

## Research Article

# Understanding Electronic Incunabula: A Framework for Research on Electronic Records

MARGARET HEDSTROM

**Abstract:** Archivists need a framework for research on electronic records issues to address one of the greatest challenges facing the profession. Historical evidence and concepts from the history of technology can inform archivists' understanding of electronic records as evolving forms of documentation. Research on electronic records must also account for the social, economic, and political aspects of organizational life that mediate how organizations and individuals adopt and use information technologies. The author provides a framework for research on electronic records issues that places electronic recordkeeping in a historical, social, and cultural context; encourages interdisciplinary research; and urges careful attention to research methodology.

*About the author:* Margaret Hedstrom is chief public records analyst at the New York State Archives and Records Administration and director of its Center for Electronic Records. She was chair of the SAA Committee on Automated Records and Techniques (1988-90) and wrote the SAA manual, *Archives & Manuscripts: Machine-Readable Records* (1984). A version of this article was presented at the Working Meeting on Research Issues in Electronic Records in Washington, D.C. in January 1991.

Research for this project was supported in part through the Research Fellowship Program in Modern Archives at the Bentley Historical Library, University of Michigan, funded by the National Endowment for the Humanities and the Mellon Foundation. The author is grateful to David Bearman, Richard Cox, Charles Dollar, Bob Frost, Sue Gavrel, John McDonald, Charles Robb, Ken Thibodeau, and Lisa Weber for comments on earlier versions.

"ELECTRONIC INCUNABULA" IS A metaphor for the current nature of electronic records as an evolving form of documentation. Incunabula, translated literally, means "out of the cradle," but in common parlance incunabula refers to the origins, infancy, or beginning of anything. In the history of printing, incunabula are the earliest printed books—generally those printed before 1500. Electronic records, as today's incunabula, present archivists with their greatest challenge in decades.<sup>1</sup>

The term "electronic incunabula" captures several themes from the framework for research that follows. Electronic records are in their infancy, and society has only begun to witness their transformative effects on documentation and communications. The shift from print to electronic communications will change the ways that organizations create and use information, much as the introduction of printing altered social practices, cultural conventions, institutions, economics, laws, and the politics of information. Change will be

evolutionary, as was the case with the introduction of printing, because profound shifts in the production and dissemination of information incorporate some traditional habits and approaches for handling information, and at the same time render obsolete some skills, professions, and institutions. Lacking any certain guideposts to predict the future, it is useful to look back at earlier shifts in paradigms.

This framework for research on electronic records uses historical evidence and concepts from the history of technology, but it is oriented toward future research. It shares common origins and inspiration with recent research agendas on the identification of records of enduring value, the management of archives, and the use and users of archives, but it differs from those research agendas in several respects.<sup>2</sup> First, electronic records are relatively new, and most archivists are unfamiliar with their nature and characteristics. Second, electronic records issues are complex and multifaceted. Consultation and collaboration with experts in other fields will be essential elements of successful research. Third, research on electronic records issues spans all archival functions and may challenge basic archival theory and practice. Consequently, the research framework proposed

<sup>1</sup>Participants in an advanced institute for government archivists, sponsored by the National Association of Government Archives and Records Administrators (NAGARA), concluded that "the archival management of electronic records is probably the most important, and certainly the most complicated, issue currently before the archival profession." National Association of Government Archives and Records Administrators, *Archival Administration in the Electronic Age: An Advanced Institute for Government Archivists* (Pittsburgh: NAGARA, 1989), iii. Electronic records problems are also discussed in Committee on the Records of Government, *Report* (Washington, 1985); National Academy of Public Administration, *The Effects of Electronic Record Keeping on the Historical Record of the U.S. Government: A Report for the National Archives and Records Administration* (Washington: National Academy of Public Administration, 1989); National Historical Publications and Records Commission, *Electronic Records Issues: A Report to the Commission*, Commission Reports and Papers #4 (Washington: National Archives and Records Administration, March 1990); U.S. Congress, House of Representatives, Committee on Government Operations, *Taking a Byte Out of History: The Archival Preservation of Federal Computer Records*, House Report No. 101-987 (Washington: U.S. Government Printing Office, 1990).

<sup>2</sup>Richard J. Cox and Helen W. Samuels, "The Archivist's First Responsibility: A Research Agenda to Improve the Identification and Retention of Records of Enduring Value," *American Archivist* 51 (Winter and Spring 1988): 28-46; Paul H. McCarthy, "The Management of Archives: A Research Agenda," *American Archivist* 51 (Winter and Spring 1988): 52-69; and Lawrence Dowler, "The Role of Use in Defining Archival Practice and Principles: A Research Agenda for the Availability and Use of Records," *American Archivist* 51 (Winter and Spring 1988): 74-86. Recommendations to develop a research agenda for electronic records are found in *Planning for the Archival Profession: A Report of the SAA Task Force on Goals and Priorities* (Chicago: Society of American Archivists, 1986), 9; and "An Action Agenda for the Archival Profession: Institutionalizing the Planning Process: A Report to SAA Council by the Committee on Goals and Priorities," 31 August 1988, 45-47.

here is interdisciplinary, more theoretical, and less specific than research agendas for functionally specific core areas of archival work.

### Research Objectives

Research on electronic records issues should produce generalizable policies, practices, methods, and applications for the management, preservation, and use of electronic records. Government agencies and large corporations have used computers extensively for recording and manipulating information since the 1960s, but archivists have made little progress toward developing specialized programs for electronic records or integrating them into core functions. There are few established and accepted standards, practices, or conceptual approaches to the management and preservation of electronic records. Social science data libraries and some traditional archives developed effective techniques for simple machine-readable data files, but those methods are not applicable to more complex forms of electronic records found in databases, office automation systems, geographic information systems, and compound documents. As the capabilities of modern information systems diverge from simple automated systems and forge sharper distinctions between electronic records and conventional paper files, some archivists question whether basic archival theories and practices can be applied to electronic records.<sup>3</sup> The absence of approaches and tech-

niques that address the wide variety of electronic records created with today's information technology is a major obstacle to development and growth of electronic records programs.

Research on electronic records issues will help the archival profession adopt a more strategic position in relation to information technology and its use by organizations. To date, archivists have responded to electronic records problems after hardware, software, and storage media have reached the market and become well established.<sup>4</sup> Archivists have had no impact on the design of information technology and little influence over its use in organizations. As a consequence, individuals and organizations have developed their own conventions for handling information in automated systems. Many current institutional practices undermine retention, preservation, and secondary use of electronic records; and most are inefficient, non-standard, and difficult to reverse. Even the word *archive* has lost much of its traditional meaning and associations. In the vernacular of data processing professionals, "to archive" means to store data off-line. A "permanent medium" is one that cannot be erased or altered even if it only lasts a few years. These new definitions do not incorporate any of the concepts that archivists normally associate with the term *archive*: to understand information in its context, to identify what is valuable, or to retain records and make them accessible as long as they have value. Archivists have literally lost control over the definition of *archive*.<sup>5</sup>

<sup>3</sup>David Bearman, *Archival Methods* (Pittsburgh: Archives & Museum Informatics, 1989), 26-27, 36-37, and 57-58; Katherine Gavrel, *Conceptual Problems Posed By Electronic Records: A RAMP Study* (Paris: UNESCO, April 1990); Charles Dollar, "The Impact of Information Technologies on Archival Principles and Methods," draft version 1.6, 18 September 1991, to be published by the University of Macerata (Italy) in 1992; Richard Kesner, "Automated Information Management: Is There A Role for the Archivist in the Office of the Future?" *Archivaria* 19 (Winter 1984-85): 162-72; and United Nations, Advisory

Committee for the Co-ordination of Information Systems (ACCIS), *Management of electronic records: Issues and guidelines* (New York: United Nations, 1990).

<sup>4</sup>Archivists often do not address technological problems until a technology is becoming or has become obsolete. For a discussion of one emerging technology, see Margaret Hedstrom, "Optical Disks: Are Archivists Repeating the Mistakes of the Past?" *Archival Informatics Newsletter* 2 (Fall 1988): 52-53.

<sup>5</sup>The data processing definition of *archive* is "a

A framework for research with ambitious goals and challenging questions encourages archivists to "think big enough" about electronic record keeping to make a difference. Whether electronic records are recognized as records, treated as an essential element of society's documentary heritage, and regarded as treasures that can preserve history and memory, depends to a large extent on the ability of archivists to influence how individuals and institutions use information technology and value its products.

One way to influence the policies and practices that govern the use of information technology is to intervene when new technologies are introduced into organizations. Such a strategic approach requires more research on emerging technologies and more emphasis on designing systems and tools that support archival and information management objectives from the outset. By concentrating on emerging technologies, archivists have the potential to build records management and archival requirements into software, applications, policies, and procedures, rather than trying to satisfy them retrospectively. Research can help archivists anticipate technology trends and forecast their likely effects on organizations and their records. This in turn would allow archivists to raise concerns about access, retention, preservation, and future retrieval before these issues have been defined as too cumbersome, too expensive, or irrelevant.

Several historical precedents suggest that

such actions can succeed. Standards for microfilm products and processes include important archival requirements, and archivists have a history of involvement with descriptive standards.<sup>6</sup> Currently, archivists and librarians are engaged in an effort to persuade the paper products industry to convert from acid to alkaline paper. Major archival repositories have made initial forays into several information technology standards, whose widespread adoption would facilitate retention and accessibility of electronic records.<sup>7</sup> Well-focused research can identify opportunities for archivists to influence standards and products, and help archivists better articulate their requirements when opportunities arise to comment on or contribute to standards development or product design.

Research on electronic records issues will help archivists account for the social, economic, and political aspects of organizational life that mediate how information technologies are adopted and used by organizations. Electronic record keeping is

<sup>6</sup>For an extended discussion of the development of standards for archival description, see "Special Section: Standards for Archival Description," *American Archivist* 52 (Fall 1989): 432-526; and "Standards for Archival Description: Background Papers," *American Archivist* 53 (Winter 1990): 24-108.

<sup>7</sup>For examples, see U.S. National Archives and Records Administration, Archival Research and Evaluation Staff, *A National Archives Strategy for the Development of Standards for the Creation, Transfer, Access, and Long-Term Storage of Electronic Records of the Federal Government*, National Archives Technical Information Paper No. 8 (June 1990); Margaret H. Law and Bruce K. Rosen, *Framework and Policy Recommendations for the Exchange and Preservation of Electronic Records*, [Report prepared by the National Computer Science Laboratory, National Institute of Standards and Technology for the National Archives and Records Administration] (March 1989); Canadian Bureau of Management Consulting, *Data and Document Interchange Standards and the National Archive* [Project No. 1-6465] (Ottawa, June 1987); Protocols Standards and Communications, Inc., *The Application of ODA/ODIF Standards* [prepared for the National Archives of Canada] (Ottawa, 1989); and Protocols Standards and Communications, Inc., *Application Portability* [prepared for the National Archives of Canada] (Ottawa, 28 December 1989).

procedure for transferring information from an on-line storage device or memory area to an off-line storage medium." Common archival definitions are: "the non-current records of an organization or institution preserved because of their continuing value," or "the agency responsible for selecting, preserving, and making available archival materials." See United Nations, ACCIS, *Management of electronic records*, 137-38. For commentary on this problem, see Lee Stout, "From the Chair," *Mid-Atlantic Archivist* 19 (Summer 1990): 3.



both a technological and a sociological phenomenon, and solutions lie in both the technical and administrative realm. Yet archivists are most familiar with the physical and technical barriers to preservation of electronic records. Equally significant changes in the ways that organizations define and carry out their missions, organize work processes, and meet the demands for inter- and intra-organizational communications are a fundamental, but neglected aspect of the problem. The technologies available to create, retrieve, store, and disseminate information are a factor in changing patterns of organizational work, documentation, and communications, but the role of technology in the process of change is not clearly delineated.

Archivists also can contribute to research on information technology and systems that support a wide range of organizational needs for usable, reliable, authentic, comprehensible, and lasting documentation. Archivists are in a unique position to contribute to information system designs because of their singular perspective on the relationship between the mission and structure of an organization, its need for records, its information flows, and document structures. A few thoughtful information professionals acknowledge that many current information systems are hampered by the narrow perspective of computer scientists, which focuses on optimal processing times and machine utilization rather than on information systems that help people do their jobs well. David Levy, a researcher at the System Sciences Laboratory, Xerox Palo Alto Research Center, contends that an essential perspective on advanced document systems is lacking in spite of a great deal of research on hardware and software tools from a computer systems perspective. According to Levy, "what is missing . . . is the set of complementary studies, in which documents—the subject matter, in effect, of our document systems—are investigated in their own right, as richly structured, cul-

tural, communicative artifacts."<sup>8</sup> The failure of new information technologies to mesh with organizational structures and information handling practices is not a recent phenomenon. In fact, the history of information processing is littered with information systems that failed because system designers concentrated on operating efficiencies and ignored organizational and cultural issues that are indispensable elements of information handling.<sup>9</sup>

In order for the archival profession to gain the maximum benefits possible from the time, effort, and other resources it invests in research, questions must be ambitious, think far ahead, and account for

<sup>8</sup>David M. Levy, "Topics in Document Research," System Sciences Laboratory, Xerox Palo Alto Research Center, Palo Alto, CA, (D. Levy 1988.8.31.1039), 1. For similar points from an archival perspective, see Luciana Duranti, "Diplomatics: New Uses for an Old Science: Part V," *Archivaria* 32 (Summer 1991): 6-24; and "Diplomatics: New Uses for an Old Science, Part VI," *Archivaria* (forthcoming).

<sup>9</sup>Critiques of office automation abound from many perspectives. Sociologists and political theorists have commented on the "macro" effects of automation on workers' skills and job satisfaction. Others have examined how a narrow perspective on office work, based in the principles of Taylorism, fails to account for the skills and knowledge needed to process information. From many perspectives, automated systems fail to work as well as their designers intended. For varying interpretations of this issue, see Harry Braverman, *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century* (New York and London: Monthly Review Press, 1974); U.S. Office of Technology Assessment, *Automation of America's Offices* (Washington, D.C.: U.S. Government Printing Office, December 1985), 125-68; Shoshana Zuboff, *In the Age of the Smart Machine* (New York: Basic Books, 1988), Chapter 4; William Bowen, "The Puny Payoff from Office Computers," in *Computers in the Human Context*, ed. Tom Forester (Cambridge, MA: MIT Press, 1989), 267-71; and Michael Hammer, "Reengineering Work: Don't Automate, Obliterate," *Harvard Business Review* 68 (July-August 1990): 104-12. Disparities between codified information handling procedures and the way workers actually accomplish their tasks are not unique to automated information systems. For an example of this phenomenon in a manual system, see Lucy A. Suchman, "Office Procedure as Practical Action: Models of Work and System Design," *ACM Transactions on Office Information Systems* 1 (October 1983): 320-28.

the social and cultural environment in which new information technology is applied. An effective program of research cannot focus only on present technologies or specific computer applications. Unless research projects recognize the broad implications of technological change and the essential changes that have taken place in organizations, the conduct of business, and record keeping, they are likely to overlook what is at stake and concentrate on issues that are time-bound and narrow in scope.

A research framework can support these objectives by promoting careful planning, sound research methodologies, and interdisciplinary approaches. Archivists have much to learn from colleagues in related fields who are concerned about related electronic records issues and much to gain from interdisciplinary research. Although there is no single discipline or organized body of "computing studies," archivists can find relevant research in the vast literature of information and library science, sociology of organizations, computer science, electrical engineering, and management science.<sup>10</sup>

The research needed to respond effectively to electronic records issues will be time-consuming, expensive, and complex, but posing broad questions about the significance of electronic records does not mean that archivists will conduct abstract research on global issues. Rather, a research framework can provide the structure for a series of smaller, practical projects that build on each others' results, contribute to an understanding of broader issues, and yield cu-

mulative results from what might otherwise be disparate efforts.

### Understanding the Context of Information Technology

Electronic records do not exist in a vacuum; rather, they emerge in response to particular conditions and needs. Research on electronic records will be influenced by the researcher's definition of information technology and assumptions about the role of technology in social and organizational change.<sup>11</sup> The research framework presented here defines information technology broadly as "the machines, processes and know-how used to create, store, manipulate, disseminate, and retrieve information." Students of technology point out that the term *technology* may be associated with physical objects, such as bicycles, washing machines, or computers; with processes or activities, such as oil refining or data processing; and with practical knowledge such as the "know-how" that goes into designing a bicycle or the knowledge required to sort a stack of cards into alphabetical order.<sup>12</sup> According to the definition used here, *information technology* may consist of physical objects that produce or process information, such as pens or computers; the processes that organize and manipulate information, such as routines to balance accounts or programs to alphabetize a list of names; and the know-how needed to organize and interpret information, such as double-entry bookkeeping or alphabetical filing rules. Machines, processes, and know-

<sup>10</sup>One anthology with an extensive bibliography of multi-disciplinary literature is Forester, ed., *Computers in the Human Context*. Comprehensive literature reviews include Rob Kling, "Social Analyses of Computing: Theoretical Perspectives in Recent Empirical Research," *Computing Surveys* 12:1 (1979): 61-110; and Kalle Lyytinen, "Different Perspectives on Information Systems: Problems and Solutions," *ACM Computing Surveys* 19:1 (1987): 5-46.

<sup>11</sup>For a recent summary of the various perspectives on computer technology, see "Editor's Introduction: Making Sense of IT," in Forester, ed., *Computers in the Human Context*, 1-15, and Parts One and Two.

<sup>12</sup>D. MacKenzie and J. Wajcman, eds., *The Social Shaping of Technology* (Milton Keynes: Open University Press, 1985), cited in Wiebe E. Bijker, Thomas P. Hughes, and Trevor Pinch, eds., *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987), 4.

how are all essential subjects for electronic records research.

A broad definition of information technology corresponds roughly to hardware, software, and the formal and implicit decision rules that govern information handling in automated systems. In information handling, the boundaries between machine, process, and know-how are constantly shifting. With the development of expanded machine instruction sets and specialized computer chips, for example, processes that used to be part of the software become fully incorporated into hardware functionality. Likewise, the line shifts between manual processes and software as each new computer program absorbs elements of know-how into an automated process.<sup>13</sup> A broad definition of information technology acknowledges the shifting line between machine, process, and know-how and helps maintain the essential connection between the manual processes of the past and their automated analogs.

Research on electronic records should avoid overly simplistic notions of the relationship between technology and social forces. One common interpretation of the role of technology in society is *technological determinism*—a perspective that attributes causality to machines or to technological process. Technologically determinist interpretations often contend that a specific machine or process was the cause of a larger social change. To argue, for example, that the printing press caused literacy because it provided a superior tech-

nology for the inexpensive dissemination of printed works is a technologically determinist argument. As Langdon Winner explains, according to this perspective, technology develops “as the sole result of an internal dynamic and then, unmediated by any other influence, molds society to fit its patterns.”<sup>14</sup>

Some analysts have adopted an equally unsatisfactory perspective that views technology as the product of raw political and economic power. This perspective, referred to as *social determinism*, compensates for technological determinism because it assumes that technical artifacts do not matter at all. Technology merely reflects the decisions of those individuals, industries, interest groups, or whoever had the power to decide which path technology ought to take.<sup>15</sup> Finally, some interpretations view technology as a completely neutral force in society that is autonomous from a social or economic context. According to this notion, technology is simply a tool under human control. One can choose to use or not to use a specific piece of technology, and the consequences of its use depend entirely on how humans apply it.

Recent research in the sociology and history of technology provides a framework to move beyond technology as a “black box” and to analyze technology as a rich social process.<sup>16</sup> Sometimes referred to as the “social construction of technology,” this new perspective considers technology to be the embodiment of human choices that influence how a machine is designed, what it is designed to accomplish, and how it is intended to accomplish its objectives. The design, development, marketing, ac-

<sup>13</sup>A redistribution of functionality also is moving software features closer to the machine. Ronald Weissman points out that features which used to be part of end-user applications are becoming features of operating systems. Ronald E. F. Weissman, “Virtual Documents on an Electronic Desktop: Hypermedia, Emerging Computing Environments and the Future of Information Management,” in *Management of Recorded Information: Converging Disciplines*, comp. Cynthia Durance, Proceedings of the 1989 International Council on Archives, Symposium on Current Records (New York: K. G. Saur, 1990), 41.

<sup>14</sup>Langdon Winner, “Do Artifacts Have Politics?” in *The Whale and the Reactor* (Chicago: University of Chicago Press, 1986), 21.

<sup>15</sup>*Ibid.*, 20–21.

<sup>16</sup>For examples, see Bijker et al., *The Social Construction of Technological Systems*; Winner, *The Whale and the Reactor*; and Bruno Latour, *Science in Action* (Cambridge: Harvard University Press, 1987).

ceptance or rejection, and interpretation of a technology are social processes shaped by rich interactions between cultural norms, economic and political power, social values, and the potential of a new machine or process.<sup>17</sup> According to this perspective, however, technologies ultimately reach "closure"—a point at which debate over the features of the technology ceases and the artifact or process stabilizes. Upon reaching closure, technical features alone may limit or eliminate possibilities, require adjustments in social and political systems, undermine deeply-held cultural values, and alter power relations.<sup>18</sup>

A perspective that considers technologies as "socially constructed" provides grounds for a cautiously optimistic view that archivists can influence key information technologies because it acknowledges that humans retain varying degrees of control over the design of technology. Many aspects of today's information technology have not reached closure. Archivists are not alone in questioning how information technology will shape their profession and the institutions they serve. Debates continue in many fields about the potential capabilities of information technology, what it should be designed to do, and how it will accommodate or transform information-handling practices.<sup>19</sup>

<sup>17</sup>Trevor J. Pinch and Wiebe Bijker, "The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other," in MacKenzie and Wajcman, eds., *The Social Shaping of Technology*, 40.

<sup>18</sup>Pinch and Bijker, "The Social Construction of Facts and Artifacts," 44–46; and Winner, "Do Artifacts Have Politics?" 20–39.

<sup>19</sup>For examples from several fields, see Alison B. Bass, "Computers in the Classroom," in *Computers in the Human Context*, 237–47; Donald P. Ely and Tjeerd Plomp, "The Promises of Educational Technology: A Reassessment," in *Computers in the Human Context*, 248–61; E. P. Krauss, "Magnetic Media and the Law," *Legal Studies Forum* 13 (1989): 301–12; Jonathan Javitt, ed., *Computers in Medicine: Applications and Possibilities* (Philadelphia: Saunders, 1986); and Julie C. Rutkowska and Charles Crook,

A perspective that views information technology as socially constructed also supports the premise that electronic records issues cannot be addressed exclusively through research on technical problems because important social and economic factors shape the decisions that individuals and organizations make about information technology. Computers are extraordinarily versatile machines, and their ability to process information once it has been reduced to a digital form places few technical constraints on potential applications. To learn much about the impact of information technology on capture, creation, and use of information, research must reach beyond the technology and examine its application and the context in which it is applied.

Research on electronic records will benefit from a clear understanding of the origins and evolutionary nature of change in information-handling technology and processes. A substantial body of literature discusses the "information revolution," but it often describes computer technology uncritically, and ascribes to it the power to transform economies from industrial- to information-based, to change the work force from manual laborers to knowledge workers, and to increase the availability of goods, services, and leisure time for all. Much of this literature speculates about the likely impact of computers, but claims of a "revolution" are seldom substantiated with evidence that distinguishes the role of computer technology from other social and economic factors.<sup>20</sup>

eds., *Computers, Cognition and Development: Issues for Psychology and Education* (Chichester, NY: Wiley, 1987).

<sup>20</sup>Examples of this perspective include Edward Feigenbaum and Pamela McCorduck, *Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World* (Reading, MA: Addison-Wesley, 1983); Alvin Toffler, *The Third Wave* (New York: Bantam Books, 1980); and John Scully, "The Relationship Between Business and Higher Education: A Perspective on the Twenty-first Century," *Communi-*



Often, the term *revolution* is misappropriated because changes in information handling are more evolutionary in nature. Formal institutions, procedures, and the know-how needed to manage and interpret information develop over an extended period of time and persist in spite of technological change. Organizations rarely have the luxury to abandon completely their old information systems, whether manual or automated, because information systems consist not only of hardware and software, but also of all the rules, procedures, skills, and habits that people have developed to handle information. Elements of information-handling practices often are carried forward from manual to automated systems, and from one automated system to the next.

If the industrial nations are in the midst of an information revolution, it is unclear how far they have traveled. James Beniger contends that data processing is the culmination, not the dawn, of a "control revolution" that began in the 1830s. Through centralized organizations and new infrastructures for transportation and telecommunications, bureaucratic organizations gained control over markets and communications. The establishment of marketing, advertising, and mass distribution, which relied on and produced massive volumes of information, was as central to this change

as development of a physical infrastructure of ports and railroads.<sup>21</sup>

Shoshana Zuboff places society at the beginning, rather than the end, of a transformation in information-handling practices. Zuboff makes a distinction between "automate" and "informate." She uses the term *automate* to refer to the replacement of human capabilities with the capacities of a machine, in contrast to the term *informate*, which refers to the capacity of information technology to change the ways organizations both process and use information.<sup>22</sup> According to Zuboff, most computer applications in factories and offices have simply automated older processes. This approach merely reproduces the logic of the assembly line and speeds up the processing of paper. When organizations "informate," they use the capabilities of computer technology to handle information in fundamentally different ways that transform work and power relations.

Ronald Weissman finds a similar shift in the world of the compound document and the personal workstation. Weissman claims that fundamental changes in software and computer architecture are transforming our notions of software and documents. New visions of computing tools result not only from massive leaps in the power of computers, but also from new ideas "about what computing can do and should be."<sup>23</sup> According to Weissman, a new computing culture has emerged, in contrast to the world of mainframe computers and MIS, which believes that "computing ought to have a

*ications of the ACM* 32 (September 1989): 1056-61. This approach, referred to as "technological utopianism," is critiqued in Rob Kling, "Reading 'All About' Computerization: Five Common Genres of Social Analysis," (August 1990), forthcoming in *Directions in Advanced Computer Systems*, ed. Doug Schuler (Norwood, NJ: Ablex Publishing Co.). Some critics question the direction and significance of the information "revolution" and the role of information technology in it. For critiques of the instrumentalist view of computing, see Stephen S. Cohen and John Zysman, "Manufacturing Matters: The Myth of the Post-Industrial Economy," in Forester, ed., *Computers in the Human Context*, 97-103; and "Mythinformation," in Winner, *The Whale and the Reactor*, 98-117.

<sup>21</sup>James R. Beniger, *The Control Revolution: Technological and Economic Origins of the Information Society* (Cambridge, MA: Harvard University Press, 1986). JoAnne Yates looks more closely at the role of records and communications systems in this process. See JoAnne Yates, *Control Through Communication: The Rise of System in American Management* (Baltimore: Johns Hopkins University Press, 1989).

<sup>22</sup>Zuboff, *In the Age of the Smart Machine*, 9-12.

<sup>23</sup>Weissman, "Virtual Documents on an Electronic Desktop," 38.

playful character and be personally empowering, more fun and far more imaginative and enriching than simple text processing, budget consolidation, or main-frame payroll processing."<sup>24</sup>

While focusing on emerging technologies, research on electronic records issues should be sensitive to social, cultural, economic, and political factors that mediate how rapidly technology can change information handling practices. A fundamental transformation of white collar work will require massive investments in software and telecommunications, changes in legal and social practices, and an infusion of new skills for handling electronic documents. Whether this transformation will take several years or several decades is a matter for debate, but it will evolve as a series of compromises between the potential of technology and the real needs and requirements of organizational life.

Lessons from the introduction of printing in the fifteenth century can help archivists gain a perspective on a profound change in information technology that may be applicable to research on the electronic incunabula facing the profession today. Because printed books are such an integral part of the material culture of Western society, we rarely stop to think about their historical origins, how they assumed the form and structure that they possess, or what the introduction of printing meant for the generations who observed the printed book as a new phenomenon. An understanding of the history of printing can help archivists discern how societies create new conventions for formatting and organizing recorded information.

The printing press created a new textual form, the printed book, that could be reproduced in multiple identical copies, but eventually the impact of the printing press stretched beyond print technology and its

immediate products. After books were printed new institutions developed, such as the book trade and eventually the lending library. The notion of literacy changed from a mystical faculty of clerics and scholars to an essential skill for survival as printed works reached deeper into the population. Books changed from treasured objects to everyday items that were bought and sold, loaned, and even discarded.

The structure and uses of books also evolved gradually during the first century or so of printing. Elizabeth Eisenstein reminds us that many of the conventions which we take for granted as part and parcel of the definition of a book—pagination, tables of contents, interspersions of graphics and text, and running headers—were invented gradually as the printed book diverged from hand-copied manuscripts.<sup>25</sup> The introduction of printing not only spread literacy, knowledge, and texts, it also changed behavior and cultural practices. Printed books, for example, were first read aloud, but eventually the practice of silent reading developed.<sup>26</sup>

History also informs us that societies accumulate recording and communications technologies, rather than replace one with another. Humans did not stop speaking when they learned to write; they did not stop writing when they learned to print, nor will they stop using paper when electronic media are widely accessible. As David Levy

<sup>25</sup>Elizabeth Eisenstein, *The Printing Press as an Agent of Change* (Cambridge, Eng.: Cambridge University Press, 1979), Volume 1.

<sup>26</sup>Roger Chartier, "General Introduction: Print Culture," in *The Culture of Print: Power and the Uses of Print in Early Modern Europe*, ed. Roger Chartier, trans. Lydia G. Cochrane (Princeton: Princeton University Press, 1987), 1–10. For similar observations about the slow transition from oral to written communications, see M. T. Clanchy, *From Memory to Written Word: England 1066–1307* (Cambridge, MA: Harvard University Press, 1979); and Hugh Taylor, "'My Very Act and Deed': Some Reflections on the Role of Textual Records in the Conduct of Affairs," *American Archivist* 51 (Fall 1988): 456–69.

<sup>24</sup>*Ibid.*, 39.

explains, "our literate culture retains and reinvests in its technological heritage, since the cost of discontinuous change is so high."<sup>27</sup> The traditions that surround information handling also have remarkable staying power, but they are not impervious to change. Even with widespread automation, people will continue to use paper documents when they provide familiar or convenient means for circulating and reading information. What is at issue is not the replacement of one form of material or one recording medium with another, but the significance of new forms of material that individuals and organizations create, using new information technologies.

The examples above support the premise that automation will bring about profound changes in the materials archivists handle and the processes used to create them. Electronic records will continue to diverge from paper records in their appearance, structure, uses, and significance as computer technology provides the capability for people to manipulate information in novel ways. Although no one knows how this new potential will be realized, archivists can frame research questions that are sensitive to the magnitude and evolutionary nature of this change and its social and cultural dimensions.

### Framing Questions: Areas for Research

The concepts and perspectives discussed above suggest five broad areas for research. This section proposes ways to frame research questions in each area, but it does not outline a detailed research agenda.<sup>28</sup>

Electronic record keeping is fertile ground for research, and the five areas are not meant to cover all potential research issues. Moreover, the research areas are not mutually exclusive. Well-designed research projects might address questions raised in several of the research areas.

#### *1. What is the relationship between activities, organizational structures, information technology, information flows, decision-making, and documentation?*

Researchers should examine electronic records in the overall context of organizations and their documentation. A priority area for research is the relationship between functions, activities, organizational structures, and information systems. Similar research has already been proposed by archivists concerned with appraisal and descriptive practices.<sup>29</sup> Archival theory posits an explicit relationship between the functions and activities of organizations and the documentation they create. In an era of rapidly changing information technology, archivists need to reexamine how information systems support organizational functions and relate to organizational structure within specific organizations and in a broader documentary context. Such research will help archivists determine where current theory can be applied and where new approaches and methods are needed to manage electronic records.

Research must also address the specific relationships between automation and the documentary requirements of organiza-

<sup>27</sup>Levy, "Topics in Document Research," 4.

<sup>28</sup>The Minnesota Historical Society, with financial support from the National Historical Publications and Records Commission, sponsored a working meeting in January 1991 on research issues in electronic records. The final report includes descriptions of ten research and development projects that constitute a research agenda. See *Research Issues In Electronic Records: Report of the Working Meeting* (St. Paul: Minnesota Historical Society, 1991), 7-22.

<sup>29</sup>For discussion of the significance of this type of research for appraisal, see Cox and Samuels, "The Archivist's First Responsibility," 40; Bruce H. Bruemmer and Sheldon Hochheiser, *The High-Technology Company: A Historical Research and Archival Guide* (Minneapolis: Charles Babbage Institute, 1989); Bearman, *Archival Methods*, 14-15; and Margaret Hedstrom, "New Appraisal Techniques: the Effect of Theory on Practice," *Provenance* 7 (Fall 1989): 1-21.

tions. Archivists would benefit from research that examines changes in communications and records when organizations automate. How does information flow through an organization and how does automation change patterns of information flow? Why do organizations decide to automate certain functions (and not automate others)? What factors influence selection and use of information technology by organizations? Do organizational characteristics such as size, type of organization, composition of the work force, complexity of work, and the power structure affect the form and content of its records? Do individual decision-making styles and corporate culture influence an organization's use of automated systems?

Research on these relationships must also account for the complex interactions between information technology, the transformation of information-handling processes, and changes in organizational structures and functions. Although researchers from a variety of disciplines contend that organizations modify their structures, alter their communication patterns, and change decision-making styles as a consequence of automation, there are widely varying explanations for these relationships.<sup>30</sup> Moreover, academic researchers rarely consider the effects of these changes on the recorded documentation that organizations create.<sup>31</sup> Participation by ar-

chivists in interdisciplinary research projects could lend an especially valuable perspective to studies of automation and organizational change.

## 2. What new forms of material do users create with information technology?

Forms of material influence many aspects of archival practice. In the broadest sense, forms of material define archivists as professionals who concern themselves with records and distinguish archivists from librarians who handle bibliographic items and museum curators who work with artifacts. The archival profession uses special arrangement, description, and preservation practices for specific forms of materials, such as maps, drawings, and motion pictures, particularly when they also constitute "special format" records.<sup>32</sup>

It is common to regard electronic records as a new form of material because of the special physical characteristics of automated records and the need for new methods for their care and handling. In the 1960s and 1970s, when most electronic records were in the form of machine-readable data files that contained numeric data, electronic records could be treated as a distinct form of material. However, as computer technology provides the capability to store almost any type of information in digitized form—words, pictures, sounds, graphics, and images—the distinction between electronic records and other forms of material becomes less meaningful, while differ-

<sup>30</sup>Reviews of this literature are found in Kling, "Social Analyses of Computing," 61–110; Rob Kling and Walt Scacchi, "Computing as Social Action: The Social Dynamics of Computing in Complex Organizations," in *Advances in Computers*, vol. 19 (New York: Academic Press, 1980), 249–327; and Lyytinen, "Different Perspectives on Information Systems," 5–46.

<sup>31</sup>The National Academy of Public Administration attempted to address this problem in a study it conducted for the National Archives and Records Administration. Unfortunately, the methodology employed was insensitive to broader issues of information flow within and between organizations and the gradual, evolutionary change in information recording and use

patterns. See National Academy of Public Administration, *The Effects of Electronic Record Keeping on the Historical Record of the U.S. Government*, Volume 2, Appendix C. The questionnaires used to determine the effects of automation on record keeping appear instead to have elicited *opinions* of users and non-users about these effects.

<sup>32</sup>Helena Zinkham, Patricia D. Cloud, and Hope Mayo, "Providing Access by Form of Material, Genre, and Physical Characteristics: Benefits and Techniques," *American Archivist* 52 (Summer 1989): 300–19.



ences among forms of electronic records become significant.<sup>33</sup> In the area of cartographic data, for example, archivists are uncertain whether digitized data used to produce the pictorial representation of a geographic area constitutes a map or an electronic record. Are there any unique aspects of a digitized "map" that distinguish it from a map that is drawn or printed on paper?

Understanding the forms of material that arise with the automation of record keeping requires more than a new taxonomy of record types. Research on this issue should be especially sensitive to the interaction between the technical capability to create, store, or display information in new formats and the meaning and values that users ascribe to that information. Only certain electronic information is created and used to support documentary requirements, and only certain electronic communications have the characteristics of authenticity, reliability, and stability to qualify as records.<sup>34</sup>

Linkages between forms of material traditionally stored on paper, such as memoranda, and similar forms of electronic documents must be explored. Networks, for example, have the capacity to transmit memoranda that can be viewed on a computer screen, but users may not define that electronic message as a record until it is printed on paper or captured and stored in an electronic "filing" system.<sup>35</sup> At the same time, electronic mail systems may capture messages that previously were transmitted over the telephone, and hence, defined as outside the realm of recorded documenta-

tion (unless an individual chose to create a written summary of important calls).

The introduction of new forms of material and the simultaneous transformation of traditional forms into something new raises a series of questions about the relationship between forms of material and archival practice.<sup>36</sup> When should new forms of material be managed differently from more traditional forms of documentation? Are there any archival principles that apply to all new electronic record types? What characteristics does an electronic memo share with a memo on paper? What does it have in common with other machine-readable records? What roles do tradition, habit, and past practice play in defining how individuals and organizations handle electronic records? For instance, do automated office systems mimic manual systems with the use of such conventions as "filing" and "paginating" documents because these practices form a bridge with paper document-based practices and terminology? Or are these superior and effective ways of organizing and handling textual information regardless of its physical form?<sup>37</sup> Answers to these questions will help archivists better

<sup>36</sup>For a systematic analysis of this problem, see Duranti, "Diplomatics: New Uses for an Old Science: Part V."

<sup>37</sup>Bruno Latour has identified the advantages of documents for presenting and transmitting knowledge. They are mobile, immutable when they move, created on a flat surface, have a scale that can be modified without change in the internal proportions; they can be reproduced, shuffled and recombined; several images of different origins and scales can be superimposed; they can be inserted into written text; and they can be merged with geometry in such a way that three-dimensional perspectives and concepts can be represented on a two-dimensional plane. Latour does not discuss the trade-offs among these various features. For example, which document technology is superior: one that produces "inscriptions" that, though fragile, are very inexpensive to reproduce and easy to shuffle and recombine? or one that produces permanent, but cumbersome documents? See Bruno Latour, "Visualization and Cognition: Thinking with Eyes and Hands," in *Knowledge and Society: Studies in the Sociology of Culture Past and Present* (Greenwich, CT: JAI Press, 1986), 20-29.

<sup>33</sup>For an initial attempt at such definitions see, United Nations ACCIS, *Management of electronic records*, 103-07.

<sup>34</sup>Physical characteristics alone do not establish the authenticity, reliability, and stability of electronic records. Even the definition of an electronic record has social and administrative dimensions.

<sup>35</sup>National Academy of Public Administration, *The Effects of Electronic Record Keeping*, 38-43.

understand the nature of electronic records and the relevance of archival theory.

### *3. Can archivists intervene at critical points in the development and introduction of new technologies?*

Research is needed in a third broad area to identify strategic opportunities for intervention by archivists into processes that affect the design and deployment of information technology and its use by organizations. The proposed research questions are based on the premise that many aspects of information technology have not yet reached closure, that outside interest groups at times can insert their views into the design process, and that archivists have something to offer to designers of information systems. Research should address the perceived need for archivists to become involved in the design of information systems while recognizing that archivists have not determined when to intervene or which tactics to use.<sup>38</sup>

Research on a series of focused questions could test these assumptions. At what critical points are decisions made about the development and introduction of new information technologies? Can archivists and records creators articulate design requirements that meet archival needs and explain them to hardware designers, software engineers, and applications developers? Which strategies and tactics successfully influence developers of information technology and designers of automated applications? How can archivists learn about significant technology trends and forecast their impact on organizations and their records? Can archi-

vists influence the ways organizations value and use electronic records through involvement in information systems design? Can archivists help organizations form sound management practices for their electronic records before new systems become established?

Answers to these questions will help archivists craft strategies that use the profession's limited resources and influence most effectively. They may identify key technologies or decisive moments in the design cycle where intervention by archivists is most effective. Research on this issue would also help archivists learn more about the technical and social aspects of information systems design and implementation.

### *4. How will changes in the supply of, demand for, and costs of storing and disseminating information change archival practice?*

Behind questions about culture, values in society, and organizational structure lie some hard questions about supply, demand, and costs. The automation of record keeping changes many of the economic dynamics that have molded the archival profession. Because paper is a very low-density storage medium and storage space has become increasingly expensive during the twentieth century, archivists take extraordinary measures to keep the volume of paper records at the minimum necessary to achieve "adequate" documentation. The appraisal function is dedicated to identifying that small percentage of records—often cited as 3 to 5 percent—that is absolutely necessary for historical purposes. Archivists at times weed superfluous materials from collections, undertake complex sampling projects, and microfilm records solely for the purpose of bulk reduction.

When information is stored in digitized form, the costs of storage per se decrease dramatically. Recent advances in storage technology make it possible to store mil-

<sup>38</sup>I have argued elsewhere that archivists are missing an opportunity to influence the values and practices associated with optical disk technology. See Hedstrom, "Optical Disks: Are Archivists Repeating the Mistakes of the Past?" 52–53. For a recent example of an organization's attempt to develop a more interventionist approach to electronic records, see United Nations, ACCIS, *Management of electronic records*.

lions of pages on a single disk at a trivial cost, and the trend toward lower storage costs is only accelerating. The costs of maintaining the information in a form that is accessible, however, quite possibly will absorb any potential savings from reduced space. David Bearman has argued that it will be prohibitively expensive to preserve electronic records in the physical custody of archives in the wake of constantly changing storage and retrieval technologies, the requirements of software systems to retrieve electronic records, and the need to migrate data from one generation of technology to another. As a consequence, archivists should not consider taking physical custody of electronic records.<sup>39</sup>

This radical recommendation is thought provoking because it encourages archivists to reexamine their role during a period of rapidly changing information technology and to reconsider basic practices designed to preserve paper records. However, to make sound decisions about the custody of archival records, archivists need to know more about the existing and potential demand for their services, the costs of meeting that demand, and the impact of major changes in information technology on access and preservation costs. Although it is impossible to calculate ultimate storage costs for records that are to be retained "permanently," the profession lacks any model to estimate the relative costs of retaining records in the various available media or to evaluate the

benefits of different media options. Archivists should conduct research on the costs, benefits, risks, and feasibility of various preservation options and strategies that can contribute directly to an informed reconsideration of the custodial role of archives.

Such an examination would quickly depart from supply and demand factors to more fundamental research on the users of archival records, their interests, behavior and motivation, and the nature of the services they demand. In his proposed research agenda for the use of archives, Lawrence Dowler encouraged all archivists to examine use and users closely and to build services around demand for records rather than supply.<sup>40</sup> Research on users will not be applicable to electronic records issues unless research projects examine potential users whose needs are not met by current archival holdings and services, including current users of electronic records who obtain access directly from the creating agencies, from social science data libraries, and from private vendors.

Archivists should study current and projected markets for electronic records to develop useful models for electronic records programs in an era of fundamental change in potential services, user expectations, and the costs of meeting them.<sup>41</sup> The experience of private-sector vendors who have captured government records in electronic form, added value to them, and redisseminated them at a profit may provide instruc-

<sup>39</sup>Bearman, *Archival Methods*, 43-47; David Bearman, "An Indefensible Bastion: Archives as Repositories in the Electronic Age," in *Archival Management of Electronic Records*, ed. David Bearman, Archives and Museum Informatics Technical Report #13 (Pittsburgh: Archives and Museum Informatics, 1991), 14-24; and Dollar, "The Impact of Information Technologies on Archival Principles and Methods," 62-66. For counter arguments, see Ken Thibodeau, "To Be or Not to Be: Archives for Electronic Records," in *Archival Management of Electronic Records*, 14-24; and Margaret Hedstrom, "Archives as Repositories—A Commentary, in *Archival Management of Electronic Records*, 25-30.

<sup>40</sup>Dowler, "The Role of Use," 74-86. See also Paul Conway, "Facts and Frameworks: An Approach to Studying the Users of Archives," *American Archivist* 49 (Fall 1986): 393-407; Bruce W. Dearstyne, "What is the Use of Archives? A Challenge to the Profession," *American Archivist* 50 (Winter 1987): 76-87; and Elsie T. Freeman, "In the Eye of the Beholder: Archives Administration from the User's Point of View," *American Archivist* 47 (Spring 1984): 111-23.

<sup>41</sup>For a proposal on criteria to evaluate the quality of electronic records user services, see Thomas E. Brown, "Machine-Readable Views" *Archival Informatics Newsletter* 2 (Summer 1988): 33-35.

tive lessons for the decades ahead. Archivists might study how private vendors identify and reach markets for electronic records, determine when and how to add value to them, and distribute data to those markets. How do purchasers of private data services evaluate the quality of the "product" they purchase? What types of value-added services are users willing to purchase? What are the implications for equity of access? These are the types of broad issues that archivists must confront because electronic records carry with them vastly reduced storage costs, coupled with new possibilities for manipulating, packaging, and adding value to records. Archivists need research that evaluates new options for delivering information to users while protecting the authenticity of electronic records.

*5. How should the requirements for management and preservation of electronic records change archives as institutions and the archival profession?*

Archivists who administer machine-readable records programs agree that the traditional methods used to appraise, process, describe, disseminate, and preserve paper records are inadequate to administer electronic records. In light of both advancing technologies and the limited success of machine-readable records programs, some archivists are beginning to question whether fundamental archival principles, such as provenance and original order, are applicable to the administration of electronic records. The concepts of original order and provenance derive from the basic archival principle of *respect des fonds*—the practice in archives of grouping together records from an administration, organization, person, or corporate body, and of segregating records of one origin from records of other origins. This principle is based on the assumption that much of the meaning and value of records derives from knowledge of the context in which the rec-

ords were created. Knowledge of the context of creation in turn can be ascertained by examining records in their original order and by studying the administrative history, organizational structure, and functions of organizations and the life history and accomplishments of individuals.<sup>42</sup>

Many archivists are aware that modern technologies undermine or complicate the application of the principles of provenance and original order. Except for the simplest data file structures, the physical ordering of data is controlled by software and is distinct from its logical order. Database management packages and the software that controls sophisticated applications, such as geographic information systems, provide users with the capability of a multitude of logical views of the data.<sup>43</sup> Networks provide users with capabilities for inter- and intra-institutional collaboration in the creation of documents and the formation of policy.

The debate over the relevance of established archival principles to electronic records has many dimensions. Some archivists question whether provenance can be used to control and describe electronic records that document multi-institutional communications, while others suggest that the principle can be applied if it accounts for fundamentally different patterns of communications and collaboration.<sup>44</sup> Archivists also disagree whether limitations on the principle of provenance relate exclusively to electronic records, or whether the

<sup>42</sup>For discussions of archival principles and electronic records, see Gavrel, *Conceptual Problems*, 13–15; and Dollar, "Impact of Information Technologies on Archival Principles and Methods," 56–61.

<sup>43</sup>Gavrel, *Conceptual Problems*, 17–29. The implications for archives of database management systems are discussed in Thibodeau, "To Be Or Not To Be," 5–10.

<sup>44</sup>Hedstrom, "New Appraisal Techniques," 17–21. For an example of an archival analysis of a multi-jurisdictional database, see Alan Kowlowitz, *Archival Appraisal of Online Information Systems*, Archival Informatics Technical Reports, Part 2, (Fall 1988).



principle has limited use for modern records, regardless of format.

Answering questions about the applicability of basic theory and practice will be the most difficult because the relevance of theories and the effectiveness of practices must be measured against the purpose of archives and the mandates for their programs. Archives could continue to serve as central repositories for the small percentage of records in any format that have enduring value, selected against the increasingly sophisticated appraisal criteria necessary to control costs. Archives could continue to distinguish themselves from libraries and other information services by defining specific types of recorded information that fall exclusively under their purview. Likewise, archivists might identify strengths and value in traditional archival theory and devise new methods that apply traditional theory to a new technological and organizational environment. For example, the need to maintain information about provenance in order to understand and interpret records might form the basis for the distinction between archival records and library materials. Archives would administer programs for access to records that are unintelligible without knowledge of the context in which they were created. The archivist's main responsibility would be to maintain linkages between context and content, with or without physical custody of records. Although research can never define for archivists what archives ought to be, research projects can help archivists evaluate the conceptual, technical, and economic obstacles to preservation of contextual information about electronic records.<sup>45</sup>

Research should also evaluate models for programs that help organizations create and

maintain adequate documentation of their history and accomplishments, ensure that sufficient and authentic records are created to hold organizations accountable for their decisions and actions, and promote preservation of electronic records with long-term value. In an era of widespread electronic record keeping, there are few good models of theories, practices, or programs that support these requirements. Social science data archives are the most fully developed institutions for preservation and dissemination of information in machine-readable form. Their history and programs provide instructive lessons about the benefits and drawbacks of possible approaches to electronic records preservation.

University campuses and government-supported research institutions around the world established specialized repositories for machine-readable data files during the 1960s and 1970s in response to a recognized need for access to disaggregate social and economic data by a community of researchers.<sup>46</sup> Early proponents of social science data archives regarded the availability of machine-readable data, use and promotion of quantitative methods, and the need to train young scholars in new research techniques as closely intertwined problems. As quantitative methods gained acceptance in the social sciences, data archives became an important component of the infrastructure for research.

Social science data archives concentrate on one form of material: social science research data commonly found in censuses, research surveys, and polling data; they cater successfully to a well-defined, although expanding, clientele.<sup>47</sup> Staff are well versed

<sup>45</sup>The research agenda developed in conjunction with the Working Meeting on Research Issues in Electronic Records elaborated on the broad questions raised here in one of its priority areas for research. See *Research Issues in Electronic Records*, 10-11.

<sup>46</sup>Kathleen M. Heim, "Social Scientific Information Needs for Numeric Data: The Evolution of the International Data Archive Infrastructure," *Collection Management* 9 (Spring 1987): 1-53.

<sup>47</sup>Carolyn Geda, "Social Science Data Archives," *American Archivist* 42 (April 1979): 158-66; Kathleen Heim, ed., *Library Trends: Data Libraries for*

in the subject areas and research methodologies that the data supports, often having acquired their training in the social sciences rather than in library or archival practices. Data archivists also are experts in data file formats and structures, survey research methods, technical processing, documentation and description, and storage issues. This constellation of expertise is valuable, because data archivists not only locate data sources, but help researchers use and interpret them.

These experiences suggest that there is a close relationship between sources, methods, uses, the value of information, and value-added services. Archivists should study the nature of this relationship, and compare the social sciences to the physical sciences, where repositories are larger and more specialized, and the arts and humanities where the few existing repositories are highly decentralized.

The experience of social science data archives raises several specific research questions. First, can archivists identify constituencies of potential users who are concerned about the availability and preservation of new forms of electronic records? How closely are researchers' concerns about the availability of data linked to the research methods they use? Could archivists mobilize potential users to urge that preservation of electronic records become part of the research infrastructure in the arts and humanities just as noted social scientists did thirty years ago? Finally, the example of social science data archives suggests the types of services that users will expect of data providers in the future, such as assistance with the interpretation of data, preparation of data packages for educational uses, and remote access to records.

Social science data archives do not offer

solutions to the most challenging questions about preservation of electronic records because they preserve data primarily for its "informational" value. Data archives methods are increasingly limited for handling records from on-line databases, automated office records, compound documents, software and software-dependent data structures, graphics, and other new types of electronic records. They do not provide models for capture, selection, or preservation of records needed for their evidentiary value, nor do they suggest ways to preserve essential contextual information. If archives are going to continue to play their traditional societal role of capturing and preserving an institutional history and memory, then there is need for considerable research on the ways that organizations make and document decisions using modern technologies. In understanding the relationship between the functions, structures, and documentation of organizational life, archivists have both the most to learn and the most to offer.

### A Note on Methodology

The final section of this research framework proposes ways to move from broad questions to specific research projects.<sup>48</sup> As other observers of archival research have noted, archivists lack the time, resources, and facilities to conduct basic research. Most archival research uses the "inductive" model, whereby a practicing archivist examines evidence and then draws conclusions from it, often in response to a specific, practical concern.

The inductive model can contribute answers to broader questions, if a series of more manageable research projects are de-

*the Social Sciences* 30 (Winter 1982), special issue on social science data libraries; and *Reference Services Review* 16: 1 & 2 (1988): 7-55, special issue on numerical and statistical data files.

<sup>48</sup>For an excellent summary of methodological concerns, see Tora Bikson, "Research on Electronic Information Environments: Prospects and Problems," unpublished paper presented at the Working Meeting on Research Issues in Electronic Records, Washington, D.C., 23 January 1991.

signed with fundamental issues in mind. Small, practical research projects can build on each others' results if they are coordinated and carefully controlled to account for the setting in which the project occurs. Archivists need to design research projects carefully and use sound methodologies to make sure that urgent and practical research needs do not eclipse more profound issues. Careful attention to research methodology and constant concern with the broader implications of research projects will help archivists overcome a narrow, time-bound perspective.

The design of a research project should begin with the development of a hypothesis that can be tested. Unfortunately, too many archival studies ask broad, open-ended questions, review current literature, and then experiment with an approach to a problem within the confines of a single institution or program. Without testable hypotheses and adequate controls over such studies, the validity of the research results and their applicability to other institutions are uncertain.

What would constitute a good hypothesis that applies the theoretical perspectives presented here? First, it must be possible to accept, reject, or modify the hypothesis and use it to answer specific questions. Taking just one example, a variety of approaches can be used to examine the issues surrounding office automation. One could construct a research project that asks "what happens to information flows, documents, and records in an office when a local area network is installed?"<sup>49</sup> This approach is problematic because the results are likely to be speculative, impossible to verify, and difficult to apply in varied institutions. A better approach would develop a hypothe-

sis that postulates likely impacts on verbal communications and written documentation. Comparing expected and actual consequences would help archivists determine whether more or less information was recorded, how use and distribution patterns changed, and which documentation and information management problems an organization encountered during such a transition. A follow-up study might test further hypotheses about temporary dislocation and long-term changes that result from automation of record keeping. The development of clear hypotheses will help focus research projects. Researchers should use literature about automation from a variety of disciplines to develop, refine, and evaluate hypotheses. The extensive literature on office automation and communications, for example, would narrow a hypothesis about the impact of local area networks on written and verbal communications.<sup>50</sup>

Even practical projects must account for the array of social, cultural, historical, and technological factors that influence records and record keeping practices. Research projects must control for factors such as structure and functions of organizations that use automated tools for record keeping, the history of information systems in the organization, past practices for managing and controlling records and information, the nature of the "computing package" being used, and the corporate culture of the organization. Looking at all of these factors will help archivists sort out cause and effect, avoid technologically determinist arguments, and identify those aspects of organizational structure, technology, and

<sup>49</sup>The National Archives of Canada and the Canadian Department of Communications participated in an inter-disciplinary project on office automation which attempted to address a series of more specific research problems. See Public Archives of Canada, *Interim Report on the PAC/DOCS Project* (Ottawa: 1985).

<sup>50</sup>For overviews of some of this literature, see U.S. Office of Technology Assessment, *Automation of America's Offices*, 125-68; Kling, "Social Analyses of Computing," 61-110; and Charles Babbage Institute, "Selected Readings in the History of Computing," Appendix B to *Resources for the History of Computing: A Guide to U.S. and Canadian Records* (Minneapolis: Charles Babbage Institute, 1987), 158-64.

behavior that should become the focus of intervention intended to promote management and preservation of archival records. Furthermore, careful attention to the environment in which the research occurs will help archivists determine which approaches and methods can be exported from one organization and applied effectively elsewhere.

To return to the example of local area networks, a researcher would need considerable data on such factors as the degree of centralization in the organization, the extent to which office activities rely on written and verbal communications, the extent to which individuals shared information before the system was installed, whether policies and corporate culture promoted or discouraged informal communications, and the effectiveness of previous record keeping and communications practices for supporting office functions. A researcher should also evaluate such technical factors as whether the network connected everyone who needed to communicate, how convenient the system was to use, and the extent to which electronic documents resembled more familiar paper documents. Without addressing all of these factors, it would be difficult to reach valid conclusions about the relationship between the communications capabilities of a local area network and patterns of documentation.

Research projects should be designed to control for the gradual nature of change that accompanies automation. Extensive research on the history of computing in organizations during the past thirty-five years shows that automation is an evolutionary force in organizations. The effects of automation are not always obvious a few months or even a few years after a new information system is installed because well-established customs for handling information and records change gradually. If one accepts the premise that automation is a gradual yet momentous change, then research must use longer time frames than

"before" and "after" studies generally permit.

Research projects should also be designed to avoid a priori assumptions about the impact of computing on organizations. Society is barraged with advertising, popular literature, images, and messages that shape perceptions of what computers are and what they can do. Too often, we accept uncritically such dictums as "computers allow organizations to process information efficiently," "computerization cuts costs," and "computers help deliver the right information to decision makers." While all of these factors may be correct in specific situations or perhaps even in the aggregate, there are enough cases where this is not true to demonstrate that such statements should be the conclusions of a research project and not part of the hypothesis. One recent study of computing literature determined that the initial assumptions of a study were the most accurate predictors of its conclusions.<sup>51</sup>

In developing research projects, archivists should also determine what types of expertise are needed. Considerable discussion has focused on the need for technical expertise to resolve electronic records keeping problems. There is little doubt that archivists need access to technical information, advice, and expertise to address many issues, but expertise in such areas as sociology of organizations, communications, economics, and marketing should not be overlooked. A realistic assessment of the need for outside expertise should encourage interdisciplinary projects.

This research framework stops short of proposing a list of specific research projects that archivists should design and conduct during the next few years. It is my hope, however, that the theoretical per-

---

<sup>51</sup>R. A. Hirschheim, "The Effect of A Priori Views on the Social Implications of Computing: The Case of Office Automation," *Computing Surveys* 18 (June 1986): 165-95.



spective and the framework presented here can provide archivists with a model to develop specific projects. By keeping the broader issues in mind, well-designed research projects could answer not only is-

sues of immediate concern to specific institutions, but also contribute answers to more challenging issues facing all archivists.