

# Digital Preservation Guidance

## How to look after your digital records

### ➤ Introduction

The vast majority of records and documents produced now are born-digital – they have always been digital - and these need to be looked after carefully to preserve them for the future. Digital records can be vitally important to maintain for either personal or business reasons, and could take a variety of forms (documents, photographs, artworks).

The work that goes into making sure digital files can survive and be accessed over time is termed **Digital Preservation**. This is an issue that affects most people in some way or another - we all rely on someone preserving information about us, and that information being reliable in the future, from health and pension records to online information and photos. Digital preservation can seem daunting, often due to the sheer amount of files produced, their intangible nature, and the risks associated with software obsolescence. Yet it's often just about using different techniques for the digital format but employing the same principles of caring for records that apply to paper.

This guidance has been produced to offer some basic advice and tips on how to care for and maintain your digital records, and the main areas you need to consider such as file formats and storage. It also includes a number of other resources that can help with digital preservation.

### ➤ What is a digital record?

There are many different types of digital record, the finished product maybe the most valuable and what you are most keen to ensure survives but you may want to consider any preparatory work that went into creating your artwork. This could include; storyboards, sketches, commissioning documents and correspondence with your clients. If you have taken a photograph or scanned something that was on paper that you have thrown away, this can also be considered a digital record.

---

### ➤ File Formats

File formats are often something that people worry about when they think about digital records and preserving them, but there are no right and wrong answers. You should choose the format which best suits your needs at the point of creation and there is no reason not to continue to preserve your digital record in the very same format in which it was created; indeed this offers you the best chance of maintaining a true representation of your work over time. Similarly the decision to use either lossy or lossless compression should be dependent on the context of your records and how they were created and used; it is rarely as simple as lossless for preservation and lossy for access copies.

When sharing or disseminating your work however, it may make sense to normalise your formats. Some examples are as follows but there are many other formats available that people will have little trouble downloading and/or viewing.

**Images:** There are many popular and well known image formats e.g. JPEG is a good format for photographs whereas GIF is a good option for graphics; GIF can preserve transparency whereas JPEG cannot, conversely GIF can only handle 256 colours so is not as good for photographs as JPEG. PNG is a good option for all images and is lossless, but file sizes tend to be larger.

**Audio:** FLAC offers lossless audio quality at a reduced file size compared with uncompressed WAV though file sizes can still be on the large side, whereas MP3 is lossy, but offers much smaller file sizes.

**Video:** There are several different 'container' formats used to contain digital video with audio, such as AVI, MOV, WMV, FLV and MP4. MP4 is a ubiquitous and versatile video container format, commonly used for video content on the web and with good support across a large number of popular media players.

**Documents:** PDF has become the de facto format for publishing and sharing documents however opting for open document formats that can be viewed in MS Office, Open Office, Libre Office etc, is also a viable option.

### ➤ **Managing file formats over time**

The National Archives maintains a registry of file formats which is utilised by our file format identification tool DROID, this tool can be freely downloaded from our website and used to scan files in order to identify what format they are. Once you are clear about what formats you have in your collection use the web to search for information about those file formats e.g. what software can be used to open them, especially if you are unable to do so with what you currently have installed. If you find you are using a particularly old version of a file format and you are worried about backwards compatibility with current software you may want to migrate to a newer version, this can often be as simple as opening and resaving your files.

Likewise when you upgrade your own software packages and you no longer have access to the version of the software that you originally used to create your files think about what this will mean with regards to accessing your files going forward. You can always check with the software vendor directly for the most accurate information about backwards compatibility for earlier file format versions, if you can't find this information yourself by searching the web. If you cannot open older files with your newer software you might consider saving additional copies of the files affected to a more generic format so that they can be accessed or edited in other software packages. As a general rule of thumb as long as you can still open your own files they are safe.

---

### ➤ **Storing your digital records**

If you don't have a digital repository in which to safely preserve your digital records or any organisational IT infrastructure at your disposal, the next best thing you can do is to keep multiple copies in different locations, a.k.a backing up.

There are many options available for storing digital records, when evaluating these different options you should keep in mind the volume and type of digital records you need to keep, what resources you have at your disposal and any privacy or security concerns you may have.

## Storage options to consider:

1. Your PC or laptop - this will likely be where you keep your main/working copy as it will be easy to access and update.
2. Email attachments – consider emailing a copy to yourself, if you have multiple email accounts even better. You may want to create folders within your inbox(es) so you can find and keep track of these backup copies more easily.
3. Removable media e.g. memory sticks or portable hard drives – these devices will connect via usb to your laptop or pc, they can be used as additional storage and most importantly can be kept in a separate location e.g. if you take your laptop home with you, leave your backup hard drive at work. If, in the unlikely and unfortunate event that one premises gets burgled or flooded, a copy of your digital records will still survive.
4. Cloud storage – as with email attachments why not use the web to store your digital records? There are numerous online storage options available e.g. Dropbox, Google Drive, Amazon Cloud Drive and Microsoft OneDrive. Many of these services provide an initial amount of storage for free and are also designed to enable sharing and collaborative work. Many services will sync with your various devices so you can access and update your documents ‘on the go’.

Keeping multiple copies of your digital records in different places may seem a simple way of ensuring their longevity; however it can quickly become unmanageable. **When keeping multiple backup copies in different locations consider the following:**

1. File naming conventions - be consistent and detailed, consider including a date and anything else that will help you to find and identify your records with ease.
2. Similarly, if you are keeping your files within a wider folder structure, think about what you are naming your folders and don't nest files unnecessarily deeply. You could create folders per project, by year and month, by client or any other logical system. Think carefully about how you to store your records, as you do not want to create a time consuming structure to maintain and consider recording file paths so you can find your files easily within their wider structure.
3. If you update/edit one copy, remember to keep your other copies in sync. If you want to keep earlier versions make sure the versioning is clear in the name of the file e.g. by appending v.1.0, v. 1.1 etc. and indicate where a file is the final version by using terms such as ‘FINAL’ or ‘PUBLISHED’.

## ➤ Managing your storage over time

Whatever storage options you choose initially, continue to review the situation going forward.

Are you running out of free space online or on a hard drive?

Have the terms of use changed for your cloud storage or email services? Are they still compatible with your needs? What would you do in the event that one of these services went down either temporarily or permanently?

The shelf life of hard drives and memory sticks is notoriously hard to predict; they can fail randomly at any time before the stage at which they start to actually wear out, and the point at which they begin to wear out depends upon how regularly they are in use.

Whatever combination of storage options you use make sure it is not too onerous; if backing up is too complicated or time consuming, the chances are you won't do it!

---

## ➤ Fixity and Virus checking

Just storing your records cannot be considered active digital preservation, to achieve this you need to regularly check the integrity of your digital records. The simplest way to do this is to virus and fixity-check your records.

Free anti-virus software is readily available, make sure you connect to the web and don't ignore prompts to update your anti-virus software. You can set up your computer to automatically run background virus scans but you can also scan on demand, usually just by right clicking and selecting your virus scanning software. The importance of protecting your digital records from becoming compromised by viruses cannot be underestimated.

Fixity checking is ever so slightly more time consuming but it is the only way you can be sure your digital records are not changing over time without your explicit intervention. Fixity checking involves the generation of checksums or hashes, which are unique strings of letters and numbers produced by particular checksum algorithms. Checksumming software scans each file byte by byte in order to produce a unique hash. Once you have generated a checksum for a file using a particular algorithm e.g. SHA 1 or MD 5, you can regenerate the checksum using the same algorithm at any point in the future; if the checksums match your file is still the same, if they don't match something has changed and you will need to investigate further. As with anti-virus software there is free checksumming software available on the web right now for you to download and start using e.g. Gizmo. The National Archives DROID tool, which is free and open source, also has a tick box option to produce checksums.

If you regularly virus and fixity check your records you will also be regularly accessing your records so you will notice, for example, if removable media is failing. If possible it is advisable to schedule this in at least every quarter.

---

## ➤ Metadata

Throughout this document it has been suggested that you might want to record certain information about your digital records to help you manage and most importantly find them, this information is commonly referred to as metadata.

You need only create a simple document, probably a spreadsheet, containing filenames and their location e.g. URL or filepath, their corresponding checksum and the date when anti-virus scans and checksum verification were last done. From there you can record as much other information as you like and is suitable to your needs e.g. dates, the software it was created in, copyright information. If you have a cataloguing schema for paper records it may well be applicable to your digital records. Likewise if you are using cataloguing software for your paper collection it may work for digital records or there may be a specific module available for digital records which you could upgrade to include. Although it may seem daunting, managing digital records requires exactly the same principles as looking after paper records – if you file things in an orderly fashion and name them well you will be well on your way.

---

## ➤ Useful resources for looking after arts records and archives

Many useful resources online – and they are free:

### **The National Archives**

Identifying file formats: DROID <http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/file-profiling-tool-droid/>

Archive Principles and Practice: an introduction to archives for non-archivists  
<http://www.nationalarchives.gov.uk/documents/information-management/archive-principles-and-practice-an-introduction-to-archives-for-non-archivists.pdf>

Digital Collections <http://www.nationalarchives.gov.uk/archives-sector/digital-collections.htm>

Digital Preservation FAQs <http://www.nationalarchives.gov.uk/information-management/projects-and-work/digital-preservation-faqs.htm>

### **OCLC Research**

You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media <http://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf>

### **Council on Library and Information Resources (CLIR)**

Born digital: Guidance for Donors, Dealers and Archival Repositories  
<http://www.clir.org/pubs/reports/pub159>

### **University of London Computer Centre (ULCC)**

Digital Preservation Training Programme (DPTP)  
<http://www.dptp.org/>

### **Digital Preservation Coalition (DPC)**

Digital Preservation Planning case study, Ed May, 2013  
[http://www.dpconline.org/component/docman/doc\\_download/548-5-a-ed-fay](http://www.dpconline.org/component/docman/doc_download/548-5-a-ed-fay)

Our digital memory accessible tomorrow:  
<http://www.dpconline.org/>

Digital Preservation handbook:  
<http://www.dpconline.org/advice/preservationhandbook>  
<http://www.dpconline.org/advice/preservationhandbook/decision-tree>

### **Library of Congress (LOC)**

Personal Archiving: preserving your digital memories  
<http://www.digitalpreservation.gov/personalarchiving/>

### **Prestospace**

PrestoSpace: working to improve and standardise practice for audio-visual material, Preservation Guide,  
<http://wiki.prestospace.org/pmwiki.php?n=Main.NavigationGuide>