Reliable Storage Media for Electronic Records
A Guide for Government Agencies

Modern computer systems use a wide variety of storage media to store and access electronic data. What media is used depends on a number of factors, but cost, speed of access and ease of use are common drivers of selection decisions. Often overlooked are concerns of long-term reliability and sustainability. Electronic records are vulnerable to degradation or loss if not maintained in an appropriate storage environment which takes into consideration media reliability and guards against technological obsolescence.

To say that media is reliable is to indicate that it can be trusted to preserve and provide access to data stored on it over time. While no storage medium can guarantee reliability and sustainability, certain media formats have distinct advantages over others in this regard. Understanding a media format's strengths, weaknesses and expected life span allows IT managers to appropriately protect the data stored on that media.

No storage media alone can ensure the preservation of electronic records. Selecting appropriate media is one part of a greater preservation strategy which includes using sustainable file formats, actively managing files over time, planning for future technology change and securing adequate resources to support preservation activities.

The below factors should play a part in the selection of any storage media for electronic records.

1) **Durability:** Durability is a factor representing the ability of electronic storage media to withstand wear and environmental conditions. Corruption (data rot) can occur as the electrical charge, magnetic orientation, or physical material degrades, causing unintended changes or loss of data.
    - **Assessing Durability:** Durability of media is commonly expressed in terms of "mean time between failures," which indicates how long a given drive/tape/disk can be expected to operate before failure.

2) **Widespread Adoption and Use:** Widespread adoption and use is a factor indicating a wide user base, meaning more stakeholders have a vested interest in keeping the storage media viable and well-supported. Widespread adoption also serves as an indicator of general media stability and generally provides a lower overall lifecycle cost of storage. The more widely a storage medium is used, the more likely it is to have long-term support to maintain it. Widely adopted technologies are typically documented and based on open standards supported by multiple hardware vendors. This reduces the chance of a medium becoming inaccessible due to one vendor going out of business.
    - **Assessing Adoption and Use:** While there is no universal benchmark that indicates something is "widely used" one should look for examples of a given technology being used by other institutions, government agencies and private corporations. Multiple manufacturers and distribution sources are also good indicators of a widely-used technology.

3) **Integrity:** Integrity is a factor indicating the ability of electronic storage media to protect against and correct data corruption. The use of parity bits, error correcting codes, checksum algorithms, physical and digital access controls, and other measures help ensure that data is not corrupted. The media format used and how data is stored on it

Printed by authority of the State of Illinois. November 2014 — 1 — AR D 153

determines which of these measures can be applied. Compressed, de-duplicated, or encrypted data is more susceptible to corruption as non-functional software or an uncorrectable error can make a large amount of data unreadable.

- **Assessing Integrity:** All storage media have listed specifications that indicate what types of integrity protection are possible, but one must also consider the source. New technologies frequently come with lofty manufacturer claims which should be considered critically until independently verified through outside testing or use.

4) **Redundancy:** Redundancy is a factor that indicates the data stored on electronic storage media is being replicated to ensure recovery of data in the event of a data loss incident. A minimum of one additional copy of any data representing electronic records must be maintained to protect against such a loss. At least one copy should be stored in a geographically separate location. Depending upon cost and performance needs, multiple types of storage may be used, such as a hard drive for the primary copy and magnetic tape for the backup copy.
   - **Assessing Redundancy:** Redundancy can be assessed by determining if you have one additional copy in a geographically separate location. IT policies should ensure that all electronically stored data will be restorable in the case of total loss of the primary storage environment.

Below is a list and descriptions of storage media formats currently recommended by the Illinois State Archives for use in storing electronic records. The list may be updated or expanded as technology warrants, so be sure to check for newer versions in the future. Questions about storage media should be directed to Kris Stenson, Electronic Records Archivist, at 217-557-1085 or kstenson@ilsos.net.

Format	Short term	Long term	Do Not Use
HDD	X	X	
SSD (Internal)	X	X	
Magnetic Tape	X	X	
Cloud Storage	X	X	
Optical (All Types)	X		
USB Flash			X
Obsolete media			X

Storage Media

Long-term Retention

The below formats are considered acceptable choices for the retention of records greater than a decade.

Magnetic Tape:

A durable recording medium which uses a plastic film coated with magnetic material to record information, magnetic tape has been used to record computer data since the 1950s. Early formats of this medium consisted of open reel tapes, but modern varieties all use a cartridge of some sort.

The most widely used current version is **LTO (Linear Tape-Open)**, which is based on open standards, as opposed to several proprietary competitors. LTO is currently in its 6th generation, with LTO-6 introduced in 2012. LTO-6 tapes have an uncompressed storage capacity of 2.5 Terabytes (TB). A number of different companies currently manufacture LTO tapes and drives, and LTO technology now accounts for close to 90% of the data tape market. LTO drives have some backwards compatibility, being able to read tape from two generations past and write to tape one generation past (an LTO-6 drive can read LTO-4, 5 and 6, and write to 5 or 6). Older versions of LTO tapes can remain viable for a few decades so long as users possess the appropriate drive, but it is best to migrate to newer versions every two generations to avoid potential loss of access.

Other current tape technologies are the Oracle Storagetek T1000X series and IBM TS1140 line. Both offer higher capacities and faster transfer speeds than LTO, but at a higher cost per GB. They are both proprietary formats, with drives and automated libraries available only from Oracle and IBM, respectively. Both companies are long-established and stable, but if either chooses to discontinue their tape technologies users will have no choice but to switch formats entirely. LTO thus remains the safer choice in terms of adoption and support.

Pros of magnetic tape:

- Durability up to 30 years (best practice migration in 8-12 years)
- High capacity, low cost compared to other storage technologies
- Widely used, mature technology
- High transfer rates, low error rates
- Low energy consumption

Cons of magnetic tape:

- Slow access time (average 50 seconds)
- Wears out faster with frequent access

Hard disk drive (HDD):

Hard disk drives store data on a stack of rapidly spinning metal disks coated in magnetic material. HDDs have been used for primary storage in computers since the early 1960s, and are used in the vast majority of personal computers and servers today. They can be internally mounted or connected externally. For stability and monitoring it is recommended that only internal HDDs be used for long-term records storage, with external drives being used for file transport or backup duties only. Due to their extensive use HDDs are inexpensive, and are available from a wide variety of manufacturers. They can be prone to unexpected failures,

however, so active monitoring, regular media refreshment and appropriate backups must be used to ensure the safety of the records stored within.
To help manage the inherent risks associated with HDD technology a RAID (Redundant Array of Independent Disks) setup should be used. RAID uses a battery of drives that are interlinked and automatically duplicate data across the drives, thus protecting content from loss. There are different levels of RAID which correspond to greater or lesser amounts of duplication, but for records preservation RAID 6 or 10 are recommended. Both involve high levels of fault tolerance, meaning one or more drives in the array could fail at once with no irretrievable loss of data.

Pros of HDD:

- Rapid access to content
- High capacity, low initial cost
- Widely used, mature technology
- Easily scalable through networking

Cons of HDD:

- Short life span (average 4-6 years, best practice migration in 3-5 years)
- High energy consumption
- Expensive for large-scale applications or for long-term content
- Higher error rate than tape

Solid State Drive (SSD):
A flash memory storage device first developed in the mid-1990s with no moving parts that typically uses the same shape, interface, and power source as standard hard drives. Data is stored in static electronic chips rather than on magnetized spinning platters. This results in much shorter time required for drive start-up, read, random access, latency and data transfer as well as reduced energy use, but at a cost up to ten times that of standard hard drives. As the price continues to drop in the coming ten to twenty years, solid state drives are expected to replace standard hard drives as the primary storage medium for laptops, desktops, servers, mobile devices, and external storage.

Pros of SSD:

- Resilient to physical shock
- Lower failure rate compared to standard hard drives
- Fast access time (<0.1 ms)
- Low energy consumption

Cons of SSD:

- High cost compared to other storage technologies
- Limited lifetime due to limited number of times a storage block can be written
- Susceptible to data loss due to power outages or long-term unpowered storage
- Maturing technology with most commercial availability beginning in 2007

Cloud storage:
“Cloud storage” refers not to a particular type of media, but a method for managing data using networked storage providers. Cloud hosting companies provide technical infrastructure which often spans across many geographical areas, providing high levels of redundancy and remote access for customers. While not a new concept, commercial cloud storage has only seen widespread adoption by both private and public entities in the last decade. There are many cloud service providers but much of the commercial market is dominated by companies like Amazon, Microsoft and Google. Cloud storage services can range from bare-bones

warehousing with minimal security and upkeep to highly customized management of data, with integrity checks, enhanced security and faster access speeds.
Cloud storage in general has shown itself to be very reliable regarding the preservation of data. Nonetheless, greater concerns arise surrounding the protection of that data from inappropriate access. Data breaches can and do happen, and network security must be a primary focus for any agency wishing to use cloud storage for their records. Only established providers with proven track records should be used, but they do not necessarily need to be one of the large corporate entities. Many smaller cloud providers actually use one of the large hosts, and simply add their own layers of services on top of the bare storage. Cloud storage providers may not automatically provide long-term preservation services such as fixity checks, audit logging or creation of additional metadata, so agencies must still plan on performing these tasks themselves or specifically contracting cloud providers to do so.

Pros of cloud storage:

- Highest level of duplication and geographic redundancy
- Easy access from multiple locations
- Keeps up with technology trends without additional investment
- Trades unpredictable maintenance costs for known subscription fee
- Can be cheaper than investing in own technology

Cons of cloud storage:

- Relatively higher risk of security breach
- Laws may prevent the storage of certain types of sensitive data in the cloud
- Some providers may not be reliable or may go out of business
- Less control over data / loss of physical custody

Short-term (10 years or less) Retention

The below formats are inappropriate for the long-term storage of electronic records, but may be used for short-term storage of records.

Optical Media:
A thin, circular, plastic disc with a reflective layer upon which data is stored in the form of pits and lands. The reflective layer typically resides on the label side of the disc facing inward although double-sided and dual-layer discs are also available. It can be engraved (read-only), dye-based (write-once), or alloy-based (rewritable). A laser is used to read data from the spinning disc based on changes in the reflection caused by the pits and lands. Several forms of optical media are widely adopted and supported; descriptions of the most common forms are provided below. Optical media is subject to damage due to scratches or breakdown of the recording dye, although proper storage and handling, regular migration to new media, and use of a gold reflective layer can mitigate these risks. Many types of writable optical media also use volatile organic dyes to store information, and can degrade over time. As cloud and network-based storage become common, optical media usage is expected to decline.

CD: The Compact Disc was originally developed in the early 1980s, evolving from the older LaserDisc format, and it is still widely supported. A standard CD is 4.7 inches in diameter and can hold up to 80 minutes of audio or 700 MB of data, although smaller and non-round shapes also exist. The most popular CD formats include CD-ROM (read-only), CD-R (write once), and CD-RW (rewritable). Both the drives and recordable media are speed rated, indicated as a multiplier of 1x (1200 Kbps). The writing speed as set by the recording software should not exceed the rated speed of the recording media to

prevent data corruption. The CD format includes strong error correction coding to prevent data loss due to scratches, fingerprints, or other environmental contaminants. CDs suffer from low capacity compared to other modern storage technologies.

DVD: The Digital Versatile Disc was developed in 1995. The DVD uses the same dimensions as a CD but offers a standard capacity of 4.7 GB or 8.5 GB for dual layer formats. Most DVD players can also read CDs. The most popular DVD formats include DVD-ROM (read-only), DVD-R and DVD+R (write once), and DVD-RW and DVD+RW (rewritable). The plus and minus formats require different recording media and drives to write. Most DVD players can read all DVD and CD formats while some can also write in all formats. Similar to CDs, both the DVD drives and recordable media are speed rated, indicated as a multiplier of 1x (10.5 Mbps). The writing speed as set by the recording software should not exceed the rated speed of the recording media to prevent data corruption. The DVD format also includes strong error correction coding to prevent data loss due to scratches, fingerprints, or other environmental contaminants. DVDs should not be used for long-term data storage, because their reliability over time has not yet been adequately demonstrated.

Blu-ray: The Blu-ray Disc was developed in 2006. Blu-ray uses the same dimensions as a CD and DVD but offers a standard capacity of 25 GB or 50 GB for dual layer formats. Most Blu-ray players can also read CDs and DVDs. The most popular Blu-ray formats include BD-ROM (read-only), BD-R (write once), and BD-RE (rewritable). Similar to DVDs, both the Blu-ray drives and recordable media are speed rated, indicated as a multiplier of 1x (36 Mbps). The writing speed as set by the recording software should not exceed the rated speed of the recording media to prevent data corruption. The Blu-ray format also includes strong error correction coding to prevent data loss due to scratches, fingerprints, or other environmental contaminants. Like DVD, Blu-ray discs have not been adequately evaluated for long-term stability.

M-Disc (Millenial Disc): A relatively new player on the optical front, M-Disc technology has only been available since 2009, and has yet to see wide-spread adoption. M-Discs use a proprietary “stone-like” material in the storage layer, sandwiched between plastic discs. Currently both DVD and Blu-ray variants are available, with capacities similar to standard varieties of those discs. Special M-Disc drives are required to write data to the discs, but standard DVD or Blu-ray drives are able to read data from them. Millenniata, the company behind the format, claims such discs have a projected lifespan of 1000 years. Independent stress tests have shown that M-discs are indeed more resistant to environmental degradation than traditional varieties, but the life span statement cannot be corroborated. Even if such claims are accurate, optical drives are already losing ground to newer storage technologies, and will likely be rendered completely obsolete within decades. It is possible that M-Disc technology will be much more widely adopted in the future, but at this juncture it is too limited to consider as a long-term storage option.

Pros of optical media:

- Convenient and portable
- Widely supported formats available
- Low energy consumption in storage

Cons of optical media:

- Not well-suited for frequent writing or for fast read access from multiple discs

- Limited capacity per disc compared to other modern storage
- Widely varying lifetime depending on use and care
- Aging technology being slowly replaced by flash media and cloud storage

Do Not Use

The below formats should not be used for the short-term or long-term storage of electronic records. If records are currently stored on such media they should be migrated onto appropriate media to avoid their permanent loss.

USB Flash Drive:
A rewritable portable data storage device developed in 2000 with no moving parts that connects to a computer using the Universal Serial Bus interface. Data is stored electrically in chips using power from the USB interface itself. USB flash drives offer many of the same performance benefits and limitations as Solid State Drives but typically have a smaller capacity. USB flash drives come in a variety of shapes and sizes from standard three inch “stick” to novelty shapes and even some barely larger than the USB plug. USB flash drives quickly displaced floppy disks as the preferred means to quickly write and transport data but may themselves be displaced as cloud and network-based storage become common. These portable devices are useful for short-term information sharing, but are far too easily lost or compromised to serve as reliable storage for electronic records.

Pros of flash drives:

- Convenient
- Durable
- Widely supported
- Fast transfer rate (up to 5 Gbps with USB 3.0)
- Low energy consumption

Cons of flash drives:

- Easily misplaced or stolen
- Limited write protection
- Limited lifetime due to limited number of times a storage block can be written

Older magnetic tape formats:
Due to lack of support for obsolete technology, agencies should avoid any tape formats which have been discontinued. Caution should also be exercised when using older variants of current tape technology such as LTO-1 or LTO-2, as backward compatibility of LTO technology only goes back two generations. Any records currently stored on older formats such as these should be migrated forward to more current versions to ensure continued accessibility.

Any size floppy, ZIP, JAZ disk
A wide array of magnetic diskette technologies have now been almost completely abandoned in favor of more current storage options. Even if still technically usable, such disks should not be used for storage of electronic records.

Additional Resources

Best practices for a digital storage infrastructure for the long-term preservation of digital files (Digitizing Contemporary Art)

Best Practices for Media Selection and Migration (University of Illinois)

Selecting Storage Media for Long-Term Preservation (UK National Archives)

FAQ about Optical Storage Media (National Archives)