Preservation in the Digital World

Conway, Paul

http://hdl.handle.net/2027.42/150192
The Commission on Preservation & Access

Preservation in the Digital World

March 1996

A private, nonprofit organization acting on behalf of the nation's libraries, archives, and universities to develop and encourage collaborative strategies for preserving and providing access to the accumulated human record.
Preservation in the Digital World

Paul Conway
Head, Preservation Department
Yale University Library

March 1996

Author's Acknowledgement

A sincere thank you to Janet Gertz, Anne Kenney, Jan Merrill-Oldham, Jim Reilly, and Don Waters for carefully reviewing a draft of this report. Constance McCormack was persistent in assembling the illustrations. I accept sole responsibility for any errors in fact. My conclusions and recommendations do not necessarily reflect the views of the Commission on Preservation and Access or its staff.
Commission Preface

This paper provides an intellectual rationale for maintaining the centrality of preservation concepts and ethics in an increasingly digital information environment. This includes a framework for preservation planning and management from two perspectives: (1) contexts for action and (2) priorities for action, pointing out similarities and differences between present practice and future needs. It suggests that many of the basic tenets of preservation management can be applied in a highly technological environment, but that some long-held principles may no longer apply. The paper also considers where the accumulated knowledge and experience of preservation management may be most effective in a digital world, and where it may be more difficult to affect change.

In its conceptualization and development, this paper has served as a basis for in-depth discussions among preservation managers and technical specialists. The Commission trusts that with its publication and dissemination, the paper will stimulate many more discussions among broader audiences as we explore together how to maintain the safety and accessibility of the world’s historical and cultural heritage as far into the future as possible for all people.

Author’s Preface

Digital images are everywhere in libraries and archives. In many cases, the quality of digital image products from demonstration projects is spectacular, while others are less satisfactory. Nevertheless, there is widespread belief that the ability to produce high-quality images will improve as the technology matures. Even though a major investment is required for digital image conversion, libraries are rearranging budgets, raising funds, and anticipating income streams to make it happen. Can any institution—library, archives, historical society, or museum—afford to squander this investment in moving from pilot project to a fully operational system? The risk of loss extends beyond the financial, however, without a serious effort to ensure long-term access to digital image files. Understanding how to adapt preservation concepts to manage risk in the midst of rapid technological ferment is what preservation in the digital world is all about.

In the public at large, preservation is a concern chiefly in the world of paper. Digital information gives us worry-free preservation, it is asserted, because an exact copy of a file can be made with the click of a button. However, those who have been involved in digital imaging projects, including university and library administrators, technology system specialists, product vendors, and research scholars, have learned quickly that digital imaging technology, in and of itself, provides no easy answers for preservation. Indeed, simply defining what preservation means in the digital imaging environment is a challenge. This report suggests a framework for applying fundamental preservation concepts, derived from the best present practices of paper and film, to the world of digital image documents so that the highest level of responsible preservation planning, management, and action can continue.
Introduction

In *Motel of the Mysteries* (1979), illustrator David Macaulay speculates about how people 2,000 years from now might interpret the cultural significance of a low-budget roadside motel buried intact under junk mail and pollution. Beyond being a wry satire on the science of archaeology, the book is a clever reminder of the danger of interpreting the past without documentary evidence. A "Do Not Disturb" sign becomes a sacred seal "placed upon the handle of the great outer door by the necropolis officials following the closing of the tomb." A television represents "the essence of religious communication."

Anthropologists and historians know that the impulses to record and to keep are a part of human nature. Truth is embedded in the symbols and artifacts that we create and then keep by choice or by accident. And yet, as we approach the end of the twentieth century, we find ourselves confronting a dilemma such as the one faced by Howard Carson, Macaulay's amateur digger: a vast void of knowledge filled by myth and speculation. Information in digital form—the evidence of the world we live in—is more fragile than the fragments of papyrus found buried with the Pharaohs.

We are living in a digital world. Computers now far outnumber office workers in many parts of the globe. We bank by phone, enjoy digitally mastered music, fax carry-out orders, and communicate with each other through keyboarded thoughts. One of the sure signs that the global village has a digital face is the high investment of money and competitive energy now being directed toward changing the Internet into the National Information Infrastructure. After only a few years of life, the World Wide Web is crowded with time-sensitive data, news summaries, chat, and multimedia entertainment. The electronic landscape changes so rapidly—and the lines between the old and the new seem drawn so sharply—that *Wired* magazine can refer to a four-year-old network service provider as a "dinosaur," and get this retort: "It's very funny that a petroleum-based product like a magazine can call an online service that has an integrated Web browser irrelevant" (Nollinger 1995, p. 204).

It has long been the responsibility of libraries and archives to assemble, organize, and protect documentation of human activity. The ethic of preservation as coordinated, conscious management, however, is a more recent phenomenon. Librarians and archivists—like the clerks and scribes who went before them—have increased the chances that evidence about how we live, how we think, and what we have accomplished will be preserved. Traditional preservation, as responsible custody, works only when this evidence has a physical form; when the value of the evidence exceeds the cost of keeping it; and when the roles of evidence creators, evidence keepers, and evidence users are mutually reinforcing.

Digital imaging technology is more than another reformatting option. Imaging involves transforming the very concept of format, rather than creating an accurate picture of a book, document, photograph, or map on a different medium. The power of digital enhancement, the possibilities for structured indexes, and the mathematics of compression and communication together alter the concept of preservation. The digital world transforms traditional preservation.
concepts from protecting the physical integrity of the object to specifying the creation and maintenance of the object whose intellectual integrity is its primary characteristic. This transformation, along with the new leadership possibilities it creates, will force librarians and archivists to transform their services and programs in turn. Just as the invention of the vacuum tube created an entirely new form of mass communication—radio—instead of simply making point to point messaging possible without wires, digital imaging technologies create an entirely new form of information. Yet, many of the core tenets of preservation developed in the analog world can be carried forward to the digital world to continue the necessary roles of stewardship and service.

A Central Dilemma

The earliest known evidence of writing—pictorial signs on sun-dried clay tablets—originated roughly 6,000 years ago. Tens of thousands of examples of Sumerian and Babylonian writing exist today in the world's major research centers. Archaeologists unearth hundreds more every year. From ancient times to the present, the entire technology of writing has undergone steady evolution. Today, we have the capacity to store detailed bit-mapped images of hundreds of books on an optical disk the size of a coaster. This capacity to record and store gives rise to one of the central dilemmas of recorded history: Our capacity to record information has increased exponentially over time while the longevity of the media used to store the information has decreased equivalently.

The graph below illustrates these changes in broad strokes. The “X” axis lays out ten “writing” surfaces in the chronological order of their widespread introduction. The “Y” axis is a logarithmic scale where each level increases in order of magnitude—a factor of ten. The dotted line is a plot of the approximate number of characters per square inch that can be conveniently recorded on any particular surface. There is relative stability in density—fewer than 150 characters per square inch—in the centuries before modern printing presses and electronic technologies caused storage densities to skyrocket. Today's typical 5-1/2 inch rewritable optical disk can
hold 650 million characters on each of its two sides—yielding an equivalent density of 51 million characters per square inch of writable surface.

The solid line—representing the life-expectancy in years of each recording medium—declines through the years. Papyrus fragments of Egyptian writing from 4,500 years ago, while quite fragile today, are still legible. Moving ahead in time, stunning examples of illuminated manuscripts and other documents from Medieval times are quite able to withstand centuries more study and admiration. A similar situation prevails with early modern book printing technologies. Herman Melville's *Moby Dick, or The Whale*, first published in 1851 on typically acidic paper, symbolizes the worldwide preservation challenges of all nineteenth- and twentieth-century publishing (Banks 1969). The introduction of wood pulp and acidic compounds to paper, required at the time by a manufacturing industry desperate to keep up with demand, started the “slow fires” of deterioration that preservation efforts are now attempting to squelch.

During the twentieth century, the permanence, durability, and stamina of newer recording media have continued to decline, with the exception of microfilm (Sebera 1990). Magnetic tape may be unreadable just thirty years after manufacture (Van Bogart 1995, p. 11). The newest recording medium—optical disk—may indeed have a longer life than the digital recording surfaces that have gone before. It is likely, however, that today's optical storage media may long outlast the life of the computer system that created the information in the first place. This is the ultimate irony of recorded history. In order to achieve the kind of information density that is common today, we must depend on machines that rapidly reach obsolescence to create information and then make it readable and intelligible (Dollar 1992).

The mix of machinery and media of the digital world increases the need for responsible preservation activity in institutions. To help discover how to meet this increased need, this report first presents a framework of primary preservation concepts as they have evolved in past decades. It then redefines and transforms these concepts into a new framework for action in the digital world.

---

**A Framework of Preservation Concepts**

In the past two decades, a community of practitioners has reached consensus on a set of fundamental preservation concepts for managing available resources in a mature preservation program (Conway 1989). At one time, advocates for the protection of cultural artifacts, including books, primary source documents, and museum objects, used the terms “conservation” and “preservation” interchangeably. Today, preservation is an umbrella term for the many policies and options for action, including conservation treatments. Preservation is the acquisition, organization, and distribution of resources to prevent further deterioration or renew the usability of selected groups of materials.

The essence of preservation management is resource allocation. People, money, and materials must be acquired, organized, and put to work to ensure that information sources are given adequate protection. Preservation is concerned with evidence—what Michael Buckland (1991) calls “information as thing”—embedded in the intellectual content of objects and in the objects themselves. Nearly endless varieties of forms and formats are preserved so that people can use them for equally varied purposes, scholarly and otherwise. Those with the responsibility to do so find the small portions of the vast sea of information—structured as collections of documents, books, photographs and films, sound recordings, and other “things”—that have
research value as evidence of thought and action well beyond the time and intentions of those who created or published them.

Cost-effective preservation action cannot take place without compromising ideal outcomes. Preservation management encompasses all the policies, procedures, and processes that together prevent further deterioration of physical objects, renew the information they contain, and increase their functional value. This distinction between the value of the content (usually text and illustration) and the value of the artifact is at the heart of a decision-making process that itself is central to effective management. Preservation management includes an ongoing, iterative process of planning and implementing prevention activities (e.g., maintaining a stable, safe, and secure environment, ensuring disaster preparedness, and building a basic collection-level maintenance program) and renewal activities (e.g., undertaking conservation treatments, replacing the content of library materials, or reformating them on microfilm).

In the world of paper and film, preservation and access are separate but related activities. It is possible to fulfill a preservation need of a manuscript collection, for example, without solving the collection’s access problems. Similarly, access to scholarly materials can be guaranteed for a very long period of time with taking concrete preservation action. Modern preservation management strategies, however, posit that preservation action is taken on an item so that it may be used. In this view, creating a preservation copy of a deteriorated book on microfilm without making it possible to find the film (usually by cataloging it and then depositing the bibliographic record in a national database) is a waste of money. Preservation in the digital world puts to rest any lingering notion that preservation and access are separable activities.

Significant financial barriers slow the design and implementation of effective preservation strategies. Some leaders in the field have suggested, however, that inadequate funding may not be the only brake on preservation success. “The major obstacle to the development and administration of preservation programs is the shortage, not of money, as many suppose, but of knowledge.” wrote Pamela Darling (1981, p. 185). “Financial constraints are serious and will become more so; but until the preservation field reaches the point at which most people know what ought to be done and how it should be done, the lack of money to do it on a scale appropriate to the need is not terribly significant.” The basic consciousness-raising task for preservation may have been accomplished for the world of paper and film. It is just beginning for the digital world.

At this critical juncture in the evolution of preservation thinking, there is a large body of knowledge upon which to draw. The literature documents some fundamental concepts that have been described, tested, and codified (Swartzburg 1995). The following nine concepts, defined below in part by the words of some of the most articulate preservation advocates of past decades, form the framework for a comprehensive preservation perspective.
Context for Action

Four concepts describe what might be called the context for preservation action.

Custody: Fifty years ago, one of the foremost and persistent advocates for quality bookbinding put his finger on the centrality of preservation to the mission of modern research libraries and archives. Preservation, wrote Pelham Barr in his most frequently cited work, "as responsible custody, is the only library function which should be continuously at work twenty-four hours a day. It is the only function which should be concerned with every piece of material in the library from the moment the selector becomes aware of its existence to the day it is discarded" (1946, pp. 218-19). Barr's allusion to the life-cycle of information sources is highly relevant today. The concept is at the center of information management theory and practice, including specifications for the disposition of government archives, the management of book collections, and the maintenance of large-scale information technology systems (Atherton 1986).

Social value: Organizations that accept preservation as central to their mission also serve a larger societal need. The fundamental value of keeping intact the documentation of people and institutions has endured through years of changing practice. Almost sixty years ago Robert C. Binkley, a pioneer of preservation microfilming, pinpointed the cultural value of preservation. "The objective of archival policy in a democratic country cannot be the mere saving of paper," he said to one of the first gatherings of professional archivists. "It must be nothing less than the enriching of the complete historical consciousness of the people as a whole" (1939, p. 168). More recently, educator Guy Petherbridge echoed the same sentiments. "The perpetuation of society as we are accustomed to conceive or idealize it is dependent to a very large extent on the preservation en masse of our accumulated group memories and consciousness stored in the form of the written, printed and otherwise recorded word or symbol" (1987, p. 1).

Efforts to explain and illustrate the overarching value of preservation have had a positive effect. Today, the widespread recognition of the value of an investment in preservation action—fueled, perhaps, by the persuasive message in the film "Slow Fires"—has resulted in national and regional planning efforts and new dollars for preservation programs. Even the word "archives," carrying the mandate to preserve and protect the secrets they contain, has entered the popular vocabulary (Cox 1993). Part of the service that libraries and archives provide to society is (or will be) their investment in converting, storing, and making available research resources in digital image form. Protecting this investment through prudent preservation action will continue to be the social role of hundreds of organizations that participate in digital library initiatives.

Structure: An organizational structure that enables preservation resource allocation is imperative. The mission of an institution and the relevance of that mission to a larger social purpose are necessary, though insufficient on their own, to mobilize preservation action. Robert Patterson suggested a mechanism for a library-wide preservation development process that was equal parts consciousness raising and strategic planning. His argument on the administrative difficulties of building an effective new program rings true today when it is read with the recognized challenges of digital image technologies in mind. "Historically, many librarians have felt that a conservation program was too technically complex a matter for them to undertake..."
Armed with some basic principles and facts, a program with realistic goals can be developed” (1979, p. 1116). The overwhelming trend in the development of mature preservation programs has been the coalescing of responsibility in a single department, which then sets about making itself, in the best Weberian tradition, essential to the organization.

Cooperation: Resource sharing and cooperation among institutions to select and preserve the best and most valuable cultural artifacts are core preservation values. The ups and downs of cooperative activities are well documented, as are the challenges of maintaining momentum when taking any action that does not immediately benefit the local institution (Gwinn 1985). The largest and most prestigious research libraries—spurred by the ready availability of external funding—may have led preservation program development in the last two decades, but advocates for peer-to-peer or regional cooperation have sounded a warning about the dangers of elitism. Vartan Gregorian reminded a large audience of preservation activists that “all of us are in the same boat. When the boat sinks, no one can claim the fact that they had a first class ticket as solace” (Morrow 1991, p. 85).

The need to organize for local action often creates a direct tension with national cooperative initiatives. Rutherford Rogers, Yale University’s librarian during the heyday of preservation program growth, called for an emphasis on local goals, informed but not controlled by national priorities. National planning, he declared, “cannot supplant local programs or relieve us of our responsibility to develop them. Instead, we must take advantage of the large support network of collective research and educational activities to develop local programs suited to local needs. a network that may serve as the basis of a true national program” (1985, p. 8). When properly managed, interinstitutional cooperative preservation programs supplement local priorities.

Priorities for Action

Five concepts together can be seen to help direct the priorities for preservation action.

Longevity: The central concern in traditional preservation practice has been the media upon which information is stored. The term “archival” used to mean “permanent” and preservation focused on infinity. Jim O’Toole reviews the transformation of the definition of longevity: “The idea of permanence as it is understood by archivists has changed considerably over time, passing from an unattainable desire to an absolute value within the realm of achievement to an extremely relative notion of little clarity” (1989, p. 23). Today, archivists and librarians evaluate storage media in terms of their life expectancies. The top priority is extending the useability of papers, films, magnetic tapes, and other media by stabilizing their structures and limiting the ability of internal and external factors to cause deterioration. The focus on external factors led to specifications for proper environmental controls, care and handling guidelines, and disaster recovery procedures. Progress on efforts to control the internal factors of deterioration has resulted in alkaline paper standards, archival quality microfilm, and more rugged magnetic media (Calmes 1987).

Choice: Preservation adds value through selection. Choice involves defining value, recognizing it in something, and then deciding to address preservation needs in the way most appropriate to that value. Selection in archives and research libraries was once thought to be primarily a one-time decision about the potential for future use, made near the time of publication or
when the documents ceased to serve the primary purposes for which they had been created originally (Atkinson 1986, p. 345). Over decades the act of preservation has evolved from saving material from oblivion and assembling it in secure buildings to encompass more sophisticated condition, value, and use assessments on the already-collected. Preservation selection in libraries has been dictated largely by the need to stretch limited resources in as wise a fashion as possible, resulting in the dictum that “no item shall be reformatted twice.” The end result is a growing “virtual” special collection of items preserved using a variety of techniques, most notably by microfilming. Selection is perhaps the most difficult of undertakings precisely because it is static and conceived by practitioners as either completely divorced from present use or completely driven by demand (Demas 1994).

**Quality:** Maximizing the quality of all work performed is such an important maxim in the preservation field that few people state this fundamental principle directly. Instead, the preservation literature dictates high quality outcomes by specifying standards for treatment options, reformatting processes, and preventive measures. The commitment to quality standards—do it once, do it right—permeates all preservation activity, including library binding standards, archival microfilming guidelines, conservation treatment procedures, the choice of supplies and materials, and a low tolerance for error. The evolution of microfilming as a central strategy for preserving brittle library materials has placed the quality of the medium and the quality of the image on an equal plane (Fox 1996). In the pursuit of quality microfilm, compromise on visual truth and archival stability is dictated more by the characteristics of the item chosen for preservation and less by cost or the present state of display technologies.

**Integrity:** The concept of integrity has two dimensions in the preservation context: physical and intellectual. Physical integrity largely concerns the item as artifact and plays out most directly in the conservation studio, where skilled bench staff use water-soluble glues, age-old hand-binding techniques, and high-quality materials to protect historical evidence of use, past conservation treatments, and intended or unintended changes to the structure of the item (Tanselle 1989). The ethics of conservation recognize that all physical treatments change the nature of the object and, therefore, are not really reversible by nature (Dureau 1986).

The preservation of intellectual integrity is based upon concern for evidence of a different sort. The authenticity, or truthfulness, of the information content of an item—maintained through careful and complete reformatting or sensitive treatment—is at the heart of intellectual integrity (Duranti 1995). Ideally, documentation supports the preservation of intellectual and physical integrity by creating a chain of evidence connecting multiple formats of the object and recording what has been or needs to be done to any one of these formats in the name of preservation. Preservation also protects and documents the relationships among items in a collection. The concepts of quality and integrity reinforce each other.

**Access:** For years, preservation simply meant collecting. The sheer act of pulling a collection of manuscripts from a barn, a basement, or a parking garage and placing it intact in a dry building with locks on the door fulfilled the fundamental preservation mandate of the institution. In this regard, preservation and access have been mutually exclusive activities often in constant tension. “While preservation is a primary goal or responsibility, an equally compelling mandate—access and use—sets up a classic conflict that must be arbitrated by the custodians.
and caretakers of archival records," states a fundamental textbook in the field (Ritzenhaler 1993, p. 1). Access mechanisms, such as bibliographic records and archival finding aids, simply provide a notice of availability and are not an integral part of the object.

After years of disciplined practice, most librarians, archivists, conservators, and scientists recognize these nine concepts of preservation action. The same learning curve that has led to the acceptance of current preservation practice must now take place for digital imaging technology in a time-frame of years instead of decades.

Technology and Organizational Change

Technologies and organizations affect each other in different ways, depending in large part upon the characteristics of technologies and how organizations choose to use them. In this regard, it is important, first, to distinguish between acquiring digital imaging technology to solve a particular problem and adopting it as an information management strategy. Acquiring an imaging system to experiment with its capability to enhance access to library and archive materials is now almost as simple as purchasing tools in a hardware store. Simply buying a technology “solution” has no long-term implications for an organization. However, when pilot imaging projects are transformed and adopted as core systems for creating and delivering electronic documents with long-term value, preservation in the digital world becomes an important organizational issue.

Microfilm technology, as it has quickly proliferated in libraries and archives, has been largely a means for achieving ends already decided upon. Microfilm was, and remains, well-suited for reproducing bulky and fragile newspapers and for distributing research materials largely unknown or unusable in their native locations. Continuous wave radio, on the other hand, created a world of new possibilities for which new corporate goals had to be formulated. “In that sense it acquired a determinative force to which organizational decision-making had to adapt,” writes Hugh Aitken. Unlike microphotography, radio broadcasting “encountered a cultural phenomenon that it has not itself created: a mass market for information and entertainment the existence of which had been barely suspected” (Aitken 1985, p. 562).

The early years of microphotography are marked by visionary descriptions of revolutionary possibilities for mass distribution of books on small glass plates rather than on paper (Luther 1959). Susan Cady (1990) suggests that any vision that sees new technologies replacing old ones ignores the centuries-old reality that new communication technology merely expands opportunities, adds diversity, and creates a more complex environment. For every new technology that today functions beside an old one (e.g., television and radio), it is possible to identify a new technology that actually was replaced by an old one (e.g., teletype by...
People cherish convenience and will ignore any technology whose characteristics (e.g., cumbersome equipment, poor image quality, weak indexing) limit that convenience.

Digital image technologies present a similar situation for libraries that radio presented more generally for the corporate world some sixty years ago. In the digital world, preservation becomes possible because preservation—as it is defined in this paper—is an intimate part of the new possibilities. In the digital world, preservation must be concerned with entire technology systems, not one or another component, such as a film or a storage disk. Like radio, television, and electric power networks, digital imaging is not one technology. It is an interlinked, open system of hardware, software, and service subsystems, each with its own components, that develops in cycles. Digital systems are characterized by multiple core subsystems (scanners and transmission devices). Data storage is but one of many peripheral subsystems, among them compression hardware and software, display technologies, and output devices.

Early students of the history of communication and technology wanted to know in what sense an individual, an organization, or a social class can be said to be in control of, or to be managing, a body of information (Innis 1950). They wrote in terms of monopolies of knowledge and the degree to which different media—clay tablets, papyrus, the printed book, LP vinyl disks—lent themselves to the formation of such monopolies. More recent scholarship suggests that understanding and influencing technological change encompasses not only managing a technology process but also managing the consequences of technology on people and organizations (Czitrom 1982). Expanding this argument further, Columbia University scholar Michael Tushman suggests that we cannot even begin to understand technological progress by examining technical characteristics alone. At critical junctures, he suggests, “organizational action (and inaction) dramatically affects the shape and direction of technological change” (1992, p. 312).

Libraries and archives are an interrelated part of this new and complex digital marketplace, but not necessarily a critical element. They need to recognize their role in the development of digital imaging technologies, as well as the new demands that these technologies will place on them as organizations.
Preservation Management in the Digital World

As librarians and archivists have experimented with the capabilities of digital imaging technologies, the concept of preservation has come to have at least three distinct meanings.

Make Use Possible: For a very small subset of valuable but deteriorated documents, digital imaging technology is a viable, and possibly the only, cost-effective mechanism for facilitating research use. A recent experiment involving digitizing oversized color maps (Gertz 1995) demonstrated that the only way to really use the maps, which have faded badly and are very brittle, is to view them on a large color monitor after they have been digitized and enhanced. Similarly, the managers of the Andrew Wyeth estate have found that reproductions of the artist’s work are most faithfully represented in digital form (Mintzer and McFall 1991).

Protect Original Items: Digital image technology can be used to create a high quality copy of an original item. By limiting direct physical access to valuable documents, digital imaging becomes a “preservation application” as distinct from an “access application.” The original order of the collection, or book, is “frozen,” much like microfilm sets images in a linear sequence. Sophisticated indexing schemes facilitate browsing and minimize the potential for damage or disruption to a collection caused by “fishing expeditions” through the published or unpublished record. Preservation via digital copying has been the most compelling force motivating archives and libraries to experiment with hardware and software capabilities.

Maintain Digital Objects: Once digital conversion of the original document has been completed, the challenge of protecting the digits from corruption or destruction becomes the preservation focus. This facet, now being called “digital preservation,” typically centers on the choice of interim storage media, the life expectancy of a digital imaging system, and the concern for migrating the digital files to future systems as a way of ensuring future access (Preserving Digital Information 1995).

A workable framework for preservation in the digital world could encompass the initial choice of the technology, the use of digital imaging technology for reproducing research materials, and the protection of the digital information for as long as that information has value to the institution and the patrons the institution serves. The process of converting library materials to an electronic form—a process which in many aspects is similar to the one used to create preservation microfilm—is distinct from any particular medium upon which the images may be stored at a particular point in time. This distinction allows for a continuing commitment to creating and maintaining digitized information while entertaining the possibility that other, more advanced systems may render digital storage media obsolete. Preservation management of digital image
files encompasses their creation, organization and indexing, storage, transmission, and continued intellectual integrity.

The following section proposes how the nine fundamental concepts of preservation described above can be transformed in response to the particular characteristics of digital information.

Context For Action

Custody: In the digital world, a library, archives, or museum cannot make a decision to adopt imaging technologies for long-term conversion and storage of research collections in digital form without a deep and continuing commitment to preservation by the parent institution. The preservation mission that was once largely the prerogative of the library must become the preservation mandate of the parent institution. The necessary financial and technological commitments to maintain image system life and migrate image system data to future generations must permeate the highest levels of the organization.

Social Value: The particular value to society of preservation in the digital world has less to do with historical consciousness and group memory—as advocates for traditional preservation have claimed—and much more to do with service to academic, scholarly, and public communities. Active use is the lifeline of the loosely connected clusters of unique (in digital form), highly valuable, and intensely consulted digital collections that form a digital library worth preserving.

Structure: Organizing for preservation in the digital world is not, first and foremost, a search for process efficiency, as has been the case with traditional preservation, but rather an ongoing process of risk management, where the cost of digital file migration is judged against the cost of failure to preserve the files in terms of the patrons who need the information. The stakeholders in this organization extend well beyond the bounds of a preservation department or the administration of a library or archives to encompass technology specialists, marketing experts, and commercial vendors.

Cooperation: As preservation in the digital world becomes less a centrally managed enterprise at the institutional level and more a mutually reinforcing shared responsibility, cooperation and institutional self-interest should reinforce each other. When approaching the challenges of preservation, institutions will contribute to the required cost and the technical expertise because there will simply be no other way to satisfy their patrons’ information needs.

Priorities for Action

Longevity: Preservation in the digital world has little concern for the longevity of optical disks, magnetic tape, and other new, fragile storage media. The viability of digital image files is much more dependent on the life expectancy of the access system—a chain only as strong as its weakest component. Digital storage media must be handled with care, but they most likely will far outlast the capability of systems to retrieve and interpret the data stored on them. Since we can never know for certain when a system has become obsolete, libraries must be prepared to migrate valuable image data, indexes, and software to future generations of the technology.
Migration is now and will remain a very challenging task, as the recent digital archiving task force has pointed out (Preserving Digital Information 1995).

**Choice:** Selection for preservation in digital form is not a one-time choice made near the end of an item’s life cycle, but rather an ongoing process intimately connected to the active use of the digital files. The value judgments applied when making a decision to convert documents from paper or film to digital image are valid only within the context of the original system. With the need to migrate data to another storage and access system comes the mandate to revisit again and again the decision to continue preserving the information. It is a rare collection of digital files, indeed, that can justify the cost of a comprehensive migration strategy without factoring in the larger intellectual context of related digital files stored elsewhere and their combined uses for research and scholarship (Conway 1996).

**Quality:** Quality in the digital world—that is, the usefulness and useability of systems—is conditioned significantly by the limitations of capture and display technology. Digital conversion places less emphasis on obtaining a faithful reproduction of the original in favor of finding the best representation of the original in digital form. Mechanisms and techniques for judging quality of digital reproductions are different and more sophisticated than those for assessing microfilm or photocopy reproductions (Kenney and Chapman 1995). Additionally, the primary goal of preservation quality is to capture as much intellectual and visual content as is technically possible and then display that content to viewers in ways most appropriate to their needs.

**Integrity:** In the digital world, both physical and intellectual integrity have their place in preservation. A commitment to the physical integrity of a digital image file has far less to do with the media upon which the data are stored than with the loss of information that occurs when a file is created originally and then compressed mathematically for storage or transmission across a network. In the domain of intellectual integrity, index information is not independent of the content of the item but rather is an integral part of the digital file. Structural indexes and data descriptions traditionally published with an item as tables of contents or prepared as discrete finding aids or bibliographic records must be preserved—as metadata—along with the digital image files themselves (Weibel 1995). The preservation of intellectual integrity also involves authentication procedures, like audit trails, to make sure files are not altered intentionally or accidentally (Lynch 1994).

**Access:** In the digital world, the concept of access is transformed from a convenient byproduct of the preservation process to its central motif. The content, structure, and integrity of the information object assume center stage; the ability of a machine to transport and display this information object becomes an assumed end result of preservation action rather than its primary goal. Preservation in the digital world is not simply the act of preserving access but also includes a description of the “thing” to be preserved. In the context of this report, the object of preservation is a high-quality, high-value, well-protected, and fully integrated version of an original source document.

Administrators who have responsibility for selecting systems for converting materials with long-term value also bear responsibility for providing long-term access to the digital versions. This responsibility is a continuing one. Decisions about long-term storage and access cannot be deferred in the hope that technological solutions will emerge. An appraisal of the present value
of a book, a manuscript collection, or a series of photographs in its original format is the necessary point of departure for making a judgment about preservation of the digital image version. The mere potential for increased access to a digitized collection does not add value to an underutilized collection. Similarly, the powerful capabilities of a relational index cannot compensate for a collection of documents whose structure, relationships, and intellectual content is poorly understood.

If libraries, archives, and museums expect to adopt digital imaging technology for purposes of transforming the way they serve their patrons and each other, then they must move beyond the experimental stage. Digital image conversion in an operational environment requires a deep and longstanding institutional commitment to preservation, the full integration of the technology into information management procedures and processes, and significant leadership in developing appropriate definitions and standards for digital preservation.

A New Framework for Effective Leadership

One view of leadership involves establishing control through the application of standards and procedural guidelines. Those who hope to exercise comprehensive control over digital imaging technology most probably will find that moral persuasion cannot often prevail in the absence of a significant market share. Others have argued that the rapid pace of technological change and the sheer complexity of information technologies render librarians and archivists helpless in influencing product developments. Yet, those who prefer to wait and see how digital imaging technology shakes out before making the administrative commitments necessary to ensure long-term preservation are shirking their responsibility to help define the terms of the debate.

Each of these perspectives ignores an imperative of leadership. "A leader is one who mobilizes others toward a goal shared by leader and followers," writes Garry Wills (1994, p. 70). Preservation in the digital world must encompass shared goals that leaders and followers elicit together. It is the responsibility of many people in many institutions fulfilling many roles of both leader and follower. Accepting the differentiation of roles is crucial if librarians and archivists are to identify which of the many facets of digital technology they can control, where their expertise is one of many important influences, and when they must accept the processes and products of the socioeconomic system of which libraries and archives are but one part.

The following table summarizes key aspects of digital imaging technology that are important for creating and maintaining access to digital image files over a long period of time. It arrays these aspects on two dimensions: the priorities for action and the degree of power that librarians and archivists can have over accomplishing the goals implied by the priorities.

**Longevity:** Librarians can exercise a large measure of control over the life-expectancy of digital image data through the careful selection, handling, and maintenance of rugged, well-tested recording and storage systems. They can influence the life expectancy of the information by making sure that local budgetary commitments are made consistently and at an appropriate level in support of long-term migration strategies. Ultimately, the library community may have little control over the evolution of the imaging marketplace, especially corporate research and development activities that have a tremendous impact on the life expectancy of the digital files being created today.
Leadership Options for Preservation Action

<table>
<thead>
<tr>
<th>Control</th>
<th>Influence</th>
<th>Accept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longevity</td>
<td>Choice of Storage Media</td>
<td>Imaging Marketplace</td>
</tr>
<tr>
<td>Choice</td>
<td>Intellectual Content</td>
<td>Migration Strategy and Timing</td>
</tr>
<tr>
<td>Quality</td>
<td>Conversion Quality Standards</td>
<td>Image Compression and Display Standards</td>
</tr>
<tr>
<td>Integrity</td>
<td>Index Content and Structural Standards</td>
<td>Metadata Interchange Standards</td>
</tr>
<tr>
<td>Access</td>
<td>Purchase of Non-proprietary Equipment</td>
<td>Backward Compatibility of Systems</td>
</tr>
</tbody>
</table>

Choice: Even while recognizing that selection decisions cannot be made autonomously or in a vacuum, librarians and archivists can choose which books, articles, photographs, films, and other materials are converted from paper or film into digital image form. Influence over the continuing value of digital image files is largely vested in the right to decide when it is time to migrate image data to a future storage and access system and in knowing when a digital file has outlived its usefulness to the institution charged with preserving it. Librarians and archivists, while not directly responsible for the capabilities of their patrons, can nonetheless help determine how digital information will be provided and help educate users about use of new technologies to meet their needs.

Quality: The commercial imaging market has transformed the principle of maintaining the highest possible quality over time to one of finding the minimal level of quality acceptable to today’s system users. Quality of the digital object, including the richness of both the image and the associated indexes, is the heart and soul of preservation in the digital world. This means maximizing the amount of data captured in the digital scanning process, documenting image enhancement techniques, and specifying file compression routines that do not result in the loss of data during telecommunication. Librarians and archivists can control standards of digital image quality, just as they have done for microfilm. They can only influence the development of standards for data compression, communication, and derivatives for display and output. Research laboratories and the imaging marketplace will have the greatest influence over improvements in the technical capabilities of digital conversion and storage systems.

Integrity: Librarians and archivists can help exercise control over the integrity of digital image files by authenticating access procedures and documenting successive modifications to a given digital file. They can also create and maintain structural indexes and bibliographic linkages within well-developed and well-understood database standards. Librarians and archivists are acknowledged experts in organizing information and so have a vital role to play in influencing the development of metadata interchange standards, including the tools and techniques that will allow structured, documented, and standardized information about data files and databases.

*Figure 9. Floppy disk. 106,200 char./in.²*
to be shared across platforms, systems, and international boundaries. However, librarians and archivists may find they are but bystanders in the rapid development of network protocols, bandwidth, or data security techniques.

Access: The digital imaging marketplace is highly competitive. Librarians and archivists can do the most to ensure continuing access to digital image files through prudent purchases of non-proprietary hardware and software components. In the present environment, true “plug and play” components are becoming more widely available; purchase decisions are the most powerful incentives for vendors to adopt open system architectures or at least provide better documentation on the inner workings of their systems. Additionally, librarians and archivists can influence vendors and manufacturers to develop new systems that are “backwardly compatible” with existing ones. This capability assists image file system migration in the same way that today’s word processing software allows access to documents created with earlier versions. The life expectancy of a given digital image system and the requirement to migrate valuable digital image files to future generations of the technology are profoundly important matters over which librarians and archivists have little or no control. It seems that the commitment of a vendor to support and maintain an old system is inversely related to that vendor’s ability to market a new system.

Conclusion

Preservation in the digital world challenges librarians and archivists, but not necessarily due to a lack of understanding of digital technologies, selection criteria, or appropriate preservation options. The technologies are becoming understood well enough. Informed choices about what to preserve in the digital world and the best ways to accomplish preservation goals both derive from this understanding.

The real challenge is creating appropriate organizational contexts for action. Cooperative strategies adopted from ongoing preservation practice are not all that is needed to address the complexities or the costs of digital technologies. Libraries and archives must do more than simply divide the preservation pie. The national digital libraries that are now under construction should be built under an umbrella of guiding preservation concepts. Additionally, new approaches to creating and maintaining digital files cooperatively must be developed. The acceptance of shared responsibility and shared funding commitments will preserve access to what are, in essence, unique electronic collections.

Will the totality of the national digital library—whether it is a single entity or many things to many people—be of such high quality and of such extraordinary value to a university, a state government, or other institution that each participant will make a long-term commitment to help support its preservation? Only if the answer to this question is yes will it be possible to build that library cooperatively. Establishing such commitments requires the kind of leadership that has set the context for traditional preservation action in the past two decades.
## List of Figures

1. Clay tablet: *Arcadian Love Song*. Early 18th century BC. "Early Incantations and Rituals" *24, Museum # YBC4643, Yale Babylonian Collection, Sterling Library, Yale University. 135x175x33 mm.

2. Papyrus: *Unidentified Fragment*. P.C.Y BR inv. 422(A)fol. Beinecke Photo File, Beinecke Library, Yale University. 215.9x228.6 mm.

3. Illuminated manuscript: *Rothschild Canticles*. 1300-1400, 182v.-183r. MS404, Beinecke Photo File, Beinecke Library, Yale University. 118x84 mm.


9. Floppy disk: 1.44 megabyte floppy disk. Preservation Department, Yale University Library.

10. Optical disk: 1.3 gigabyte magneto-optical disk. Project Open Book, Yale University Library.
BIBLIOGRAPHY

Preservation


Digital Preservation


---

**Microphotography**


**Technology, Culture, and Libraries**


