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

[138]See 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.

[139]See 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.

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

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

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

[143]Barbara 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.)