Forbes.com

Electronic Evolution

Saving Dying Data

Arik Hesseldahl, 09.12.02, 12:00 PM ET

For **Bryan Bergeron**, inspiration struck when he wanted to run an old computer program.

He wrote his first computer program more than 20 years ago on a S100 system, a 1970s-vintage computer built around microprocessors like Zilog's Z80 and **Intel**'s (nasdaq: INTC - news - people) 8080. Back then, programs were written and saved to cassette tapes.

"A few years ago, I found the tape and realized I couldn't run the program," he says.

A Boston-based medical doctor who teaches at Harvard University's Medical School and the Massachusetts Institute of Technology, Bergeron later ran into a similar problem with a medical program he had once written for the Commodore 64 that he wanted to demonstrate to a colleague. He had also saved that program on cassette tapes but didn't have a machine that could read the tapes. He quickly turned to **eBay** (nasdaq: EBAY - news - people) and bought an old C64 in an auction.

- Everything Is Obsolete Except Ideas
- Technology For The Ages
- Technology Past Its Prime

It turned out that others he knew had experienced similar problems. The relentless march of computing progress has mercilessly left untold terabytes of old software, writings, computer games and academic and professional research in its wake. Much of that data can't be recovered without the exact computing environment that it was created on–even if the tapes or diskettes it has been stored on still function properly.

For the average computer user, it may not be much of a problem. But for those early adopters among us who typed endless letters, personal journals and term papers on **Apple Computer** (nasdaq: AAPL - news - people) IIs, Commodores and other systems that predate more modern computing environments, it's a daunting problem when we finally realize we need to recover the data.

For corporate America, it's just part of a bigger problem concerning the management of huge troves of irreplaceable digital data that may have cost millions to accumulate and can cost billions' worth of lost business if it's irretrievably lost.

Digital data, both personal and corporate, are becoming increasingly precious, as nearly every kind of information–from home movies to customer research data–is stored somewhere on a hard drive. The World Trade Center disaster was a stark example of how vulnerable digitalized data are to loss and destruction. Some companies had redundant copies of their data off-site, and companies like **SunGard Data Systems** (nyse: SDS - news - people) and **IBM** (nyse: IBM - news - people) can help other companies recover from disasters. Storage companies like **EMC** (nyse: EMC - news - people) lately have taken to pushing their storage technologies as an important part of disaster recovery plans.

To Bergeron, it goes deeper than that. Last year he published a book called Dark Ages II: When *Digital Data Die* (Prentice Hall, 2002), in which he argues that society's current preference for preserving information in digital formats exposes that knowledge to the inherent weaknesses of technology. Computers fail. Power fails. Storage disks can be damaged or destroyed by human error, fire, floods or war.

"It's almost like we're writing on tissue paper," Bergeron says. "Valuable data can be lost before anyone even notices it's gone."

The good news is that for the most part, technology is on your side. It just takes some effort to keep important data alive so they can be accessed in the future.

Jim Porter, president of **Disk/Trend**, a Mountain View, Calif.-based research firm that focuses on the data storage industry, points out that in most cases, data stored on a PC hard drive or server are relatively safe. Assuming proper maintenance and a good operating environment, computer servers on average are designed to run for 1 million hours–or 114 years–before suffering a hardware failure, he says. The hard drives in personal computers are designed to work for between 200,000 and 300,000 hours, or between 20 years and 30 years.

Of course, it doesn't always work out that way in reality. Bergeron's book details the experience of a technician at a hospital who took a job running an antiquated system that, among other things, was the sole backup for several years' worth of research data. One ill-advised tug on a six-disk storage system trashed years of research data that had been irregularly backed up to a defective tape drive. Three months and $40,000 later, only a quarter of the data was salvaged–and then only by a data recovery firm.

Similar stories are frighteningly common. Since publishing his book, Bergeron has learned there is only one existing set of CDs containing the computer blueprints for the government's abandoned superconducting supercollider project that was intended to perform massive particle-physics experiments. Most copies were poorly stored. Future historians will have a hard time doing research on the project if the original isn't available.

In another case, data collected from the Viking spacecraft that landed on Mars in 1976 were lost after tapes containing the information were found to be degraded. When rediscovered later, reams of data collected at a substantial cost to taxpayers were gone.

Then there's the case of floppy disks in all their various forms. The ubiquitous 3.5-inch floppy disk is only now starting to be phased out on current generations of PCs. But there's still a lot of data stored on them. Before that, there were the thin 5.25-inch floppies that were the standard in the mid-1970s. Prior to that, there was the 8-inch floppy disk, which entered use in 1973 and was only finally decommissioned in 1999.

Porter says those disks should generally last well, given proper storage. "The key is to store them in any place where a human being is able to live comfortably," he says. Provided that compatible hardware that still works exists–which is often a big "if"–old storage media can be surprisingly robust. But if there ever was a time to transfer data from these formats, which are slowly becoming extinct, the time is now.

Bergeron also warns that "rust never sleeps" and has devoted an entire chapter to it in his book. Magnetic storage media like floppies are based on a thin film of metal oxide deposited on a plastic or metal disk. The same thing goes for magnetic tape. "It's all just rust," he says. "Rust can flake off."

And the now-popular CD-R discs available from your neighborhood office-supply store are no more permanent. Bergeron says the information stored on those discs is based on a dye that doesn't last as long as the information stored on CD-ROM disks sold commercially. If you rely on CD-R to preserve important data, he recommends re-copying the contents of a disc to a new CD-R disc every two years and discarding the old one.

He also recommends saving written documents, like those created in Microsoft Word, twice: once in its original format and once in plain old ASCII text format. One way or the other, the files should be accessible, assuming the media still works.

For the really important stuff, go analog. Companies such as **Epson** and **Tektronix** (nyse: TEK - news - people) that make printers have recently started selling paper and ink intended to help printed text and images last longer. But pick the important stuff and store it somewhere safe. After all his efforts to back up personal data on various formats that can be read by a computer, Bergeron falls back to paper as the last resort.

"I have a filing cabinet," he says. "You can always make a print."

More From Forbes

Redundancy Saves 10.15.01
Good intentions fade as life gets back to normal. Make a personal technology disaster plan–now.

Shattered, Not Broken 10.15.01
Fiduciary Trust's backup plan worked.

Coping With The Collateral Damage 09.26.01
Disaster plans were put to the test Sept. 11. But getting back to business will take more effort.

Mastering Business Disaster 09.20.01
There's more that goes into planning for the worst than backing up computer data.

Disaster Of The Day: Cable Cut On The Net 11.21.00
You never know how much you depend on an undersea data transmission cable–until it's gone.