Title: Comparing methods of generating 3.5 inch floppy disk forensic images

Author: Nicholas Schiller

Author Affiliation: WSU Vancouver, The Electronic Literature Lab

Abstract: I propose comparing methods of generating forensic images of 3.5 inch floppy disks in order to evaluate methodologies for use in media archeology labs. Many key works of electronic literature (including the bulk of Eastgate Systems, Inc. early publications) were released on 3.5 inch floppy media. I will use both Kyroflux and Superdrive floppy disk controller units to generate forensic images and also generate images using BitCurator suite of forensic software and using legacy computing hardware and software.

Gathering data on the quality of images created by these disparate methods and also on the workflows involved and the ease and practicality of employing them will produce useful information for other media archeology labs examining how the field of floppy disk forensics has advanced.

The results of these tests should show useful comparison data between the quality of the images created from identical media, the range of image types that can be created using each technique, and the usefulness for online access and emulation each forensic methodology and platform provides.

Submission Formats and Guidelines

Full papers: 2,000-4,000 words in PDF format A4, suggested format for citations is MLA. Full papers are due by May 7th and will be published open access on the conference site and the ELMCIP database. Full paper presenters will also participate in a live roundtable discussion of their papers (5-minute presentation plus discussion). Accepted abstracts will be grouped into peer panels, and peers will be asked to give feedback on each other’s papers.

Comparing methods of generating 3.5 inch floppy disk forensic images

Nicholas Schiller

Washington State University Vancouver

The Electronic Literature Lab, Associate Director

schiller@wsu.edu

Introduction

The 3.5 inch (88.9mm) floppy disk was a common storage format on personal computers sold in the final two decades of the twentieth century. Floppy disks store data on a round piece of magnetic material affixed with adhesive to backing material and also to a central metal spindle. Each of these parts presents a point of failure. Even disks that have been carefully stored in cool, dry locations are subject to degradation and data loss (bitrot) over time. The magnetic film can come free from its backing, the spindle can detach from the disk, and the adhesive can fail. Experts put the expected lifespan of a floppy disk's data integrity most likely between five to twenty years.(Levy; Rothenberg) Fortunately, floppy media were largely replaced by optical media (such as CD-ROM and DVD-ROM) or flash based storage that are expected to maintain their data integrity for much longer (20-200 years for a CD-ROM(Byers). This should provide a small respite to allow us to save works of electronic literature that currently reside only on floppy disk before the next preservation crisis

Eastgate Systems, Inc, an early commercial publisher of electronic literature, distributed at least 24 works of elit on floppy disk between 1990 and 2001. Dene Grigar has a collection of Eastgate's hypertext and electronic literature housed at The Electronic Literature Lab (ELL) in Vancouver, Washington. As the Associate Director of ELL, part of my responsibilities entail seeing to the preservation and continued access to the works in this collection. This project will take on the task of digitizing works currently preserved on floppy disk and comparing alternative methods of migrating digital content off of precarious media an Into sustainable long-term storage solutions.

Using the ELL's collection of works distributed by Eastgate Systems, Inc on floppy disk, this project compares techniques for migrating digital content off of floppy disks and on to more sustainable long-term storage. Currently there are multiple solutions available to archivists and curators that can achieve this goal. Content can be captured in the form of the digital files saved on the disk, an image recorded of the disk as a whole, or as magnetic flux recordings (raw stream files) that are generated by specialty tools designed for floppy disk preservation.

The starting point for the comparison will be a hardware and software preservation solution that is well accepted in the field, the KryoFlux (*KryoFlux Products & Services Ltd.*) I will examine the features that make the Kryoflux a current best practice while also examining the known issues, limitations, and barriers to use that the Kryoflux device brings with it. I will also explore three possible alternatives to KryoFlux device in the hope that one or more of them will offer the benefits of the KryoFlux while allowing us to bypass the limitations.

The four floppy preservation methods to be explored Includee, one, using network-connected vintage computers to read floppy disks using original disk drives and transferring content off of floppy disks and onto network storage. Two, using the Bitcurator (*Https*) suite of forensic tools to capture disk images of floppies using contemporary hardware. Three, using the KryoFlux magnetic flux recorder, currently one of the accepted best practices for capturing floppy desk images. Four, using the SuperCard Pro, a similar device to the Kryoflux that provides similar steam flow image recording with a different software licensing model. Small production runs have meant the Supercard Pro has not always available for purchase. Recent availability means that archives and libraries have a new option to consider for making magnetic flux recordings of their floppy disk collections.

Four Methods

Networked Vintage Computers or "Rosetta Machines"

The first method for migrating the contents of the floppy disk collection to long-term storage media involves using some of the sixty-one vintage computers in the Electronic Literature Lab's collection to read the floppy media and then transfer the Individual files saved on the disks over a network to a contemporary storage device, such as a USB flash drive or a portable hard drive. This method has the advantages of being built in to the infrastructure of the lab. Undergraduate research assistants can be trained to perform this migration without having to first be trained on delicate hardware or command line interfaces (CLIs). It has the restriction of not scaling. By 'scaling' I mean that it cannot be employed by anyone who does not have access to a large collection of working vintage computers and the technical support infrastructure to keep them functioning and networked together.

The Covid-19 pandemic interfered with using this method as a. practical test. ELL is a working lab with several research activities running concurrently at any given time and when the ELL researchers were sent home to shelter in place, several of the lab's vintage computers went home with them, enabling continued progress with their research. Unfortunately, this meant that once I was vaccinated and ready to return to work in-person with the collection in ELL, the computers I needed to read the floppy disks, specifically the 400k or 800k 3.5 inch Macintosh SuperDrive floppy disk drives, they were not in the lab connected to the network. So testing this method will be delayed until work can take place In the lab, safely.

Describing this kind of preservation method, Doug Reside has labeled computers like these "Rosetta Machines" (Kirschenbaum et al.) meaning, for example, that a Macintosh computer with a SuperDrive floppy disk drive capable of reading 400k, 800k, and 1.40mb floppy disks while also connecting to a network with a still-standard RJ-45 network jack can give a researcher access to content on a floppy disk written for Apple System Software 6 while also giving access to shared drives on a present-day network. This works In the same way as the Rosetta Stone gave researchers who can read Ancient Greek letters or Demotic script access to Egyptian hieroglyphs. The special stone that contained the same text in three iterations unlocked an unknown language. Because the ELL has these special machines, these Rosetta Machines, researchers can read the 800k floppy disk AND the RJ-45 physical network protocol, it allows the researcher with access to the network to transfer content directly from 1988 storage media to 2021 storage media.

The Rosetta Computer method can be attempted again in the future. We certainly can copy files from floppies onto the network, once the network and its nodes are reassembled in place. However, the need for this method to be used may have been rendered moot by the successful implementation of other methods to be mentioned later in this paper.

Bitcurator Suite, Guymager, and Contemporary USB Floppy Drive

Bitcurator is a suite of open-source digital forensics tools packaged into an Ubuntu Linux environment (*BitCurator*). Installing the most recent long-term distribution version of Bitcurator on a machine in ELL provided us with a number of powerful tools in a single package. The most relevant of these tools is the Guymager disk imaging tool. Using this on a surplus PC along with an inexpensive USB floppy drive purchased for under twenty-five American dollars from Amazon, ELL had a nearly-free method of capturing disk images.

In addition to the low-cost barrier to implementation, Bitcurator and Guymager offered two other distinct advantages over the other options in this project. One, it allowed for the automated inclusion of metadata with the disk images, greatly improving the potential for useful organization at a very early stage of the project. Two, it automates the inclusion of hashes for the newly created disk images, an important step in maintaining digital archives. These features are expanded on in The Archivists Guide to KryoFlux (Allen, Jennifer) which is absolutely a necessary and vital document for anyone doing floppy disk preservation to have read. Additionally, it has a software-based write-protection feature that can be very useful. For example, external drives can be set to read-only by default, an important safety step that protects valuable content from being unintentionally over-written.

Drawbacks to using Guymager include the disk image format produced by the software is not as convenient for our purposes as other methods. Images created using Guymager were saved in a .e01 format that requires a specific reader (included in Bitcurator) to access. Other methods use a .img format that is immediately recognizable as a disk image by Macintosh operating systems and emulation platforms. Creating disk images in a directly usable format removes a step from the workflow and makes the process more efficient. Another drawback is that the process of creating disk images was slower and less reliable than either of the magnetic-flux imaging solutions. After imaging about a dozen floppy disk, the drive I used stopped creating readable images. This is most likely caused by dirty read heads, but as my access to the lab is restricted, I was not able to stop the entire imaging process to troubleshoot this one aspect. After the USB floppy drive stopped reliably capturing .e01 images, I stopped using the Guymager and Bitcurator method.

So while the Bitcurator method did not end up being a working solution for this project in ELL, it is the lowest-cost solution, it does work, and it does provide some workflow advantages by simplifying metadata entry and hash creation.

KryoFlux Forensic Floppy Disk Controller

The experts who contributed to writing the documentation for The Archivists Guide to KryoFlux deserve special mention. Jennifer Allen, Elvia Arroyo-Ramirez, Kelly Bolding, Faith Charlton, Patricia Ciccone, Yvonne Eadon, Matthew Farrell, Allison Hughes, Victoria Maches, Shira Peltzman, Alice Prael, Scott Reed, and Dorothy Waugh, in writing documentation for a particular piece of hardware, have also set out workflow for digitizing floppy disk content for everyone who is approaching the project, regardless of the hardware used.

The KryoFlux provides a superior solution to Bitcurator and Rosetta Machines in the following ways. First, it allows for pure flux recordings to be saved. These are direct magnetic flux recording taken from the drive and do not require compatibility with a floppy disk's format (Purity) or File Allocation Table in order to be created. It may be useful here to take a step back and consider three different aspects of a floppy disk standard. One is the physical size of the disk. This is obvious and important because one cannot fit a 5 ¼ inch floppy into a 3.5 inch floppy drive. The second aspect is the density of the magnetic medium on the disk's read surface and, more importantly for our needs, the speed at which the drive spins. Until the early 1990s, Apple used a variable speed floppy drive, which allowed for greater flexibility and read accuracy than the single-speed 1.44 mb floppy drives used by IBM. After the 1990s, Apple adopted the IBM standard single-speed, making one 3.5 inch floppy standard, but also making disks encoded in their previous variable-speed standard orphaned and unsupported by future floppy drives. Archivists and curators need to be aware of these three elements of floppy disk compatibility, physical size and shape, speed and density, and file system. To employ another linguistic metaphor, we can think of the first as physical form a text is encoded on, say a stone tablet, a scroll, or a bound codex. The disk type (density, speed, and sector configuration) of a floppy drive is roughly equivalent to the alphabet or system of glyphs the text is written in. Finally, the file system of the disk is similar to the language and grammar used to write it down. In order to read the "text" we have to know whether to read right-to-left or left-to-right (top-bottom, etc.), we have to know the characters or glyphs used to encode meaning, and we have to know the language in which the meaning is encoded. Otherwise the meaning remains opaque to us.

The 'magic' of the KryoFlux device lies in how it records the magnetic flux read by the drive's read head, rather than the sequence of ones and zeros encoded by the drive's standard rules, this means its disk images are density, speed, and file system agnostic. We do not need a variable speed drive to capture images off of a disk that was manufactured and encoded according to a variable-speed standard. Using a KryoFlux board we can use a common 1.44mb IBM standard floppy drive and read (potentially) any 3.5 inch floppy disk. Whether or not that disk complies with the 1.44mb IBM standard.

A second key feature of the KryoFlux device is that its software allows us to write disk images from flux recordings. This means we can capture an image of a disk without knowing whether it is single or variable speed and then make disk images from the first flux (Allen, Jennifer) capture without having to run the original floppy through the drive again. Once we have flux recordings for the floppies in our collection, we can create disk images to use with emulators or even write new floppy disks to use in vintage computers without submitting the valuable originals to additional wear and tear.

The KryoFlux software runs on Linux and Windows computers and on OSX 10.11, which is more flexible than the other solutions listed so far. It also has a hardware write protect switch, which is more secure than the software switch offered by Bitcurator.

The potential barriers and issue with implementing Kryoflux as a floppy disk migration solution have little to do with the features and powerful options of the tool. The primary issue many may have with Kryoflux lies in its pricing structure. The board itself is relatively inexpensive. Around one hundred euro for just the board, 140 euro for the board with cables and a power supply. The software license is free for personal, non-commercial work. However, the license specifically forbids use in educational institutions. In 2019 I was quoted 3,699 euro price for a license for institutional use. This price-point makes using a KryoFlux for institutional projects untenable for many.

For this project, to ensure compliance with the license, I have used only personal collections of floppy disks, I have excluded anything that is owned by the university, and I have used personally owned computers. I also purchased the device myself, avoiding using research or lab budgets for the purchase. These steps may not be feasible in other contexts, leaving a choice between breeching intellectual property norms or using a lower-quality solution to capture disk images.

A second barrier to wider-spread implementation of the KryoFlux is the precise care that must be taken both with the physically fragile device and also with the installation of the accompanying software. Again, I strongly recommend anyone interested in floppy disk preservation to visit the Archivists' Guide to KryoFlux (Allen, Jennifer) on GitHub. Nearly all of the issues one may encounter are detailed there with solutions. One issue that I did not see fully addressed in the guide is that when connecting a KryoFlux to a workstation running Bitcurator, some extra care must be taken in the install. The Java applet that contains the Bitcurator software runs as the ROOT user, and the Ubuntu OS that Bitcurator is built on top of does not, by default, have a root user account set up. This leads to account permissions issues, issues accessing folders created with the KryoFlux software, and issues running the software. I was able to work around this by creating a root account and logging in to the Bitcurator desktop environment using that account, but it is rather klugey and inelegant and it also requires a bit of knowledge of the Linux file structure in order to navigate between folders.

In summary, barriers to wider implementation of the KryoFlux include a very heavy price tag for use in academic settings and a fragile hardware and software ecosystem that requires a bit of system administration background in order to comfortably navigate.

Super Card Pro (SCP) Floppy Disk Controller

The Super Card Pro (SCP) is a very similar piece of technology to the KryoFlux. It captures magnetic flux images of floppy disks and can create disk images from the flux data. In terms of workflow it is a direct replacement for the KryoFlux. There are subtle differences in what the software allows the user to do with the flux images, but for most purposes they are highly similar pieces of hardware. The SCP allows the user to record flux images on a computer and then to write that image to floppy disk. The SCP also has a micro-SD card socket, allowing the disk images to be recorded directly to the card, this is a feature missing from the KryoFlux device. The SCP has drivers only for Windows computers, it does not work with either Macintosh or Linux operating systems. The installation and setup of the SCP software is more streamlined than the KryoFlux installation.

Unfortunately, the SCP does not allow the user to save raw flux recordings. In order to make an image using SCP, you must specify which Disk Type for which you are creating an image. Selecting an image type that does not match the disk's actual type leads to an unreadable image. Thus, using the SCP requires an accurate knowledge of the disk type, whereas with the Kryoflux a flux recording can be made and then many disk images created with different outputs (KryoFlux software refers to disk types as "outputs") until a successful match is found.

One clear advantage that the SCP holds over the KryoFlux is that once you have purchased the hardware use of the software is free without exceptions. Academic and commercial uses of the SCP software do not require additional licensing. For this reason, while I purchased the KryoFlux personally, the SCP was purchased and it is owned by the lab.

Over the course of the project I have used the KryoFlux to record twenty five works of electronic literature. I have re-scanned five of these with the SCP. This disparity is due to my failure to note that A: the SCP software only runs on Windows computers and B: In the Winter of 2020 when I was setting up the Bitcurator workstation in ELL, I replaced the original Windows partition with Ubuntu. This left me with no Windows machines in the lab when I was available to make my images and required the need to return at a later date with a Windows laptop on which to run the SCP. On this return visit I focused on very specific questions about the SCP's viability as a KryoFlux alternative in University settings.

My first question was "Can the SCP record flux images and write them to both harddisk and to floppy? The answer to this question is an unqualified yes. I made a copy of a 1.4mb Macintosh floppy disk version of Tim McLaughlin's "Notes toward Absolute Zero" on a 1.4mb floppy disk (formatted for Macintosh, although that should be irrelevant) and successfully opened the new floppy on a Macintosh LC II. I also attempted to do the same with 400k or 800k Macintosh floppy disks without success. Research in support forums suggests this should be possible if I either write to different floppy media or use additional third party software to edit the flux image before writing. My second question was "can the SCP record raw flux recordings when the disk format Is undetermined?" and the answer to this question Is no. Disk Images made using the SCP require a disk format type to be selected and Inaccurate selections result In an unusable Image. My third question Is a more general "can the SCP replace the KryoFlux board In situations where the license creates a cost prohibition?" My answer to that Is a slightly qualified, yes. The SCP lacks the ability to make free flux recordings, which are excellent to have for long-term preservation. However, If one's goal Is to create disk Images for use In emulation or to create floppy disks (with some effort) to use In vintage computers In place of the originals, the SCP does an excellent job at both of these tasks. The developers of the SCP have posted plans for future development of additional features as well, but It Is yet to be seen If these will come to fruition.

Conclusions

My work with the Rosetta Machines and the Bitcurator Installation did not reveal any new insights into either of these floppy disk preservation techniques. I was not able to capture files from Rosetta Machines due to the pandemic lockdown and I was only able to confirm that Bitcurator Is Inexpensive, It works, and cheap floppy drives can be unreliable.

I was able to add some Insight with KryoFlux, beyond confirming what has already been well documented In the field. While I was not the first person to encounter the user account Issue between Ubuntu-based Linux Installs or the solution of creating a new root user account, documenting this work-around In the community may save time and frustration for users who are Integrating KryoFlux boards with workstations running BitCurator.

Use of the Super Card Pro as an alternative to KryoFlux boards and software Is another useful addition to our knowledge base. When much of the current documentation was written, Super Card Pro devices were not available for sale. Now that they are a real option for labs who have a need for creating magnetic flux recordings of floppy drives without an expensive license for academic use, preservation activities may go forward more efficiently.

Works Cited

Allen, Jennifer. "The Archivists' Guide to KryoFlux." *GitHub*, https://github.com/archivistsguidetokryoflux/archivists-guide-to-kryoflux. Accessed 5 May 2021.

BitCurator. https://bitcurator.net/bitcurator/. Accessed 6 May 2021.

Bitcurator.Net. https://bitcurator.net/. Accessed 6 May 2021.

Byers, Fred R. "4. How Long Can You Store CDs and DVDs and Use Them Again? • CLIR." *CLIR*, Oct. 2003, https://www.clir.org/pubs/reports/pub121/sec4/.

Kirschenbaum, Matthew G., et al. *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Council on Library and Information Resources, 2010.

KryoFlux Products & Services Ltd. https://www.kryoflux.com/. Accessed 6 May 2021.

Levy, David M. "Heroic Measures: Reflections on the Possibility and Purpose of Digital Preservation." *Proceedings of the Third ACM Conference on Digital Libraries - DL '98*, ACM Press, 1998, pp. 152–61. *DOI.org (Crossref)*, doi:10.1145/276675.276692.

Purity, Sonic. *Working with Macintosh Floppy Disks in the New Millennium*. https://siber-sonic.com/mac/newmillfloppy.html. Accessed 6 May 2021.

Rothenberg, Jeff. "Avoiding Technological Quicksand: Sections 1-3 • CLIR." *CLIR*, Jan. 1998, https://www.clir.org/pubs/reports/rothenberg/introduction/.