Optical Media and Floppy disks

Basic concepts & related tools

Agenda overview

○   ○   ○   ○

1 Optical Media Basics

2 Media specifications– How does optical media work?
Tools & Hardware

3 Floppy Disk basics

4 Media specifications– How do floppy disks work?
Tools & Hardware

Optical Media formats

- CDs
    - CD-R
    - CD-RW
    - Audio CD
- DVDs
    - DVD-R
    - DVD-RW
- Blu-ray
- MiniDisc
- LaserDiscs

Optical media is a catch-all term used to describe those shiny, prismatic discs that we all know and love: CDs, DVDs, Blu-Ray, and to a lesser extent, MiniDiscs and LaserDiscs. Most of the optical media in the world is commercially produced, but for a long time CDs and DVDs were the average consumer's preferred method of storing and transporting digital data.

Optical disc structure

- Polycarbonate layer
- Data layer
  - Variation in encoding technology across different subtypes (R-discs, DVD, RW,etc)
- Metal reflective layer
- Lacquer (metal protective) layer
- Optional surface layer

Structurally, CDs consist of a base polycarbonate layer, which is covered with a layer to which data is written. This data layer consists of dye (for write-once discs) or metallic alloys (for rewritable discs), and is covered with a thin lacquer layer.

On CDs, the top layer is the only data side, and scratches on this side are far more likely to affect the disc's readability. DVDs have the same base polycarbonate layer, but write data to a metal "reflector" layer (commercial DVDs) or a metallic alloy layer (for rewritable discs). Both sides of a DVD are polycarbonate and are therefore harder to scratch.Data on optical disc is recorded and read by laser. Data is written to recordable discs in "pits," tiny depressions, and "lands," or the spaces between these depressions. These binary states represent digital zeros and ones. Rewritable discs rely on a binary in the metallic alloy data layer between a reflective,"crystalline" state and a more matte, "amorphous" state.
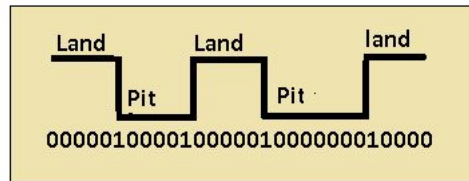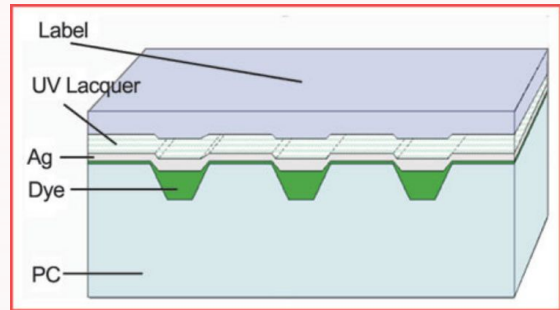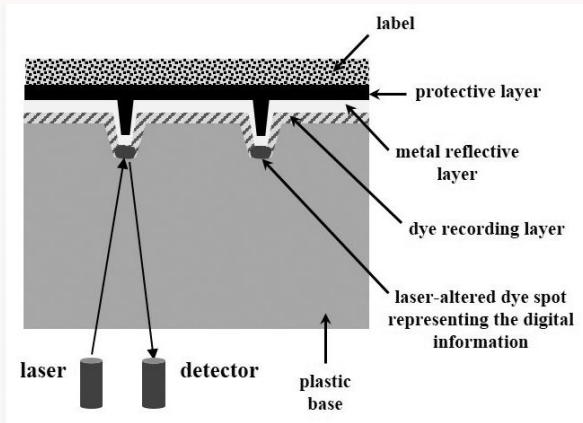
Storage Capacity on average

- CD-R: 700 MB

- CD-RW: 570 MB

- DVD±R (single-layer, one-sided): 4.7 GB

- DVD±R DL (dual-layer, one-sided): 8.54 GB

- DVD±R DS (single-layer, double-sided): 9.4 GB

COMPACT disc DIGITAL AUDIO

Optical Media

It is crucial to understand the logical layout of optical media before attempting any preservation activities. As the first optical media standard was IEC 60908 (1982) for audio storage and playback, future standards reflected a media-centric approach in how they structure the disc.

A simplified model of a compact disc's logical structure is as such: a series of sessions, each containing a series of tracks. A track is, as its name implies, designed to be one discrete track of audio. For non-audio data (e.g. CD-ROM), any number of filesystems may be contained within a single track. These tracks are arranged in a linear series and bounded by a lead-in (which contains the table of contents – locational and descriptive metadata – for the following tracks) and a lead-out. This collection comprises one session. Early CDs were designed as if only one session would be on a disc; this was expanded in 1990 to provide for multiple sessions.

Data is written onto optical discs as a spiral of data from the centre to the outer edge. This spiral is divided into sections called sessions (typically only one), which can contain one or more tracks. Tracks are made up of sectors of data which are typically 2048 bytes in size.

all depend on this concept of "pits and lands." Information is stored on optical media by essentially creating small valleys, called "pits," in plastic (or dye in the case of rewritable discs, depicted in the top right corner). The areas between the pits are called lands. This resembles the "on" or "off" quality of bit cells in floppy disk

encoding, right?

For the most part, optical media drives, be they CD players, Blu-ray players or whatever, are more concerned with a change of state, that is to say going to or from a pit or a land, which is demonstrated through the image on the bottom right.

Not all optical media uses pits and lands exactly like that. LaserDiscs for instance use Pulse Width Modulation to encode an analog video signal on to the disc. Pulse Width Modulation is used in lots of technologies from dimmer-switches to servo motors.

Optical Media

IEC 60908 (1982) - The "Red Book" for audio CDs

ISO9660 (1988) - CD-ROM filesystem standard

table of contents	audio	audio	audio		table of contents	iso9660	joliet	hfs	
lead-in	track 01	track 02	track 03	lead-out	lead-in	track 01			lead-out
session 01					session 02				

Image Source: Duryee, Alexander. "An Introduction to Optical Media Preservation." The Code4Lib Journal Syndication. April 16, 2014.

The familiar compact disc comes in a wide range of formats, almost all of which conform to one of the series of proprietary manufacturer specifications referred to as the "Rainbow Books," some of which have become official international standards. The Rainbow Books were largely authored by Philips and Sony, occasionally with other corporate partners.

first optical media standard was the **Red book** for audio CDs. The first CD standard and used as the basis for all other Rainbow Book standards. Developed to hold up to 74 minutes of audio (later extended to 80). The first discs to use this standard were CD-DA (Compact Disc Digital Audio), or CD-Audio, introduced in 1982. Standardized as IEC 60908 in 1987.

This standard is pretty flexible, you have sessions and within sessions you have tracks. But within anyone track, you could have multiple filesystems.

Yellow Book (1988): Defines data CDs. Though the Yellow Book was not published until 1988,CD-ROM(Read Only Memory) discs conforming to the standard were introduced in 1985 to carry software and video games.

Also known as ISO9660, this is the standard filesystem or method of organization for data on CD-ROMs, is built on top of a system designed for recording music. The filesystems are broken down into sectors, in the ISO 9660 file system, these are 2048 bytes.

Yellow book has a subset of books; green, beige, white, and blue which define a

whole subset of OM, including video cd's, multimedia cd's etc

The logical structure of optical discs has been heavily influenced by the original Red Book CD standard, which houses multiple audio "tracks" in one "session." The same vocabulary is used when discussing subsequent disc storage formats. Today, discs may be comprised of one or more sessions, each of which contains one or more tracks. Tracks store either audio or data organized within filesystems.

File systems and extensions

- **HFS (1985)**: an Apple file system that supports longer filenames and Mac-specific metadata. Often contains some unique Mac-specific data as well as pointing to the underlying data.
- **Joliet (1995)**: a Microsoft extension to ISO 9660 that supports Unicode filenames (allowing for non-Latin scripts), longer filenames, and deeper directories.
- **UDF (1995)**: a universal filesystem that supports longer filenames and more file attributes. Often associated with DVD-Video, but also appears on data CDs and DVDs as a "bridge format" along with ISO 9660, Joliet, and HFS+. Written by the Optical Storage Technology Association to replace ISO 9660.
- **HFS+ (1998)**: an Apple file system that supports even longer filenames, and in Unicode. A replacement for HFS.

Data discs

While discs are not partitioned as hard drives are (using a partition table), they may—and often do—contain multiple filesystems that point to largely the same underlying data. The typical base filesystem is ISO 9660, a standard for data storage that all major operating systems can access. In order to provide access to allOS, ISO 9660 is limited in filename length (originally 8 characters) and size (4 GB) as well as directory depth. Layered on top of this base filesystem are OS-specific extensions and filesystems, which point to the same data on the disc while providing enhanced features for users of that OS.

- **HFS (1985)**: an Apple file system that supports longer filenames and Mac-specific metadata. Often contains some unique Mac-specific data as well as pointing to the underlying data.
- **Joliet (1995)**: a Microsoft extension to ISO 9660 that supports Unicode filenames (allowing for non-Latin scripts), longer filenames, and deeper directories.
- **UDF (1995)**: a universal filesystem that supports longer filenames and more file attributes. Often associated with DVD-Video, but also appears on data CDs and DVDs as a "bridge format" along with ISO 9660, Joliet, and HFS+. Written by the Optical Storage Technology Association to replace ISO 9660.
- **HFS+ (1998)**: an Apple file system that supports even longer filenames, and in Unicode. A replacement for HFS.

DVD

VIDEO_TS

- VOB (Video Object): contain the video and audio streams, as well as subtitles and menus. VOB is the container format; the video is encoded in MPEG-2 in standard definition (PAL or NTSC); and the audio may be one of various codecs (LPCM, AC3, DTS) at a sample rate of 48 kHz.73
- IFO (Information): contains playback information essential to the DVD-Video functioning as a whole.
- BUP (Backup): backups of the IFO files.

In DVD-Video ("authored") discs, video data is organized in the proprietary DVD-Video specification built atop a UDF filesystem. DVD-Video discs include a VIDEO_TS (Video Title Sets) folder, which stores the following files:

●VOB (Video Object): contain the video and audio streams, as well as subtitles and menus. VOB is the container format; the video is encoded in MPEG-2 in standard definition (PAL or NTSC); and the audio may be one of various codecs (LPCM, AC3, DTS) at a sample rate of 48 kHz.
●IFO (Information): contain playback information essential to the DVD-Video functioning as a whole.
●BUP (Backup): backups of the IFO files.The entire disc should be imaged as one file, as these files may not be separated or menu functionality will be lost.

However, the VOB files may be concatenated into a single video file for access. Note that any corresponding AUDIO_TS folders will typically be empty, if not completely absent, as the audio data is present in the VOB files within the DVD-Video specification. AUDIO_TS folders were intended to carry audio information in the DVD-Audio format, which was a similar specification for carrying high-quality audio atop a UDF file system. It enjoyed less success and was described as "extinct" as of 2007; however, AUDIO_TS folders are still present on many DVD-Video discs. (Some suggest this structure is a legacy of early DVD players that require the presence of an AUDIO_TS folder—even if empty—to play.)

Identifying the dye in a CD-R

Table 1: identifying the dye in a CD-R

Dye and dye colour	Reflective layer	Appearance from the non-labelled side of the disc with reflected light	Appearance from the non-labelled side of the disc with light transmitted through the disc
cyanine (blue)	gold	green	blue
phthalocyanine (light green)	gold	gold	light green
cyanine (blue)	silver alloy	blue or greenish-blue	blue or light blue
phthalocyanine (light green)	silver alloy	light green	light green*
azo (dark blue)	silver alloy	dark blue or light blue	dark blue or light blue

* Because the phthalocyanine dye is very light green, the label on the top of the disc may change the appearance of the dye. In this case, the colour of the disc with reflected light is generally a better indication of the dye type than the colour of the disc with transmitted light.

The type of dye in a CD-R can sometimes be determined by transmitting light through the disc and viewing the colour (Table 1). However, the presence of a thick, dark label may make identification impossible with transmitted light. In these cases, reflective light can be used, although the colour of the metal reflective layer may alter the appearance of the dye colour. Note also that some discs have pigmented bases, such as black or possibly a variety of other colours; these colours do not indicate a different dye or metal layer than what is shown in Table 1. The type of dye in a disc can also be determined by consulting the manufacturer or their literature.

Care and Handling– What to do

- Handle discs by the edge or by the throughhole
- Use a non-solvent based felt-tip permanent marker to label side of the disc
- Keep dirt or other foreign matter from the disc
- Store discs upright (book style) in plastic cases
- Leave discs in their packaging or cases to minimize the effects of environmental changes
- Store discs in cool, dry, dark environment in which the air is clean
- Remove dirt, foreign material, fingerprints, smudges, and liquids by wiping with a clean cotton fabric in a straight line from the center of the disc toward the outer edge.
- Use CD/DVD-cleaning detergent, isopropyl alcohol, or methanol to remove stubborn dirt or material.

As with all other information carriers, storage and handling of optical media are critical factors in determining how long they survive. The following recommendations will help maximize longevity.

Store discs vertically in standard-sized jewel cases.
Hold discs by the centre hole and the outer edge between the forefinger and the thumb.
Avoid touching the disc surface with bare hands, as fingerprints will interfere with readability.
Label discs with a water-based permanent marker on the clear inner hub. Do not apply adhesive labels of any type.

Care and Handling– What NOT to do

- Touch the surface of the disc.
- Bend the disc.
- Use adhesive labels.
- Store discs horizontally for a long time (years).
- Open a recordable optical disc package if you are not ready to record.
- Expose discs to extreme heat or high humidity.
- Expose discs to extremely rapid temperature or humidity changes.
- Expose recordable discs to prolonged sunlight or other sources of ultraviolet light.
- Write or mark in the data area of the disc (the area the laser "reads").
- Clean by wiping in a direction going around the disc

And what not to do….

Conservation risks

1. Scratches on the Laser-reading side of CDs and DVDs
2. Adhesive labels
3. Environmental storage conditions–irregular and/or unstable fluctuations in temp/humidity
4. Lack of protocols in the acquisition process
5. Lack of standardization in acquisition, hardware/software used to record on discs
6. Decomposition of the organic layer of the disc
7. "bit rot" or "flipped bits"

Data degradation is the gradual corruption of computer data due to an accumulation of non-critical failures in a data storage device. The phenomenon is also known as **data decay**, **data rot** or **bit rot**.

How to clean optical media

- Remove loose debris with a compressed air duster or a soft, non-abrasive tissue or cloth, wiping from the centre of the disc outward in a radial direction. Never wipe in a circular direction.
- Try cleaning with a dry cloth first, before using any cleaning -solutions.
- Do not wipe in a direction going around the disc.
- Wipe from the center of the disc straight toward the outer edge.
- Avoid using paper products, including lens paper, to wipe the disc.
- Avoid using anything abrasive on the surface of the disc.
- If the disc has a heavy accumulation of dirt, try rinsing it with water first.
- Use commercially available water-based detergent formulated for cleaning the surface of optical discs.
- Use isopropyl alcohol or methanol, as an alternate to water-based detergents, to clean the disc surface.
- To remove fingerprints, apply a small amount of dishwashing liquid to the disc surface and gently wipe it with a wet soft cloth. Afterwards, rinse the disc in distilled water and carefully blot it dry to avoid the formation of water spots.
- Cleaning procedures can scratch the disc surfaces if not performed correctly.

CDs or DVDs do not require routine cleaning. It is best to clean the disc only when it is absolutely necessary, specifically:

- before storing, when surface contamination is visible
- before recording, when surface contamination is visible
- before playing, to prevent surface contamination from being "flung off" while the disc is spinning in the disc drive
- when readability (playability) is impaired and surface contamination is visible

In general, avoid using organic solvents. Harsher solvents (acetone, benzene) will dissolve the polycarbonate and damage the disc beyond repair. Mild solvents (isopropyl alcohol, methanol), however, may be used. These mild solvents evaporate quickly and will not dissolve the polycarbonate.

Other solutions that are not harmful are water-based lens cleaners or water-based detergents (with mild soap) formulated for cleaning CDs or DVDs.

The polycarbonate substrate is a relatively soft and transparent type of plastic. Each time a disc is wiped, rubbed, treated with some solution, or otherwise manipulated for cleaning, that substrate, and thus the disc itself, is at risk of scratching or contamination.

Storage Rec's

Source	Media	Temperature	Maximum Temp. Gradient	Relative Humidity (RH)	Maximum RH Gradient
ISO TC 171/SC Jan. 2002	CD-R CD-ROM	+5°C to 20°C (41°F to 68°F)	4°C /hr (7°F /hr)	30% to 50%	10% /hr
IT9.25 and ISO 18925 February, 2002	CDs DVDs	-10°C to 23°C (14°F to 73°F)		20% to 50%	Cycling no greater than: ±10%
NARA, FAQ About Optical Media, April, 2001	CDs DVDs	68°F (20°C)	+/- 1°F /day (+/- 0.6°C /day)	40%	5% /day
National Archives of Australia, April, 1999	CDs	18°C to 20°C (64°F to 68°F)		45% to 50%	10% /24 hrs
Library Technical Report Nov.-Dec. 1997	CDs	-10°C to 50°C (16°F to 122°F)		10% to 90%	
DVD Demystified, Second Edition, Jim Taylor, 2001	DVD-R DVD-ROM	-20°C to 50°C (-4°F to 122°F)	15°C /hr (27°F /hr)	5% to 90%	10% /hr
	DVD-RAM	-10°C to 50°C (16°F to 122°F)	10°C /hr (18°F /hr)	3% to 85%	10% /hr
	DVD+RW	-10°C to 55°C (14°F to 131°F)	15°C /hr (27°F /hr)	3% to 90%	10% /hr
National Library of Canada, 1996	CDs	15°C to 20°C (59°F to 68°F)	2°C /24 hrs (9°F /24 hrs)	25% to 45%	5% /24 hrs
Media Sciences, Inc. Jerome L. Hartke July 2001	CD-R	10°C to 15°C (50°F to 59°F)		20% to 50%	

Optical discs will perform well within a wide range of temperature and relative humidity conditions. Discs kept in a cooler, less-humid environment and not subjected to extreme environmental changes should last longer. Optical discs stored in an optimal environment will outlast discs that are not.

The life expectancy of optical discs depends on many factors, some controllable by the user, others not.

Factors that affect disc life expectancy include the following:

- manufacturing quality
- condition of the disc before recording
- quality of the disc recording
- handling and maintenance
- environmental conditions

Environmental factors can affect the rate of disc degradation. In each of the three basic disc types, environmental forces will degrade the data layer much faster than the polycarbonate substrate layer (the clear plastic that makes up most of the disc). Because degradation of the data layer will render the disc useless well before the polycarbonate begins to deteriorate, the relative degradation rate for the polycarbonate layer is not used for life expectancy considerations. Physical mishandling of the disc is usually the cause of polycarbonate layer damage. The polycarbonate may also flex or bend if stored for a long period of time in a nonvertical position.

Environmental storage conditions

- Limit exposure to pollutants and light as much as possible. Note that the negative effects of light (such as fading of the dye layer) are not a problem if discs are returned to their cases promptly after use.
- Suitable temperature and RH conditions are described in ISO Standard 18925
- The recommended RH range for extended storage is 20% to 50%, with RH never falling below 10%. The recommended temperature range is from -10°C to 23°C, with the temperature never exceeding 32°C.
- Storage of discs in a cool and dry area will significantly increase media longevity.

Limit exposure to pollutants and light as much as possible. Note that the negative effects of light (such as fading of the dye layer) are not a problem if discs are returned to their cases promptly after use.

The recommended RH range for extended storage is 20% to 50%, with RH never falling below 10%. The recommended temperature range is from -10°C to 23°C, with the temperature never exceeding 32°C.

Storage of discs in a cool and dry area will significantly increase media longevity.

Why do anything? No time to wait!

- "…recordable CDs and DVDs are currently not in a position to serve as a permanent storage solution for digital data."From: "How Long Is Long-Term Data Storage?" Por Barry M. Lunt; Brigham Young University" https://www.imaging.org/site/PDFS/Reporter/Articles/REP26_3_4_ARCH2011_Lunt.pdf

- Optical discs that come from a glass matrix (Glass Master), also known as "embossing", have a significantly longer life than recordable optical discs burned with household equipment.

- Artist created DVDs or those sold by Galleries are probably burned in a non-commercial, DIY way, thus would require special attention at this point.

Tools– Hardware

- Recommended: DVD±R/W drives that have backwards read/write compatibility with CD-R/W and can read CD-ROM. These machines have built-in write blockers or access to removable write blocking hardware, which is best practice for rewritable discs.
- Depending on a collection's size, mass optical transfer may be preferred
  - Nimbie Disc Auto-Loader + Iromlab (Image and Rip Optical Media Like A Boss)
  - Iromlab: https://github.com/KBNLresearch/iromlab
  - https://nypl.github.io/digarch/transfers/nimbie-transfers.html
  - Assessing High-volume Transfers from Optical Media at NYPL

Optical disc and their drives are becoming rarer as trends in hardware, production, and storage shift to file-based media. The discs themselves may not last long enough to become obsolete; both the discs and their signals are susceptible to data loss, and their failure is difficult to predict.

Recommended: DVD±R/W drives that have backwards read/write compatibility with CD-R/W and can read CD-ROM. These machines have built-in write blockers or access to removable write blocking hardware, which is best practice for rewritable discs.

It is difficult to write to discs without intensive intervention, but it appears operating systems can be modified to disable disc writing software; forum posts suggest that Windows XP could

Blu-ray discs use the same UDF filesystem found on some CDs and most DVDs, meaning that imaging strategies will be quite similar save the hardware. Luckily, Blu-ray drives are backwards compatible with most CDs and DVDs. Any discs of non-standard size (greater or less than 12 centimeters in diameter) will require a different type of drive. Some esoteric disc formats may be unreadable with modern drives—for example, DVD±R/W drives may not be able to read DVD-RAM discs

Nimbie: Automated Blu-ray/CD/DVD loading systems with 100-disc capacity and USB 3.0 interface for multiple applications
IROMLAB: writen by Johan van der Knijf, provides a simple and straightforward way to save the content of offline optical media from the KB collection. It currently supports the following carrier types:

- CD-ROM
- DVD (includes DVD-ROM, DVD-Video and DVD-Audio)
- Audio CD
- 'Enhanced'(Blue Book) CD
- Mixed mode CD
- Philip CD-Interactive

Internally Iromlab wraps around a number of widely-used software tools. Iromlab automatically detects the carrier type, including whether it contains data, audio, or a combination of both. The content of data sessions is extracted and saved as an ISO image using IsoBuster. For audio sessions all tracks are ripped to WAV or FLAC format with dBpoweramp.

Tools– Software

Linux tools that can be used to **probe** disc contents:
- disktype
- cd-info
- cdrdao disk-info

Imaging Software
*Depending on the structure and contents of the disc
- Single data track: Guymager (per Tate rec)
- Data discs: FTK Imager
- CD-DA (Compact Disc Digital Audio): Exact Audio Copy
- DVD-Video discs: IsoBuster Pro
- Multi-track: cdrdao

IsoBuster

AD AccessData
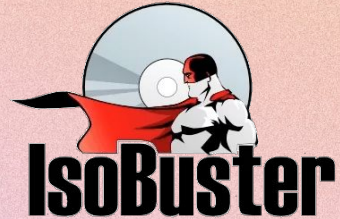
It is essential that prior to imaging the structure of the optical disc is identified, as this dictates the approach you will use.

The important things to identify are:
• The number of sessions on the disc
• The number of tracks on the disc
• Whether these sessions and tracks contain data or audio

There are various Linux tools that can be used to probe disc contents, including:
• disktype
• cd-info
• cdrdao disk-info

these tools identify the structure of the optical disc, save their output as txt files and save them in an artwork folder.

Imaging tools:
IsoBuster Pro is recommended to image DVD-Video discs. IsoBuster, unlike FTK Imager, offers users the ability to specify the creation of a single disc image instead of segmented (multiple)disc images, which is crucial for discs containing large video files. In exploring the imaging of more complex discs, the Pro version of IsoBuster (licensing cost: $39.95)might be considered for its support of multisession discs and file extraction ability from all filesystems.

Exact Audio Copy is recommended for extracting Audio CDs (CD-DA/Red Book). These discs cannot be imaged because of their propensity for error, so must be extracted using software that accounts for inconsistency. Exact Audio Copy is a Windows tool that accomplishes this task in a graphical user interface, and is used by Yale and recommended by AVP for this purpose

cdrdao is a Linux command-line tool which can be used to capture BIN/CUE disk images of complex optical media like multi-session and multi-track discs. This tool creates the equivalent of a CUE file in the TOC format, which can be converted to CUE using the toc2cue command

Types of capture

- **Physical (raw)** image capture: A bit-for-bit copy of all the data on the disc (2352 bytes per sector). This type of image will include both the 2048 bytes per sector of "user data"—the files accessed and changed by the user—and the layer of error correction that makes up the rest of each sector. This approach is most useful with complex discs, such as those with multiple sessions (CD-Extra) or multiple tracks (CD-i).
- **Logical(userdata)**image capture: A copy of all the user dataonthedisc(2048 bytes pe rsector). This type of image will not include the error correction in each sector. Note that DVD sectors are only 2048 bytes, all of which are error-corrected user data; other DVD data is inaccessible with consumer drives. CD sectors are 2352 bytes.
- **Targeted** file extraction: Extraction of the files stored on the disc, without preserving the multiple filesystems and extensions that may also be present. This approach is most effective for discs serving as contemporary carrier mechanisms (for example, if a donor wants to transfer specific files and would otherwise do so over a file transfer protocol).

Types of image file

ISO Disk Image File Format: A logical image format that stores only the 2048 bytes of "user data" per sector, which includes filesystems and their extensions. Though the name implies an ISO 9660 filesystem, this format is file system-agnostic and does not conform to the ISO (or any) standard.

Raw Formats (DD, IMG, BIN): Physical image formats that store the exact 2352-byte sector suences on the media; "any image that is solely an uncompressed sector-by-sector copy of stored bits (with no additional headers/footers/structure/metadata)." It is sometimes possible to rename a raw image file with a .iso extension to open it in programs that do not recognize raw images.

- Raw formats can be standalone files (DD) or navigated with sidecar files. The **BIN/CUE Format** is one example of the latter; the binary file (BIN) stores all 2352 bytes per sector, while a cue sheet (CUE), or a sidecar text file, describes the layout of tracks within the consumer drives.

EW-E01: Physical image that was formerly proprietary but has been reverse engineered into an open specification

WAV/CUE format: A single audio file (WAV) for the entire disc, combined with a cue sheet (CUE), or a sidecar text file describing the layout of track within the image. This format is not a physical disc image, but an extraction of the audio content across a disc.

ISO Disk Image File Format: A logical image format that stores only the 2048 bytes of "user data" per sector, which includes filesystems and their extensions. Though the name implies an ISO 9660 filesystem, this format is file system-agnostic and does not conform to the ISO (or any) standard.

Raw Formats (DD, IMG, BIN): Physical image formats that store the exact 2352-byte sector suences on the media; "any image that is solely an uncompressed sector-by-sector copy of stored bits (with no additional headers/footers/structure/metadata)." It is sometimes possible to rename a raw image file with a .iso extension to open it in programs that do not recognize raw images.

- Raw formats can be standalone files (DD) or navigated with sidecar files. The **BIN/CUE Format** is one example of the latter; the binary file (BIN) stores all 2352 bytes per sector, while a cue sheet (CUE), or a sidecar text file, describes the layout of tracks within the consumer drives.

EW-E01: Physical image that was formerly proprietary but has been reverse engineered into an open specification

WAV/CUE format: A single audio file (WAV) for the entire disc, combined with a cue sheet (CUE), or a sidecar text file describing the layout of track within the image. This format is not a physical disc image, but an extraction of the audio content across a disc.

disk image format matrix (somewhat old- AVP)

https://docs.google.com/spreadsheets/d/18t-fU8ZO20Pgio6-QyPYHP3BR07lqL226gg4vRZPl84/edit#gid=1603236156

Recommended Image Capture Strategy (generally speaking)

- **Physical (raw)** ISO disc images of CDs, except for CD-DA. Physical images (capturing 2352 bytes per sector) are appropriate for bit-level CD preservation, because CD sectors consist of 2352 bytes—2048 bytes of user data and 304 bytes of other data, including error correction.
- **Logical (user data)** ISO disc images of DVDs. Logical images (capturing 2048 bytes per sector)are appropriate for bit-level DVD preservation, because DVD sectors only consist of the 2048 bytes of logical user data.
- **WAV/CUE files of CD-DA discs.** The single audio file (WAV) per disc retains a 1:1 disc:file relationship and allows bit-level media refreshment while sidestepping processing decisions about file naming for multiple tracks on a disc with scattered metadata. The CUE sheet retains the layout of tracks within the image and allows the archivist to break down the preservation file into its original component tracks.

Floppy disks

A blast from the past

don't copy that copy

Floppy disk drives

Digitally shaped analog signal called a "flux" is either
"played" or "recorded" onto a magnetic floppy disk
using a read/write head.
The head is attached to another motor called a
"stepper" that allows it to move along the surface of
the disk into 40-80 fixed positions.
On most drives, the motor rotates at 300 RPM for
"double density" or 360 RPM for "high density" drives.
The head detects these changes in the magnetic flux
of the disk as it spins, amplifies them, and sends then
to an attached controller board. The controller board
refines this signal into digital data that is sent to the
host computer.

A floppy disk drive consists of a spindle motor attached to a mechanism that grabs the
center of a disk and rotates it at certain speed which is measured in RPM, or rotations
per minute. A digitally shaped analog signal called a "flux" is either "played" or
"recorded" onto a magnetic floppy disk using a read/write head. The head is attached
to another motor called a "stepper" that allows it to move along the surface of the disk
into 40-80 fixed positions. On most drives, the motor rotates at 300 RPM for "double
density" or 360 RPM for "high density" drives. The head detects these changes in the
magnetic flux of the disk as it spins, amplifies them, and sends then to an attached
controller board. The controller board refines this signal into digital data that is sent to
the host computer. Disk drive sizes are most commonly 8", 5.25", or 3.5", although
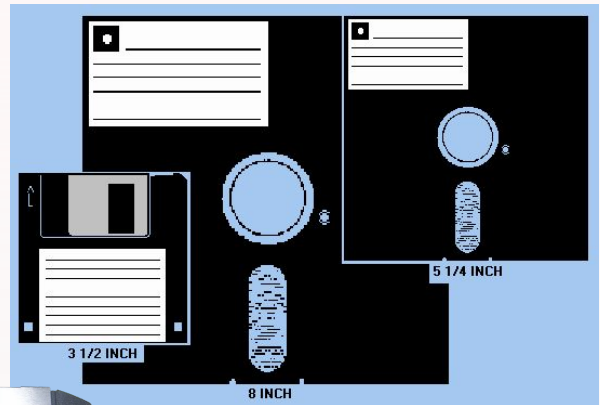some other odd sizes exist.

https://www.cs.mcgill.ca/~rwest/wikispeedia/wpcd/wp/f/Floppy_disk.htm

Floppy Disks types

- 8" released in 1971
- 5.25" released in 1976
- 3.5" released in 1980

Other types of magnetic storage media:

- Zip disks
- Jaz disks

Floppy disks, also known as **floppies** or **diskettes** were ubiquitous in the 1980s and 1990s, being used on home and personal computer ("PC") platforms such as the Apple II, Macintosh, Commodore 64, Amiga, and IBM PC to distribute software, transfer data between computers, and create small backups. Before the popularization of the hard drive for PCs, floppy disks were often used to store a computer's operating system (OS), application software, and other data. Many home computers had their primary OS kernels stored permanently in on-board ROM chips, but stored the disk operating system on a floppy, whether it be a proprietary system, CP/M, or, later, DOS

The 3 broad categories of floppy disks, based on size of the disk, were all released within about a decade of one another. Most common is the 3.5" double-sided 1.44 MB disk
8"floppies fell out of favor relatively quickly in favor of the 5 1/4" versions

All floppies work on the same principle: a magnetic disc with a hole in the middle is inside a case, and a disk drive reads the magnetic data off the disc.

Drives can be either "double sided" meaning the can use the front and back of the disk at the same time, or "single sided" meaning only one side is used. Since most disks are double-sided, you can typically flip the disk over and use the other side on

single-sided drives. Double density and high-density disks have different coercivity levels and cannot be interchanged, or the data will not be able to be written properly.

Worth mentioning other types of magnetic storage media; such as zip disks, very similar to floppy disks. Again they are plastic discs coated with magnetic material. The difference between them is that zip disks can store much more. The one shown stores 100MB and you can get them up to 250MB.
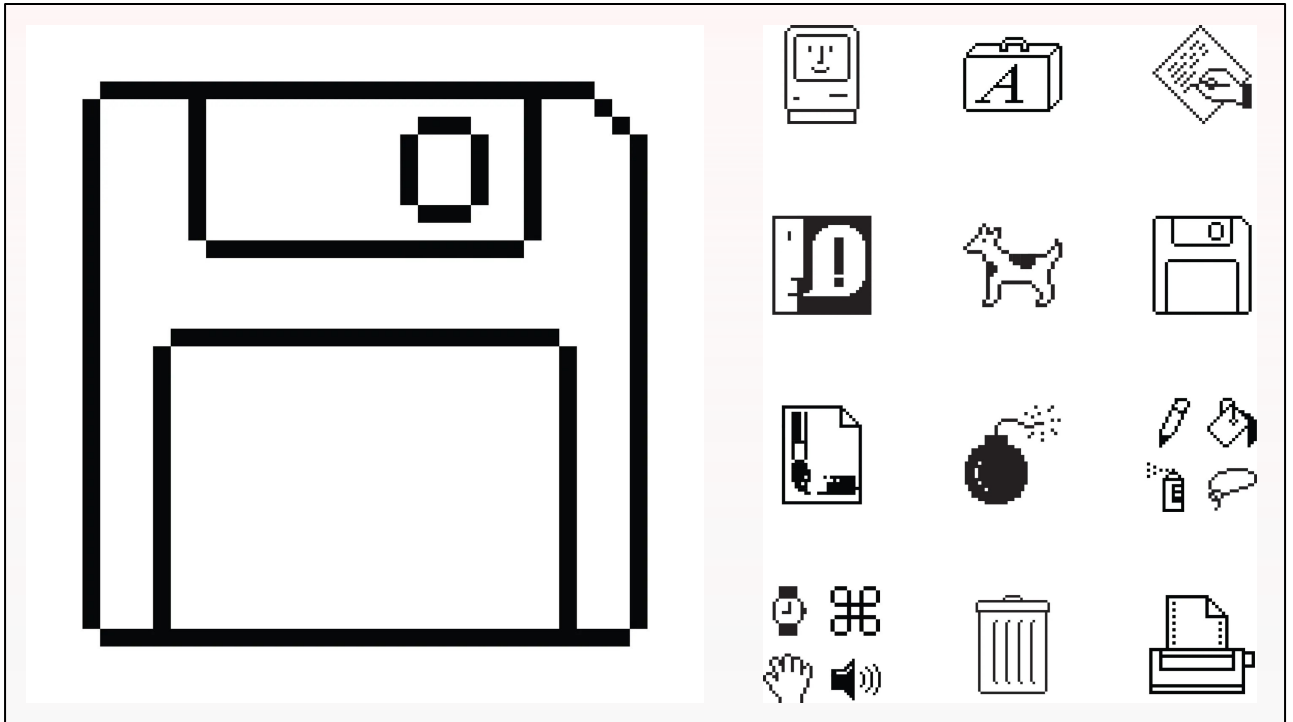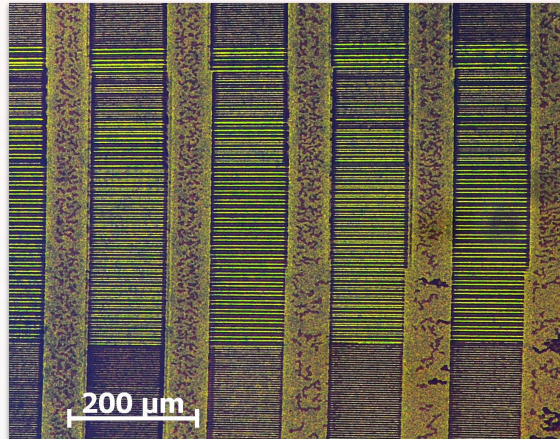
While we associate floppy disks with the shape of their housing, the save icon, they are actually disks, hence the name. Inside that plastic shell we're all familiar with is a disk that looks like this.

Floppy Disk structure

Magnetic Media

- DD: 2 µm magnetic iron oxide
- HD: 1.2 µm cobalt-doped iron oxide
- ED: 3 µm barium ferrite

200 µm

The 3.5-inch floppy disk can easily be compared to a cassette tape. The floppy disk and cassette tape use a thin plastic base material coated with magnetic iron oxide on both sides. Both can record information immediately, erase and write data, and can be very inexpensive and easy to use. One of the major differences from a cassette tape is its shape. While the cassette tape's magnetic tape is a long thin ribbon, the floppy is arranged like a disk. In this way, the floppy disk is similar to a vinyl record: the tracks are arranged in concentric rings, also called sectors, which permit the software to jump from file to file without needing to go through the files in between. The read/write heads move to the correct track, providing direct access storage

Data is encoded on a disk by arranging metal particles on the disk. The image on the left shows the head, which uses a magnetic field to align the magnetic particles.
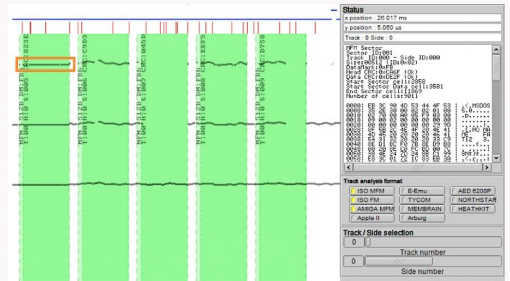
To read data, the magnetic particles in the media induce a tiny voltage in the head as they pass under it.

Just like videotape!

Magnetic Flux

A drive head detects and creates polarity changes in the magnetic flux on the surface of the disk, which is interpreted by the drive as a binary 0 (no change) or a 1 (polarity change) over a specific period of time.

Because the data recorded by one disk drive must be able to be read in another as well as the limitations inherent in analog magnetic recording, the flux stored on the disk has to be recorded so that it can be read back in exactly same way it is written.


A digitally shaped analog signal called a "**flux**" is either "played" or "recorded" onto a **magnetic floppy disk** using a read/write head.

A floppy disk consists of a mylar-coated magnetic surface that is rated to store a certain number of flux reversals (from positive to negative, or negative to positive) within a defined area of the surface at a certain speed, at a certain TPI (tracks per inch, which is how close they can be to each other), at a certain coercivity (which is how strong the signal is).

A drive head detects and creates polarity changes in the magnetic flux on the surface of the disk, which is interpreted by the drive as a binary 0 (no change) or a 1 (polarity change) over a specific period of time. Because the data recorded by one disk drive must be able to be read in another as well as the limitations inherent in analog magnetic recording, the flux stored on the disk has to be recorded so that it can be read back in exactly same way it is written. This is called "clocking" and means that we have to have set timing patterns for the flux transitions to occur.

Floppy disk structure

- Encoding
- Tracks
- Sectors
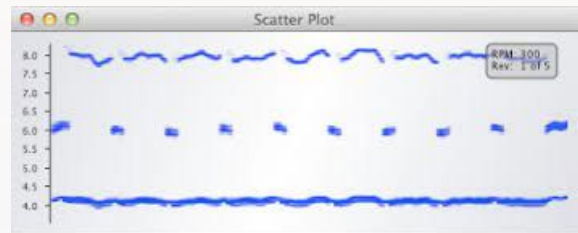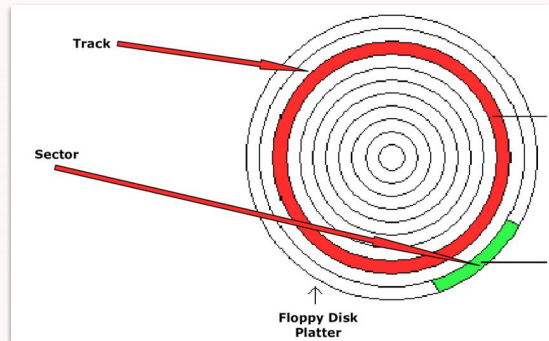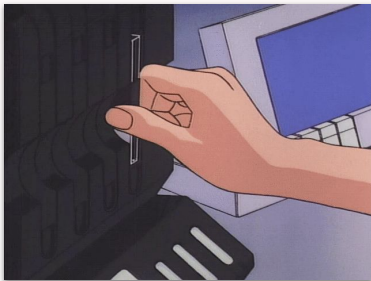
So how is that information organized?

First there's the "disk surface" which is just which side of the disk you're on (certain types of floppies allow for reading and writing on both sides, some don't). Tracks, which are concentric circles going around the disk. How many tracks a disk holds depends on how the disk was encoded, but I believe the maximum is around 80 tracks per disk. Sectors are smaller chunks of a track. You can see that represented on the sort of pie chart image in the bottom left corner.

The scatter plot on the right shows how the floppy drive's read/write head "sees" a particular track on a disk. One full rotation around the disk, represented here as a straight line. The repeating patterns you see on the top two lines of this graph are sectors.

Encoding- Because of the need for proper clocking, data is encoded so that it cannot violate the rules. The different encoding types include GCR (Apple, Commodore), FM (Atari 8-bit, many other early systems), or MFM (Amiga, Atari ST, PC, many others). These encoding rules, whether implemented in firmware or software, dictate that the data is written in a way that it can be read back exactly the same each time.

Tracks- The concentric circles on the disk that the flux reversals are recorded on are

called tracks. The stepper motor can move and stop at different tracks on a disk, typically around 40 in total for single/double density, and 80 for high density.

Sectors- On most disk systems, tracks are further divided into sectors, which are sections of each track typically divided by "sync" marks. A sync mark is a special pattern on the disk that tells the controller where the data starts. Different drives store data at different "bit densities" (essentially clock speed rates of the read/write hardware). However, most systems are locked to a specific density on all tracks.

Floppy Disk parts and structure

1. **Paper ring:** The magnetic disk is between the two paper rings.
2. **Hub**: the metal centre of the magnetic disk that holds the disk together while it spins.
3. **Magnetic disk**: The disk is covered with iron oxide that can be magnetized. When data is saved on the floppy disk, a recording head creates a magnetic pattern on the iron oxide in a form that the computer can read and access the data anytime
4. **Shutter**: folded over one edge of the disk which goes inside the computer first. Inside the computer, the shutter slides over and the data can be read through the rectangular slot
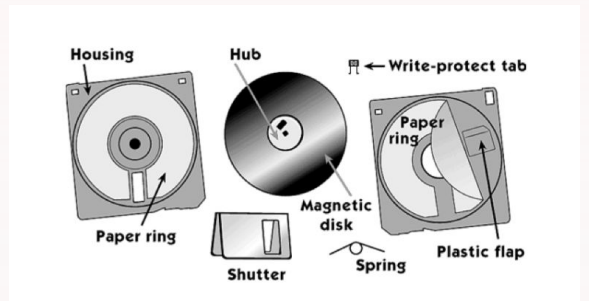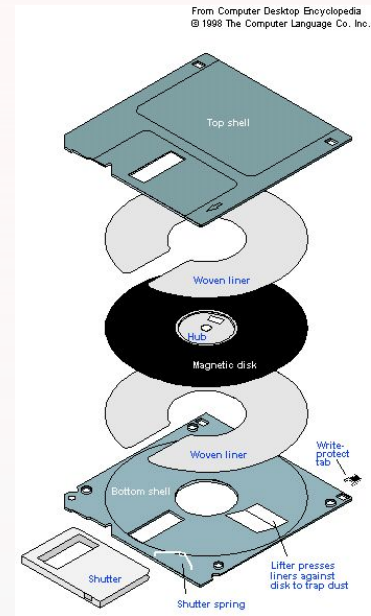
When taking apart a 3.5-inch floppy disk, you will find two colouredplastic squares: the housing that holds the other smaller parts of the diskette, also known as the disk jacket.

Floppy Disk parts and structure

5. **Write protect tab**: in the upper right corner of the diskette. When the hole is slid open, the floppy disk is locked and the computer will not allow you to add or erase

6. **Read/Write Head**: It is located on both sides of the floppy disk. The heads are not directly opposite of each other in an effort to avoid interaction between write operations on each surface. One head is used for reading and writing while the wider head is used to erase a track prior to begin writing

11. **Boot Sector:** The boot sector is the area where the Master Boot Record (MBR) is located; it is the first sector of the disk. The MBR is the program that runs when the computer starts

From Computer Desktop Encyclopedia
© 1998 The Computer Language Co. Inc.

Top shell

Woven liner

Hub

Magnetic disk

Woven liner

Write-protect tab

Bottom shell

Lifter presses liners against disk to trap dust

Shutter

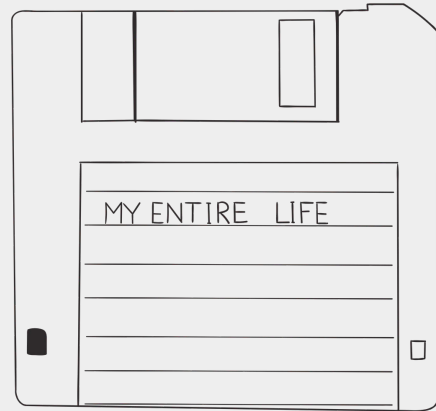Shutter spring

Conservation risks

- In order to create these disk images, floppy disks need to be kept in good physical condition to be read.
- As with any physical media, the way that floppy disks are stored and handled can have a significant impact on their condition.
- Physical damage caused by improper care as well as natural degradation over time pose a threat to the archivist's ability to preserve and digitize diskettes.
- Over time, magnetic media slowly demagnetizes itself due to stresses like thermal cycling and alternating magnetic fields nearby or even adjacent disks and movement of those disks", As bits lose their magnetic orientation, data on the disk becomes corrupt
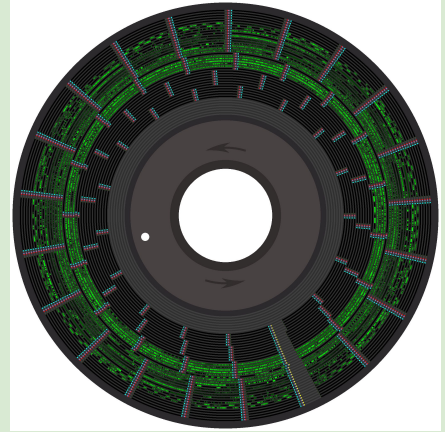
While there is no hard data about the longevity of magnetic computer media, their lifespan is estimated to be less than 10 years. Disks and diskettes are vulnerable to strong magnetic fields, which can demagnetize the magnetic particles that store the data. Diskettes are especially susceptible to loss of data from debris entering the cartridge - a small scratch can remove a single "bit" that could make the entire diskette unreadable. Finally, each time magnetic media are used, a small amount of magnetized material is worn from the disk; frequently used disks may eventually become unreadable due to this data loss.

Storage Recs

- Floppy disks should be stored in a cool, dry place below 23°C and below 50% relative humidity, as hotter and more humid conditions reduce the lifespan of the disks
- Particulate matter can scratch the magnetic media, resulting in permanent loss of disk contents. Over time, dust or debris may adhere to the surface of the disk causing read errors, such as causing the head to skip along the surface of the disk
- Though the plastic disk jacket of 3.5-inch floppies affords some protection, care should be taken to not bend or twist disks
- Storing disks on top of each other, jamming into boxes, or placing heavy objects on top of them can cause warping in the jackets making the disk spin unevenly, resulting in erratic reading and writing

MY ENTIRE LIFE

23 C is 73 F

- Floppy disks should be stored in a cool, dry place below 23°C and below 50% relative humidity, as hotter and more humid conditions reduce the lifespan of the disks
- Particulate matter can scratch the magnetic media, resulting in permanent loss of disk contents. Over time, dust or debris may adhere to the surface of the disk causing read errors, such as causing the head to skip along the surface of the disk
- Though the plastic disk jacket of 3.5-inch floppies affords some protection, care should be taken to not bend or twist disks
- Storing disks on top of each other, jamming into boxes, or placing heavy objects on top of them can cause warping in the jackets making the disk spin unevenly, resulting in erratic reading and writing

Tools– Hardware: Disk drives

The physical disk drive used can have an impact on how easily the disk is read.

Due to mechanical features such as the head alignment of the drive, images of the same disk captured using different disk drives can vary widely in quality.

For this reason, having multiple older disk drives on hand is useful as certain drives are better equipped at reading particular disk formats.

Recommended: Panasonic JU-257 and a Teac FD-55GFR

Legacy technology is often difficult to acquire, time consuming and complicated to set up, and difficult to use. While a legacy machine with an internal disk drive may read certain disk formats very well, it lacks the flexibility of modern tools designed for digital archivists that can analyze the disk format and recover the data, as well as presenting additional challenges such as the lack of modern ports to export data.

The physical disk drive used can have an impact on how easily the disk is read. Due to mechanical features such as the head alignment of the drive, images of the same disk captured using different disk drives can vary widely in quality. For this reason, having multiple older disk drives on hand is useful as certain drives are better equipped at reading particular disk formats.

sometimes all that is needed is the persistence to make multiple attempts with the same configuration to achieve the desired results.

Tools– Hardware: Floppy disks drive controllers

- <u>Kryoflux</u>
- Device Side Data <u>FC5025</u>
- DiscFerret
- CatWeasel
- RetroFloppy (Apple only)

The obsolescence of hardware and software required to access floppy disks is a serious challenge for their preservation and digitization.

Because of the fragile nature of floppy disks, and the prevalence of read errors with this media, in the field of digital forensics, and digital preservation, floppy drive controllers have become very popular.

A floppy drive controller connects to a floppy disk drive and, among other things, controls the rate that the drive reads the media. By slowing down the rotations of a disk, a floppy drive controller can help reduce the amount of read errors encountered when moving media off of the disk, and limit wear and tear on the disk and the drive.

The most popular floppy drive controller in cultural heritage is the Kryoflux. The kryoflux has built in write blocker, which prevents inadvertently writing data to the floppy disk. It also comes with custom software for operating the kryoflux hardware,and disk imaging the floppy disks.

kryoflux- Samples magnetic flux transition timings on magnetic media (particularly floppies). Assists in recovering degraded media, copy-protected media, most floppy formats from the 80's and 90's. Announced in 2009/2010, particularly promoted to video game preservation communities

FC5025- USB 5.25" floppy controller plugs into any computer's USB port and enables you to attach a 5.25" floppy drive. Even if your computer has no built-in floppy
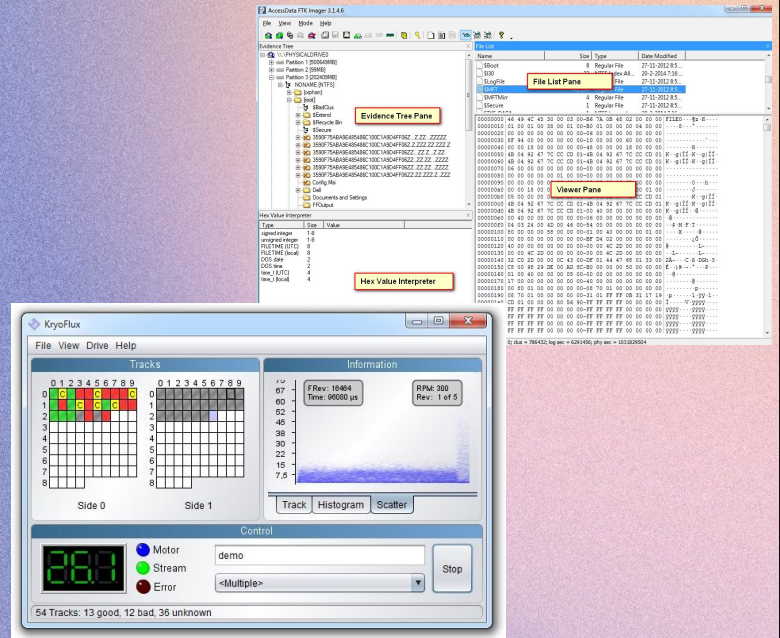
controller, the FC5025 lets you read those old disks. And it's not just for IBM PC disks – it also understands formats used by Apple, Atari, Commodore and TI, among others. The FC5025 is intended for 5.25" disks only, not 3.5" or 8" disks.

Focusing on Kryoflux

Tools– Software

For Imaging:

- Kryoflux GUI
- FTK Imager
- Guymager

Identifying the formats of floppy disks that need to be imaged is a challenge that can prevent the disk from being read and imaged. "the formatting of the disk appeared ultimately to have the greatest significance in determining whether or not an image could be captured." competing computer system developers from the 1980s implemented many proprietary technologies such as disk encodings, file systems, and file formats, making digital information far more enmeshed in the particular technologies that created them than is the case today.

Trial and error required when attempting to read and image disks

Kryoflux- is a hardware and software system that can read, convert, store, and write contents of various legacy disk formats. The system consists of a control board and a Java application, DiskTool Console (DTC). DTC refers to the command line interface. Kryoflux records data from an attached drive as a flux data stream. Kryoflux can also output disk images in a variety of sector formats.

Guymager- Developed by Guy Voncken as a simple, performant alternative to existing commercial forensic imaging tools.

Multithreaded compression, verification handling. Ideal for running multiple hardware acquisitions simultaneously.

FTKImager (CLI is free)- Tool built by AccessData for dd-style and forensic imaging.

Forensic packaging stores additional metadata and checksums about image, file system, and capture environment.

Kryoflux & FC5025 GUIs
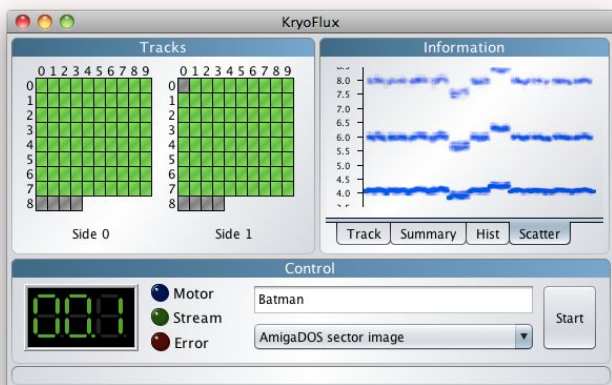
The software that comes with the kryoflux drive controller allows you to capture the data off of the floppy disk as raw stream files. These are basically just the zeroes and ones that make up the data you're hoping to preserve. The upside to this method is that you will get a complete copy of all of the information on your disk, just as it would have been read off of a floppy drive. The downside to this is that the data must be interpreted. An operating system can't just open raw stream files. There are now several emulators that can work with kryoflux raw stream data, but I would say the safer route would be to keep the raw stream data, and create a more typical disk image.

To do this, you have to match the raw stream data that you initially captured off of the disk against an encoding. A filesystem. Individual tracks and sectors will match certain parameters due to their encoding, so it's just a matter of knowing, or learning the way the disk was encoded. Thankfully, once you have the raw stream data you can use the kryoflux software to test it against different types of encoding until you're successful. Also if all of your floppy disks are from the same computer or were used at the same workplace, than they're all likely encoded the same way.

Once you've identified the floppy's file system and created the disk image, depending on how the disk was encoded, a current operating system may not be able to open it. So you may have to explore emulation or virtualization of an operating system to be able to access the data.