Perspective

Archival Theory and the Preservation of Electronic Media: Opportunities and Standards Below the Cutting Edge

FREDERICK J. STIELOW

Abstract: The archival mission presupposes stewardship, which can provide the field with a mechanism to carve a distinctive niche within automated environments. The author warns against the untutored embrace of the "cutting-edge" of new technology. Instead, he suggests the need for study and reliance on standards, as well as the adoption of a process-oriented view for preservation management to deal with rapid technological change. He discusses the effects of automation on archival theory, presents common sense guidelines, and includes a brief analysis of specific magnetic and optical storage media. He also proposes a new archival law: With each new storage medium, archivists must reexamine their theory and expect to meet new preservation challenges. Original version presented before the Society of Ohio Archivists, 11 April 1991, with funding under a grant from the Online Computer Library Center, Inc. (OCLC). Revisions were facilitated by support from the Commission on Preservation and Access. An earlier version of this essay was published in the *Ohio Archivist* 22 (Fall 1991): 3–7.

About the author: *Frederick Stielow is now the executive director of the Amistad Research Center at Tulane University. Before earning a dual Ph.D. in history and American studies and an M.L.S. with an archival concentration, he served two years as a data processing specialist and later was head of archives at the University of Southwestern Louisiana. Stielow went on to spend ten years teaching archives, information technology, and preservation at the Catholic University of America and the University of Maryland. He has written or edited six books, including the Leland Prize-winning* Management of Oral History Sound Archives, *and more than sixty articles.*

A PATTERN HAS EMERGED in starting presentations on the preservation of electronic materials: Disasters! In 1975, the U.S. Census Bureau discovered that only two computers on earth can still read the 1960 census. The computerized index to a million Vietnam War records was entered on a hybrid motion picture film carrier that cannot be read. The bulk of the National Aeronautics and Space Administration's research since 1958 is threatened because of poor storage. These tales are akin to Jorge Luis Borges's short story in which the knowledge of the world is concentrated in one mammoth computer—and the key is lost.[1]

The essential question for the Information Age may well be how to save the electronic memory. Answers will need to come from several arenas. Manufacturers and technologists must keep preservation in mind during their testing and production. But governments and individuals at the institutional and buyer's level will need to remind manufacturers and technologists of their duties. Ultimately, information managers will have the basic responsibility for deciding what will be preserved, and here it may be well to call on the expertise and traditional mission of archivists.

According to Don Wilson, archivist of the United States, archivists now face a rare "window of opportunity" on the care and preservation of electronic records. The archival field must empower itself to meet this challenge by stepping forward and proving itself with theories and standards.[2]

Archivists know that they are already enveloped by the computer revolution. We deal with automation as a tool for office management and for the control and description of collections. Archival automation also refers to machine-readable files and in the near future may take on wholly new meanings with the potentials of optical character recognition and even bar coding systems. Still, not all archival programs need to be totally proactive or to have the mission and resources of the National Archives. Individual archivists are circumscribed by the pragmatic realities of budgets and the types of media actually used in their institutions. As soon as different types of information storage enter their domain, however, archivists automatically incur a new burden: studying and trying to understand the physical nature of the media. The archival mission for automation does not stop with administrative uses and tool skills; it extends to stewardship and inherent professional responsibilities for building a documentary heritage.

Whatever the problem, archival recourse should begin from preexisting traditions and strengths. No single answer exists for all archives or media—only general principles that must be reanalyzed over time. Any such evaluation should include the longitudinal vision and holism of historicism, which set archivists apart from other information managers. This perspective provides a useful counterbalance to the short-term blinders of systems analysis—the mode of thought that dominates the computer world and allowed for the original preservation oversights. But we cannot rest on passive models. Our theories and responses must be sufficiently proactive and flexible to deal with the dynamic nature of the new storage vehicles.[3]

Archivists must deal with reality. For ex-

[1]These examples have appeared repeatedly: See, for example, Kenneth Tibodeau, "Keynote Address," Preservation of Electronic Records Conference, National Archives, 19 March 1991, a presentation that heavily influenced this paper.

[2]Don Wilson, "Welcoming Remarks," Preservation of Electronic Records Conference.

[3]Frederick Stielow, "Information Technology and Archival Theory," JELIS, forthcoming, from a paper presented at the May 1992 Mid-Atlantic Regional Archives Conference.

ample, the financial concerns of the computer industry do not necessarily serve the preservation of records. The computer industry has vested interests in producing new and proprietary products with little continuity or thought of preservation—the obverse of an archival perspective. Too many buyers have picked up the industry's habits and the excitement of "cutting-edge" products with little or no thought about time and future consequences. The results are technology-driven selections fraught with rapid obsolescence, compatibility problems, million-dollar mistakes, and "vaporware." In such a context, preservation devolves to "use it or lose it," the archival recourse to maintaining Smithsonians of outmoded equipment or high reconversion costs.

Solid archival practice and theory emerge away from the cutting edge. They flow from a consumer- and information-driven perspective. The historical lessons are clear: Electronic preservation has a chance of success only at the place where standards exist and where we can reasonably project some constancy over time.

Archival Theory and Electronic Preservation

Historically, archival theory developed from physical abilities to read and analyze information, which is inextricably linked to the paper or parchment on which it was written (human-readable records). This approach is no longer sufficient. Stare as you might, the floppy disk is indecipherable to the human eye. Thus, such traditional archival skills as paleography and diplomatics must be rethought. Although contextual analysis remains, physical legitimation of signatures and concepts of originality are quite moot for instruments that can readily be manipulated and copied in an undetectable fashion. Those who wish to understand authentification for electronic signatures must now possess other types of knowledge, including document authentification codes and basic encryption standards—such as DES and RSA, which are now used to safeguard military and financial transactions. Even the sacred precepts of original order and provenance must be reconsidered in light of "virtual records," "multimedia documents," and "groupware," in which elements of a text may have been drawn from multiple data sources created by many authors at different times and places.[4]

Such reassessment does not deny the validity of earlier archival hypotheses or approaches. Rather, it builds on them, with a healthy reexamination for enhancement. Theoretically, preservation arises from the doctrine of stewardship, but electronic records alter a portion of this role. Concepts of artifactual or media-specific intrinsic values disappear while security consciousness may be heightened. The guarded nature of archival repositories with closed stacks, vaults, and climate controls emerges with extra merit for the off-line storage and safeguarding of computer files against attack.[5]

The burden of preservation is actually an advantage that may provide archivists with a mechanism to guarantee their niche within

[4]Luciana Duranti, "Diplomatics: New Uses for an Old Science," *Archivaria* 28 (1989): 7–24, and Jane Turner, "Experimenting with New Tools," *Archivaria* 30 (1990): 91–103. Both show other areas of applicability for the auxiliary sciences. In February 1992, the comptroller general of the U.S. Commerce Department formally approved the validity of electronic signatures. My information on encryption stems partially from participation as the only archivist at the second Computers, Freedom, & Privacy Conference in March 1992. Ken Sister, *Information Security in Financial Services* (New York: Stockton Press, 1991).

DES stands for the secret Data Encryption Standard put forth by the National Bureau of Standards in the 1970s and employing a 64 bit block cipher; RSA (Rivest, Shamir, Adleman) is a newer 100 bit asymmetric block cipher.

[5]Frederick Stielow, "Archival Security," in James Gregory Bradsher, ed., *Managing Archives and Archival Institutions* (Chicago: University of Chicago Press, 1988), 207–16.

modern information management. Security transfer and backup storage are venues that can tie the archives directly into the active information flow within an institution. Instead of the wait for deposits near the end of the life cycle, electronics allows a routinization of archival deposits. Valuable data can be automatically conveyed to the archives at data entry, at some key phase in the life cycle, or as part of normal backup procedures. This approach also coincides with the common doctrine that effective preservation lies with "front-end" controls from the moment of creation. (The alternative is far more expensive—conservation repairs later in the life cycle.) Once sequestered, the archivist can also begin to think from a more expansive information resources management perspective about enhancing values and perhaps even about marketing new information goods.[6]

Electronic preservation itself is rooted in well-established principles of redundancy and sequestration, with precedents that date back to the origins of writing in ancient Sumer and Egypt. Typical procedures involve the production of a sacrosanct master record. The master is used to generate user copies, which are then disbursed and periodically recopied. As with all media, such replication and extra storage space come at a price. Preservation implies added expense and the corresponding call for cost-benefit evaluations. Electronic preservation elevates such concerns and adds new ones. Archivists of machine-readable files must be attentive not only to costs for storage, recording media, and handling, but also to potentially high charges for hardware, software, and their maintenance.

In addition, computer records bring new information categories for preservation. Unlike the unity of a paper record, a computer record typically consists of two physically distinct sections. The first is storage flat file of raw data. The second is a logical structure file that controls the data's visual or intellectual representation, such as the row and column definitions in a spreadsheet. Both sections of the record should be captured. At present, the logic file is too often ignored for long-term preservation purposes.[7] Issues and concerns such as these promise to revolutionize archival practices. Archivists must anticipate a restructuring of descriptive theory. The new structure will include the capturing of process information and automatic item-level description. A benefit will be the prospect of enhanced retrieval from subitem-level descriptions of the "meta-data" such as data elements or field titles (for example, the categories of *To: From: Date: Subject:*).[8]

The benefits of assessing logical structures lend credence to structural appraisals for both digitized data and raster-scanned maps or photographs. For example, we could argue for additional analysis in terms of the best storage media for types of records. Not

[6]Douglas Finlay, "Archives: Old Records Meet New Technologies," *Administrative Management* 47 (1986): 37–40. This article provides one of the few examples of such thoughts outside archival literature.

[7]Harold Naugler and others from SAA's Committee on Archival Records and Techniques (CART) have made similar points in discussion. The present ways to save the structures in a value-neutral format are by inference from data element dictionaries, which should be constructed in keeping with Information Resources Dictionary System (IRDS) standards, and through software-independent notational structures, such as Standard Generalized Markup Language (SGML) for published and Structured Query Language (SQL) for relational databases. For all practical purposes, the major alternative is maintaining the originating software. For the few computer "ancients," we could also pose questions on saving the documentation for the layout of the wired boards or the switches that were flipped on earlier generations of computers.

[8]For one exploration of such changes, see David Bearman, "Archival Methods," *Archives and Museum Informatics Technical Report* no. 9 (Pittsburgh: Archives and Museum Informatics, 1989). Although a good deal of work is being done on this section of archival theory, whole areas remain to be debated: e.g., to parallel discussions in artificial intelligence with an exchange on the dialectic between library-based, "neat" descriptive categories versus the "fuzzy" and relativistic nature of an archival description.

all computer records need to be digitally saved. Some may be better routed to paper or to something like computer-output microfilm (COM)—perhaps with the readily indexable, or "strippable," headers retained online.[9] Hence, we can add a new compositional or morphological analysis to evaluate:

- the format of the information to distinguish between fixed communications (memorandums, letters, directives) and more mutable databases with the potential for mathematical analysis and new combinations.
- the ease and utility of removing indicative pointers to make separate indexes to the contents of the document (especially for electronic mail and internally produced communiques).
- the degree to which the logical structure cannot be replicated in another media (such as computer-aided design, hypertext, intelligent systems, and spreadsheets).

Similar hypothesizing dismisses the originating storage medium from ongoing preservation concerns. Storage is theoretically incidental to a process in which the information is truly displayed (human-readable) only when engaged in the computer's active memory (RAM). The originating medium lacks informational values and is of interest only as a potential vehicle of transfer into the archives. This aspect of electronic preservation thus frees one segment of preservation theory from a focus on an information/media continuum (e.g., the letter on paper or the bound book) to an information-independent model.

Note: The preceding observation does not liberate archivists from the demands of understanding the nature of machine-readable records—rather, they are freed only from preserving the information in one particular format. As Charles Dollar pointed out long ago, archivists actually take on the added burden of a technical appraisal and sampling to ensure readability of the information at the point of transfer.[10]

Legal issues also remain. The archivist must be cognizant of the 1986 Computer Fraud Act and must also pay particular attention to the division between physical ownership and copyright. Assuming the archives owns both latter rights, there would be no problem in making another copy onto any media. (One would, however, need to be especially careful when marketing a product.) The Copyright Act of 1976 also allows the library or archives as physical owners to make *a* copy "in kind" for preservation purposes (Title 17 U.S. Code: Copyright). Whether copying for preservation purposes to a new media is legally permissible is an interesting question that may ultimately turn on archivists' own policy development. Other legal questions on retention scheduling and authentification remain, but, again, archivists should note that their expertise in such issues can further define their positions within modern information management.

Archivists, too, must be concerned with technological knowledge about the media and machinery under their control. At present, this suggests keeping up with two broad categories of automated storage: magnetics, currently the dominant medium, and optics, which is now entering the scene. Such expertise does not call for an engineering degree. Again, this new responsi-

[9]This paper does not deal specifically with appraising information values, which can be seen in such works as Michael Anderson, "The Preservation of Machine-Readable Data for Secondary Analysis," *Archives* 17 (1985): 79–93, and Bearman, "Archival Methods." The National Research Council, *Preservation of Historical Records* (Washington, D.C.: National Academy Press, 1988) had made recommendations for the off-line printing for preservation of all retainable electronic records. This "flat" view is inconsistent with archival appraisal and the values of some of the logical structures.

[10]Charles Dollar, "Appraising Machine-Readable Records," *American Archivist* 41 (1978): 423–30.

bility is accompanied by new opportunities. In most cases, few institutional "techno-jockeys" outside of the mainframe environment will have taken much account of the storage media and preservation concerns.[11]

Magnetic Storage Media

The two magnetic formats of general interest are digitized computer tapes and analog recordings. Digitized computer tapes have been the underlying model medium for much of the preceding discussion. For preservation, analog recordings, especially sound tapes and video, are much more troublesome. For, unlike digitized records that can be copied exactly, each new generation or recopying of a videotape or audiotape produces a loss in signal. Videotapes are even more problematic than audiotapes, which have a longer history and standard formats. Videotapes are quite short-lived (no matter what the television commercials say about permanent memories) and appear in several noncompatible versions.[12]

Any "permanent" solution to the problem of preserving analog recordings will likely arise through digitization—converting the analog wave structures into binary code. At this time, however, digital redress is somewhat limited by the amount of storage capacity required. The practical choices are either optical storage (which will be discussed later in this article) or such new vehicles as digital audio tape (DAT).

DAT employs helical scanning technology. This technology borrows from videotape concepts and, at least for the moment, is also a prime illustration for avoiding the cutting-edge syndrome. Helical scanning gathers information by rotating drums with two or more heads that scan the moving tapes in diagonal stripes. Such scanning provides initially high-quality reproduction and almost unbelievable compression ratios. Sampling at a rate as high as 48,000 times a second, DAT can consume the equivalent of 2.8 megabytes of storage in a minute. Unfortunately, the initial equipment is sensitive to movement and even microscopic particles can quickly throw it out of line. Moreover, according to hearsay evidence from engineers in the Association of Recorded Sound Collections (ARSC), DAT longevity is quite suspect. DAT was reportedly partially withheld from the market following an accidental discovery that its original usable life was a month or so. Even now, it may last only one to three years.[13]

In contrast, long-term storage of magnetic tape for fixed-head recordings is acknowledged to be fifty to one hundred years, given proper storage conditions. Works from the early 1980s, such as Ford Kalil's *Magnetic Tape Recording in the Eighties* and Sidney Geller's *The Care and Handling of Computer Magnetic Storage Media,* provide a good starting point for understanding the proper treatment of magnetic tapes. The basic rules of thumb are commonsensical: Block the introduction of stray magnets and reduce particulate contamination by banning eating, drinking, smoking, and pencils. (The graphite pieces and eraser shavings may attach to the read heads and scratch—felt-tip pens should be used instead.) Tapes should be of good quality with a gamma ferric oxide emulsion layer on a thick (1mm to 1.5mm) mylar base. They should be

[11]Note that earlier forms of machine-readable storage, such as punch cards and paper tape, may still be in storage. For an introduction to the preservation literature, see Mary Bowling, "Literature on the Preservation of Nonpaper Materials," *American Archivist* 53 (Spring 1990): 340–48.

[12]Frederick J. Stielow, *The Management of Oral History Sound Archives* (Westport: Greenwood Press, 1985); see the chapter on conservation.

[13]Philips is introducing DCC, a new linear-scan DAT similar to traditional fixed-read mechanisms but with a radically accelerated compression schema that will also play analog tapes. Andrew Pollack, "Another New Wave in Sound," *New York Times,* 23 December 1990, p. F-9.

evenly wound under playback tension, stored tails out, and exercised annually or biannually, preferably on a tape winder/cleaner or device that does not engage the read heads.[14]

Even though each play is an attack, the frequent engaging of data is, ironically, a major guarantee that it will be around in the near future. Still, damage to magnetic tapes does occur far more quickly from improper conditions during playback than during storage. Equipment maintenance is an essential element in magnetic preservation. Another recommendation is for filtered, positive-pressure air conditioners to be used to reduce particles in the air. The environment is crucial and, for both storage and playback, should be stable, with nonfluctuating temperature and humidity levels. The most frequently quoted figures for temperature advocate a range of from 60 to 70 degrees Fahrenheit (65° is optimal), and for humidity from 35 to 50 percent relative humidity (40 or 45 percent is optimal).[15]

Yet, given the volatile state of the technology, how likely is it that the equipment will exist to read those fifty-year-old, well-maintained tapes? Instead of thinking about permanent storage media, archivists may need to consider a dynamic program. Unlike the two-hundred-year cycles of scribal monks in the Dark Ages, archivists may be recopying records to newer standardized media every five to ten years. In addition, the technology is changing so rapidly that even the standards do not stand still. For example, the open-face reel with ferric oxide particles—the longest-standing storage constant in this area—is itself effectively replaced in archives by "streaming" to the "square reel" 3480 cartridge with a chromium dioxide emulsion. But since 1987, a new industry standard has arisen, and the 8mm quarter-inch cartridge is now used in a quarter of a million machines.[16]

As the following section indicates, magnetic storage technology itself may not survive in anywhere near its current state, and this may be to the benefit of preservation. The physical wear and tear on current tapes as they pass the read/write head means that each use of the tape is an attack. Add to that the threat of head crashes. Traditional magnetic storage is also labor-intensive; it requires ongoing attention for proper climate and periodic exercising, or the tapes will rapidly degrade. Such intrusive factors naturally lead to a search for other media that (like the paper records of yore) are more resistent to playback and environmental factors.

Optical Storage Media

Optical disks are the new mass storage rage. They offer a physically tough and relatively climate-neutral medium that can be exactly copied without generational loss. The acceptable range for storage and playback conditions, for example, extends from 50 to 120 degrees Fahrenheit and from 10 to 90 percent relative humidity. Because

[14]Sidney Geller, *Care and Handling of Computer Magnetic Storage Media* (Washington, D.C.: National Bureau of Standards, 1983); Ford Kalil, ed., *Magnetic Tape Recording for the Eighties* (Washington, D.C.: National Aeronautics and Space Administration, 1982), especially Appendix B, "A Care and Handling Manual for Magnetic Tape Recording."

[15]Ibid.; Geller recommends lower temperature and humidity standards for long-term storage to reduce the exercise interval, but notes that the recordings should be stabilized at least one day in the playback environment before they are read. Such recommendations are still open for debate, as are questions such as the propriety of reel storage within neutral plastic bags to reduce potential desiccation of the binder and the preferred nature of the backing layer for computer tapes.

[16]Thomas Weir, *3480 Class Tape Cartridge Drives and Archival Data Storage* (Washington, D.C.: National Archives, 1988). These cartridges typically hold 200 megabytes and current products appear to have decent longevity, but chromium dioxide is nowhere as stable as ferric oxide particles. Arlin Raedeke, "Technology Update: 8mm Helical Scan Tape," *CD-ROM Professional* 5 (March 1992): 81–83, this article also notes the superiority of the 8mm quarter-inch cartridges over DAT, with its more limited 3.81mm-width face.

optical disks employ a laser reading mechanism, there is also no wear during use and no real danger of head crashes.

Play-only CD-ROMs are fairly well established and standardized, but they still have problems. Proprietary searching software and problems with networking may still bedevil the user and retrieval problems do occur, even with Yellow Book guides for error detection, SCSI (Small Computer Systems Interface), and High Sierra or ISO 9660 standards for the arrangement of materials on the disk. Indeed, if you had not heard—disasters. Stories circulate about disk rot and the inability to address certain sections of the disk over time. A fungus is reportedly capable of blocking reading, and reports from England described fogging as label inks ate through to the metal layer of the disk. Although dramatic, such difficulties are minimal compared with the nagging question of longevity. The industry's claims have escalated from early ten-year predictions to Sony's or Digipress's more reliable Century Disk with its one-hundred-year predicted life. These figures await confirmation from independent sources. Ten to twenty-five years seems to be the best current "guesstimate" for most current products.[17]

Instead of published CD-ROMs, archivists will be more interested in less established WORM or WO (write once–read many) technology. With 600 megabytes in the typical 5.25-inch CD version and the promise of gigabytes in the near future, WORM technology provides an intriguing prospect. It may have the potential for replacing microfilm, even for legal retention. Although it may sound like looking a gift horse in the mouth, archivists must be wary of this still quite costly option. As the Canadian National Archives has already discovered, standards are lacking for the original technology, especially for the large 12-inch disks. Eighteen months into their venture, the Canadians' pilot project was somewhat jolted when their supplier announced its upgrade to a new, noncompatible disk format.[18]

Only at the CD level can archivists feel somewhat safe under the ISO continuous composite format and currently developing Orange Book guides. The former arrived with the development of M-O (magnetic-optical) erasable disks and has since been extended back to newer WORM products.[19] But again, the question arises: Will the current formats last even ten years? The archivist just thinking of taking this route may want to hold off to be certain of the survival of WORM technology. The market niche has remained fairly small and is well under the levels projected since its introduction in 1984. Major manufacturers once announced that they would stop production to concentrate on more marketable but far less archivally attractive (and currently more expensive) rewritable disks.[20]

[17]*CD-ROM Product Guide* (Parsipanny, N.J.: Bureau of Electronic Publishing, 1990); William Saffady, *Text Storage and Retrieval Systems* (Westport, Conn.: Meckler, 1989); Denis Oudard, "Archival Technology: The Evolution of Century Disk," *CD-ROM Professional* 4 (November 1991): 41–46. Oudard provides a good description of this project. (We might note that like so many other background articles, the article comes from the vendor.) In regard to standards, a student in April 1991 asked to borrow an old Bibliophile data disk from my office. She was unable read the non-ISO 9660 and called the company, which informed her they no longer even had the software for this ancient version from 1986. *Tempus fugit.*

[18]John J. Hay, "Write-Once or Not Write Once," *Optical Information Systems* 11 (1991): 24–25. In addition to that journal, *Byte, PC-World,* and other computer journals, one can also consult specialized CD-ROM serials, such as *CD-ROM End User, CD-ROM Librarian,* and *CD-Rom Professional.* Rewritable players should be designed to the ISO 9660 standard to read their WORM cousins.

[19]This emerging CD standard should not be confused with the U.S. National Security Agency's Orange Book for automation security.

[20]Roger Blais, "Institutional Perspective: National Archives of Canada," Preservation of Electronic Records Conference. I am aware of at least one other

Although some dye polymers are entering the market as relatively stable recording strata, the rewritable or erasable M-O disk is making the most waves. Not a true optical medium, M-O digitization occurs on a highly resistant ablative (corrective) stratum. During this stage, a laser heats individual bits to the Curie point, where they can easily be turned on with a weak magnetic impulse. Such technology offers obvious benefits for easy direct downloading from magnetic tapes, for bigger projects that can readily absorb a queue of 30,000 pages, and for the retrieval of pictorial or map images. (Note, however, that the retrieval of images is not for preservation purposes but for publication and for the convenience of users.) During manufacture, it is possible to introduce phase-encoded protocols (PEP) and standard format protocols (SFP) to convert such disks, or portions of them, to WORM. However, some experts have challenged the reliability of the codes that supposedly prevent overwriting, and, more important, these media are still new. Although the technology is tough, environment-independent, and not subject to head crashes, the industry currently promises only a five- to ten-year life span for such encoding.[21]

The basic cautions remain. Plan. Do not move into optical data storage until you are truly ready. Then, use only that equipment advertising compatibility with standards. And consider the arguments that some observers have raised—that optical data storage is a transitional medium which, like eight-track audiotapes of yore, may disappear within ten years.

Technology and Standards

The keys to success for electronic preservation begin to emerge in the form of careful purchases, commonsense procedures, demands for standards, and an understanding of the strange acronyms that appear in this article. The ideal configuration is an *open system*—a hardware- and software-independent platform in which the components parallel the connectivity of an audio stereo system. At a minimum, the system must provide the interoperability offered by common command languages. These emerge from transportable operating systems (those that are not machine-dependent or proprietary, such as DOS, OS-2, and Unix) plus the standards issued by the American National Standards Institute (ANSI), the Consultative Committee on Telephone and Telegraph (CCITT), and the International Standards Organization (ISO).[22]

archival WORM project in which the data are no longer retrievable. Note also the appearance of 3.25-inch optical disks and the increase in scanning speeds, as well as a troublesome movement from the Unix base to a different standard; Andrew Young, "The CD-ROM Standards Frontier: Rock Ridge," *CD-ROM Professional* 4 (November 1991): 53–56. As I revised this article in April 1992, I also noticed JVC announcements for a new WORM, which supports Orange Book and ISO 9660 standards (the basic questions for such media) and costs around $12,000.

[21]Barry Cinnamon, *Optical Disk Document Storage and Retrieval Systems* (Washington, D.C.: Association for Information and Image Management, 1988); information also from a variety of vendor brochures, including CanonFile 250, Bow Industries' R-50 Series Magneto-Optical Storage Sub-System, and 3M's Rewritable Optical Disk Cartridge Reference Manual. Among other possible preservation concerns in regard to true format compatibility, some questions on unintended overwriting due to failures in write-protect protocols have been raised. On archival uses for optics, see David Bearman, "Optical Methods," *Archives and Museum Informatics Technical Report* no.1 (Pittsburgh: Archives and Museum Informatics, 1987); Pamela Cruse, "Multifunction Optical Offers Versatility," *CD-ROM Professional* 4 (July 1991): 72–75; and John Nairn, "Proposed Disk Description Protocol," *CD-ROM Professional* 4 (March 1991): 56–57, which discusses compatibility with the new multimedia standards.

[22]Operating systems are the layer of software that communicates with the computer and translates storage locations and commands between standard programs (fourth-generation languages—such as database management systems, spreadsheets, and word processors) and the central processing unit. ANSI controls Z39 and is a subset of ISO; the United Nations' CCITT

A little knowledge can go a long way. Even name recognition helps and lends credibility, through the undeniable power of jargon. Archivists should become acquainted with such esoterica as publishers' use of SGML (Standard Generalized Markup Language) to replicate printer's marks and IRDS (Information Resources Dictionary System) for unifying data element dictionaries across an institution. Archivists who need to should be exposed to the ODA/ODIF office document standards and ANSI's X12 EDI (electronic data interchange) protocol for the exchange of standardized business forms. The OSI (open systems interconnection) model of the International Standards Organization is crucial, but it is already under some archival purview through MARC (machine-readable cataloging) and the library world's Z39 bibliographic standard.[23]

A select section of the archival community is somewhat belatedly working to insert archival and preservation concerns into the standards process and to translate current developments for the archival profession. By 1990, for example, the Society of American Archivists (SAA) had launched its own standards committee, which is trying to clear up potential confusion and to publish applicable standards in a somewhat understandable fashion.[24]

A few larger institutions preceded SAA and are in the unenviable position of sitting on the standards boards and pushing at the cutting edge on the world stage. The Canadian National Archives, for example, has been quite active, especially with data standards. The U.S. National Archives was responsible for commissioning significant works on magnetic storage and on the role of standards for preserving electronic records, especially through its mid-1980s creation of an Archival Research and Evaluation Staff. Independent archivists are also leaving their imprint, notably Richard Kesner and the iconoclastic David Bearman with his *Archives and Museum Informatics.*[25]

At a more practical level, archivists in general are responding to these new conditions. Archivists are studying and attending continuing education courses. The profession's new incoming members also tend to be more computer-literate, especially those from programs that meet the 1988 SAA Graduate Education Guidelines.

At the institutional level, the central challenge remains inserting archivists, or at least a consciousness of preservation and standards, into the normal process for hardware and software selection. This concept is the backbone of electronic preservation. A commonsense hint is to purchase products only from vendors that advertise compliance with standards and to be sure that the product itself complies. For example, a database management system must offer SQL (Structured Query Language) compatibility; an electronic mail package must comply with the X.400 standard. Fortunately, such needs can be met more readily now than at anytime in the past. Consumer demands have reduced the exclusivity of

works on X.400 and ISDN. For a further introduction to the acronym forest, see Victoria Walch, "Checklist of Standards Applicable to the Preservation of Archives and Manuscripts," *American Archivist* 53 (Spring 1990): 324–39.

[23]SGML, for example, has standard codes to duplicate printer's marks for the construction of a text; these codes are software-independent and readily translated from one program to the next. SQL commands to define database structures can be similarly "filtered" to allow the same database to be manipulated by different DBMSs.

[24]Personal files, SAA Standards Committee.

[25]Charles Dollar, *A National Archives and Records Administration Strategy for the Creation, Transfer, Access, and Long-Term Storage of Electronic Records of the Federal Government* (Washington, D.C.: National Archives, 1990); Margaret Law and Bruce Rosen, *Framework and Policy Recommendations for the Exchange and Preservation of Electronic Records* (Washington, D.C.: National Institute of Standards, 1989); National Archives of Canada, *Data and Document Interchange Standards and the National Archives* (Ottawa: National Archives of Canada, 1987).

proprietary systems and pushed manufacturers and suppliers toward upward compatibility and open systems. Marketplace competition seems to be working, and computer prices for power and storage are steadily declining. Indeed, a likely side benefit of recopying cycles will be reduced transfer and storage costs as users move to more powerful and compact systems.[26]

Problems will still confront archivists especially those with the need or desire to be at the cutting edge. These individuals and institutions usually operate in the arena of proprietary systems, away from the safety of standards. They also pay a dollar premium for equipment and face the perils of more "bug-ridden" products. But more conservative archivists can also expect difficulties. Initially, all archivists are left to the mercy of industry pronouncements on longevity and production that may bear little resemblance to truth. The "hard-wired" folks themselves are not known for their ability to communicate; jargon, technical specifications, and their version of English can create a bewildering maze that the neophyte may be unable to navigate. Moreover, as in other areas of preservation, the results of practical research are often confined to a very small group of insiders.[27]

The basic dilemma, however, is the fluidity and rapid generational turnover in the market. Newer types of attractive software with elaborate logical structures, like hypertext and expert systems, constantly appear with no accompanying standards. Ultimately, purchasing remains a disturbing art form: practical necessity means one must automate, but the purchaser knows that the product will soon cost less and will undoubtedly be replaced by a more powerful version.

Conclusion

Archival stewardship for the information technologies provides archivists with both responsibilities and a rare window of opportunity. *Carpe diem.* Archivists can use preservation to gain a toehold within automated information management environments. Moreover, while archivists must study the makeup and tendencies of any medium under their purview, their traditional archival skills and thought can play a positive role.

The preservation of machine-readable files does alter portions of archival theory and can lead archivists to proactive positions. In contrast to the continuity over time of human-readable records, machine-readable files are subject to rapid cycles of obsolescence. Archivists face a constant evolution of standards, equipment, and storage media, as well as the lure of exciting products that may quickly fade from the scene. Archivists simply cannot afford to wait at the end of the computer industry's conveyor belt for whatever comes their way; they

[26]Along with higher compaction ratios for magnetics and optics, one can also predict the addition of revolutionary new storage media in the near future, for example, the possibilities from industry work on solid-state technology and crystalline storage in ceramics. Billy Allstetter, "Bacteria Could Lead to Drives with the Equivalent of 5000 Platters," *Byte* 17 (January 1992): 32, discusses research on a bacteria in salt marshes that produces a protein capable of storing a gigabyte per cubic centimeter. Currently available options include optical tape, with its laser-encoded dye polymer format that offers 5,000 megabytes of storage per cubic inch and accelerated age testing at the 15-year level; see David Owen, "Optical Tape Provides High Density Low Cost Storage," *CD-ROM Professional* 5 (March 1992): 73–75. Archivists will need to keep a wary eye on the General Agreement on Trade and Tariffs (GATT) discussions that are the basis for international standards.

[27]The problems in communicating preservation advances to the practitioners and the frequent delays in publishing basic findings are at once intriguing and vexatious. Some examples of "folkloric" wisdom just now emerging for analog recordings include the rapid forwarding of stored audiotapes prior to replay, which in some way reduces print-through, and the temporary restoration for immediate recopying of faded or separating tapes by baking at a constant 130° Fahrenheit in a convection oven.

must assert themselves and their concerns for enduring documentation from the time of purchase.

Preservation in this new age must respond from good management and not from a focus on conservation repairs. Archivists are buying into a process and a set of artificially enforced standards, not a permanent product. Thus, they must be prepared for short-term cycles and ongoing recopying and reappraisal with each new media base. History suggests that archivists should not embrace any of these successor formats as panaceas. Rather, the new law is that with each new storage medium, archivists must reexamine their theory and expect to meet new preservation challenges.[28]

[28]To examine some of the current thoughts on managing information in the electronic environment from the technology side, such as document image processing, intelligent document management, and optical character-recognition systems, see the articles in "The Paperless Office," *Byte* 16 (April 1991): 156–241; "Managing Infoglut," *Byte* 17 (June 1992): 244–90. Although they now include standards, such technology-driven views still rarely account for human operators or the idea of a documentary heritage.