From floppy disks to deep freeze: what's the best way to store data?

28 JUN 2017 1

Security threats

Previous: News in brief: Wimbledo... Next: New Petya ransomware: eve...

by Bill Camarda 0

A key aspect of security is maintaining reliable access to the data you *thought* you owned. That's our excuse for sharing NPR's sweet story about the XFR Collective, a New York-based team of volunteer archivists and preservationists working to transfer old VHS videotapes into digital formats. It's also our excuse for sharing a few long-term solutions to data preservation that might someday solve the problem once and for all – because data rot is a problem virtually all of us have – or will have).

XFR's volunteers meet weekly in a Tribeca loft filled with "racks of tape decks, oscilloscopes, vector scopes and wave-form monitors" to painstakingly digitize cassettes from the 1980s and 1990s. (As they note, transferring video isn't plug-and-go; much tweaking and troubleshooting can be required to get it right. That's why they've only managed to transfer 155 tapes so far – a subnanoscopic percentage of the billions of tapes individuals recorded back then.)

XFR "partners with artists, activists, individuals, and groups to lower the barriers to preserving at-risk audiovisual media – especially unseen, unheard, or marginalized works."

According to NPR, they often digitize videos from:

> *...people of color, queer people, immigrants, artists and activists... old videos of police brutality... weddings or old public access TV.*

Whatever the content, once it's digitized, it becomes publicly available via the Internet Archive. (See some of what they've saved here.)

What about your tapes? If XFR Collective isn't an option, plenty of paid services still digitize old videotapes – or you can do it yourself using directions from CNET or Wikihow.

But you'd better get on it. Per Moving Image Preservation of Puget Sound:

> *Tape manufacturers predicted 20 to 30 years of life expectancy, but media lifespan depends greatly on environmental conditions... Format obsolescence contributes to the crisis... Umatic and VHS tapes are no longer manufactured and BetaSP will soon be discontinued.*

> *Machines to play these formats... are becoming more scarce as are the skills to maintain and repair them.*

MIPoPS shares a page of resources and best practices for digitizing important old analog video while you still can. It's ideal for archivists, "heritage organizations," and anyone with lots of content to protect.

Of course, it's not only videotape that's at risk. Entropy is relentless, and anything recorded on magnetic or optical media will eventually suffer the fate of Ozymandias. Even if the medium remains intact, formats and interfaces become obsolete and disappear. Preserving data for the long term is a discipline worth more attention than we can give it here, but a few tips might be helpful:

1. Keep track of how long media is likely to last – but remember that the statistics are controversial projections, and many folks won't be so lucky. According to CNET:

   > *The general consensus is that CD-Rs should last 30 to 50 years, DVD-Rs less than that, and CD-RWs and DVD-RWs even less. Similarly, tapes and hard disks can be expected to be readable for 10 to 30 years, while portable disks, USB thumb drives, and other solid-state storage devices may survive for half that time, maybe.*

   Back in 2005, The New York Times reported that 3.5" floppies have "an estimated life span of 10 years if stored in a cool, dry place with average care and use". If you've still got any, we'll bet they're older than that!

2. With this in mind, regularly copy data to new media, especially if it's approaching its expiration date. (And make sure anything

you haven't saved is "in a cool, dry place," not your attic or garage!) If you're *really* serious, PC World suggests considering "write-once BD-R HTL (High To Low) [which] can last for 100 to 150 years given a relatively mild environment" or "Milleniatta's M-Disc BD-R and DVD+R write-once discs [which] use an even more stable data layer that is rated for 10,000 years" based on French and US government testing.

3. Move away from physical formats that are becoming obsolete. For example, many people who used to back up their data on Zip drives, Syquest cartridges and 1.44MB floppy drives no longer have access to these. Even interfaces can be an issue: external devices often used serial or parallel ports that no longer ship standard on computers (though desktop PC and ExpressCard laptop adapters can still be found). Make sure you've migrated your data before you dispose of an old device or format.
4. A common related issue: data trapped on a working hard disk in a dead PC or laptop. The Guardian serves up some useful guidance on installing the drive in an external USB enclosure, and restoring from there.
5. Migrate data from obsolete programs, or at least make sure you have the tools to do so when necessary. Millions of people still have content trapped in ancient word processing formats such as PFS: Write or Multimate. Tools for viewing such data or move it into "living" software include Quick View Plus and FastLook; for some formats, the free LibreOffice productivity suite or XNView image viewer might be all you need.
6. TechRepublic offers some useful high-level advice on planning a long-term strategy for protecting your data here.

The future: some radically new solutions

All this is great as far as it goes, but as the amount of data we're generating continues to soar, we're likely to need some radically

new solutions — especially if we want our data to last longer than some stray DVD-R. Here are three of our favorites:

Analog micro-etching: The Long Now Foundation – which specializes in trying to envision the long-term future and solve the problems it might present – ran a full conference on super-long-term data storage. The solution it found promising enough to test: analog micro-etching onto nickel disks. Eight years later, they had a prototype: a disk containing information in about 1,500 human languages, plus translations of the Book of Genesis in each. Since the information is analog, it's readable directly by humans (though they *will* need a microscope).

The Arctic World Archive: Officially opened on March 27 in Norway's Svalbard Arctic region, the for-profit Arctic World Archive is already housing key documents from Brazil, Mexico, and Norway — safe, theoretically, from natural disaster and warfare. According to a report in The Verge, data is actually imprinted on special film, in huge high-density greyscale QR codes – and the archive is completely disconnected from the Internet to protect against hackers and ransomware. (Unfortunately, however, it's located near the Global Seed Vault, which is already getting hammered by global warming.)

Best of all: *DNA*. According to Science Magazine, researchers have been making breathtaking progress since the first attempts to store data in DNA molecules back in 2012. DNA is "ultracompact, and it can last hundreds of thousands of years if kept in a cool, dry place. And as long as human societies are reading and writing DNA, they will be able to decode it – not something you can say with confidence about videocassettes or QR codes. Plus, new technologies are making its data storage capacity almost infinitely scalable. Columbia University's Yaniv Erlich and New York Genome Center's Dana Zielinski have partnered on an approach they say can hold 215 petabytes (215 million gigabytes) in a single gram of DNA.

As Science reports, speed and cost are still big issues, but “the system could, in principle, store every bit of datum ever recorded by humans in a container about the size and weight of a couple of pickup trucks”. Now all we have to do is figure out where to park them.

floppy disk NPR VHS XFR Collective