LOST magazine

Are We Losing Our Memory? or
The Museum of Obsolete Technology

by Alexander Stille

Running out of time at the National Archives.

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 cylindrical disc begins to turn and from its large wooden 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 traveled the country by train during his legendary come-from-behind victory in the election of 1948. Instead of capturing sound on magnetic tape, the device stored its data on coils 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 modern media that have long vanished from circulation. But the laboratory is more than a curious rag-and-bone shop of technologies past; 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 The problem of technological obsolescence — of fading words and images locked in odd-looking, out-of-date gizmos — is an even bigger problem for the computer age than for the new media produced in the first half of the 20th century.

One of the great ironies of the information age is that, while the late twentieth century will undoubtedly have recorded more data than any other period in history, it will also almost certainly have lost more information than any previous era. A study done in 1996 by the Archives concluded that, at current levels, it would take approximately 120 years to transfer the backlog of nontextual material (photographs, videos, film, audiotapes, and microfilm) to a more stable format. "And in quite a few cases, we're talking about media that are expected to last about 20 years," said Charles Mayn, the head of the laboratory. Decisions about what to keep and what to discard will be made by

default, as large portions will simply deteriorate beyond the point of viability.

Mayn is a tall, thin man with gray hair, a soft-spoken, gentle manner, and the neat, understated, conservative dress of a computer engineer of the 1950s — the time of his youth. A self-described "science weenie," he is more comfortable fiddling with the interior of a machine than talking about himself. He plays down his own considerable ingenuity in rebuilding or reinventing many of these machines, rigging up pieces of the original items with 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 parts such as audio and visual players. In his spare time, Mayn has been known to scour junk shops and yard sales in the Washington area, looking for old castaway Dictabelts or movie projectors that have been consigned to the dustbin of history.

A short distance down the laboratory workbench from the Edison phonograph are some 18-inch glass discs, precursors of the long-playing vinyl record — rapidly becoming a relic itself. The U.S. army used the disks to record enemy broadcasts throughout World War II. They play on a machine called the Memovox, which has a turntable that changes speed as the record plays, slowing down to compensate for the quicker rotation of the disc as the stylus approaches the center, so that the needle always moves at a constant speed in relation to the groove in which it sits. It was an ingenious invention, but it didn't catch on, perhaps because it required rather complex internal machinery. A glass disk — marked "Germany, October 24, 1941, 11:55 p.m." — lies shattered on its turntable. "Luckily, the glass generally breaks in fairly clean pieces so we are often able to put them back together," Mayn explained. The Archives possesses some 70,000 of these foot-and-a-half military recordings, each of which has a playing time of about two hours. It would take a researcher who worked without interruption for eight hours a day approximately 48 years to listen to this collection in its entirety. "A lot of them 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 said.

On the wall are the internal organs of a film projector from the 1930s; the old heads have been mounted to play together with modern reels. "Twenty-eight different kinds of movie sound-tracking systems were devised during the 1930s and 1940s, trying to improve the quality of sound tracks," Mayn explained. "Most of them are unique and incompatible." This particular one used something called "push-pull" technology, in which the sound signal was split onto two different tracks. The technology was meant to cancel out 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 said. In the mid-1980s at a theater in downtown Washington, he was able to actually use this reconfigured projector to show several reels of push-pull film containing the trials of top Nazi leaders at Nuremberg. And the lab has transferred some 1800 reels of push-pull tape onto new negatives.

Potentially, the computer age appears to offer the historian's 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," said Steve Puglia, a preservation and imaging specialist whose laboratory is just down the hall

from Mayn's. In the 1980s, the Archives stored 250,000 documents and images on optical disks — the cutting edge of new technology at the time. "I'm not sure we can play them," said Puglia, explaining that they depend on computer software and hardware that is no longer on the market.

In fact, there appears to be a direct relationship between the newness of technology and its fragility. A librarian at Yale University, Paul Conway, has created a graph going back to ancient Mesopotamia that shows that while the quantity of information being saved has increased exponentially, the durability of media has decreased almost as dramatically. The clay tablets that record the laws of ancient Sumer are still on display in museums around the world. Many medieval illuminated manuscripts written on animal parchment still look as if they were painted and copied yesterday. Paper correspondence from the Renaissance is faded by still in good condition while books printed on modern acidic paper are already turning to dust. Black-and-white photographs may last a couple of centuries, while most color photographs become unstable within 30 or 40 years. Videotapes deteriorate much more quickly than does traditional movie film — generally lasting about 20 years. And the latest generation of digital storage tape is considered safe for about ten years, after which it should be copied to avoid loss of data.

Digital technology — 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 said, "and up to that point, you don't realize there are any errors. Analog technology" — used in vinyl records or electromagnetic tapes — "deteriorates more gracefully. The old wax cylinders of the original Edison phonograph sound faded and scratchy, but that are still audible." Mayn picked up some tiny plastic digital audiotapes that fit neatly in the palm of his hands. "People love these things because they are so small, compact, and lightweight and store tons of data, but as they put larger and larger amounts of data on smaller and smaller spaces, the technology gets more precise, more complex, and more fragile." He bends the little data tape in his hand. "We have a lot of these from the late 1980s and even the mid-1990s that can't be played at all."

The National Archives and Records Agency (NARA) was created during the 1930s on the optimistic premise that the government could keep all of its most vital records indefinitely, acting as our nation's collective memory. Now, as it drowns in data and chokes on paper, the agency is facing the stark realization that it may not be able to preserve what it already has, let alone keep up with the seemingly limitless flow of information coming its way.

The numbers are so huge as to be almost comical. The Archives is currently custodian to 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 its budget so, ironically, the more information it keeps the less money it has to spend on making it available to the public. Because other government agencies are generally not required to hand over their records for permanent storage for some 30 years, the Archives is only just beginning to grapple with the extraordinary explosion of information over the last generation.

Space has been a problem at the National Archives from before it opened on November 8, 1935 in a

grand neoclassical structure on Pennsylvania Avenue, down the street from the White House. That building was supposed to have a handsome internal courtyard, but the nation's first archivist had the space filled in for more stacks. These, too, quickly proved inadequate, so the high-ceilinged floors were chopped in half, creating 21 short floors of stacks. An archivist much over six feet tall would risk a concussion navigating this rabbit warren of seemingly identical corridors and shelves. Here you can see the information explosion in tangible terms. Six rows of shelves on a single floor hold all of the documents generated by the U.S. Supreme Court in its first 140 years of life, while it takes the rest of the floor, the equivalent of about half a city block, to house the papers from the last 60 years. One term of the Supreme Court now generates as much paper as 40 years did in the early 19th century.

With nowhere left to store all the paper, the Archives built new headquarters in College Park, Maryland, which opened in 1994. Although the third-largest government building and about half the size of the Pentagon, Archives II is already approaching its storage capacity. Despite predictions some 20 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 average, the Archives receives about 1.5 million cubic feet of new records, of which about one-third are kept for storage.

In theory, computer technology should be more helpful with the storage of textual documents than with the audio and video records of Mayn's dynamic media lab. But so far, it has only compounded the problem. In 1989, a public interest group trying to get information about 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 all their computer files and electronic mail. Because government offices use different kinds of computers, software programs, and formats, just recovering this material has proved to be a logistical nightmare. It took the National Archives two and a half years (and its entire electronic records staff) just to make a secure copy of all the electronic records of the Reagan White House. And it may take years more to make most of them intelligible. "They are gibberish as they currently stand," said Fynette Eaton, who worked at the Archives' Center of Electronic Records before moving over to the Smithsonian Institution.

The beauty of digital technology is that it reduces everything to a series of zeroes and ones — a simple, seemingly universal mathematical language — but unless one has the software that gives meaning to those zeroes and ones, the data is meaningless. The problem of deciphering Egyptian hieroglyphs may look like child's play compared with recovering all the information on the hundreds of major software programs that have been discarded during the astonishing transformations of the computer revolution.

The losses from the first decades of the digital age are likely to be considerable. The federal government, with its multitude of departments, agencies, and offices, is a dense thicket of incompatible computer languages and formats — many of them old and obsolete. Many of the records of the National Military Command Center are stored in a database management system (known as NIPS) that IBM no longer supports and that the National Archives has difficulty translating into readable form. The Agent Orange Task Force has been unable to use herbicide records written in NIPS format.

For several years a disturbing rumor circulated that the data from the United States Census of 1960 had been lost. According to the story, the information lies locked on obsolete 36-year-old computer tapes that can no longer be read by today's machines. The Archives continues to reassure the public that the material has been safely copied to more modern media, but because census data must be kept private until 72 years after its collection, the rumor will probably persist until independent researches can view the material for themselves in the year 2032. Meanwhile, later census surveys are still at risk. "Bureau of Census files prior to 1989 threaten to eclipse the NIPS problem," the Archives reported to Congress a few years go. "The Bureau reported to us ... that they have over 4000 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." Each computer tape can store 75,000 pages of information so that, if the data cannot be recovered, the Census Bureau might lose up to 300 million pages of data.

Because of the problems posed by reconstructing obsolete hardware and software, the Archives issued an order that government agencies were free to print out their email onto paper for permanent storage. The Archives may be faced either with mountains of computer data it cannot interpret or an avalanche of paper of unprecedented volume. But Scott Armstrong, a journalist who helped bring the initial White House email suit, has protested the Archives' directive. "It makes no sense," said Armstrong. "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 are doing the exact opposite. They are already drowning in paper, but they are still telling people to print out their records onto paper. If the government had dedicated the energy it has spent fighting the email lawsuits into modernizing its record-keeping operations, it would have gone a long way to solving its problems."

Although "the era of big government is over," as President Clinton declared, the era of big government data banks is only just beginning. Ken Thibideau, the head of the Electronic Records Division at the Archives, insisted that Armstrong and others underestimate the immense technological difficulties in trying to recover email from thousands of different government computers. Between 1989 and 1996, the Electronic Records Division took in 25,000 new records. The email from the Reagan-Bush White House suddenly buried it in an avalanche of 200,000 files, just as the State Department prepared to hand over 1,250,000 electronically stored diplomatic cables. And this represents just the tip of the iceberg — the period from 1972 until 1975. Since then, the State Department has been averaging about a million messages a year. Meanwhile, in recent years, the White House has been pumping out an average of six million electronic files a year.

These expected additions could well lead to a crash of the Archives' computer system. "We designed a new system to handle maybe 10,000 messages a year. You cannot scale up our system to deal with a million messages a year," Thibideau said. His office ran an experiment, trying to copy one single storage tape of the Clinton email. The Archives' computer churned and ground for some 50 hours but failed to copy the entire tape. "We can normally copy a whole tape with up to 200 megabytes in about ten or 15 minutes," Thibideau explained. The reason the computer had so much trouble with the White House tape is that email systems are not designed with long-term storage in mind. Given the

state of current technology, the computer insisted on treating each individual email message as a single file that had to be opened and closed in order to be copied from one tape to another. It takes far longer to copy 100,000 one-page messages than to copy a thousand 100-page messages even through they may use up the same amount of space on the tape.

Thibodeau said the Electronic Records Division was looking at sophisticated storage devices coming onto the market, but these present problems of their own. "There's a new kind of tape that can hold 200 times the volume as the kind of data tapes we are using, the same plastic cartridge we use," Thibideau said. "So it would be great in terms of space. But as we talk to people who use this technology, we have not talked to anyone who has successfully taken the tape out of the silo and read it on a different machine." The extreme precision and miniaturization of the new technology is such that each machine produces tapes that are unintentionally customized to fit the particular alignment of the laser beams that encode and read information. It's as if you were stamping a record with grooves that were thousands of times smaller than on an LP and using a stylus that needed to land just right in order to play back the record. "When you get to these highly dense media, your tolerance for error is extremely small, "said Thibideau. "A slight misalignment of the head is sufficient to guarantee that you will never read the tape other than on a machine that has the same misalignment. And if you are in the archive business, if we can't take a tape from another system and read it on ours, then it's no good. "

Ironically, the downsizing of government has actually magnified the information crisis. "When a government agency downsizes, usually the first thing they do is get rid of record keepers and clean out the storage closet," Mayn said. "We suddenly get a call telling us to pick up a trailer-sized truckload of records." When the Pentagon closed Northrop Air Force Base, it decided to turn over its huge motion picture storage warehouse to the Archives, doubling in a single stroke the Archives' video holding. At the same time, the Archives is having to do more with fewer resources. Factoring in for inflation, the budget of the Electronic Records Division has fallen by about 15 percent and its personnel have been cut by ten percent during a period when the volume of new data has increased tenfold. The staff of Mayn's dynamic media laboratory has been cut from 16 to nine in the last decade. Everyone seems to want to keep everything, but nobody wants to pay to keep it.

The problem, in Mayn's view, is that nobody inside or outside government is making the tough decisions about what to keep and what to discard. "I'm not a historian, but personally I have my doubts about some of the stuff we are trying to keep," Mayn said. "Do we really need hundreds of different films on the workings of the M1 tank?" he asked. "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 jungles and the battlefields to film the combat. "Each of these people was told to shoot hundreds or thousands of feet of film," Mayn said. While much of this film is of genuine historical interest, the total quantity would take several lifetimes for a technician to copy or for a researcher to study. Because much of this material will eventually deteriorate beyond the point of intelligibility, Mayn believes that the choice of what to keep will be made by default. "We will keep those things that researchers happen to have requested and that get copied onto new media," he said.

The sorting out of the information explosion may resemble the process that determined the 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 made at different times. But many of the works we regard as fixtures of our culture (including Plato) were lost for centuries and are known to us only because of a copy or two that turned up in medieval monasteries or in the collections of Arab scholars. Some works of undoubted greatness did not survive at all: Sophocles is known to have written some 120 plays, of which we possess only nine.

There is not likely to be a modern Sophocles in the databases of the Department of Agriculture or the Census Bureau. The greater risk, instead, is of such a vast accumulation of records that the job of distinguishing the essential from the ephemeral becomes more and more difficult. The Archives of the future may resemble the "Library of Babel" that Jorge Luis Borges imagined nearly 60 years ago, an infinite library that contained every conceivable book in the universe. There were books that consisted purely of a repetition of a single letter of the alphabet and others in which all the pages except for 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 for a coherent, complete book.

Back to Top

AUTHOR BIO:

ALEXANDER STILLE is the author of *Excellent Cadavers: The Mafia and the Death of the First Italian Republic and Benevolence*; *Betrayal: Five Italian Jewish Families Under Fascism*; and *The Future of the Past.* He is a frequent contributor to *The New Yorker* and lives in New York City.

Buy Alexander Stille's books through Amazon at the LOST Store.

Where loss is found.