

Remembrance of Things Past

I RECENTLY CLEANED OUT MY father-in-law's safe deposit box. There wasn't much in it: just a diamond ring that hadn't been worn in more than 30 years, and two birth certificates—one for him, and one for my recently deceased mother-in-law.

Years ago, a family's safe deposit box might hold a treasure trove of goods and documents. Opening a box, you might expect to find jewels, stock certificates, or the deed for some long-forgotten property. But that time is long past. These days, we use bits inside

a computer's memory bank, not tokens of irreplaceable paper, to keep track of our life's records and our net worth. Few people hold stock certificates; information about stock ownership is kept in brokerage accounts. Few officials insist on seeing an original birth certificate; a fax or a photocopy will suffice. Even interest in gold and jewels seems to be faltering: in the 1960s, my father-in-law told me, his father gave him a gold watch—as something to sell if he were ever out of cash and needed to eat. Such was the mind-set of people who lived through the Great

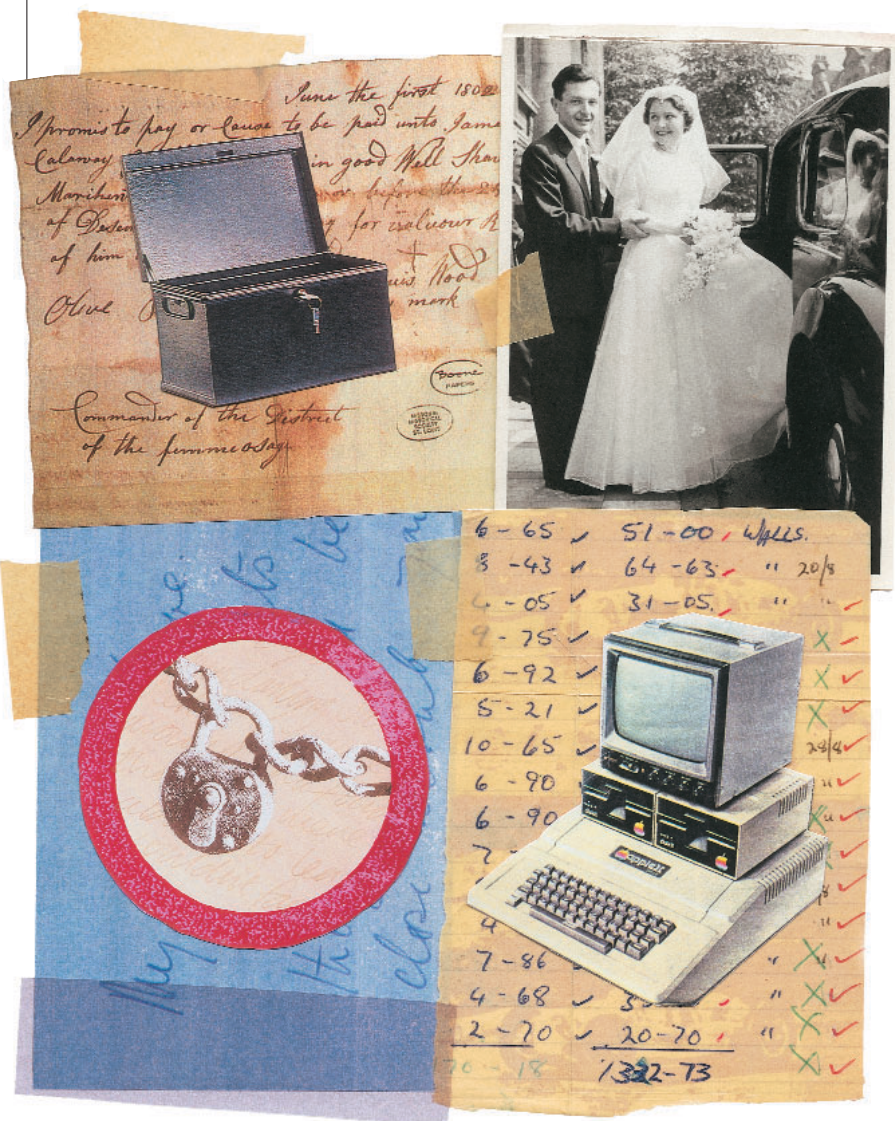
Depression. But these days, few people buy jewels for their investment potential. Instead, jewelry and gold is mostly bought for enjoyment and show.

Today it is data, more than money, that is the lifeblood of our society. And yet more than three decades into the "Information Age," data is something that we still don't quite understand how to steward. Data is not physical, not something that you can lock away today and hope you'll be able to access in 10 or 20 years. Large collections of data are almost impossible to safely maintain—especially over long periods. At the same time, data is just as difficult to dispose of properly. Indeed, individuals and businesses now have so much data in so many different formats on so many different computers that we are all heading for our own individual data catastrophes.

I once bought 10 used computers from a store that was going out of business. The machines were old and slow, but I didn't care—I wanted them for parts and software tinkering. I took them home, and just before I wiped their hard drives I decided to see what was on them.

I couldn't believe what I had stumbled upon. One computer had been a file server for a medium-sized law firm; with a few keystrokes I retrieved from its hard drive letters to clients, court filings and employee records. Another machine had been used by an organization that was delivering mental health services, and a third by a stockbroker: it had records of trades and account numbers, and more. Were I less scrupulous, I suppose that I could have had a lot of fun—and perhaps caused a lot of mischief—with the information that I had unwittingly purchased.

It's easy to chide the now-defunct store for failing to protect its customers, but the sad truth is that removing sensitive information from



modern computer systems is hard to do. As Oliver North learned during the Iran-Contra hearings, hitting "delete" is not enough. Instead, to properly clean, or "sanitize," a hard disk, it is necessary to overwrite every single block of storage. This can take hours, and even then it doesn't guarantee true erasure; readily available soft-

Data is now our lifeblood. Yet as we store information in many different formats on many different computers, we are heading for our own individual data catastrophes.

ware tools can recover information after a disk has been "formatted." Most people don't bother sanitizing their computers before they throw them away; they just toss and pray.

My story isn't unique. Over the years there have been news reports of used computers turning up with records from the federal witness-protection program, pharmacies and police departments. And it's likely to be a growing problem: according to a 1997 study by researchers at Carnegie Mellon University, some 325 million computers will be obsolete by the year 2005. And that means a lot of potentially damaging information on the loose.

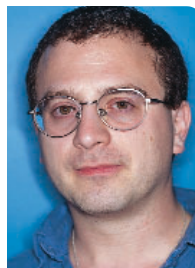
But at the same time that we are doing a poor job disposing of our data, we are doing an equally poor job of holding onto it.

In my basement, for instance, I have a collection of eight-inch floppy disks. These disks hold all of the papers and letters that I wrote in high school on the first computer that I ever owned. Alas, that machine has long since departed from the face of this planet. I doubt that I will ever be able to read those disks again, and I don't have a copy of the documents anywhere else.

The MIT Artificial Intelligence Laboratory had the same problem with a large collection of magnetic tapes made in the 1970s and '80s. Even the National Archives has had problems with computer records: you can't just leave them in a box. Instead,

you need to copy them every three or four years from older computers to newer computers. Failure to do so risks losing the data as the magnetic medium deteriorates.

This endless cycle of copying is the approach that I now take with my home computer. On my computer there are three electronic folders that contain all



the digital data from the last two decades of my life that I truly care about. There are three gigabytes of e-mail stretching back to 1983, another gigabyte of articles, letters and papers that I've written, and one more gigabyte of programs that I've coded, photographs I've taken, financial records and electronic keepsakes. Every time I get a new computer, I painstakingly copy this data from one machine to the next.

Organizing this data store over the past two decades has been a major challenge. But even after I got all of my directories set up, a continuing problem was software churn. For example, today's Microsoft Word can't read the letters that I wrote on my Macintosh in 1994 with WriteNow. Similarly, today's e-mail programs can't access the mailboxes of my old e-mail files, even though the messages themselves are stored as pure text. As a result, on those occasions that I need to go back and search for things, I usually end up using Unix and Linux tools that are comfortable working with pure text files, rather than the fancy Windows-based applications that can't handle even minor variations in file formats.

Another fear of mine is losing the data due to some sort of hardware failure. Like most computer users, I don't do a particularly good job of backing up. In all of last year, I made but a single tape backup. Instead, I protect my data by using multiple hard disks. The computer is set up so that every piece of information is recorded simultane-

ously onto two matched hard drives; if one drive fails, I still have a copy. As a second level of backup, at the end of each day my computer automatically copies the files that I've modified since the beginning of the month to a third hard drive. This archive has saved me on numerous times when I have accidentally deleted an important file. Even with these safeguards in place, though, I still manage to lose information from time to time.

All of this is a lot of work, but that's the price I pay to make sure my data is safe. Unfortunately, as hard disks have become more and more reliable, many people and organizations have forgotten the need to constantly back them up. In the old days, when hard drives might be expected to fail about once a year, you had a clear incentive to do your backups. Now that disks fail only every five or 10 years, keeping up sound data practices seems like busywork.

A growing number of companies are now trying to help businesses and individuals deal with these data issues. Retro Box, based in Columbus, OH, picks up a company's aging computers, properly sanitizes the hard drives and then helps to redeploy the computers within the corporation, sell them on the open market or donate them to charity. Several online backup companies, such as SkyDesk (www.backup.com) and Connected (www.connected.com), will back up your files over the Internet to their own data vaults. Of course, if you actually use one of these companies, you need to trust them more than you trust yourself.

I'm sure that over the next 20 or 30 years we'll finally get the hang of this data thing. Years from now, when my grandchildren go to clean out my safe deposit box, they'll probably sit down at a computer terminal somewhere, have their eyes scanned by some kind of biometric reader and transfer the data from my data vault to theirs. Either that, or they'll just hit "delete" and wipe it all away.