



Books

Publish And Perish

Elisabeth Eaves, 12.01.06, 12:00 PM ET

Nothing is safe. Not your e-mails, digital photos or Word files. Not old newspapers or books. When it comes to storing information, everything will disappear into digital obsolescence or crumble to dust.

Even White House e-mails, important blueprints and influential works of 20th-century literature--the very artifacts that you'd expect would be carefully preserved--are at risk of being lost forever.

The National Archives and Records Administration, the agency responsible for preserving the federal government's documents, realized in the 1990s that it couldn't cope with the digital era using its old electronic storage system of magnetic tapes. The White House under the Bush Administration alone will generate as many as 100 million e-mails. Copying them would take years. NARA has contracted Lockheed Martin to build a federal digital archive, but the system won't be ready until at least September of 2007.

The Library of Congress, meanwhile, ditched most of its original newspaper collection after transferring the content to microform, which uses a machine to read film. But Nicholson Baker, in his book *Double Fold: Libraries and the Assault on Paper*, says the medium is at least as iffy as paper: Some early acetate films "shrink, buckle, bubble or stick together in a solid illegible lump," he writes. In the '80s libraries switched to polyester-based films. But some types of polyester films are prone to spots, others attract fungus and another suffered "complete image loss" when exposed to the high heat of common microform readers.

Books aren't safe either. Librarians say that most works published between the mid-1800s and the mid-1980s are disintegrating thanks to the high acid content in their paper. A rescue treatment known as mass deacidification is commercially available, mainly from Pennsylvania-based Preservation Technologies. But it's expensive, says Thomas Teper, head of preservation at the University of Illinois library. As a result, he says, "no U.S. library has been deacidified completely." By the 1980s most publishers had started using acid-free paper, at least for hardcover editions. But Dianne van der Reyden, head of preservation at the Library of Congress, has a new worry over the rising use of recycled paper. "Every time it's recycled, it becomes weaker," she says.

The dream of preserving all human knowledge is an ancient one, dating back at least to the Library of Alexandria, which began assembling papyrus scrolls circa 300 BC. But Alexandria burned down, and as knowledge grew exponentially, the possibility of uniting it once again grew more distant--until the advent of computers, when our capacity to store words and images suddenly became vast.

While digital technology promised huge amounts of virtual warehouse space, though, our data are not all safe and accessible. Some computer scientists have dubbed this era a "digital dark age" because we may end up with no record of it. Part of the problem is the breakneck pace of technological change, which results in alarming cases of obsolescence. Several years ago, U.S. Navy engineers noticed that diagrams of the USS Nimitz, a nuclear-powered aircraft carrier, had been subtly transformed by new software. Over at NASA, early spaceflight data stored on digital tape had deteriorated irreversibly by the 1990s. Of course, untold numbers of people have experienced the personal calamity of losing the contents of their home computers, thanks to hard-drive crashes.

Alexander Rose, the executive director of the futurist Long Now Foundation, worries about the impermanence of digital information. "If you save that computer for 100 years, will the electrical plugs look the same?" he asks. "The Mac or the PC--will they be around? If they are, what about the software?" So far there's no business case for digital preservation--in fact, for software makers like Microsoft, planned obsolescence *is* the plan.

"The reality is that it's in companies' interest that software should become obsolete and that you should have to buy every upgrade," Rose says. We could be on the cusp of a turning point, though, in the way businesses and their customers think about

digital preservation. "Things will start to change when people start losing all of their personal photos," Rose said.

So what, if anything, can be done? In the short term, at least, open-source software and nonproprietary file formats--like .txt, .xml and .html--give you the greatest chance of migrating your documents forward as technology changes. As for the historical record, the Internet Archive, a nonprofit organization that collaborates with the Library of Congress and the Smithsonian, is going some of the distance to save us from a dark age. It captures Web pages before they disappear and stores them in its searchable Wayback Machine. Co-founder Brewster Kahle says that if no one recorded all the material originating online, "we'd live in the perpetual present, in which any organization could change history by taking down the Web page."

Gathering information is one thing, saving it another. To keep its digital files accessible, the Internet Archive has to move them to a new system every three years, Kahle says, and the organization is beginning large-scale data swaps with foreign libraries. "The real answer for digital preservation is diligence and don't just have one copy," he says. "You can be faced with institutional instability, government instability, geographic instability."

Massive book-scanning projects like the one launched by Google may help preserve literature by making it more accessible. "The broader the access to any resource, the more likely it is to survive," says Rose. That's because someone is more likely to notice, and raise the alarm, when a format stops working. On the other hand, "as the digitization projects proceed, the desire for universities to hold onto their physical books may decrease," says Kahle. Space-strapped libraries could decide to send old books to the dump, just as they have done with hundreds of thousands of historic newspapers. And scanned books are as vulnerable to technological change and obsolescence as other digital formats.

Can anything last forever? The Long Now Foundation is micro-etching its 15,000-page Rosetta Project, an archive of data on human languages, onto a 3-inch metal disk it hopes will last at least 10,000 years. But we still may not have improved on 4,000-year-old technology. Asked what the most permanent medium is, Kahle doesn't miss a beat: "The clay tablets of the Babylonians. Their libraries are readable to us today."

[Click here for more on our special report on books.](#)