

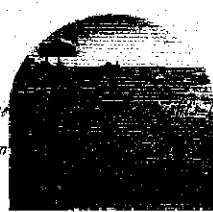
Timeless and
Durable for Life



Tilley Hats!

The Tilley Hat blocks UV Rays, repels rain, ties on, floats, won't shrink or mildew and comes in every size. Best of all, if your Tilley Hat ever wears out it will be replaced FREE.

Tilley Endurables
FREE CATALOGUE 1-800-363-8737
www.tilley.com



Spring in Palmilla

PALMILLA
A RESORT

Las Cabezas, Mexico • 1-800-637-2226 • www.palmliflavors.com

SHAMROCK BRACELET

Sterling Silver \$165
14k Gold \$595

1-800-235-0471 5/4 65.75

ROGER NICHOLS *Jewelry*
17 NW Irving Avenue • Bead, OR 97101 Value Box

Since 1989

UPTON TEA IMPORTS

Purveyor of the World's Finest Teas

1-800-234-8327 Free catalog listing over 220 varieties of garden-irish, loose tea

www.nptontea.com

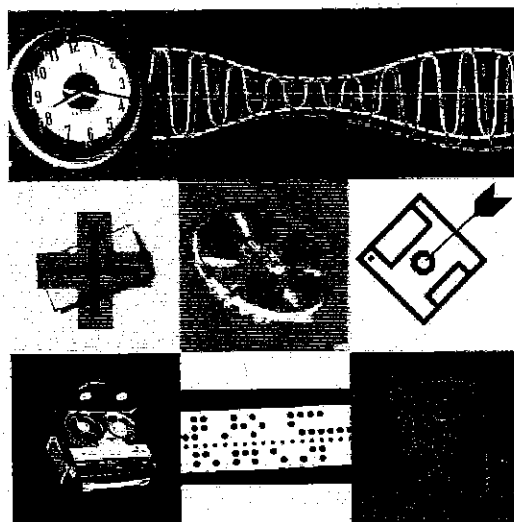
231 South St. * Hopkinton, MA 01748

HIGH-TECH DEPT.

OVERLOAD

There's just no way to save all the information of the information age.

BY ALEXANDER STILLE



IN a temperature-controlled laboratory in the bowels of the vast new National Archives Building, outside Washington—nearly two million square feet of futuristic steel-and-glass construction—an engineer cranks up an old Thomas A. Edison phonograph. A cylinder disk begins to turn, and from the phonograph's large metal horn we suddenly hear the scratchy oompah-pah of a marching band, striking up a tune at a Knights of Columbus parade in July of 1902. Nearby sits an ancestor of the modern reel-to-reel tape recorder: it's the very machine that recorded President Harry Truman's famous whistle-stop speeches as he travelled around the country by train in 1948 to win an extraordinary come-from-behind election victory. Instead of capturing sound on magnetic tape, the device stored its data on spools of thin steel wire as fine as fishline. Now some of the wire has rusted, and it occasionally snaps when it is played back through the machine.

This laboratory, in the Department of Special Media Preservation, is a kind of museum of obsolete technology, where Archives technicians try to tease information out of communications devices that vanished from circulation long ago. But the lab is more than a curious rag-and-bone shop of old gizmos; in many ways, it offers a cautionary vision of the future. The problem of technological obsolescence—of fading words and images locked in odd-looking, out-of-date equipment—is an even bigger problem for the computer age than it was for the media produced in the first half of the century.

One of the great ironies of the information age is that, while the late twentieth century will undoubtedly record more data than have been recorded at any other time in history, it will also almost certainly lose more information than has been lost in any previous era. A study done in 1996 by the Archives concluded that at current staff levels it would take approximately a hundred and twenty years to transfer the backlog

ANDY MARTIN

of nontextual material (photographs, videos, film, audiotape, and microfilm) onto a more stable format. "And in quite a few cases we're talking about media that are expected to last about twenty years," Charles Mayn, the head of the laboratory, says. Decisions about what to keep and what to discard may be made by default, as large portions of the archive simply deteriorate beyond the point of viability.

Mayn is a tall, thin man with gray hair, a gentle manner, and the neat, understated dress of a computer engineer of the nineteen-fifties—the time of his youth. A self-described "science weenie," he is uncomfortable talking about himself and plays down his own considerable ingenuity in rebuilding many of the old machines with the help of modern parts in order to get them to play back intelligible sounds and images. His particular laboratory is dedicated to "dynamic media"—things with moving components, such as movie cameras and audio and video players. In his spare time, Mayn has been known to scour junk shops and yard sales in the Washington area, looking for old Dictabelts or movie projectors.

A short distance down the laboratory workbench from the Edison phonograph are some eighteen-inch acetate disks, precursors of the long-playing vinyl record. The United States Army used them to record enemy radio broadcasts throughout the Second World War. They play on a machine called a Memovox, which has a turntable that changes speed as the record plays, slowing down to compensate for the quicker rotation of the disk as the stylus approaches the center. It was an ingenious invention, but it didn't catch on, perhaps because it required rather complex internal machinery. A glass disk, used on yet another recording device, marked "Germany, October 24, 1941, 11:55 P.M.," lies shattered on the turntable. "Luckily, the glass generally breaks in fairly clean pieces, so we are often able to put them back together," Mayn explains. The Archives has some seventy thousand of these military recordings, each of which has a playing time of up to two hours. It would take a researcher who worked without interruption for eight hours a day approximately forty-eight years to listen

to this collection in its entirety. "A lot of the disks may contain a lot of nothing—airwave noise, shortwave whistles—but you may have to listen to the whole thing to figure that out," Mayn says.

On one wall are the internal organs of a film projector from the nineteen-thirties. The old heads have been mounted to play with modern reels. "Twenty-eight different kinds of movie soundtracking systems were devised during the thirties and forties in an effort to improve the quality of sound," Mayn explains. "Most of them are incompatible." This particular model used something called "push-pull" technology, in which the sound signal was split onto two different tracks. The technology was meant to cancel noise distortion, but the two tracks

must play in near-perfect synchrony. "If it is played back properly, it is better than a standard optical track, but if it is played back even a little bit improperly it is far, far worse," Mayn says. About fifteen years ago, he was actually able to use this reconfigured projector to show several reels of push-pull film from the Nuremberg trials, and the lab has transferred some eighteen hundred reels of push-pull tape onto new negatives.

Potentially, the information age appears to offer the historian a Holy Grail of infinite memory and of instant, permanent access to virtually limitless amounts of information. But as the pace of technological change increases so does the speed at which each new generation of equipment supplants the last. "Right now, the half-life of most computer technology is between three and five years," explains Steve Puglia, a preservation and imaging specialist, whose laboratory is just down the hall from Mayn's. In the nineteen-eighties, the Archives transferred some two hundred thousand documents and images onto optical disks—the cutting edge of new technology at the time. "I'm not sure we can still play them," Puglia says, because they depend on computer software and hardware that are no longer on the market.

There also appears to be a direct relationship between the newness of a technology and its fragility. Paul Conway, a librarian at Yale University, has created a graph going back to ancient Mesopotamia which shows that, while the quantity of information being saved

has increased exponentially, the durability of media has decreased almost as rapidly. The clay tablets that record the laws of ancient Sumer are still on display in museums around the world, and many medieval illuminated manuscripts written on animal parchment still look as if they were painted or copied yesterday, whereas modern books printed on acidic paper as recently as the nineteen-eighties are already turning to dust. Black-and-white photographs will last a couple of centuries, but most color photographs become "unstable" within thirty or forty years. Videotapes deteriorate much more quickly than traditional movie film does—they last about twenty years. And the latest generation of digital storage tape is considered to be safe for only ten years.

Digital technology, which is based on incredibly precise mathematical coding, either works perfectly or doesn't work at all. "If you go beyond the limits of the error rate, the screen goes black and the audio goes to nothing," Mayn says, "and up to that point you haven't realized there were any errors. Analog technology"—used on vinyl records and on electromagnetic tapes—"deteriorates more gracefully." The old wax cylinders of the original Edison phonograph sound faded and scratchy, but they are still audible. Mayn picks up a tiny plastic digital audiotape that fits neatly in the palm of his hand. "People love these things because they are so small, compact, and lightweight, and can store tons of data, but as larger and larger amounts of data are crammed into smaller and smaller spaces the technology gets more precise, more complex, and therefore more fragile. We have a lot of these tapes from the late nineteen-eighties that can't be played at all."

THE National Archives and Records Administration, or NARA, was established during the nineteen-thirties, on the optimistic premise that the government could keep all its most vital records indefinitely, and thus act as our nation's collective memory. Now, as the agency drowns in data and chokes on paper, it is facing the stark realization that it may not be able to preserve what it already has, let alone keep up with the incessant flow of information coming its way.

The numbers are so huge as to be al-



most comical. The Archives is currently the custodian of four billion pieces of paper; 9.4 million photographs; 338,029 films and videos; 2,648,918 maps and charts; nearly three million architectural and engineering plans; and more than nine million aerial photographs. Storage consumes nearly half of the agency's budget, so, ironically, the more information it keeps the less money it has to spend on making this information available to the public. Because other government agencies are generally not required to hand over their records for permanent storage until thirty years after their creation, the Archives has only recently begun to grapple with the extraordinary explosion of information produced during recent decades.

Space has been a problem at NARA since before it opened, on November 8, 1935, in a grand neoclassical building on Pennsylvania Avenue, down the street from the White House. That building was designed to have a handsome internal courtyard, but the nation's first archivist had the space filled in to accommodate more stacks. These, too, quickly proved inadequate, so the original seven stories of high-ceilinged rooms were chopped into thirds, creating twenty-one short floors of stacks. Any researcher who is more than six feet tall thus risks a concussion navigating this warren of scarily identical corridors and shelves. Six rows of shelves on a single floor hold all the documents generated by the United States Supreme Court in its first hundred and forty years of life, while it takes the rest of that floor—the equivalent of about half a city block—to house the Court's papers from the last sixty years.

By 1993, the Archives had outgrown its original home and a new headquarters was built in College Park, Maryland. Although it is the third-largest government building, half the size of the Pentagon, Archives II, which opened in 1994, is already approaching the limits of its storage capacity. Despite predictions made some twenty years ago about the "paperless" office, most government agencies are still printing out their computer files and producing

more paper than ever. Each year, on the average, the Archives receives about one and a half million cubic feet of new records, of which about a third are stored. At this rate, the space for paper records at Archives II is expected to run out by 2007.

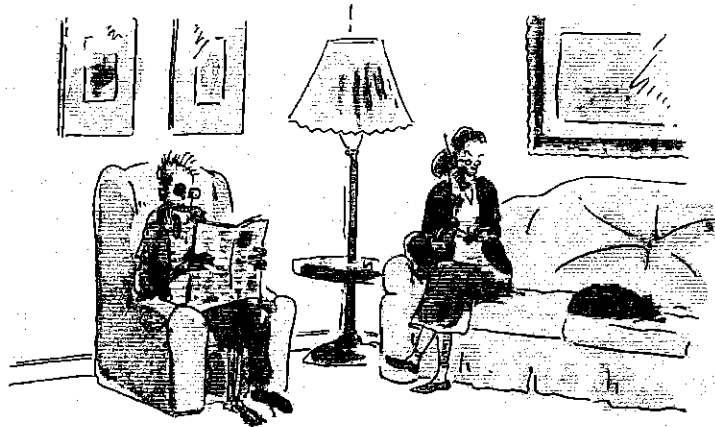
In theory, computer technology should be more helpful in the storage of text than it is in storing the audio and video records of Mayr's dynamic media lab. So far, however, computers have only compounded the problem. In 1989, a public-interest group trying to obtain information on the Iran-Contra scandal successfully sued the White House to prevent it from destroying any electronic records. The result is that all federal agencies must now preserve computer files and electronic mail. But it took the Archives two and a half years (and its entire electronic-records staff) just to copy the electronic records of the Reagan White House, and "they are gibberish as they currently stand," says Lynette Eaton, who worked at the Archives' electronic-records center before moving over to the Smithsonian Institution.

The problem is that the multitude of departments, agencies, and offices that constitute the federal government use different computers and software programs—many of them old and obsolete. Many of the records from the Vietnam War era are stored in a database-management system (known as NIPS) that I.B.M. no longer supports, and

which the Archives has difficulty translating into readable form. The Agent Orange Task Force, for example, has been unable to use herbicide records written in the NIPS format.

For several years, a disturbing rumor has circulated that the data from the United States census of 1960 were lost. According to the story, the information exists only on obsolete thirty-nine-year-old computer tapes, which can no longer be read. The Archives has reassured the public that the material has been safely copied onto more modern media, but, because census data are kept confidential until seventy-two years after their collection, the rumor will probably persist until 2032, when the data are made available to researchers. Meanwhile, later census surveys are still at risk. A few years ago, the Archives informed Congress that the Bureau of Census "reported to us . . . that they have over 4,000 reels of tape, containing permanently valuable data, which are difficult, if not impossible to use because they are in CENIO"—Census Input/Output—"format or because the files have been compressed on an ad hoc basis."

Because of the problems posed by outdated computers and software, the Archives issued an order in 1995 that government agencies were free to print out their E-mail for permanent storage. But Scott Armstrong, a journalist, who helped bring the initial White House E-mail suit, has protested that direc-



Shanahan

"He refuses to let me get on with the grieving process."

tive. "It makes no sense," he says. "If your basement were flooded, the first thing you would try to do is turn off the flow of water, and then start worrying about mopping up. The Archives is doing the exact opposite. It is already drowning in paper, but it is still telling people to print out their records onto paper. If the government had dedicated the energy it has spent fighting the E-mail lawsuits into modernizing its record-keeping operations, it would have gone a long way toward solving its problems." Under court order, NARA is formulating a new directive, to be announced next month.

ALTHOUGH it may be true that "the era of big government is over," as President Clinton has said, the era of big-government data banks is just beginning. Ken Thibodeau, the director of electronic-records programs at the Archives, maintains that Armstrong and others disregard the immense technological difficulties in trying to recover E-mail from thousands of different computers. Between 1989 and 1996, the Electronic Records Division took in twenty-five thousand new records. Then the E-mail from the Reagan-Bush White House, delivered four years after the fact, suddenly buried it in an avalanche of two hundred thousand electronic files, just as the State Department prepared to hand over a million two hundred and fifty thousand electronically stored diplomatic cables. Since then, the State Department has been averaging about a million messages a year. Meanwhile, President Clinton's office is pumping out E-mail at the rate of more than six million files a year.

This kind of volume could crash the Archives computer system. "We designed a system to handle maybe ten thousand messages a year," Thibodeau says. "You cannot scale up our system to deal with a million messages a year." The problem is that most E-mail programs were not written with long-term storage in mind. So, in the current state of technology, the Archives computers must treat each individual E-mail message as a separate file, which has to be opened and closed in order to be copied from one tape to another.



Thibodeau says the Electronic Records Division was looking at new, sophisticated storage devices coming onto the market, but that these presented problems of their own. "There's a kind of tape that can hold two hundred times the volume of the plastic cartridge we use," he says. "So it would be great in terms of space. But, as we've been talking to people who use this technology, we haven't talked to anyone who has successfully taken the tape out of the silo and read it on a different machine." The extreme precision of the new miniaturization technologies is such that each machine produces tapes that are unintentionally customized to a particular alignment of the laser beams that encode and read information. "A slight misalignment of the head is sufficient to guarantee that you will never read the tape except on a machine that has the same misalignment," Thibodeau points out. Because of this difficulty, the Archives has turned to the San Diego Supercomputer Center to find a new solution that is not yet on the market.

Ironically, the downsizing of government has actually magnified the information crisis. "When a government agency downsizes, usually the first thing it does is get rid of the record-keepers and clean out the storage closet," Mayn says. "We suddenly get a call telling us to pick up a trailer-size truckload of records." When the Pentagon closed Norton Air Force Base, for instance, it decided to turn over its huge motion-picture-storage warehouse to the Archives, thus doubling in a single stroke the Archives' video holdings. At the same time, the Archives has to do more with fewer resources. The budget of the Electronic Records Division fell by about fifteen per cent and its personnel was cut by ten per cent during a period when the volume of new data has increased tenfold. Since 1996, the budget has almost doubled, though the workload stands to increase a thousandfold.

The problem, in Mayn's view, is that nobody inside or outside government is making the tough decisions about what to store and what to throw away. "I'm not a historian, but, personally, I have my doubts about some of the stuff we're trying to keep," he says. "Do we really

need dozens of different films on the workings of the M1 tank? I can see keeping a few as a sample, but I'm not sure we need the entire collection." At the height of the Vietnam War, the Pentagon routinely sent hundreds of men with cameras out into the jungle to film the combat. "Each of these people was told to shoot hundreds or thousands of feet of film," Mayn says. While much of this film is of genuine historical interest, the total quantity would take one technician several lifetimes to copy. Mayn believes that, because much of this material will eventually deteriorate, the choice of what to keep will be made by default. "We will keep those things that researchers happen to have requested and that consequently get copied onto new media," he says.

In the end, the sorting out of the information explosion may repeat the process that determined which books we now possess from antiquity. The works of authors such as Homer and Virgil survived intact because of their enduring popularity and the multiple copies that were therefore made at different times. But many of the works that we regard as "fixtures" of our culture (including Plato) were lost for several centuries and are known to us only because one or two copies turned up in medieval monasteries or in the collections of Arab scholars. Undoubtedly, a good many great works did not survive at all: Sophocles is known to have written more than a hundred and twenty plays, of which we have only seven.

The danger now is not that some modern Sophocles will be totally lost in the databases of the Department of Agriculture or of the Census Bureau but, rather, that such a vast accumulation of records makes it nearly impossible to distinguish the essential from the ephemeral. The Archives of the future may resemble "The Library of Babel," which Jorge Luis Borges imagined nearly sixty years ago—an infinite library that contained every conceivable book in the universe. There were books that consisted solely of the repetition of a single letter of the alphabet, and others in which all the pages except one were blank; the discovery of an intelligible sentence was cause for jubilation. Eventually, after many centuries, the librarians of Babel were driven to despair in their unfulfilled quest to find a coherent, complete book. ♦