

Case study and recommended procedure: Digital Preservation in Small Museums (UK)

Imogen Smith, 2023

Introduction

About this study

Due to the lack of a specified minimum standard in, the purpose of this case study is to compile and share some basic guidance and provide our own minimum standard of practice for small museums and similar organisations. Similarly, the Digital Preservation Coalition (DPC) - the global authority on digital preservation education - say “do not try to aim for “best”, aim for “good enough”, and expand upon this in many days’ worth of reading and training materials. Therefore, we aim to give smaller collecting institutions a voice on digital preservation, a matter most often discussed by larger organisations.

In this study, we share some of the problems we encountered and overcame and offer a Recommended Procedure with detailed steps on how to achieve a minimum standard of digital preservation to contribute to ACE Museum Accreditation and make lighter work of future tasks. Our aim is not to supersede training and resources provided by the DPC or other authoritative sources such as the [Digital Preservation recommendations for small Museums from the Canadian Information Network](#), but to distil and accompany their provision into a more consumable format for organisations similar to us, working with UK standards for collections management.

We documented our progress, training and reading while exploring digital preservation for our funded project: Bath’s Underground History Unlocked. The project involved the digitisation of an extensive, previously inaccessible archive of Digital Betamax videos. The lack of accessibility and reliable longevity of the tapes, and the small size of our team, were strong motivation for developing the easiest and most future-proof digital preservation workflow for us as soon as possible. This came with some challenges and lots of learning that we hope will be useful for similar organisations.

About the Museum of Bath Stone

There are a few important things to share about the Museum of Bath Stone for the purpose of this study. If you would like to know more about the museum, beyond this, please visit our website.

Founded upon a 21st century engineering project during which archaeologists produced an extensive record of Combe Down’s lost underground landscape, the museum cares for a unique collection. Almost entirely digital, the records capture 18th and 19th century quarrying, where very little other material recording this history has survived. The collection also includes objects (archaeological/geological items) and analogue archives (photographs and paper records produced during the Stone Mines Stabilisation Project). We aim to manage this collection according to the Spectrum 5.0 Standard in order to achieve and maintain accredited status. This has been managed by two paid staff, assisted by a small team of volunteers and a small board of Trustees but we are growing our team.

As is common among similar organisations, many operations will need to be carried out by volunteers without digital preservation knowledge or volunteer teams without IT specialists. Therefore, woven into this study you will find a set of **golden rules** to help working at this museum or, a comparable organisation with similar goals and characteristics.

Case study

This study is not going to include basic information about digital preservation, such as a definition for the phrase or any other jargon used. If you are seeking for this sort of information, it is useful to take a look at the DPC website: <https://www.dpconline.org/digipres/what-is-digipres> and resources on this Collections Trust webpage: <https://collectionstrust.org.uk/digital-isnt-different/digitisation-resources/digitisation-future-proof/>. This case study instead covers the main challenges we encountered and solutions that influenced our digital preservation procedure.

Golden rule: Complete the Novice to Know How training first

If there's anything that's worth doing to help you on the way to a minimum standard of digital preservation, it is this training. You will need to dedicate 1-3 workdays to it and, while you can [download the entire course in PDF form](#), it is useful to record your immediate reactions while learning in notes. It is free, easy to sign up and not too much hassle to get an extension for accessing the course if you need it and the team behind it are friendly, responsive and helpful: <https://www.dpconline.org/digipres/train-your-staff/n2kh-online-training>.

The Museum of Bath Stone already cared for a digital archive before the beginning of the digitisation project and, while there was not yet a documentation or preservation procedure in place yet, the collection had been kept tidy and searchable meaning items could easily be found, and their provenance was known. However, the presence of existing collections still produced a couple of challenges. Firstly, it meant that navigating digital preservation guidance felt more time consuming than anticipated with most guidance on the process of 'Ingest' seeming irrelevant. You may ask: why read guides on appraising, acquiring and ingesting digital archives if your museum computer is already full of it? It was easy to feel similarly about doom and gloom copyright and rights management resources. However, our museum is open to acquiring more material in the future, so we have established an ingest workflow *now* ready for any new incoming digital acquisitions.

Golden Rule – Manage expectations

This applies to you, colleagues and co-workers (if you have any), trustees, volunteers and, vitally, potential donors and lenders who want to contribute to the collection. It is unlikely that these people will have any experience of the challenges surrounding digital preservation. This is one of the reasons why it is useful to not be in a hurry to use digital archives. Even if a donor really wants to see their material used in an exhibition as soon as possible, that is not a reason for you to bypass digital preservation tasks. If you do, it might be harder to achieve the minimum standard of preservation that you are aiming for later on.

On fixity checking, not even N2KH provides a recommended procedure, it all depends so much on the scale and type of organisation implementing digital preservation. The DPC provides a brilliant resource for as many people as possible, not guidance tailored to different needs. Fixity and checking and replacing files were the part of digital preservation that was the steepest learning curve but the solution was to trial and get used to the software needed.

It can also be tricky deciding which copies of the archive to use for access and which to leave undisturbed as back-ups. The archive is too large to store locally on the PCs themselves, so we needed an alternative location for access. Making a selection meant balancing risk, against ease of access. The first step was deciding who is going to be accessing the collection, how often and for what purposes.

We decided that, since files in Dropbox are at least risk of loss or damage, they should be the 'master' version, accessed as little as possible, and the on-site hard drive copy should be for general access and use.

Golden rule - Plan your work so you don't need to rush

The best thing you can do for the sake of preserving your digital material is take a long-term view of your digital preservation goals, assess what is achievable soon, how you might be able to improve later on and how you can take steps in the present to make future tasks easier. During first steps, try to avoid being in a hurry to make use of material as this can produce more problems than it solves.

Without IT expertise on hand, it was difficult to bridge the gap between knowing the functionality and limitations of a piece of software/hardware and being sure it will do the job it required to do. For example, selecting a cloud storage provider felt like going around in circles! We have chosen a well-known and trustworthy provider, Dropbox, and if we encounter issues, or discover a better alternative, we can change over. We did not find that choosing external hard-drive products was as much of a challenge, but we did bump into a few difficulties with formatting file systems.

The bigger challenge is still on the horizon: the recommended procedure works for us but if something goes wrong, for example a file format becomes inaccessible while in storage, we will need to look for further help to solve the problem. The only solution is to be prepared to do further training and make sure any policies, plans and guides reflect that the procedure we set up is going to need updating.

Be prepared to ask for help

If you are lucky enough to be working or volunteering with someone who has experience of digital preservation, they can probably help you. If not, there are lots of people who work in heritage who can and are willing to help you if you reach out. In particular, if you share our goal of achieving accreditation, your local Museum Development Officer(s) are likely to be able to put you in touch with people who can help. While they may not have experience with collections management, it can be useful to speak to IT specialists if you know or work with any. This is another reason why it's best not to rush – you can get the help and guidance you need but it may take a week or two to receive replies to questions.

At the Museum of Bath Stone, we strive to reduce our environmental impact wherever possible. We were aware that digital preservation procedures would produce waste and pollution of some kind – it's unavoidable – and were especially concerned with finding a 'green' cloud storage option. However this matter was not explored by N2KH and does not appear much in other resources except [this DPC page](#) (which ought to be read by anyone starting out with preserving a digital collection). In the end, other factors had to take precedence over this, such as cost and usability, a reflection on our small team and limited budgets.

With a view to opening the door to necessary updates, I encourage any reader to get in touch with recommended improvements which, as long as they are achievable, we will implement at the Museum of Bath Stone. Similarly, if you need more help than this study provides, you can contact us and we will do our best to signpost you to relevant guidance. Please send comments, questions, advice and guidance to info@museumofbathstone.org.

Recommended Procedure

Appraisal & Ingest

1. Complete the **Novice to Know How** training.
2. Have a detailed conversation with any potential donor about their material so that you can answer the following questions confidently:
 - How much material (individual items and total data) is in the potential donation?
 - What type of material is it, what file formats does it consist of?
 - How useful and how interesting could it be for our audience?
 - How relevant is it to our collections policies and existing collection?
3. Keep 1 or 2 spare external hard drives to hand for transferring donated files
 - Make sure they are empty or quickly emptiable and contain no important backups
 - They ought to be replaced after 3 years (this is when they start to degrade and put your digital assets at risk)
 - Name them something logical and label them with their name
 - Example: “MoBS 8TB 1” and “MoBS 8TB 2” are our two large backup drives which we keep in the original packaging that is clearly labelled
 - Watch out for how they are formatted as this affects whether different types of computers can read/write to a drive. Formatting drives to the ‘exFAT’ format means they will work for Mac and Windows PCs. Formatting the drive can put your digital assets at risk so it’s best to get it right when they are new so you don’t need to change it later.
 - Help if you’re using a Mac:
<https://www.youtube.com/watch?v=k1bBiEP34TM>
 - Help if you’re using a Windows PC:
<https://www.thewindowsclub.com/format-a-drive-in-exfat-windows-and-mac>
4. If the potential donation contains any file formats that you are not familiar with, research them **before** taking on the material
 - Example: ‘.psd’ is a proprietary image file format created when using Adobe Photoshop and unless you have that software, you will not be able to access files of this format.
 - Help on identifying and ‘normalising’ unusual formats
 - PRONOM technical registry: <https://www.nationalarchives.gov.uk/pronom/> (click ‘Help’ then ‘Overview’ for a basic introduction)
 - DROID file format identification tool: <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/> (you will come across this if you do the recommended training)
 - Be aware: disposing of material that you don’t understand because it is in an unfamiliar format can endanger access to other material as some files rely on other files in order to be accessible. This is also why it’s important not to be in a hurry to take on and use new digital donations!

5. Ensure you have sufficient anti-virus protection
 - Due to the nature of material and donors involved with small museums and archives, you are unlikely to need any more protection than whatever your computer comes with as standard.
 - Windows 10 and newer PCs come with 'Defender' which you can use to scan external devices as well as locally stored files
 - Mac computers are built to be very secure
 - It is better to be safe than sorry, if you are unsure that you have an adequate level of protection against computer viruses/'malware' for the type of donations you take on, you will need to find some funds to pay for anti-virus software such as McAfee or Norton or research the cheapest option.
 - DPC and TNA recommend creating a dedicated Transfer PC, but smaller organisations are unlikely to have the resources to do so. Here, we are lucky enough to have 2 PCs and a laptop. When taking on new material from donors, we use a computer that does not have anything important stored on it without backups and while copying new digital archives, we keep it disconnected from the internet and other devices in the Museum.

6. When you copy files from the donor's storage onto yours, beware of losing metadata that comes as part of your files
 - Relevant DPC section <https://www.dpconline.org/handbook/tool-demos/command-line-intro/command-copy>
 - Decide what metadata you want to preserve (we mainly just have image files in the digital archive so want to retain 'timestamps' like 'date created' and 'date modified' for folders and files, some digital cameras also record an author and location which can be useful).
 - **Test your copying method.** Just using Ctrl.+C and Ctrl.+P (for Mac its cmd+C and cmd+P) to copy files often does not retain folder metadata and this is similar for some copying/'synchronisation' software such as FreeFileSync.
 - If you are using a Windows PC, you can use your N2KH training to use Robocopy. If you are using a Mac, it's more of a challenge to preserve the metadata, you can try compressing files into a .zip but this may put different types of files at risk so **test it**.

7. Create a destination folder for newly acquired collections
 - Name it something unique, helpful and searchable (the easiest thing to do is use whatever ID and Title you were going to give the collection in the inventory/catalogue)
 - Store it in a file structure where it is easy to find and has the correct level of accessibility for the people who need to see it and is secure from those who ought not to be able to change or move it.

8. Create a simple, easy to use Digital Assets Register
 - DPC guidance: <https://www.dpconline.org/docs/miscellaneous/training/1679-digital-asset-registers-getting-started/file>
 - Example: see Appendix 1
 - Use this to record top-level collection records, just the main 'destination' folder or a collection as mentioned above, not sub-folders or individual files.

9. Create a Verifiable File Manifest

- Use your DPC training notes
- This won't come into play until you are in the file checking stage, but you need to create it and fill it in as soon as possible after taking on new material.
- You can use DROID to create the vital 'checksums' that you will use to verify the preservation of your digital assets in the future. Alternatively, you can use Checksum by Corz.
- Example, our File Manifest:

File path	Catalogue ID (where applicable)	Title	Original Checksum	Fixity Check
-----------	------------------------------------	-------	-------------------	--------------

10. Create a basic guide with checklists and links to further reading so that your successors know how to carry out the procedure you have established.

Storage and back-ups

1. Ensure you know the total amount of data you need to be able to store, to start with we needed 2TB. Try to estimate how much that might increase over 3 or 5 years and, if possible, ring-fence funds for expanding storage capacity.
2. Create back-up copies of your entire archive
 - Using the copying guidance from the Ingest section, make 2 identical copies and store them using the **3-2-1 rule** (standard practice for anyone working with digital data)
 - In summary: keep **3** copies, use **2** different types of storage, keep **1** copy in a different location. You can read more about the 3-2-1 rule here: <https://www.seagate.com/gb/en/blog/what-is-a-3-2-1-backup-strategy/>
 - To make the best use of this rule, I recommend:
 - 1 copy on a hard-drive kept at the Museum
 - 1 copy on a hard-drive kept at the Museum, or if possible, at another location such as a neighbouring Museum
 - 1 copy in Cloud storage
 - N2KH encourages this, stating "don't put all your eggs in one basket"!
3. Choose your storage hardware
 - Vitally, note that optical discs like CDs and DVDs are not a reliable enough storage media and are largely obsolete now, computers rarely come with disc drives.
 - Western Digital have a wide range of external hard drives: <https://www.westerndigital.com/en-gb/products/hdd/external-hdd>
 - Some popular alternatives are reviewed here: <https://uk.pcmag.com/hard-drives/7890/the-best-external-hard-drives>
 - If possible, keep archive back-ups under lock and key in a fire-proof cabinet.
4. Choose your cloud storage
 - If you have no IT expertise on hand, you will need to find a user-friendly storage product. Popular, user-friendly storage is provided by Google, Microsoft, Amazon and many other large companies

- We chose to use Dropbox because it is sufficiently user-friendly, we can afford the subscription (there are several different Plans, all at least £10 per month), when uploading, it doesn't change the metadata we want to preserve
 - If there's a free trial, use it and try uploading/downloading/previewing files and watch out for packages that have lots of features you won't use
 - Some alternatives are <https://icedrive.net/> and <https://file.io/>
 - Ensure you save the password for the account for your chosen cloud storage website – it would be tragic for your successor to not be able to access the Cloud back-up!
5. Decide which copy is your access copy and which back-up is the most secure and reliable. The access copy will be at the greatest risk of data loss as the people accessing the files may accidentally move or delete items. The reliable back-up is at least risk and can be used later to replace lost files.
- Our 'access' copy is on one of the physical large hard drives which we keep locked in a cabinet in the Museum.
 - Our back-up is also on a large physical hard drive is kept off-site
 - Our most secure back-up is in Dropbox. Cloud storage providers create several copies of your files, often stored in servers in different countries, making your files much more secure here than on physical storage media.
6. When you have decided where everything is going to be stored, create a spreadsheet just for digital archive storage locations, ours is a sheet in our Digital Assets Register workbook. This means that you can always keep track of your back-ups. For example:

Container ID	Location	Format	Capacity	Access level	Dependencies	Date acquired	Replace/update date
OneDrive - Combe Down Stone Legacy Trust	Cloud (accessible via Museum PCs)	Cloud	1TB	Staff, Trustees & trained volunteers only	Internet connection, Microsoft Office subscription	2022-11-01	re-subscribe to Microsoft Office package 2022-11-1

7. Use Cloud storage to back up a copy of your Digital Assets Register, File Manifest and, if you want, any catalogue report files. Again, make sure these are stored in folders with sensible names!
8. Create a basic guide with checklists and links to further reading so that your successors know how to carry out the procedure you have established.

Checking and replacing files

1. Check your archive periodically for 'fixity'
- If you have completed the Novice to Know How Training, you will know that just backing-up your files is not enough, you must periodically check that they are unchanged. Our policy is to fixity-check the digital collection annually
 - Once a year, we check files on the physical hard drives only since Cloud storage providers usually carry out all the necessary backing-up, checking

and replacing from their end. Plus, to check files in Cloud storage we would have to download the entire massive amount of data from the cloud.

- We aim to implement a system of checking the Cloud back-up small parts at a time once we have recruited a dedicated digital collections volunteer.
- Working through the entire digital archive, use DROID to produce a checksum for each collection and save this in a CSV file (save this somewhere findable with a sensible, future-proof name). Alternatively, use Checksum by Corz as per N2KH training.
 - Copy and paste the data from the checksum column (the column title will be the type of checksum such as 'SHA256') into your File Manifest in the 'Fixity Check' column.
 - If you're using Microsoft Excel, use 'Conditional Formatting' to highlight where the checksums do not match, if the new checksum for a file does not match the original one, it means something about that file has changed.
 - Here's how to use 'Conditional Formatting' in Excel:
<https://www.ablebits.com/office-addins-blog/how-to-highlight-duplicates-excel/>
 - You may wish to save this version of the file manifest as evidence that you carried out your periodical check.

2. Replace problem files

- If your check showed up files that have been changed, try opening them to check why their checksum might have changed.
- If you can't even open the file, download a copy from your Cloud storage and try that. If even this won't open you've got a bigger problem with multiple possible causes such as:
 - You don't have the right software needed to access the file
 - The file itself was already corrupted before being backed-up
 - The file has been changed by someone accessing it (e.g., typing in a word document)
 - There's something wrong with both the file and the computer on which you are trying to access it
 - There is a problem with your hard drive
- You will likely need to seek advice from an IT professional to fix the problem
- If your Cloud stored copy is fine, copy it onto the hard drive to replace the problem copy
- Once you have replaced problem files it might be worth running the same check again if you have time. If the files you replaced are still showing up as changed and you can't access them, I recommend seeking advice from an IT specialist and/or local digital preservation specialist such as at a record office, university or large museum.

3. Improve your process

- Once you have gone through the whole process once, you can make a note of how to make it quicker and easier in 6 months' time when you do it again.
- Create an easy-to-read guide on how to carry out a fixity check at your organisation. You never know, you might not be around the next time it needs to be done.
 - Make sure that before you do any of this, you are happy with your folder names and structure. If any of it is changed after you put all that data in the

File Manifest, you risk making a big headache for yourself when trying to compare checksums later on!

Appendix 1: Digital Assets Register column headings

	A	B	C	D	E
1	Essentials				
2	Collection Title	Responsible person	Accession ID (if applicable)	Size on disk	Size

F	G	H	I
Location			
Location (1)	Location (2)	Location (3)	Folder file path

J	K	L
Use		
Catalogue ID (if accessioned)	Retention	Rights status

M	N	O
Preservation		
Formats included	Preservation risk note	Dependencies

P
Note
Note

Appendix 2: Further Reading

- **Canadian Information Network:** [Digital Preservation Recommendations for small museums](#)
- [The DPC Digital Preservation Handbook](#)
- [Novice to Know How free training programme](#)
- The National Archives: [Digital Preservation Workflows](#)
- Digital Preservation Coalition: Digital Preservation [Topical Note 2: Back-Up & Storage](#)
- Digital Preservation Coalition: Digital Preservation [Topical Note 3: Digital Images](#)
- Digital Preservation Coalition: [Technology Watch](#) Guidance Note, 2021: Preserving Moving Images, Data Type Series from Artefactual Systems & DPC
- DPC blog article: [‘Memories from the Anthropocene: digital preservation in a time of climate crisis’ by William Killbride](#)
- DPC blog article: [‘It’s not easy being green: Evaluating the impact of digital preservation’ by Elizabeth Thurlow](#)
- DPC & TNA: [‘Digital preservation for small business: an introductory guide’](#)
- Install ‘Checksum’ by Corz: <https://corz.org/windows/software/checksum/>
- Install DROID: <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>