Back-Up and Storage

Digital Preservation Topical Note 2

DigitalPreservationCoalition

What is digital storage?

Digital files are made up of a series of zeros and ones, known to as 'bits', which are arranged in sequences to represent different kinds of information, from text and images to geographic data, video or software. A key part of digital preservation is storing and providing access to those bits.

Digital storage provides a mechanism for recording digital information and retrieving it using a computer. Many kinds of digital storage are available, including optical technologies such as CDs, magnetic media such as hard disk drives, digital tape or flash storage. These storage technologies can be employed in different numbers and configurations to deliver different results. Storage can be as simple as a handheld memory stick, or as complicated as a set of servers with multiple hard drives that are installed in a purpose-built computer room and managed round the clock by dedicated staff.

What are the most common risks to digital storage?

Digital storage can be a powerful tool, but it can also be imperfect. These are just some of the risks facing any digital storage:

- Bit rot: storage devices can have a limited lifetime as short as 4-5 years, and faults may occur earlier which may damage your files.
- Obsolescence: Storage solutions can become obsolete if they are no longer produced and supported by manufacturers or become incompatible with the computer hardware they plug-in to. For example, new computers no longer have floppy disk drives and most do not have CD/DVD drives.
- Human error: whether intentional or otherwise, human error is one of the biggest threats to your files.
- Disaster: catastrophic disaster, whether it's theft, fire or flood could destroy all copies of your files kept in a single location.

This Digital Preservation Topical Note was produced with the kind support of the National Archives of Ireland

How can you address these risks and ensure you keep the bits?

Choosing the right storage to meet your needs requires a trade-off between managing storage risks and the cost of the solution. It's important to think about what type or what mix of storage to use, how many copies to keep, where to place those copies, and how everything will be checked and managed. These are some of the approaches that can help:

- *Fixity checking* can provide you with confidence that your files have not been damaged or changed over time.
- Having a managed back-up programme will help address data loss, this should utilise different media types to also avoid problems with obsolescence.
- At least one copy of data should be held in a different geographical location to allow retrieval in the event of a disaster.
- Actively managing the security of storage is key, this will usually include controlling who can access, change, and delete digital files.

Key Term: **Fixity or Integrity Checking**

A process of automatically verifying that all files are present and none have become damaged. A fixity checking tool will create a list or manifest of files. This will include generating a number (called a checksum or hash) that represents the structure of the file. By periodically regenerating the manifest and checksums, any lost, deleted or damaged files can be detected.

Who is responsible for digital storage?

Digital storage cannot be placed on a shelf and forgotten. It requires attention to make sure it is still functioning. It requires effort to check the data is still live and undamaged and it will need refreshing to new storage media at the end of its life. It is important that responsibility for this is defined so necessary tasks are completed.

Storage can be managed by individuals, but in most organisations this function is carried out by specialists. This may be the IT department of your organisation, or maybe a third party, as in the case of cloud storage. Although they may undertake most of the technical tasks, it is important that departments actively manage information about how they use storage. This can include recording information on how folders on a group drive are structured or documenting usernames and passwords for services that are used (e.g. DropBox). It is also useful to understand how storage is managed at your organisation and to be aware of any relevant data security or access policies.

Where should you keep your bits?

Hand held media such as CDs or memory sticks are quick and easy to use, but won't hold up at scale and are unreliable over the long term. Managed storage improves the reliability and dependability of your storage. Utilising dedicated storage servers with specialist staff to manage them is a far better approach for working at scale. Managed storage of this type provided by a third party is typically called the 'cloud'. Simple storage and file syncing from your desktop computer is also provided by services such as Dropbox. Larger scale solutions are also available, both for back-up or as full Digital Repository services.

It is important to be aware of the basics of how storage is managed at your organisation to ensure files are easy to find and are being backed-up. It is best to avoid using internal drives on laptops and PCs, instead storing files in the appropriate location in shared storage which will have a managed back-up program. This will also often bring efficiencies in working and reduce duplication.

For more information on Digital Preservation visit the DPC Website: https://www.dpconline.org