, several hundred dollars. Ongoing connect charges in the United States are minimal. Most archivists should experience little trouble obtaining an electronic mail address because the majority of campuses are already wired for network connections, as are federal and state agencies and many private organizations and corporations, especially those affiliated with scientific research and development. In fact, several hundred archivists[163] already participate on BITNET in the network list *Archives and Archivists.* Once hardware and communication software are in place, the archival profession can become an Internet participant.

Recommendation 1: Archivists should begin monitoring and responding to scholars' intellectual activities conducted on networks.

Besides the standard methods for keeping current on research trends, archivists should participate in scholarly electronic conferences. To participate, one signs up, or "subscribes," to a conference. Because thousands of conferences exist, archivists should use conference lists and compiled directories to select those that involve subject areas most closely approximating the holdings of their repository. For instance, a repository strong in women's history sources may subscribe to the lists devoted to women's and gender studies. An institution noted for its collection of pre-Civil War holdings may choose a conference devoted to eighteenth century America. Social welfare archives may sign up for conferences related to social work, social activism, and family studies. Those with strong collections of utopian records may select the Shaker conference. Repositories noted for their holdings on the arts may join the many conferences on theater, film, and drama.[164]

One way scholars use these conferences is to exchange information about source materials related to research topics. In an effort to participate in these dialogues, NARA's Center for Electronic Records began monitoring several scholarly conferences in 1991. The conferences offer the center a forum for responding to several dozen additional inquiries each month from scholars and librarians relating to the center's holdings. One center staff member currently spends about thirty minutes each day monitoring four BITNET Listservs on topics related to government documents,

[163]As of June 1992, approximately 440 archivists subscribed to the Bitnet Archives and Archivists listserv.

[164]Examples of electronic conferences are from Kovacs, *Directory of Scholarly Electronic Conferences,* 3rd rev.

electronic data sets, social science data lists, and the Vietnam War.[165]

These conferences not only provide a means for keeping up with trends in scholarly research but also provide a mechanism for establishing a presence on the networks by attaching a name and institutional affiliation to each communication. As simplistic as this sounds, a more substantive involvement with networks can occur only when archivists are familiar with the arena's discourse and techniques and when the archival profession is established as a network participant. We therefore recommend as an initial action that archivists establish a presence on the Internet by participating in network conferences.

Recommendation 2: The archival profession should identify and implement archival methods appropriate to new forms of scholarly communication.

Establishing a presence on the networks is a necessary first step. But in addition to conference participation, the archival profession should pursue archival methods responsive to changes in scholarly communication. These new archival practices and techniques include: providing access on the Internet to source materials in machine-readable form, initially as bit-mapped images; documenting the activities of network-mediated scholarship; and establishing archives that operate in the Internet/NREN environment.

2 (a): The archival profession should make source materials available on the Internet. The archival profession should make sources directly available to scholars via research and education networks. The sources should include both records that originate in electronic form and those created in nonelectronic forms. Since the transfer of nonelectronic records to machine-readable form is a formidable undertaking, this discussion focuses primarily on conversion strategies.

Converting nonelectronic sources to machine-readable form is justified for several reasons. First, the scholarly expectation that full-text materials should be available online as a research convenience is unmistakably evident and growing.[166] Indeed, electronic document delivery represents the undisputed standard for the information field. Second, beyond convenience, conversion of source materials to machine-readable form is essential for analyses that rely on computational processing. Third, with increasing frequency, the types of questions posed by researchers require entire electronic libraries of sources, instead of a single collection, available for computational processing. From this perspective, the larger the corpus of converted collections, the greater the research value.

As further justification, in the absence of an archival role in the conversion of source materials, the commercial sector is certain to prevail. This is not to suggest that many types of conversion projects would not be more suitable as commercial sector undertakings. But as the transition to the online era proceeds, archivists have the responsibility to ensure that publicly available records remain so when converted to machine-readable form and to alert citizens to the danger of losing the right of free access through inaction.

The proposal to convert source materials to machine-readable form is neither radical nor original. Many leaders in the library profession argue that conversion is one of

[165]Conversation between Avra Michelson and Ted Hull of NARA's Center for Electronic Records, 26 August 1991; also Ted Hull, "NNXA Reference Report," NARA Center for Electronic Records, June 1991, draft.

[166]Shrinking travel allocations also may spur requests for online access, if the cost of geographically dispersed archival research exceeds academic budgets.

the most important actions librarians can take to establish a comprehensive record of scholarship.[167] As discussed earlier, some libraries are already performing pilot conversions. Further, the Commission on Preservation and Access recently released several reports recommending that preservation microfilming include the generation of digital images.[168] In an alternative approach, Cornell University Library, in conjunction with Xerox Corporation and the Commission on Preservation and Access, demonstrated the feasibility of directly converting text to digital form, avoiding the costs associated with microfilming.[169]

In arguing that archivists should convert nonelectronic holdings to machine-readable form, we are not suggesting that it is either feasible or desirable to convert all records. The volume of archival holdings is simply too great, and many holdings do not warrant the investment. Rather, our point is that it is time to begin breaking the tie with the printed past and establishing a connection with the machine-readable future.

Converting source materials to machine-readable form entails the resolution of many issues that are beyond the scope of this paper. However, we would like to comment on a few basic archival questions related to conversion: What should be converted? What electronic form should conversion result in? What kind of new descriptive devices are necessary to facilitate the independent use of electronic versions of source materials?

What should be converted? Most repositories periodically, if not regularly, microfilm deteriorating collections of enduring value. Applying current technology, microfilm preservation projects could be expanded or transformed to digital conversion projects through the development of several funded, model programs. The benefit of establishing digital conversion programs based on preservation microfilming is that many procedures in place for microfilming are also suitable for imaging. First, materials for preservation microfilming typically are selected because they are in need of preservation attention and are deserving of wider access. These two elements are adequate criteria for the current selection of collections to be digitized.[170]

[167]See, for instance, Smith, *The Librarian, the Scholar, and the Future of the Research Library,* 71–72. Clifford Lynch also recommends conversion of source materials to digital form in "Achieving the Promise: A Proposed Strategic Agenda for Libraries and Networked Information Resources in the 1990s," unpublished paper presented at the Networks for Networkers II Pre-Conference, Chantilly, Virginia 17–19 December 1990, 18 (Also published under that title in *Networks for Networkers: Critical Issues for Libraries in the National Network Environment,* edited by Barbara Evans-Markuson with Elaine W. Woods [New York: Neal-Schuman Publishers, forthcoming]).

[168]See Donald J. Waters, *From Microfilm to Digital Imagery* (Washington, D.C.: Commission on Preservation and Access, June 1991), and Michael Lesk, *Image Formats for Preservation and Access* (Washington: D.C.: Commission on Preservation and Access, July 1990). These reports explore microfilming as a means to achieve digitization.

[169]The Cornell project, co-managed by Anne R. Kenney and Lynne K. Personius, involves the direct conversion of one thousand volumes of brittle books to digital form. Half of the volumes are mathematical books, some of which are handwritten or contain formulas and graphic images. The Cornell project uses Xerox hardware that is capable of producing both digital output and enhanced print output from a digital copy. This collaborative effort has produced meaningful data on costs, procedures, and models associated with digitization programs useful to the archival profession. See Kenney and Personius, "The Future of Digital Preservation."

[170]Other categories of records also may make good candidates for conversion even though they are not deteriorating. In selecting records primarily to provide greater access, other factors should be considered, including the nature and extent of use of the records, the institutional visibility or impact afforded by the conversion, the type of image required for use, the volume and condition of the records requiring conversion, special labor costs, and the extent to which conversion can be accomplished through scanning, optical character recognition (OCR), or manual input. But we think it would be a mistake for the archival profession to expend much effort at this point on refining selection criteria until the results of a number of digital conversion projects can be analyzed. Fur-

Second, the document preparation processes used with microfilming are largely compatible with digital conversion.[171] This means that handling procedures in place for preservation microfilming can essentially be applied to digitization. Third, microfilming and digitizing can be intertwined technical processes. That is, while it is technically possible to generate a microfilm copy as output from a digitized collection, it is also possible to generate a digitized copy of a record set from microfilm output. This means that it is possible to create microfilm and then digitize the output, or digitize directly and then generate microfilm. As such, repositories concerned with the longevity of digital storage mediums, or their ability to move digital data from one generation of technology to another, can continue to rely on microfilm for preservation purposes and still convert records to machine-readable form.

Repositories that plan to microfilm are encouraged to establish pilot digital programs that draw on many structures already in place for preservation microfilming. The archival profession needs tested models to establish the most cost-effective procedures for administering ongoing conversion programs. Pilot projects should provide sufficient technical and programmatic guidance and an awareness of how digital sources are used, to equip the profession with the ability to implement large-scale digital conversions.

What electronic form should digital conversions result in? The profession's assessment of appropriate electronic forms will probably change over time. The overriding concern, however, must be to identify the kinds of representations patrons need. Do they need a facsimile image of documents? A stream of straight ASCII text that can be manipulated? ASCII text encoded with tags that identify document structures and formats? Although the electronic forms that patrons need depends on the type of research they are conducting, very little is known about the actual use of electronic documents for different types of research.

Trends in the technology suggest that in the future the archival profession should be able to provide access to electronic sources both as bit-mapped images and encoded text. But current limitations make large-scale encoding of text an unrealistic undertaking. For many reasons, the existing methods of performing ASCII conversions, manual key entry or automatic optical character recognition (OCR) are inadequate. For example, the cost of performing key entry with great volumes of materials is prohibitive, and OCR processes are unreliable with handwritten script and unusual type fonts. In contrast, bit-mapped conversions, which result in image representations (like facsimiles, but potentially of far greater resolution), are readily attainable with today's technology. Further, automatically converting bit-mapped images of modern printed documents to ASCII is typically considered a straightforward process (equivalent to OCR). If desired, encoded text can be generated from the ASCII version, provided the relevant structural information has been retained.

ASCII and encoded text differ from bit-mapped images in that the latter cannot be searched and computationally processed without considerable programming. It is highly probable, however, that software designed to encode text automatically will improve and reduce in cost during this decade. If this happens, it may be feasible to justify large-scale textual encoding. Until

ther, the Commission on Preservation and Access has contracted with Margaret Child to reconsider current criteria used to select source materials for preservation microfilming. Presumably the archival profession will find the results of this study relevant to digital conversion efforts as well.

[171]See Archival Research and Evaluation Staff, *Optical Digital Image Storage System: Project Report* (Washington, D.C.: National Archives and Records Administration, 1991), 6; and Kenney and Personius, "The Future of Digital Preservation," 9.

then, sources not amenable to OCR should be converted to bit-mapped images. But since bit-mapped images will not satisfy the research needs of certain scholars, archivists should monitor advances in OCR and structure-encoding software.

What kind of new descriptive devices are necessary to facilitate independent use by researchers of electronic versions of source materials? Digital versions of large archival collections will need specialized finding aids, descriptors, navigational aids, or informational hooks to facilitate their independent use.[172] Developing these finding aids and navigational tools represents a key challenge for the information profession. Nonetheless, it would be ill-advised to convert unstructured and voluminous collections to machine-readable form, or to make collections that originate in machine-readable form available for independent use, without addressing the need for a descriptive system suitable to the electronic environment. As a further complication, standard bibliographic approaches to retrieval are proving an inadequate method for locating and managing remote electronic text banks. But metadata, data about data that archivists typically collect about a body of records, may serve as the basis for a supplementary descriptive system to complement existing bibliographic information. Administrative histories, accession records, and other contextual data used to establish the provenance of a collection may prove very useful in retrieving information from electronic sources in the absence of human intermediaries.[173]

It is encouraging to note that contextual information accreted to each document in records originating in machine-readable form is likely to be greater than in their print counterparts. For example, e-mail messages interchanged on the Internet identify the sender and institution, the receiver(s) and institution(s), the date and time of transmittal, and the subject of the communication. Archival intervention into the design phase of software could result in the accumulation of other metadata that would be useful for both accountability and retrieval purposes. We therefore endorse the National Historical Publications and Records Commission's proposal to research the implications of capturing and retaining data, descriptive information, and contextual information in electronic form, and we speculate that the findings of this research can also advance the development of descriptive systems suitable for independent use by end-users.[174]

2 (b): Archivists should develop and implement a strategy for documenting network-mediated scholarship as a new phenomenon of scholarly communication. A key finding of this report is the substantial level of scholarly activity being conducted outside the purview of traditional archival practice. Network-mediated scholarship raises two very different but related documentation issues for the archival profession. The first is the need to document the origin and administration of research and education networks themselves. The second is the need to document the

[172]For a justification of the need for new access tools, see Clifford A. Lynch and Cecilia M. Preston, "Internet Access to Information Resources," in *Annual Review of Information Science and Technology (ARIST),* vol. 25, edited by Martha E. Williams (Amsterdam: Elsevier Science Publishers B.V., 1990); and Lynch, "Achieving the Promise," 24–25.

[173]Charles Robb, at the Kentucky Department for Libraries and Archives, is developing a locator system for statewide information using metadata to complement existing bibliographic information. This approach may be useful in developing descriptive systems that provide access to information within a collection as well. See Charles Robb, "Networking Metadata in Kentucky," unpublished paper presented to the National Association of Government Archivists and Records Administrators, Chicago, July 1990.

[174]This recommendation is part of *Research Issues in Electronic Records,* 10–11. Charles Dollar also argues that archivists should define metadata elements in his report, *The Impact of Information Technologies on Archival Principles and Methods,* 98–100.

programmatic use of these networks for the advancement of scholarship and learning. As an approach and process, documentation strategy[175] represents a tool that archivists can use to address these documentation problems—e.g., to identify the key agents operating in the network environment, to determine the universe of documentation that exists, and to develop recommendations for preserving documentation of enduring value.

The large number of agents and the global scope of activities associated with research and education networks suggests that archivists may want to collaborate and seek multi-institutional funding for documentation projects. At least three types of archival repositories are well-positioned to initiate such projects: (1) college and university archives, because network research and education efforts originate largely in academia; (2) government archives, because government is a key partner in most academic-based collaborative research projects and network-mediated education programs (either as a funder, research associate, or network administrator); and (3) discipline history centers (such as the American Institute for Physics, the Beckman Center for the History of Chemistry, and the Babbage Center), as these centers, by definition, explore a universe of documentation and are heavily devoted to science and technology, disciplines in which network-mediated scholarship is currently the most pervasive.

The documentation effort should identify key representatives to participate in strategic discussions, such as those from the Internet and scholarly communities, academic computing centers, private industry, and government research laboratories. A goal of the effort should be to clarify the principal records-creating agents and the activities that warrant preservation. The project report should include a statement on the nature of electronic archival records and the relationship of these sources to non-electronic documentation.

This recommendation involves a certain urgency because existing documentation tends to be transient. In fact, compilers of several network directories report that at least a half dozen recent scholarly electronic conferences are already defunct, as are more than a dozen electronic newsletters and journals.[176] Some argue that these efforts become inactive when moderators switch jobs and no longer possess the equipment or time to continue in that role or when the interest in a once-timely topic, such as the Gulf War, dissipates. Instead of papers removed to an attic for storage, the records of a defunct electronic conference typically take the form of a mass of bits abandoned on a campus mainframe computer or file server, awaiting a purge of the file by a systems administrator in a routine cleanup. Given this situation, academic computing staff represent key contacts for campus archivists concerned with network files. State archivists also should be concerned with the transient nature of network communication because network-mediated distance education programs are under way in most state departments of education.[177] In summary, archivists at insti-

[175]Two seminal essays that together provide an intellectual foundation for the concept of documentation strategy, as well as an examination of procedures and case examples are: Helen Willa Samuels, "Who Controls the Past," *American Archivist* 49 (Spring 1986): 109–24; and Larry Hackman and Joan Warnow-Blewett, "The Documentation Strategy Process: A Model and a Case Study," *American Archivist* 50 (Winter 1987): 12–47.

[176]Correspondence via Bitnet on 23 August 1991 between Avra Michelson and Diane Kovacs, compiler of *Directory of Scholarly Electronic Conferences;* also, a list of defunct electronic journals and newsletters appears in Michael Strangelove, *Directory of Electronic Journals and Newsletters.*

[177]See two reports by Barbara Kurshan: *Statewide Telecommunication Networks: An Overview of the Current State and the Growth Potential* (Roanoke, Va.: Educorp Consultants, December 1990), and with Marcia Harrington, *Statewide Education Networks: Survey Results* (Roanoke, Va.: Educorp Consultants, April 1991). Both are available through Bitnet from the author (Kurshan@vtvm1.bitnet).

tutions that support online scholarly communication are urged to seek funding for programs to identify and preserve valuable records related to the administration of networks comprising the Internet and network-mediated scholarship.

2 (c): The archival profession should support the development of archives designed to operate on global networks. The growth in network-mediated scholarship suggests that the archival profession needs to define its role in relation to the development of archives designed to operate in the global network environment. The need for archival operations on research and education networks is already widely recognized by the network community. For example, program planning in the network community involves archival concerns. At a biannual meeting of the Coalition for Networked Information (CNI), many subcommittees reported on work that entailed the resolution of archival functions in a network environment.[178] Although separate from the archival profession, CNI represents a group that is identifying issues related to the archiving of network resources.

Further, the development of electronic network archives is already evident. Most moderators of scholarly electronic conferences maintain an archives of the conference's transactions accessible via the network.[179] Others are capturing subject-oriented transactions across research and education networks and making the archives available on the Internet.[180] Still others are exploring commercial models for preserving both volume and breadth in network transactions.[181] Those involved in network archiving communicate with one another through electronic conferences about such issues as data compression algorithms, information filtering techniques, and file transfer protocols.[182] This means that seminal models for microarchiving within a network environment are already in place, while those for archiving on a grander scale are either on the drawing board or being prototyped, each established apart from the work of the traditional archival profession.

Archivists must not underestimate the significance of these actions. The future of the archival mission in relation to electronic communication is being defined by a set of agents wholly separate from the work of the traditional archival profession. Further, the scope of the new archival agents is apt to grow as NREN evolves into a piece of the backbone used in the conduct of official government business.[183] The appropriate role for the archival profession in this arena remains undefined, but the key questions are clear. Can the archival profession establish the political authority necessary to improve the archival methods used in conjunction with research and education network transactions, and can it rise to the

[178]Observation by Avra Michelson at the CNI spring meeting, 18–20 March 1991, Washington, D.C.

[179]Correspondence by Avra Michelson with Diane Kovacs, compiler of the *Directory of Scholarly Electronic Conferences,* on 23 August 1991 via BITNET.

[180]For instance, Edward Vielmetti at MSEN in Ann Arbor, Michigan, collects and makes available descriptions of network resources publicized on the networks [emv@msen.com]; Nathan Torkington at the Computing Services Center in Wellington, New Zealand, maintains a publicly accessible electronic archives of text on information management captured from network exchanges [gnat@kauri.vuw.ac.nz].

[181]Vielmetti has developed commercial models for archiving select network transactions.

[182]The key electronic conference where these issues are discussed is comp.archives.admin moderated by Edward Vielmetti.

[183]The U.S. Office of Personnel Management recently released guidelines already effective for the acceptance of electronic signatures in the conduct of official government business. With the issue of electronic signatures resolved, the use of networks in official government business can be expected to increase rapidly. See U.S. Office of Personnel Management, *Federal Personnel Manual,* Chapter 293, Subchapter 6, Installment 39, 1 April 1991 (Washington, D.C.: U.S. Government Printing Office, 1991).

challenge of defining an archival practice suitable not only for electronic records but also for a new communication medium?

Part II: Establishing a Strategy for the Future Usability of Electronic Records

No discussion of information technology trends can ignore the issues surrounding the storage and use of electronic records themselves. Although this subject has been discussed in the archival literature,[184] our focus here is on the scholarly research perspective. This article has concentrated on the near-term effects of information technology on current scholarly practice and products. It is equally important, however, to consider how new ways of producing records (whether they are of scholarly origin or not) will affect future users of those records. In particular, how will the creation of electronic records affect future scholars when they use such records in their research? What current technology trends bear on the ways these future scholars will perform their research and—by implication—on the ways future archives will have to serve them?

One of the main advantages of electronic information is that it is usually digital, which ensures that it can be copied and transmitted without loss or degradation. Yet, ironically, the preferred media on which this digital information is stored—disk, tape, and even CD-ROM—have far shorter shelf lives than acid-free paper or microfilm. Moreover, these media tend to become unusable long before they reach their ultimate age limits. As technology evolves, it quickly reaches a point where older media can no longer be accessed by existing equipment. It is only somewhat facetious to express this irony by saying that digital data lasts forever—or five years, whichever comes first. There is no theoretical problem with storing digital information on archival media, including microfilm, but such media are not in popular use, nor does evidence suggest that they will become so. This problem has a straightforward, though cumbersome and relatively expensive, solution: to "update" or "migrate" data, that is, to copy the data from one medium to another as media wear out or become obsolete. Although various technology trends (including the continued development of optical storage devices such as CD-ROM) may improve the longevity of media, the overall trend of continued improvement and replacement of media implies that the problem of obsolescence is unlikely to disappear in the foreseeable future.

Despite this problem, it is axiomatic that the records produced by governments, organizations, individuals, and researchers themselves will become increasingly "electronic" over the next few decades. This implies that scholars of the not-so-distant future will be confronted increasingly with electronic records as both the primary and secondary source materials for their research. Moreover, the current first generation of such records will have unique historical significance, representing the most drastic change in the form and conception of records since the introduction of print-

[184]A selection of the key literature includes David Bearman, *Archival Methods,* Archives and Museum Informatics Technical Report, no. 21 (Spring 1989); Advisory Committee for the Co-ordination of Information Systems (ACCIS) *Management of Electronic Records: Issues and Guidelines* (New York: United Nations, 1990); U.S. House, Committee on Government Operations, "Taking a Byte out of History: The Archival Preservation of Federal Computer Records," House Report 101-978 (Washington, D.C.: Government Printing Office, November 1990); *Research Issues in Electronic Records* (St. Paul, Minn.: published for the National Historical Publications and Records Commission, Washington, D.C., by the Minnesota Historical Society, 1991); David Bearman, ed., *Archival Management of Electronic Records,* Archives and Museum Informatics Technical Report no. 13 (Pittsburgh: Archives and Museum Informatics, 1991); Margaret Hedstrom, "Understanding Electronic Incunabula: A Framework for Research on Electronic Records," *American Archivist* 54 (Summer 1991): 334–54.

ing, or even of writing.[185] Yet at the current rate of technological change, electronic documents (and the programs that produce and access them) typically become obsolete and unusable in a distressingly short time. How can the loss of this unique generation of records be prevented? How will scholars be able to understand and analyze these documents decades from now? How can archives hope to preserve such documents in a form scholars will be able to use?

Furthermore, media longevity is only a part—and in many ways the easier part—of the problem. Migrating data can keep them "accessible," but to be usable they must be more than just accessible: they must also be interpretable. The data stored on digital media are simply binary digits (bits), which cannot be interpreted without a translation of the codes they represent and an understanding of the structure in which they are placed on their media. Migrating data may solve the media longevity problem, but by itself it does not solve the larger problem. Like an illiterate monk dutifully copying text in a lost language, migration may save the bits but lose their meaning. Even if we assume that the media longevity problem can be solved, what technology trends bear on whether electronic records will be interpretable in the future?

This issue is often referred to as that of *software-dependent records*, though there is somewhat more to the problem than this term suggests. Software-dependent records are electronic documents that can be read only by using some particular piece of computer software (that is, a program). Examples of software-dependent records include documents created with word processing or electronic publishing programs, spreadsheets, databases, geographic information systems (GISs), and hypertext/hypermedia. Though a data file for such a document may be saved on some medium (such as a disk), the file can be properly interpreted only by its software; the document itself is accessible (and in some cases may come into existence) only by running the software.[186] This can be thought of as the problem of "preserving" electronic documents. However, in this case, "preservation" means more than simply preserving media; unlike printed records, electronic records require software and hardware in order to be accessed and interpreted.

The obvious way to access a software-dependent document is to run the software that produced it. However, programs themselves quickly become obsolete, and running obsolete software is currently very difficult. Any given program works only on certain computers and only with certain system software. This means that accessing a document may actually require the user to run this entire hardware and software environment. In fact, what is typically meant when a document is called "software-dependent" is that it can be accessed only by running the entire hardware and software environment in which it was created. The problem is that such environments become obsolete in the blink of an archival eye, and maintaining them in working condition beyond that time is a complex, costly, and ultimately futile task.[187] Preserving elec-

[185] See Jay David Bolter, "Text and Technology: Reading and Writing in the Electronic Age," *Library Resources and Technical Services* 31 (January/March 1987): 12–23.

[186] In a very real sense, *all* electronic documents are software-dependent. Simple text and numeric files are not typically referred to as "software-dependent" only because they are encoded and stored in fairly straightforward ways that currently are considered obvious (e.g., simple sequences of ASCII codes representing characters). Yet even these cannot be accessed or interpreted without hardware and software that can understand their encoding.

[187] For several discussions of this issue, see David Bearman, *Collecting Software: A New Challenge for Archives and Museums,* Archives and Museum Informatics Technical Report no. 2 (Pittsburgh: Archives and Museum Informatics, 1987, reprinted 1990); and Coalition for Networked Information Director Paul Evan

tronic documents in a way that will allow future access to their form and meaning is therefore not straightforward.

There appear to be two general approaches to providing meaningful future access to software-dependent documents. Either they must be transformed in some way that makes them independent of the software that created them, or they must be saved along with some kind of description of their associated software sufficient to allow accessing them as was originally intended. The first approach might be facilitated by the development of standards for various kinds of documents, whereas the second approach might be facilitated by the development of formal models of computation. Several technology trends bear on each of these approaches.

Software-dependent documents might be preserved in a usable form by transforming them so that they become "software-independent" in some way. For each recognized category of program now in use (word processing, database, spreadsheet, etc.) a standard data file format might be defined, along with a standard set of functions that any such program can perform. For example, most word processing programs provide functions for displaying pages of text, footnotes, and chapter headings. In principle, a data file for a document from any such program could be transformed into some standard format, and its behavior could be duplicated by some standard program.[188] This transformation process would have to be repeated periodically as the standard itself evolved. Standardization trends such as those discussed above may help make this possible. However, there may always be programs whose behavior cannot be duplicated by any standard or which do not even fit into the recognized categories of programs (e.g., word processing or database). As noted above, standards generally lag behind the advancing technology; until computer science becomes far better formalized (that is, based on firm, theoretical underpinnings), there will always be programs that defy the most well-conceived efforts at standardization. Policies in various organizations may attempt to force the use of programs that conform to standards, but current trends of technological innovation make enforcement difficult because users find it hard to resist new capabilities, whether they are standard or not.

Even aside from standardization efforts, a "natural migration" of documents occurs as the programs on which they depend evolve through successive versions. New versions of programs often provide upward compatibility to allow old documents to migrate into the required updated forms. It may be possible, as has been suggested,[189] to rely to some extent on this phenomenon to keep documents accessible. The effectiveness of this approach, however, is limited by the fact that periodic upheavals occur in software paradigms. Two examples of

Peters, "The Machine Aspects of Preservation," unpublished paper (ca. 1990).

[188]SGML is an attempt to provide a standard for this kind of text, though it is generally recognized that even a standard for text will not magically remove all the incompatibilities among existing word processing formats. Another example of this approach that has been discussed in the literature involves relational databases. The argument has been made that a database produced by *any* relational database management system (RDBMS) can be transformed into a standard form that can be used by any other RDBMS. See the National Archives and Records Administration's response to the recommendations in "Taking a Byte Out of History," and Kenneth Thibodeau, "To Be Or Not to Be: Archives for Electronic Records," in *Archival Management of Electronic Records,* edited by David Bearman, 1–13. Although this may be true to a large extent, it is a relatively atypical example; relational database systems are one of the very few higher level applications for which a formal (mathematical) computational model exists. Most other common applications, such as word processing, spreadsheets, hypertext/hypermedia, or GISs are not nearly this well formalized.

[189]Dollar, *The Impact of Information Technologies on Archival Principles and Methods,* Chapters 1-4, draft version.

such upheavals are the change from simple textual tables to spreadsheets and the change from hierarchical databases to relational databases. Such upheavals make it difficult enough to transform documents that are crucial to the daily functioning of organizations; transforming old documents that are no longer in use may require more effort than most organizations are willing to spare.

The alternative to transforming software-dependent documents into software-independent form is to interpret them by somehow using the software that they depend on, despite its being obsolete. Interpretation does not necessarily require actually running the software. If a complete description existed of how a program interprets its data files in accessing a document, it would not be necessary to save the software itself (or its environment). The document could be accessed by following this description, effectively recreating the *behavior* of the software. In most cases, unfortunately, such complete descriptions of software exist *only* in the form of the software itself. Computer science is not yet very good at describing what complex software does.[190]

Interpreting a software-dependent document by using the software it depends on therefore requires either being able to run the software that has been saved along with the document (by effectively recreating its environment), or interpreting the software without running it (effectively recreating, or emulating, its behavior). The former option requires saving vast (though finite) documentation for the software and its environment, including detailed technical descriptions of any required hardware and all of its components.[191] The latter option requires a more sophisticated computational theory than is currently available, i.e., an understanding of the semantics of what programs do at the human level of information processing and how they do it. Without such a theory, it remains impractical to interpret software except by running it in its original hardware and software environment.[192] Current trends toward improving the formal specification of systems and environments may facilitate the former option, whereas trends toward modeling human level computational processes may facilitate the latter.[193] Finally, it should be noted that the overriding trend toward increased computational power may enable the performance of tasks that now appear unthinkable, just as we now routinely perform computations that were unthinkable a decade or two ago. Such future tasks might include automatically decoding lost file structures, transforming obsolete document formats through successive generations of standards, or recreating the behavior of ar-

[190]There are exceptions to this, such as the relational database case discussed above. In general, however, current formal descriptive techniques cannot capture the "human level" semantic behavior of programs. What is required is a computational theory, not of how programs *work*, but of what they *do* for their users; i.e., a theory of human information processing that describes such things as how humans create and use documents and how humans interact with each other to perform research.

[191]Although this is a huge task, it may not be insurmountable: These environments could not exist in the first place if they did not already possess such technical descriptions. Furthermore, many of these descriptions are already in patent or copyright offices, where they might be accessible for this purpose.

[192]Recreating the behavior of a program by figuring out what it was intended to do and building a new program that does what the original program did is sometimes called "reverse engineering." It is widely recognized as a difficult task.

[193]Advances in computational theory may enable future generations of scholars to understand how we viewed and manipulated our documents far better than we understand it ourselves. The present is, after all, only the dawn of the information age, and the organizing principles of the new "computation" paradigm are only beginning to emerge. Future scholars may have a far better formal (i.e., mathematical) understanding of computation and human information processing; this would provide them with a theoretical framework that could explain any kind of software-dependence and allow them to reconstruct past capabilities at will.

chaic computational environments from imperfect documentation. These computational possibilities may well allow future generations of scholars to derive the equivalent standard form of obsolete software-dependent documents in their archives or to reproduce the behavior of the software that produced them at will.

In the context of scholarly research and information technology, the issue of software-dependent records can be phrased in terms of two questions: "How can access to software-dependent documents be provided to future scholars?" and "What technology can help to provide this access?"

To answer the first question, one must articulate certain assumptions about what kinds of access future scholars are likely to need to such documents and what they will do with them after they have accessed them. The software used to create a software-dependent document determines the capabilities available to its author for viewing and manipulating it. How accurately must scholars be able to reproduce these capabilities? Is it enough to preserve the content of such a document without its form? Is it enough to preserve its content and form without being able to recreate the way its author saw it?[194] These questions require making assumptions about the kinds of research future scholars will perform, which can be informed by analyzing trends in scholarly practice, as undertaken above. Given such assumptions, how would alternative software-dependent records management policies constrain or enhance the capabilities of future scholars in performing their research using software-dependent documents?[195]

To answer the second question, one must articulate other assumptions about the technological future (while recognizing that all such assumptions are speculative).[196] In particular, what do current technology trends imply about future capabilities for accessing software-dependent records?

Saving data files for software-dependent documents is a necessary but insufficient step toward making them usable. As discussed above, data can be migrated to new media to keep them readable, but data must be more than just readable to be usable: They must also be interpretable. Is there some way to transform such documents before saving them in archives, so that they can be used *without* their software? If so, what would this sacrifice in terms of being able to recreate the author's original capabilities? Alternatively, is there some practical way of saving the software with each document (in particular, without maintaining obsolete hardware/software environments) so that the software itself can be used in the future to access the document? If solutions to these problems are not found and implemented soon, much of the first generation of electronic documents—representing a unique historical event in the evolution of records—will be irretrievably lost.

To summarize, there appear to be two general approaches to solving this problem, as discussed in the archival literature: Transform each document and save it in software-independent form, or save the software for each document in some way

[194]Margaret Hedstrom suggests that "The solution to preservation of electronic records lies somewhere between the present approach of preserving only data values and the need to retain all of the functionality of an active records system. There are tremendous advantages to retaining the descriptive, search, retrieval, and manipulation functions of some automated systems. The ability to retain more complex electronic records and more of the useful functionality of automated systems, however, will remain beyond the control of archivists if they continue to utilize only the tactics [that] have been employed in the past." "Archives: To Be or Not to Be: A Commentary," in *Archival Management of Electronic Records,* 28.

[195]This article raises this question without attempting to answer it. Our point is that the assumptions that underlie any answer must be made explicit.

[196]The archival literature on this subject has not yet generally articulated such assumptions.

that allows it to be used in the future to interpret the document.

The solutions that have been proposed in the literature for both approaches (e.g., translating documents into one of a few current standard forms or keeping hardware/software environments running for as long as possible) appear to be based on implicitly conservative assumptions about future technology. It seems likely, however, that inevitable advances in computational theory and computational power will produce a vastly more capable future, enabling better, longer-range solutions to one or both of these approaches. This analysis has implications for the actions that should be taken now to ensure the preservation of these records. We see the following recommendation as a necessary step toward deciding on such actions.

Recommendation 3: The archival profession should establish an evolving policy on the management of software-dependent records, informed by an assessment of the kinds of access future scholars will require to such records and a realistic assessment of the computational capabilities that will be available in the future.

Because of the short effective life of most electronic media and the rapidity with which software-dependent documents tend to become obsolete and unusable, this recommendation has an urgent aspect: Electronic records of enduring value that are not appropriately preserved will soon be lost to posterity.

The archival profession should take steps to ensure that its evolving software-dependent records management policy considers the ways that future scholars are likely to use these records and the ways that future technology is likely to facilitate this use. Assessments, such as the one we have undertaken here, which attempt to analyze trends in scholarly practice and information technology should be used to attempt to project future needs and capabilities that are *realistic*, i.e., neither wishfully grandiose nor unimaginatively chained to the past. These projections should be used to produce evolving policies aimed at the moving target that is the future.

Evolving trends in scholarly practice should be sought out by the archival profession, in an attempt to coordinate the development of archival policies with the perceptions and projections of those scholars who represent the leading edge of change in scholarly research practice. This coordination might be achieved through scheduling paper sessions or panel discussions on evolving scholarly practice, to be presented at archives and library science conferences and at conferences in various scholarly disciplines. Workshops, journals, or network discussions might also be organized on this subject, soliciting input from scholars while establishing the archival profession as a focal point for this inquiry.

Similarly, archivists should seek out evolving trends in technology, with particular emphasis on formalisms and standards for representing various kinds of documents and on formal models of computation and human information processing, which ultimately may make it possible to describe the behavior of software in ways that will allow it to be emulated in the future. In this endeavor, archivists should actively engage the computer science community as a partner, for example by organizing sessions or panels on these subjects at both computer science and archives conferences.

Finally, archivists should engage in an ongoing effort to understand the most likely future uses of software-dependent records, and they should articulate their assumptions about future scholarly practice and future computational capability as a prerequisite for proposing archival policies on the management of software-dependent records.

Part III: Recognizing and Rewarding Leadership

Recommendation 4: The archival profession should reward activities that advance archival practice with information technology, electronic records, and electronic communication.

The archival profession must respond to the changing patterns of scholarly communication and the emergence of a new communication medium. Leadership capable of guiding the archival profession should be cultivated by promoting graduate education programs, collaborative projects, and professional coalitions targeted at advancing archival operations in global network environments. The Society of American Archivists and the field's other professional associations should recognize and reward excellence in research, pilot projects, collaborative associations, and programmatic implementations related to the management of electronic records, the use of information technology to improve archival practice, and the establishment of archival methods suitable to modern communication mediums.

BIBLIOGRAPHY

Adman, Peter. "Computers and History." In *Information Technology in the Humanities: Tools, Techniques and Applications.* edited by Sebastian Rahtz, 92–103. New York: Halsted Press, 1987.

Advisory Committee for the Coordination of Information Systems. *Management of Electronic Records: Issues and Guidelines.* New York: United Nations, 1990.

Alberico, Ralph. "The Development of an 'Information Superhighway'." *Computers in Libraries* 10 (January 1990): 33–35.

Alexander, Michael. "But Can You Read It Like a Book?" *Computerworld* 24 (19 November 1990): 18.

Anderson, Paul A. "Using Artificial Intelligence to Understand Decision Making in Foreign Affairs: The Problem of Finding an Appropriate Technology." In *Artificial Intelligence and National Security,* edited by Stephen J. Cimbala, 133–46. Lexington, Mass.: Lexington Books, 1987.

Archival Research and Evaluation Staff. *Optical Digital Image Storage System: Project Report.* Washington, D.C.: National Archives and Records Administration, 1991.

Arms, Caroline, ed. *Campus Strategies for Libraries and Electronic Information.* Bedford, Mass.: Digital Press, 1990.

———. "Libraries and Electronic Information: The Technological Context, Part Two." *EDUCOM Review* 24 (Fall 1989): 34–43.

Arms, William Y., and Thomas J. Michalak. "Carnegie Mellon University." In *Campus Strategies for Libraries and Electronic Information*, edited by Caroline Arms, 243–73. Bedford, Mass.: Digital Equipment Corporation, 1990.

Association for Computers and the Humanities, the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing. *Proposal for Funding for an Initiative to Formulate Guidelines for the Encoding and Interchange of Machine-Readable Text.* Prepared for the National Endowment for the Humanities, 1988.

Baerson, Kevin M. "Marines Put Library Online." *Federal Computer Week* 5 (2 September 1991): 1, 4.

Bailey, Charles W., Jr. "Electronic (Online) Publishing in Action . . . *The Public-Access Computer Systems Review* and Other Electronic Serials." *Online* 15 (January 1991): 28–35.

———. "Intelligent Multimedia Computer Systems: Emerging Information Resources in the Network Environment." *Library Hi Tech* 8 (1990): 29–41.

Bantz, David A. "The Values of the Humanities and the Values of Computing." In *Humanities and the Computer: New Directions*, edited by David S. Miall, 27–37. Oxford: Clarendon Press, 1990.

Bearman, David. *Archival Methods*. Archives and Museum Informatics Technical Report no. 21. Pittsburgh: Archives and Museum Informatics, 1989.

———, ed. *Archival Management of Electronic Records.* Archives and Museum Informatics Technical Report no. 13. Pittsburgh: Archives and Museum Informatics, 1991.

———. *Collecting Software: A New Challenge for Archives and Museums.* Archives and Museum Informatics Technical Report no. 2. Pittsburgh: Archives and Museum Informatics, 1987, reprinted 1990.

———. *Towards National Information Systems for Archives and Manuscript Repositories: The NISTF Papers.* Chicago: Society of American Archivists, 1987.

———. "User Presentation Language in Archives." *Archives and Museum Informatics* 3 (Winter 1989–90): 3–7.

Bell, Rudolph M. "User Perspectives and Requirements: Creator of Non-Bibliographic Databases Has to Share with Others." Unpublished paper presented at the Library of Congress Network Advisory Committee Meeting, Washington, D.C., 29–31 March 1989.

Bikson, T. K., and S. A. Law. *Electronic Mail Use at the Bank: A Survey and Recommendations.* Washington, D.C.: World Bank, Information, Technology, and Facilities Department, September 1991.

———. *Electronic Information Media and Records Management Methods: A Survey of Practices in United Nations Organizations.* A RAND Note (N-3453-RC). Santa Monica, Calif.: RAND Corporation, 1991.

Blake, Virgil L. P., and Rene Tjoumas, eds. *Information Literacies for the Twenty-First Century.* Boston: G.K. Hall, 1990.

Bolter, Jay David. "Text and Technology: Reading and Writing in the Electronic Age." *Library Resources and Technical Services* 31 (January/March 1987): 12–23.

Booker, Ellis. "Vive le Minitel." *Telephony* 215 (8 August 1988): 24.

Bortnick, Jane. "Information Technology Overview." *CRS Review* 11 (July/August 1990): 1–3.

Bourcier, Danielle. "Interpretation in Legal Expert Systems: From Private Conviction to the Formalisation of Rules." In *Interpretation in the Humanities: Perspectives from Artificial Intelligence,* edited by Richard Ennals and Jean-Claude Gardin, 215–28. Library and Information Research Report no. 71. Cambridge: University Press, 1990.

Bourlet, Caroline, and Jean-Luc Minel. "A Declarative System for Setting Up a Prosopographical Database." In *History and Computing,* edited by Peter Denley and Deian Hopkin, 186–91. Manchester, England: Manchester University Press, 1987.

Boyce, Bert B., and Kathleen M. Heim. "The Education of Library Systems Analysts for the Nineties." In *Computing, Electronic Publishing and Information Technology: Their Impact on Academic Libraries*, edited by Robin Downes, 69–76. New York: Haworth Press, 1988.

Brent, Edward. "Is There a Role for Artificial Intelligence in Sociological Theorizing?" *American Sociologist* 19 (Summer 1988): 158–66.

———. "New Approaches to Expert Systems and Artificial Intelligence Programming." *Social Science Computer Review* 6 (Winter 1988): 569–78.

Brilliant, Richard. "How an Art Historian Connects Art Objects and Information." *Library Trends* 37 (Fall 1988): 120–29.

Brown, Thomas E. "The Society of American Archivists Confronts the Computer." *American Archivist* 47 (Fall 1984): 366–82.

Brunner, Theodore F. "Data Banks for the Humanities: Learning from Thesaurus Linguae Graecae." *Scholarly Communication* 7 (Winter 1987): 1, 6–9.

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

Carpenter, Edwin H., and Rick D. Axelson. "Statistical and Graphical Research Methods: State of the Art." *Social Science Computer Review* 7 (Winter 1989): 503–34.

Carvalho, Joaquim. "Expert Systems and Community Reconstruction Studies." In *History and Computing II*, edited by Peter Denley, Stefan Fogelvik, and Charles Harvey, 97–102. Manchester: Manchester University Press, 1989.

Casetti, E., S. Amarel, A. Blumstein, R. H. Day, A. H. Levis, S. R. Rosenthal, J. Sacks, and C. Werner. *Regarding the Feasibility and Desirability of Conferences on "The Methodological Research Frontiers and the Social Sciences."* NSF Award No. OJR-8406230. Final Report to the National Science Foundation, 10 September 1986.

Catlett, Charles. "The NSFNet: Beginnings of a National Research Internet." *Academic Computing* 3 (January 1989): 18–21, 59–64.

Cimbala, Stephen J., ed. *Artificial Intelligence and National Security.* Lexington, Mass.: Lexington Books, 1987.

Clement, John R. B. "Increasing Research Productivity Through Information Technology: A User-Centered Viewpoint." Version October 19, 1989.10.17 prepared for publication.

Cline, Nancy. "Information Resources and the National Network." *EDUCOM Review* 25 (Summer 1990): 30–34.

Coalition for Networked Information. *Mission Statement* (March 1990).

Conklin, Jeff. "Hypertext: An Introduction and Survey." In *Computer-Supported Cooperative Work: A Book of Readings*, edited by Irene Greif, 423–75. San Mateo, Calif.: Morgan Kaufmann Publishers, 1988.

Conway, Paul. "Facts and Frameworks: An Approach to Studying the Users of Archives." *American Archivist* 49 (Fall 1986): 393- 407.

———. *Partners in Research: Towards Enhanced Access to the Nation's Archives.* Washington, D.C.: National Archives and Records Administration, forthcoming.

———. "Research in Presidential Libraries: A User Study." *Midwestern Archivist* 11, no. 1 (1986): 35–36.

Corbin, John. "The Education of Librarians in an Age of Information Technology." In *Computing, Electronic Publishing and Information Technology: Their Impact on Academic Libraries*, edited by Robin Downes, 77–87. New York: Haworth Press, 1988.

Council on Library Resources. "Communications in Support of Science and Engineering, A Report to the National Science Foundation from the Council on Library Resources." Washington, D.C.: Council on Library Resources, August 1990.

———. "Setting Library Policies and Priorities in Research Universities: A Special CLR Grant Program," n.d.

Cunnyngham, Jon. "Group to Form Foundation for Electronic Publishing." Unpublished paper. Columbus, Ohio: Foundation for Electronic Publishing, ca. 1989.

Dain, Phyllis, and John Y. Cole, eds. *Libraries and Scholarly Communication in the United States: The Historical Dimension.* Beta Phi Mu Monograph, no. 2. New York: Greenwood Press, 1990.

David, Paul A., and W. Edward Steinmueller. *The Impact of Information*

Technology Upon Economic Science. CEPR Publication no. 121. Stanford, Calif.: Center for Economic Policy Research, December 1987.

Davis, M., S. Rosenschein, and N. Shapiro. *Prospects and Problems for a General Modeling Methodology.* N-1801-RC. Santa Monica, Calif.: The RAND Corporation, June 1982.

Davis, Paul K. "A New Analytic Technique for the Study of Deterrence, Escalation Control, and War Termination." In *Artificial Intelligence and National Security,* edited by Stephen J. Cimbala, 35–60. Lexington, Mass.: Lexington Books, 1987.

DeLoughry, Thomas J. "Scholarly Journals in Electronic Form Seen as Means to Speed Pace of Publication and Promote Dialogue." *Chronicle of Higher Education* 35 (22 March 1989): A11, A16.

Denley, Peter. "The Computer Revolution and 'Redefining the Humanities'." In *Humanities and the Computer: New Directions,* edited by David S. Miall, 13–25. Oxford, England: Clarendon Press, 1990.

Denley, Peter, and Deian Hopkin, eds. *History and Computing.* Manchester, England: Manchester University Press, 1987.

Denley, Peter, Stefan Fogelvik, and Charles Harvey, eds. *History and Computing II.* Manchester, England: Manchester University Press, 1989.

"Despite Budget Cuts, Campuses Attempt to Maintain Computing Services." USC Center for Scholarly Technology Newsletter. (October 1991).

Dillon, Andrew, John Richardson, and Cliff McKnight. "The Effects of Display Size and Text Splitting on Reading Lengthy Text from Screen." *Behaviour and Information Technology* 9 (May–June 1990): 215–27.

———. "Human Factors of Journal Usage and Design of Electronic Texts." *Interacting with Computers* 1 (1989): 183–89.

Dollar, Charles. *The Impact of Information Technologies on Archival Principles and Methods.* Macerata, Italy: University of Macerata Press, 1992.

Doran, Jim. "A Distributed Artificial Intelligence Reading of Todorov's *The Conquest of America: The Question of the Other*, by Tzvetan Todorov, 1985." In *Interpretation in the Humanities: Perspectives from Artificial Intelligence*, Library and Information Research Report no. 71, edited by Richard Ennals and Jean-Claude Gardin, 143–68. London: British Library Board, 1990.

Doty, Philip, Ann Peterson Bishop, and Charles R. McClure. "The National Research and Education Network (NREN): An Empirical Study of Social and Behavioral Issues." In *Information in the Year 2000: From Research to Applications* [proceedings of the 53rd American Society for Information Science Annual Meeting], edited by Diane Henderson, 284–99. Medford, N.J.: Learned Information, 1990.

Dowler, Lawrence. "Availability and Use of Records: A Research Agenda." *American Archivist* 51 (Winter/Spring 1988): 74–86.

———. "Conference on Research Trends and Library Resources, February 22–23, 1990." Unpublished draft report. Cambridge, Mass.: Harvard University, Widener Library, n.d.

Downes, Robin, ed. *Computing, Electronic Publishing and Information Technology: Their Impact on Academic Libraries.* New York: Haworth Press, 1988.

Dyer, Alan, and Kate Milner. "An Examination of Hypertext as an Authoring Tool in Art and Design Education." In *Humanities and the Computer: New Directions*, edited by David S. Miall, 137–48. Oxford: Clarendon Press, 1990.

Ehringhaus, Michael. *The Electronic Jour-*

nal: Promises and Predicaments. Technical Report no. 3, ED316732. Syracuse, N.Y.: Syracuse University, School of Education, February 1990.

Engelbart, D. C., and W. K. English. "A Research Center for Augmenting Human Intellect." *American Federation of Information Processing Societies Proceedings* (Fall Joint Computer Conference) 33 (May 1974): 395–410.

Ennals, Richard. *Artificial Intelligence: Applications to Logical Reasoning and Historical Research.* New York: Halsted Press, 1985.

Ennals, Richard, and Jean-Claude Gardin, eds. *Interpretation in the Humanities: Perspectives from Artificial Intelligence.* Library and Information Research Report no. 71. Cambridge: British Library Board, 1990.

Evans, Nancy H., Denise A. Troll, Mark H. Kibbey, Thomas J. Michalak, and William Y. Arms. *The Vision of the Electronic Library.* Mercury Technical Report Series, no. 1, Pittsburgh: Carnegie Mellon University, 1989.

Fargues, Jean. "Remarks on the Interrelationships Between Artificial Intelligence, Mathematical Logic and Humanities." In *Interpretation in the Humanities: Perspectives from Artificial Intelligence*, Library and Information Research Report 71, edited by Richard Ennals and Jean-Claude Gardin, 265–83. Cambridge: University Press, 1990.

Fayen, Emily Gallup. "Loading Local Machine-Readable Data Files: Issues, Problems, and Answers." *Information Technology and Libraries* 8 (June 1989): 132–37.

Freeman, Elsie T. "In the Eye of the Beholder: Archives Administration from the User's Point of View." *American Archivist* 47 (Spring 1984): 111–23.

Friedman, Edward A., James E. McClellan III, and Arthur Shapiro. "Introducing Undergraduate Students to Automated Text Retrieval in Humanities Courses." In *Humanities and the Computer: New Directions*, edited by David S. Miall, 103–12. Oxford: Clarendon Press, 1990.

"The Future of Research Libraries." Unpublished minutes of conference sponsored by the American Academy of Arts and Sciences, 8 June 1990.

Gardin, Jean Claude. *Artificial Intelligence and Expert Systems: Case Studies in the Domain of Archaeology.* New York: Halsted Press, 1988.

———. "The Future Influence of Computers on the Interplay Between Research and Teaching in the Humanities." *Humanities Communication Newsletter* 9 (1987): 17–21.

Garson, G. David. "The Role of Inductive Expert Systems Generators in the Social Science Research Process." *Social Science Microcomputer Review* 5 (Spring 1987): 11–24.

Gaunt, Marianne, and Robert Hollander. "A Proposed U.S. National Center for Machine-Readable Texts in the Humanities." *ACH Newsletter* 11 (Fall 1989): 14–15.

Gillikin, David P. "Document Delivery from Full-Text Online Files: A Pilot Project." *Online* 14 (May 1990): 27–32.

Gordon, Ann D. *Using the Nation's Documentary Heritage: The Report of the Historical Documents Study.* Washington, D.C.: National Historical Publications and Records Commission, 1992.

Gould, Constance C. *Information Needs in the Humanities: An Assessment.* Stanford, Calif.: The Research Libraries Group, 1988.

Gould, Constance C., and Mark Handler. *Information Needs in the Social Sciences: An Assessment.* Mountain View, Calif.: Research Libraries Group, 1989.

Gould, Stephen B. "Computing and Telecommunications in the Federal Government." *CRS Review* 11 (July/August 1990): 12–15.

Graham, Peter S. "Research Patterns and Research Libraries: What Should

Change?" *College and Research Libraries* 50 (July 1989): 433–40.

Grand Challenges: High Performance Computing and Communications, The FY 1992 U.S. Research and Development Program. 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.

Green, Kenneth C. "A Technology Agenda for the 1990s." *Change* 23 (January/February 1991): 6–7.

Green, Kenneth C., and Skip Eastman. *Campus Computing 1990.* Los Angeles: University of Southern California, Center for Scholarly Technology, 1990.

Greenberg, Saul. "An Annotated Bibliography of Computer Supported Cooperative Work." *SIGCHI Bulletin* 23 (July 1991): 29–62.

Greif, Irene, ed. *Computer-Supported Cooperative Work: A Book of Readings.* San Mateo, Calif.: Morgan Kaufmann Publishers, 1988.

Hackman, Larry and Joan Warnow-Blewett. "The Documentation Strategy Process: A Model and a Case Study." *American Archivist* 50 (Winter 1987): 12–47.

Harnad, Stevan. "Scholarly Skywriting and the Prepublication Continuum of Scientific Inquiry." *Psychological Science* 1 (November 1990): 342–44.

Hedstrom, Margaret. *Archives and Manuscripts: Machine-Readable Records.* SAA Basic Manual Series. Chicago: Society of American Archivists, 1984.

———. "Archives: To Be or Not to Be—A Commentary." In *Archival Management of Electronic Records*, edited by David Bearman, 25–30. Archives and Museum Informatics Technical Report no. 13. Pittsburgh: Archives and Museum Informatics, 1991.

———. "Understanding Electronic Incunabula: A Framework for Research on Electronic Records." *American Archivist* 54 (Summer 1991): 334–54.

Heim, Kathleen M. "Social Scientific Information Needs for Numeric Data: The Evolution of the International Data Archive Infrastructure." *Collection Management* 9 (Spring 1987): 1–53.

Hewitt, Joe A. "Altered States: Evolution or Revolution in Journal-Based Communications?" *American Libraries* 20 (June 1989): 497–500.

Hickerson, Thomas H. *Archives and Manuscripts: An Introduction to Automated Access.* SAA Basic Manual Series. Chicago: Society of American Archivists, 1981.

Hockey, Susan. "An Historical Perspective." In *Information Technology in the Humanities: Tools, Techniques and Applications,* edited by Sebastian Rahtz, 20–30. Chichester, England: Ellis Horwood Limited, 1987.

Hockey, Susan, Jo Freedman, and John Cooper. "Computers in the Study of Set Texts." In *Humanities and the Computer: New Directions*, edited by David S. Miall, 113–22. Oxford: Clarendon Press, 1990.

Hudson, Valerie M. "Using a Rule-Based Production System to Estimate Foreign Policy Behaviour: Conceptual Issues and Practical Concerns." *Artificial Intelligence and National Security*, edited by Stephen J. Cimbala, 109–31. Lexington, Mass.: Lexington Books, 1987.

Hunter, Karen. "Academic Librarians and Publishers: Customers Versus Producers or Partners in the Planning of Electronic Publishing." In *Computing, Electronic Publishing and Information Technology: Their Impact on Academic Libraries*, edited by Robin Downes, 35–47. New York: Haworth Press, 1988.

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

———. "Very Large Neural Networks for

Word Sense Disambiguation." Unpublished paper presented at European Conference on Artificial Intelligence, Stockholm, August 1990.

Johnson, J., 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.

Joseph, Earl C. "Twenty-First Century Information Literacies and Libraries." In *Information Literacies for the Twenty-First Century*, edited by Virgil L. P. Blake and Renee Tjoumas, 7–16. Boston: G.K. Hall, 1990.

Kahn, Robert E., 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*, Reston, Va.: Corporation for National Research Initiatives, March 1988.

Kaufman, Paula, and Angie LeClercq. "Archiving Electronic Journals: Who's Responsible for What." *Information Issues* 2 (Fall 1990).

Kenney, Anne R., and Lynne K. Personius. "The Future of Digital Preservation." In *Advances in Preservation and Access*, Volume 1, edited by Barbara Buckner Higginbotham and Mary E. Jackson. Westport, Conn.: Meckler Press, 1992.

Klemperer, Katharina. "New Dimensions for the Online Catalog: The Dartmouth College Library Experience." *Information Technology and Libraries* 8 (June 1989): 138–45.

Kovacs, Diane K. *Directory of Scholarly Electronic Conferences.* 3rd ed. Kent, Ohio: Kent State University Libraries, August 1991. (Electronic document available via Internet.)

Kurshan, Barbara. *Statewide Telecommunication Networks: An Overview of the Current State and the Growth Potential.* Roanoke, Va.: Educorp Consultants, December 1990.

Kurshan, Barbara, and Marcia Harrington. *Statewide Education Networks: Survey Results.* Roanoke, Va.: Educorp Consultants, April 1991.

Lancaster, F. W. "Electronic Publishing." *Library Trends* 37 (Winter 1989): 316–25.

Larsen, Ronald L. "Survey and Feasibility Study to Develop Pilot Project Proposal for Distributed Electronic Federal Depository." Unpublished proposal to Coalition for Networked Information, 29 October 1990.

Leavitt, Michael R. *The Analyst and Technology—2000.* Prepared for the U.S. Intelligence Research and Development Council, January 1991.

Lehmann, Stephen, and Patricia Renfro. "Humanists and Electronic Information Services: Acceptance and Resistance." *College and Research Libraries* 52 (September 1991): 409–13.

Lesk, Michael. *Image Formats for Preservation and Access.* Report of the Technology Assessment Advisory Committee to the Commission on Preservation and Access. Washington, D.C.: The Commission, July 1990.

Lewis, David W. "Economics of the Scholarly Journal." *College and Research Libraries* 50 (November 1989): 674–88.

Library of Congress. "American Memory." *Library of Congress Information Bulletin* (26 February 1990): 83–87.

Lindquist, Christopher. "Ferret Lovers Unite and Download." *Computerworld* 25 (12 August 1991): 1, 88.

Loughridge, Brendan. "Information Technology, the Humanities and the Library." *Journal of Information Science* 15 (July–September 1989): 277–86.

Loup, Jean. *National Research and Education Network: Overview and Summary.* Washington, D.C.: Association of Research Libraries, July 1990.

Lucier, Richard E. "Knowledge Management: Refining Roles in Scientific Communication." *EDUCOM Review* 25 (Fall 1990): 21–27.

Lynch, Clifford A. "Achieving the Promise: A Proposed Strategic Agenda for Libraries and Networked Information Resources in the 1990s." In *Networks for Networkers: Critical Issues for Libraries in the National Network Environment*, edited by Barbara Evans-Markuson with Elaine W. Woods. New York: Neal-Schuman Publishers, forthcoming.

———. "The Growth of Computer Networks: A Status Report." *Bulletin of the American Society for Information Science* 16 (June/July 1990): 10–11.

———. "Library Automation and the National Research Network." *EDUCOM Review* 24 (Fall 1989): 21–26.

Lynch, Clifford A., and Cecilia M. Preston. "Evolution of Networked Information Resources." *Proceedings of the Twelfth National Online Meeting, May 7–9, 1991, New York, N.Y.,* edited by Martha E. Williams, 221–30. Medford, N.J.: Learned Information, 1991.

———. "Internet Access to Information Resources." In *Annual Review of Information Science and Technology (ARIST)*, vol. 25, edited by Martha E. Williams, chapter 7. Amsterdam: Elsevier Science Publishers B.V., 1990.

Lytle, Richard H. "An Analysis of the Work of the National Information Systems Task Force." *American Archivist* 47 (Fall 1984): 357–65.

MacCafferty, Maxine, and Kathleen Gray, eds. *The Analysis of Meaning: Informatics 5.* London: Queens College Oxford, 1979.

Maher, William J. "The Use of User Studies." *Midwestern Archivist* 11, no. 1 (1986): 15–26.

Margulius, David L. "C'est la France, C'est Minitel." *PC-Computing* 2 (January 1989): 194–96.

Martin, Susan K. "Library Management and Emerging Technology: The Immovable Force and the Irresistible Object." *Library Trends* 37 (Winter 1989): 374-82.

McAdams, Alan, Thomas Vietorisz, William Dougan, and James Lombardi. "Economic Benefits and Public Support of a National Education and Research Network." *EDUCOM Bulletin* 23 (Summer/Fall 1988): 63–71.

McKnight, Cliff. "Using the Electronic Journal." In *Scholarly Communication and Serials Prices: Proceedings of a Conference Sponsored by the Standing Conference of National and University Libraries and the British Library Research and Development Department, 11–13 June 1990,* edited by K. Brookfield. New York: Bowker-Saur, 1991.

McKnight, Cliff, John Richardson, and Andrew Dillon. "Hypertext Authoring: Some Basic Issues." *Humanities Communication Newsletter* 11 (1989): 25–29.

McLean, Alice, Musick, Robert Morrissey, and Donald A. Ziff. "ARTFL: A New Tool for French Studies." *Scholarly Communication* 8 (Spring 1987): 1, 6–9.

Matters, Marion, ed. *Automated Records and Techniques in Archives: A Resource Directory.* Chicago: Society of American Archivists, 1990.

Miall, David. "An Expert System Approach to the Interpretation of Literary Structure." In *Interpretation in the Humanities: Perspectives from Artificial Intelligence*, Library And Information Research Report no. 71, edited by Richard Ennals and Jean-Claude Gardin, 196–214. Cambridge: University Press, 1990.

Miall, David S., ed. *Humanities and the Computer: New Directions.* Oxford: Clarendon Press, 1990.

Michelson, Avra. *Expert Systems Technology and Its Implications for Archives.* National Archives Technical Information Paper, no. 9. Washington, D.C.: National Archives and Records Administration, March 1991.

———. "Forecasting the Use of NREN by

Humanities Scholars." Paper presented at the panel "New Constituencies for the NREN," 27 March 1992, National NET '92, Washington, D.C.

Mills, William deB. "Rule-Based Analysis of Sino-Soviet Negotiations." *Social Science Computer Review* 8 (Summer 1990): 181–95.

"Mission, Themes, and Strategies." Washington, D.C.: Coalition for Networked Information, 1 June 1990.

Morris, Dilys E. "Electronic Information and Technology: Impact and Potential for Academic Libraries." *College and Research Libraries* 50 (January 1989): 56–64.

Morton, Herbert C., and Anne J. Price. *The ACLS Survey of Scholars: Final Report of Views on Publications, Computers, and Libraries.* Washington, D.C.: American Council of Learned Societies, Office of Scholarly Communication and Technology, 1989.

National Archives and Records Administration. "Management, Preservation and Access for Electronic Records with Enduring Value." [Response to recommendations in *Taking a Byte Out of History: The Archival Preservation of Federal Computer Records* (Twenty-fifth report by the Committee on Government Operations: House Report 101-978)]. Unpublished paper 1 July 1991.

———. *Optical Digital Image Storage System: Project Report.* Washington, D.C.: National Archives and Records Administration, Archival Research and Evaluation Staff, March 1991.

"The National Research and Education Network: A Policy Paper." rev. Washington, D.C.: EDUCOM Networking and Telecommunications Task Force, March 1990.

Naugler, Harold. *The Archival Appraisal of Machine-Readable Records: A RAMP Study with Guidelines.* Paris: General Information Programme and UNISIST, Unesco, 1984.

Nichol, Jon, Jackie Dean, and Jonathan Briggs. "Logic Programming and Historical Research." In *History and Computing*, edited by Peter Denley and Deian Hopkin, 198–205. Manchester, England: Manchester University Press, 1987.

Nora, Simon, and Alain Minc. *The Computerization of Society: A Report to the President of France.* Cambridge, Mass.: MIT Press, 1980.

North Carolina State University, The Libraries. "The North Carolina State University Libraries and the National Agricultural Library Joint Project on Transmission of Digitized Images: Improving Access to Agricultural Information." Unpublished paper, 1990.

NREN: The National Research and Education Network. Washington, D.C.: Coalition for the National Research and Education Network, 1989.

Olsen, Mark, and Louis-Georges Harvey. "Computers in Intellectual History: Lexical Statistics and Analysis of Political Discourse." *Journal of Interdisciplinary History* 18 (Winter 1988): 449–64.

Osburn, Charles B. "The Structuring of the Scholarly Communication System." *College and Research Libraries* 50 (May 1989): 277–86.

Ostrow, Stephen E., and Robert Zich. In *Research Collections in the Information Age: The Library of Congress Looks to the Future*, edited by John Y. Cole. Washington, D.C.: Library of Congress, 1990.

Pagell, Ruth A. "Primary FTDBs for the End User: New Roles for the Information Professional." *Online Review* 13 (April 1989): 143–54.

Pemberton, Jeffrey K. "*Online* Interviews Harry Boyle on CAS's New License Policy . . . Effects on Searching/Prices." *Online* 12 (March 1988): 19–25.

Perritt, Henry H., Jr. "Government Information Goes On-Line." *Technology Re-*

view 92 (November/December 1989): 60–67.

Peters, Paul Evan. "The Machine Aspects of Preservation." Unpublished paper, ca. 1990.

Potter, Rosanne G., ed. *Literary Computing and Literary Criticism: Theoretical and Practical Essays on Theme and Rhetoric.* Philadelphia: University of Pennsylvania Press, 1989.

Pugh, Mary Jo. "The Illusion of Omniscience: Subject Access and the Reference Archivist." *American Archivist* 45 (Winter 1982): 33–44.

Rahtz, Sebastian, ed. *Information Technology in the Humanities: Tools, Techniques and Applications.* Chichester, England: Ellis Horwood Limited, 1987.

Research Issues in Electronic Records. St. Paul, Minn.: Published for the National Historical Publications and Records Commission, Washington, D.C., by the Minnesota Historical Society, 1991.

Robb, Charles. "Networking Metadata in Kentucky." Unpublished paper presented to the National Association of Government Archivists and Records Administrators, Chicago, July 1990.

Roberts, Michael M. "Implementing the NREN—The Prologue Is Over." *EDUCOM Review* 25 (Fall 1990): 9–10.

Rogers, Gil. "Teaching a Psychology Course by Electronic Mail." *Social Science Computer Review* 7 (Spring 1989): 60–64.

Rogers, Sharon J., and Charlene S. Hurt. "How Scholarly Communication Should Work in the 21st Century." *Chronicle of Higher Education* 36 (18 October 1989): A56.

Rothenberg, J. "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 and Sons, August 1989.

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

Ruller, Tom. "Managing and Appraising GIS Data: Issues and Strategies." Unpublished paper presented at the 1991 Society of American Archivists annual meeting, Philadelphia.

Samuels, Helen Willa. "Who Controls the Past" *American Archivist* 49 (Spring 1986): 109–24.

Sarin, Sunil, and Irene Greif. "Computer-Based Real-Time Conferencing Systems." *Computer-Supported Cooperative Work: A Book of Readings*, edited by Irene Greif, 397–420. San Mateo, Calif.: Morgan Kaufmann Publishers, 1988.

Schick, James B. M. *Teaching History with a Computer: A Complete Guide.* Chicago: Lyceum Books, 1990.

Schofield, Janet Ward, Debra Evans-Rhodes, and Brad R. Huber. "Artificial Intelligence in the Classroom: The Impact of a Computer-Based Tutor on Teachers and Students." *Social Science Computer Review* 8 (Spring 1990): 24–41.

Scholarly Communication: Notes on Publishing, Library Trends, and Research in the Humanities, nos. 1–8 (June 1985–Spring 1987).

Schrodt, Philip A. "Pattern Matching, Set Prediction, and Foreign Policy Analysis." In *Artificial Intelligence and National Security*, edited by Stephen J. Cimbala, 89–107. Lexington, Mass.: Lexington Books, 1987.

Schurer, Kevin. "Artificial Intelligence and the Historian—Prospects and Possibilities." In *Interpretation in the Humanities: Perspectives from Artificial Intelligence*, Library and Information Research Report no. 71, edited by Richard Ennals and Jean-Claude Gardin, 169–95. Cambridge: University Press, 1990.

Schurer, Kevin, Jim Oeppen, and Roger Schofield. "Theory and Methodology: An Example from Historical Demography." In *History and Computing II*, edited by Peter Denley, Stefan Fogelvik, and Charles Harvey, 130–42. Manchester: Manchester University Press, 1989.

Schwartz, Karen D. "Agencies Use Software to Dig Up Links Among Data." *Government Computer News* 19 (15 October 1990): 60.

Seiler, Lauren H. "The Future of the Scholarly Journal." *Modern Language Journal* 74 (1990): 1–9.

Shaughnessy, Thomas W. "Scholarly Communication: The Need for an Agenda for Action—A Symposium." *Journal of Academic Librarianship* 15 (May 1989): 68–78.

Shaw, J. C. *JOSS: Conversations with the Johnniac Open-Shop System.* P-3146. RAND Corporation, 1965.

Shaw, J. C. "JOSS: A Designer's View of an Experimental On-Line Computing System." In *American Federation of Information Processing Societies Conference Proceedings* 26 (1964 Fall Joint Computer Conference), 455–64. Baltimore, Md.: Spartan Books, 1964.

Shuman, Bruce A. *The Library of the Future: Alternative Scenarios for the Information Profession.* Englewood, Colo.: Libraries Unlimited, 1989.

Slatin, John M. "Text and Hypertext: Reflections on the Role of the Computer in Teaching Modern American Poetry." In *Humanities and the Computer: New Directions*, edited by David S. Miall, 123–35. Oxford: Clarendon Press, 1990.

Slatta, Richard. "Historians and Microcomputing, 1989." *Social Science Computer Review* 7 (Winter 1989): 446–57.

Smith, Eldred. *The Librarian, the Scholar, and the Future of the Research Library.* Contributions in Librarianship and Information Science, no. 66. New York: Greenwood Press, 1990.

Sperberg-McQueen, C. M., and Lou Burnard, eds. *Guidelines for the Encoding and Interchange of Machine-Readable Texts.* Chicago: Text Encoding Initiative, October 1990.

Strangelove, Michael, and Diane Kovacs. *Directory of Electronic Journals, Newsletters, and Academic Discussion Lists*, edited by Ann Okerson. Washington, D.C.: Association of Research Libraries, March 1992.

Stutt, Arthur. "Argument Support Programs: Machines for Generating Interpretations." In *Humanities and the Computer: New Directions,* edited by David S. Miall, 71–81. Oxford: Clarendon Press, 1990.

Tanenbaum, A. S. *Computer Networks*, 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall, 1988.

Taylor, Betty W., Elizabeth B. Mann, and Robert J. Munro. *The Twenty-First Century: Technology's Impact on Academic Research and Law Libraries.* Boston: G.K. Hall, 1988.

Tenopir, Carol, and Man Evena Shu. "Magazines in Full Text: Uses and Search Strategies." *Online Review* 13 (April 1989): 107–18.

Thaller, Manfred. "The Historical Workstation 'Project'." Unpublished paper delivered at the seventeenth International Congress of Historical Sciences, Madrid, 29 August 1990.

Thibodeau, Kenneth. "To Be or Not to Be: Archives for Electronic Records." In *Archival Management of Electronic Records*, edited by David Bearman, 1–13. Archives and Museum Informatics Technical Report no. 13, Pittsburgh: Archives and Museum Informatics, 1991.

Thorson, Stuart, and Kristi Andersen. "Computational Models, Expert Systems, and Foreign Policy." In *Artificial Intelligence and National Security*, edited by Stephen J. Cimbala, 147–58. Lexington, Mass.: Lexington Books, 1987.

Trainor, Richard H. "History, Computing

and Higher Education." In *History and Computing II*, edited by Peter Denley, Stefan Fogelvik, and Charles Harvey, 35–42. Manchester: Manchester University Press, 1989.

Triebwasser, Marc A. "The Electronic Library: The Student/Scholar Workstation, CD-ROM and Hypertext." Unpublished paper presented at the annual meeting of the American Political Science Association, Washington, D.C., 31 August–4 September 1988 (ED 307 896).

Troll, Denise A. *Library Information System II: Progress Report and Technical Plan.* Mercury Technical Report Series, no. 3. Pittsburgh: Carnegie Mellon University, 1990.

United Nations, Administrative Committee for the Coordination of Information Systems, Technical Panel on Records Management. *Electronic Records Guidelines: A Manual for Policy Development.* New York: United Nations, 1989.

U.S. Congress, Office of Technology Assessment. *Linking for Learning: A New Course for Education.* OTA-SET-430. Washington, D.C.: U.S. Government Printing Office, 1989.

U.S. House, Committee on Government Operations. *Taking a Byte Out of History: The Archival Preservation of Federal Computer Records.* House Report 101-978. Washington, D.C.: Government Printing Office, November 1990.

Van Orden, Richard. "Content-Enriched Access to Electronic Information: Summaries of Selected Research." *Library Hi Tech* 31 (1990): 27–32.

Waters, Donald J. *From Microfilm to Digital Imagery.* Report of the Yale University Library to the Commission on Preservation and Access. Washington, D.C.: The Commission on Preservation and Access, June 1991.

Weber, Robert. "The Clouded Future of Electronic Publishing." *Publishers Weekly* 237 (29 June 1990): 76–80.

———. "Libraries Without Walls?" *Publishers Weekly* 237 (8 June 1990): S20–S22.

Weiskel, Timothy C. "University Libraries, Integrated Scholarly Information Systems (ISIS), and the Changing Character of Academic Research." *Library Hi Tech* 6 (1988): 7–27.

White, Mary Alice. "Imagery in Multimedia." *Multimedia Review* (Fall 1990): 5–8.

———. "The Third Learning Revolution." *Electronic Learning* 7 (January 1988): 6–7.

Wild, Martyn. "History and New Technology in Schools: Problems, Possibilities and the Way Forward." In *History and Computing II,* edited by Peter Denley, Stefan Fogelvik and Charles Harvey, 20–34. Manchester: Manchester University Press, 1989.

Williams, Daniel. "New Technology for Coordinating Work." *Datamation* 36 (May 15, 1990): 92–96.

Wilson, Flo. "Article-Level Access in the Online Catalog at Vanderbilt University." *Information Technology and Libraries* 8 (June 1989): 121–31.

Woodsworth, Anne. "Computing Centers and Libraries as Cohorts: Exploiting Mutual Strengths." In *Computing, Electronic Publishing and Information Technology: Their Impact on Academic Libraries,* edited by Robin Downes, 21–34. New York: Haworth Press, 1988.

Woodsworth, Anne, Nancy Allen, Irene Hoadley, June Lester, Pat Molholt, Danuta Nitecki, and Lou Wetherbee. "The Model Research Library: Planning for the Future." *Journal of Academic Librarianship* 15, no. 3 (1989): 132–38.

Yavarkovsky, Jerome. "A University-Based Electronic Publishing Network." *EDUCOM Review* 25 (Fall 1990): 14–20.

———. *Technology and the Research Environment of the Future: The Impact of the Information Science Revolution on the Research Environment of the Future*, Albany: New York State Library and The

Nelson A. Rockefeller Institute of Government/SUNY, February 1989.
Zarri, Gian Pierro. "Artificial Intelligence and Information Retrieval: A Look at the RESEDA Project." In *The Analysis of Meaning: Informatics 5*, edited by Maxine MacCafferty and Kathleen Gray, 166–72. London: Queens College Oxford, 1979.