MEDIA PRESERVATION INITIATIVE SURVEY: THE WHITNEY'S WORKS ON FLASH DRIVES

Introduction

When the Media Preservation Initiative first launched, initial surveys of the Whitney's media collection focused on works whose media carriers are considered the most at-risk formats. Despite being some of the newest technologies in the collection, flash memory drives were quickly identified as a high priority. Though the collection contains many older, more obsolete media formats such as analog videotape, flash drives, despite their current ubiquity and physical durability, can exhibit unpredictable internal behaviors that make them unsuitable for long-term storage. As the files on these devices had not previously been backed up or copied, an early goal of MPI was to ensure that these artworks were safely transferred onto the Whitney's servers.

In November and December of 2018, Savannah Campbell, MPI's Video and Digital Media Preservation Specialist, began an assessment of the USB flash drives in the collection. The primary aim was to assess the functionality and condition of the carriers and transfer the works of art housed on them onto the Museum's server and subsequently have them ingested through the newly developed Archivematica digital preservation pipeline.

About flash

Flash memory is utilized in a wide array of storage devices, including USB flash drives (also known as thumb drives), both external and internal solid-state drives (SSDs), and memory cards. Flash carriers store data through electronic charges. In binary terms, the computer reads an electronic charge as "0" and no charge as "1". These carriers are non-volatile, meaning they are able to store electronic charges without the need for an external power source.

Flash drives are typically used for the short-term storage and transfer of data. Their small size facilitates portability, making them convenient to use for transporting and sharing data from one machine to another. As flash memory is entirely electronic, these carriers do not contain any moving parts, hence the terminology "solid-state". While the lack of moving components makes flash devices more physically durable than spinning disk hard drives (HDDs), there are many disadvantages of flash carriers that make them treacherous for long-term archival storage.

Risks of using flash drives for long-term storage

Even though flash devices use electronic charges to store data, they do not require any additional power (they don't need to be plugged into an electrical socket to charge/function, unlike many HDDs). This means that data will stay on the device, but also that the data has the potential to leak over time, since the electrical charge will inevitably and randomly dissipate (in the same manner that a laptop that was powered off and left unplugged for a month will not power on again until it has been recharged). When memory leaks from a flash device, a form of data corruption known as bit flipping occurs (the charged 0s will turn into uncharged 1s), potentially rendering files unreadable.

An additional archival concern with flash is its potential for read disturb errors. Flash data is both read and written via electronic pulse. Read disturb can occur when a user attempts to read a file and the electronic pulse inadvertently charges nearby cells, causing them to flip from 1s to 0s, thus bit flipping in the opposite direction that a memory leak would cause. Again, these flipped bits can cause enormous problems, especially for art conservation. The data on the drive can become corrupt and completely unreadable, and at the very least the data can be unintentionally altered in ways that would make the digital object different from what the Museum received from the artist. The use of a write-blocker cannot prevent these kinds of changes from being made to flash-based carriers. In terms of retaining the authenticity of artist-provided files, HDDs are much more reliable than flash drives.

Flash devices also run the risk of failing outright and spontaneously becoming unreadable. Unlike with HDDs where the data has physically been written to the magnetic disk, with solid-state carriers such as flash drives, the electronic charges can suddenly vanish without a trace. While it can be possible to recover data from a failed HDD using digital forensics techniques, it is nigh impossible to do so with an SSD.

For all of these reasons, USB flash drives and other solid-state carriers are not recommended for long-term archival use.

Flash drives at the Whitney: the numbers

As part of this survey, 40 flash carriers (38 USB Flash Drives, 1 SD Flash Card, and 1 MicroSD Flash Card) containing files for time-based media artworks were examined in the Museum's

Media Lab. Some drives contained multiple works, and in other cases more than one drive contained files for the same work. In total, the drives collectively contained files for 58 works of art from 29 different artists. All of the assessed components were received by the Museum between 2005 and 2017, with the majority of the flash drives dating from 2012–2015.

The assessment workflow

The condition assessments of the 40 flash carriers were carried out in the Museum's Media Lab on an iMac Pro workstation. All of the flash drives were assessed using a Tableau Forensic USB 3.0 Bridge write-blocker, to prevent any changes from being written to their files. A write-blocker is a device that allows the user to open and view files on a drive, but prevents any changes from being written to it. This is an archival best practice for working with digital media carriers as it ensures the files received by the institution are not accidentally modified.

If the flash drive was successfully mounted to the iMac Pro via the write-blocker, the next step in the condition assessment was to open the drive in Finder, take a screenshot of the drive's contents, and save the screenshot as documentation. Any other non-media files on the drive, such as artist-provided installation instructions and still images of the work, were also saved to the Museum's server.

Savannah then performed quality control checks on all media files stored on the drives. This entailed viewing all video files from beginning to end, checking for any errors or corruption. To capture technical metadata, MediaInfo reports were made for each video file. Then, both MD5 and SHA-256 checksums were created from the files on the flash drives.

After successful QC and checksum creation, all media files were transferred to the Museum's server, and the MD5 and SHA-256 checksums were verified to make sure they matched the checksums generated from the original flash drives. Verifying the checksums in this way ensures that the digital file on the Museum's server is a bit-for-bit copy of the digital file received from the artist or gallery.

Summary of survey findings

The results of the assessment were as follows:

— 34 flash drives mounted and the files on them were successfully moved onto the art storage server
— 1 flash drive containing executable software files mounted successfully, files were temporarily moved to an external hard drive
— 5 of the flash drives failed outright, either because the drives themselves failed to mount or the files on them were corrupt and unreadable

Next steps

The results of this assessment yielded the following three courses of action for MPI project staff to take next for the artworks examined:

— Contact artists, galleries, and foundations about works where new files need to be requested (due to corrupt flash drives) and works where further research is needed (for example, in cases where files were not labeled and it was it clear which formats were intended for exhibition and which were the artist's native file). This could include asking artists to fill out an artist questionnaire.
— One of the flash drives contained executable files for a software-based work and it sparked a discussion on storage concerns for works on this nature. Internally, MPI project staff and the Whitney's IT department should discuss how to handle software-based art and executable files. Should these files live on the same server as all the other video, image, audio, and text files? Or should they be stored in their own partition?
— MPI will begin moving the files assessed in this survey through the new digital preservation workflow. The ultimate goal is that the files will be archivally stored on Archivematica and available for Museum staff to access through ResourceSpace.

Resulting changes to Museum policy

Given the findings of this survey, two new policies were put in place at the Whitney for acquiring new artworks on digital media carriers.

— The Museum no longer accepts artworks submitted on USB flash drives or flash memory cards. Only files that are delivered on external spinning disk hard drives (HDDs) can be acquired. While spinning disk hard drives can also fail, they are more secure for long-term storage, easier to recover data from, and more reliable to disk image and run checksums on.
— During the acquisition process, the Museum requests that artists specify the native file format of the work as well as their preferred exhibition format. Ideally, the artist or gallery will provide the Museum with both a native file and an exhibition file for each work, as well as documentation about their production process.

Report prepared in December 2018 by Savannah Campbell, Project Video and Digital Media Preservation Specialist.

Sources Consulted

"Inside a Flash Drive," How USB Flash Drive Works? https://howflashdriveworks.wordpress.com/what-is-a-flash-drive/

Jacobi, Jon L. "Death and the unplugged SSD: How much you really need to worry about data retention," PCWorld. May 13, 2015. https://www.pcworld.com/article/2921590/death-and-the-unplugged-ssd-how-much-you-really-need-to-worry-about-ssd-reliability.html

Neary, David. "A Solid State of Affairs: Memory revolution or technological Flash in the Pan?" New York University. 2014. https://www.nyu.edu/tisch/preservation/program/student_work/2014fall/14F_1807_Neary_A3.pdf

Newman, Jared. "Leaving unpowered SSDs in a warm room can kill your data fast," PCWorld. May 11, 2015. https://www.pcworld.com/article/2920727/leaving-unpowered-ssds-in-a-warm-room-can-kill-your-data-fast.html

Rosenthal, David. "The Medium-Term Prospects for Long-Term Storage," DSHR Blog. October 13, 2016. https://blog.dshr.org/2016/12/the-medium-term-prospects-for-long-term.html

Sheward, Mike. "Rock Solid: Will Digital Forensics Crack SSDs?" Infosec Institute. January 5, 2012. https://resources.infosecinstitute.com/ssd-forensics/

"USB Flash Drive Versus SSD: Revealing the Differences," Premium USB. July 12, 2011. https://www.premiumusb.com/blog/usb-flash-drive-versus-ssd-revealing-the-differences