Family History Archives: More Research on Permanent Media

Dr Barry Lunt
Professor, Cybersecurity
Electrical and Computer Engineering
Brigham Young University

Abstract

Paper and ink-based records, followed by photographs printed on photographic paper, have provided personal and institutional historians access to original documents dating back hundred and even thousands of years. However, the age of digital documents has raised the alarming specter of a "digital dark age" – an era characterized by a complete absence of original sources due to the lack of permanence of digital data.

Research at BYU has culminated in the first truly permanent digital data storage option, which at least has addressed the persistence issue. Needed are additional permanent media, and this need is being met by research on a permanent solid-state storage medium and a permanent ½-inch tape storage medium, both at BYU. This paper provides an update on progress on this research.

Background

The persistence of records is the *sine qua non* of personal and institutional history – if ink and paper only persisted for a decade, there would be no personal or institutional histories. Written records would lose nearly all their value, and librarians and historians would be hard pressed to provide much of any significant value to civilization – at least at an affordable price.

Thus it is very reasonable that some historians have pointed out this alarming situation when it comes to digital records, referring to our era of computer records as a "digital dark age"[1,2], or "digital doomsday"[3], and efforts to preserve what we have seem to be sounding alarms[4,5], using terms like "time bomb" and "fading memory".

Some historians and archivists have termed data persistence to be only half a solution, and it is acknowledged that without the appropriate readback hardware and software, reading old digital records is difficult, and can appear to be impossible. But if the data itself has not persisted, all further discussions about how to read this data are even worse than merely academic – they are completely superfluous. But more on this later.

The state of storage persistence was outlined at the time in a paper by this author at Archiving 2011[6], with the summary data shown in Table 1. With the exception of the Millenniata recordable optical discs, the numbers are very alarming, and deserving of the terms being used to describe it. The only solution presently being used is migration – periodic monitoring of the status of the data, then migrating the at-risk data to newer media. This labor- and cost-intensive effort can only be justified by a few institutions, yet the value of the data being preserved is inestimable.

Media	Life Expectancy of Data
Magnetic tape	10-50 years
Magnetic hard-disk drives	1-7 years
Flash drives and Solid-state drives	10-12 years
Recordable optical discs	1-25 years
Millenniata recordable optical discs	1,000 years (advertised)

Table 1: Life expectancy for data stored on today's media.

Round 1: A Permanent Optical Disc

Stamped optical discs (CDs, DVDs, and Blu-ray discs or BRDs) are projected to have extraordinarily long lifetimes – up to 1,592 years according to one report[7]. This indicates that the problem of data persistence is not one of a lack of materials nor manufacturing processes that can produce digital data with a long lifetime, but rather one of engineering. A *recordable* optical disc was needed which could have a similar lifetime.

In any attempt to improve something, the first step is to identify the causes of the problem. This has been well reported in previous research; in reference #8, the authors identify three of the major failure mechanisms[8]. The fourth failure mechanism is dye degradation, and is widely known in the recordable optical disc industry; it was reported on in reference #9 and elsewhere.

Accordingly, knowing these factors as the primary failure mechanisms, a completely new type of recordable optical disc was researched and developed. It uses no reflective layer, and has no organic dyes at all. The recordable portion of the media is an optical stack as shown in Figure 1. These materials are inorganic, and as such, they do not degrade with time. Marks are made in the recording layer as shown in Figure 1 and Figure 2, where it is seen that the Data Recording Layer 2 has been removed in specific places (where the image dark spots are), thus recording the bits of data as though etched in stone.

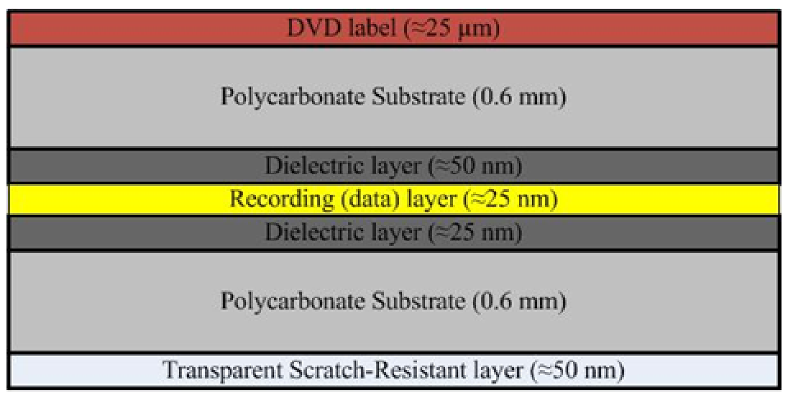

Figure 1: Optical stack of the Millenniata disc, showing all of the layers.

This disc has been arduously tested against the very best “archival-quality” and standard-quality discs available, by an independent test agency, and the results have been published[10]. Figure 3 is a graph from that published report, showing that after 24 hours of accelerated aging testing (at elevated temperature, humidity and full-spectrum light) all the other discs tested were either unreadable

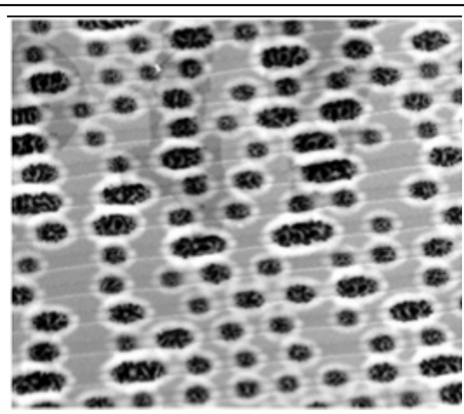

Figure 2: Scanning electron micrograph (SEM) of the marks on the Millenniata disc, showing they are both physical and optical.

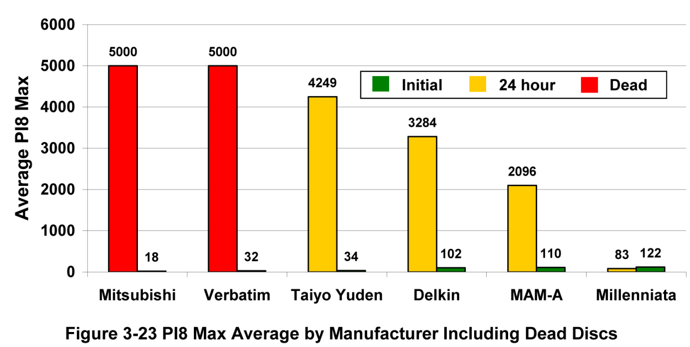

Figure 3: Test results from the Naval Air Warfare Center Weapons Division (China Lake, CA).

("Dead") or had thousands more errors than the spec allows (PI8 Max should not exceed 280), while the Millenniata disc was completely unaffected by this test, and remained well below the spec of 280.

Round 2: Is There Any Hope for the HDD?

The magnetic hard-disk drive (HDD) is still a major part of storage for today's digital data. Unfortunately, loss of data from HDDs is legendary, and many people have personal stories to tell regarding this risk (including this author).

As with optical discs, the first step to developing a permanent digital storage medium is to understand the causes of failure of existing media. In the case of HDDs, the causes are quite well known, and are all essentially mechanical. With head flying heights in the realm of 10 nm and track densities greater than 18,000 tracks/inch, there is no room at all for mechanical variation. Unfortunately, mechanical wear is inevitable in anything that is constantly spinning, and HDDs are relatively limited in their lifetimes. Presently, our research group has not conceived of a way to make HDDs dramatically better for archival purposes. And due to the fact that HDDs are mechanical, something permanent would have to be non-mechanical, and a non-mechanical HDD seems to be an oxymoron. Future developments on holographic storage have not proven successful, so that isn't an option either.

Round 3: A Permanent Solid-State Storage Medium

Solid-state data storage has a long history, starting with the static random-access memory (SRAM) and dynamic random-access memory (DRAM) of the late 1960s. Unfortunately, these memory types are extremely non-permanent because they are volatile (the data disappears or "evaporates" when power is lost). Also from the 1960s, ROM (read-only memory), PROM (programmable read-only memory), and EPROM (erasable programmable read-only memory) all found wide market acceptance, because these memory types were non-volatile. The most recent development in non-volatile memory was over 20 years ago, when "flash" memory appeared on the market. It has come to dominate this market, in formats of "jump drives", or "USB memory sticks" (or other names).

Unfortunately, this storage medium is inherently non-archival, as it stores the data as the charge on a floating gate, as shown in Figure 4. This floating gate is surrounded by oxide, which is an excellent insulative material, and thus the charge remains on the gate for a prolonged period. However, even excellent oxides are not perfect, and the charge will eventually deplete. It is well known in this industry that the storage time for the data is from 8 to 12 years, which is long enough for most computer needs, but clearly not long enough for archival purposes.

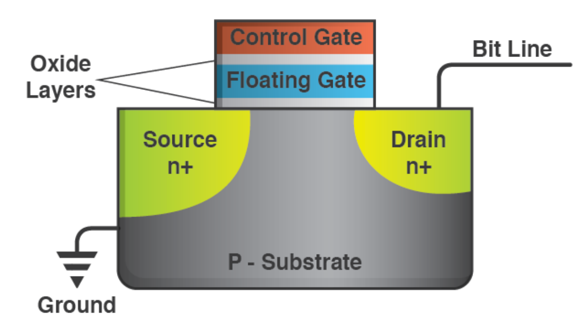

Figure 4: The basic storage cell of "flash" memory.

PROMs are quite permanent, as the data on them can be retained almost indefinitely, but only if a solution can be found to the problem of dendrites (see Figure 5). Even today, the factors which cause dendrites to grow are not fully understood, but when they grow, they compromise the data that has been stored.

The basic storage cell of PROM is a fuse. An unblown fuse is usually regarded as a "1", while a blown fuse is generally regarded as a "0". In order to make PROMs into a permanent form of data storage, a device must be invented which will completely prevent dendrites. If such a device can be designed, the data stored in PROM fuses will be permanent, with projected lifetimes well beyond 100 years, possibly greater than 500 years.

Figure 5: A programmed PROM storage cell, showing the remains of the blown fuse and a dendritic growth emerging from those remains.[11]

Figure 6 is an SEM of a PROM structure designed and tested in our research several years ago. In this case, after the fuse is blown, it appears there is no material left at all, thus making it very difficult for a dendrite to grow.

We have researched multiple materials, and there is one that has shown itself to be ideal. This material is amorphous carbon, which is inexpensive, non-toxic, does not form dendrites, is extremely stable, and is compatible with today's processes for manufacturing integrated circuits. Research continues into the testing of this option, while a commercial version of the product is also under development and is expected in 1-2 years. This research will result in the development of a PROM that is dendrite-free, and thus more permanent than anything ever made in solid-state storage. The presently-estimated lifetime of the data for this solid-state storage device is over 1,000 years.

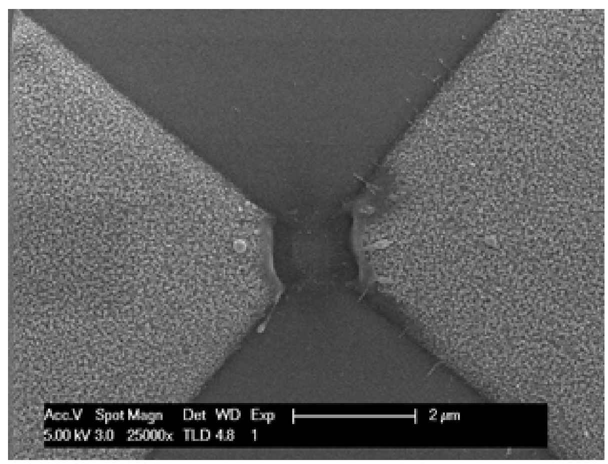

Figure 6: Blown fuse of material "A", with no material remaining.

The potential density for this technology is equal to that of today's flash memory, meaning that after initial production problems are solved, this technology would allow memory sticks with 32 Gigabytes of storage, and more as the integrated circuit industry continues to improve along Moore's Law.

Round 4: A Permanent Tape Storage Medium

Magnetic tape has been around for over 50 years, and has shown itself to be the medium of choice for backups and long-term storage. However, as discussed in the Background section at the start of this paper, it still has a rather limited life when considered for archival purposes.

Our research is presently focused on using the materials researched and patented for the Millenniata disc, since we already know their durability and writeability. The present plan is to leverage this material, plus the tape drives and cartridges of the ½-inch tape industry, plus the recording and readback technology of the optical disc industry, to produce an optical tape that is as permanent as the Millenniata disc. Based on present calculations, the density of this storage would be equivalent to today's ½-inch magnetic tape, meaning that a cartridge of this tape should be capable of storing multiple TB of data. It is estimated that this product could be made into a commercial product within four years.

Round 5: Format and Hardware Obsolescence

There is no question that, once the data persistence problem has been solved (and we are confident that it WILL be solved), the next main issue is assuring a way to read the data far into the future. Many are the examples of obsolete hardware or obsolete formats that make it very difficult today to read yesterday's floppy discs, reel-to-reel magnetic tape, and other storage technologies of the past. Probably the best answer to this issue lies in history.

If you were an author in western Europe in the early middle ages, and you wished to write a book which would endure, the first thing you'd have to do is choose a medium which would persist – both the paper and the ink. As discussed before, data persistence is the *sine qua non* of data archiving.

After choosing a medium which would persist, your next challenge is to choose a language that is unlikely to become a forgotten language far in the future. This would the most likely way to assure that your book will be readable in its original form long after you have left this life. Given the situation in western Europe in the early middle ages, your choice of language would probably have been Latin, since it was spoken over most of the Roman Empire, and since there already existed many thousands of documents in that same language.

Why is it that today, when no country nor ethnic group on the entire planet claims Latin as their spoken or written language, academia continues to teach this language? It is simply because there still exists today a very large body of original texts which have persisted and which historians wish to be able to read.

Arguing historically, the same is bound to be true of digital data storage formats. Yes, there are many obsolete formats around today, but none of them were ever very widely adopted, from a world-wide perspective. That is not true, however, of optical discs, which in their hayday were the most widely-adopted digital data storage format in the history of the world. There are hundreds of billions of discs, all over the planet; and there are billions of disc drives, also all over the planet. If the data on these discs does not persist, format and hardware obsolescence are moot and irrelevant concerns. However, if this data does persist, and technology continues to improve as the decades move on, it is essentially guaranteed that historians will preserve a way to read these, and we have many historical examples that support this argument.

Even though the wax cylinder, the vinyl disc, and the 8-mm video tape have all become obsolete formats, there are libraries and companies around the world that preserve devices capable of playing these antiquated storage media, and of converting them to more modern formats. Because the data (music, videos) on these media have NOT faded into non-existence, we have preserved ways to play them back – and this is highly likely to continue into the future.

And the same is true of USB sticks – there are hundreds of billions of them, and if the data on them persists, historians will preserve a way for them to be read. History argues this point very powerfully as we look at the holdings of the great libraries of the world.

Conclusion

One solution to permanent digital data storage has been created and is widely available. Other solutions are being researched, and it is only a matter of time before those solutions will also become available. As these solutions become more widely adopted, the problem of data persistence will be more widely solved.

And because the solutions are based on very widely-adopted formats, history would argue that the ability to read these solutions far in the future will be preserved by those who preserve such abilities today – academia, libraries, and historians.

References

1. Kurt D. Bollacker, "Avoiding a Digital Dark Age", *American Scientist*, vol 98 #3 (May-June 2010), p. 106-114.
2. Computer History Museum, "Digital Dark Age – Revolution Preview", (video at: http://test.xtremedump.com/content/digital-dark-age-revolution-preview).
3. Tom Simonite, Michael Le Page, "Digital doomsday: the end of knowledge", *New Scientist*, #2745 (Feb 2, 2010).
4. Margaret Hedstrom, "Digital preservation: a time bomb for Digital Libraries", http://www.uky.edu/~kiernan/DL/hedstrom.html.
5. David Talbot, "The Fading Memory of the State", *Technology Review*, July 2005, pp 44-49.
6. Barry M. Lunt, "How Long Is Long-Term Data Storage?", *Proceedings of Archiving 2011*, Salt Lake City, Utah, USA, May 2011.
7. Navale, Vivek, "Predicting the Life Expectancy of Modern Tape and Optical Media", *RLG DigiNews*, Aug 15, 2005, 9:4; also at www.rlg.org/en/page.php?Page_ID=20744#article3
8. Shahani, Chandru J., Basil Manns, Michele Youket, "Longevity of CD Media: Research at the Library of Congress", Preservation Research and Testing Division, Washington, DC.
9. Slattery, Oliver., Richard Lu, Jian Zheng, Fred Byers, Xiao Tang, "Stability Comparison of Recordable Optical Discs- A study of error rates in harsh conditions," *Journal of Research of the National Institute of Standards and Technology*, 109, 517-524, 2004 .
10. Svrcek, Ivan, "Accelerated Life Cycle Comparison of Millenniata Archival DVD", Life Cycle and Environmental Engineering Branch, Naval Air Warfare Center Weapons Division, China Lake, CA; Nov 10, 2009.
11. Pecht, Michael, "Handbook of Electronic Package Design", p. 597.