

Digital Preservation **Handbook**

Getting Started



Illustrations by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Gold sponsor



Silver sponsors



Bronze sponsors



Reusing this information

You may re-use this material in English (not including logos) with required acknowledgements free of charge in any format or medium. See [How to use the Handbook](#) for full details of licences and acknowledgements for re-use.

For permission for translation into other languages email: handbook@dpconline.org

Please use this form of citation for the Handbook: Digital Preservation Handbook, 2nd Edition, <http://handbook.dpconline.org/>, Digital Preservation Coalition © 2015.

Contents

Getting Started 4

 Introduction..... 4

 Get to Know Your Organisation and Your Data..... 4

 Where next? 7

 Resources 8

 Case studies..... 10

 References..... 11

Getting Started



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

This section is for you if you have yet to start digital preservation or have just begun to do so. It provides a rapid introduction to a number of approaches that will support you in getting started, moving towards using other sections of the Handbook, and in building confidence and skills.

The section developed out of the "Getting Started in Digital Preservation" workshops run by the Digital Preservation Coalition. It supports 'learning by doing' and assumes a minimal level of prior knowledge. As you progress you will benefit from dipping into the resources and case studies, other topics and sections in the Handbook, and the [Glossary](#) for any unfamiliar terms.

Digital preservation can seem like a daunting prospect. It can help to map out the skills and resources you have and the materials you want to preserve. That way you start with what you know rather than what you don't. The first step in digital preservation is almost always to undertake a rapid assessment. This will have two or three components:

- knowing the practical capacity of your organisation;
- understanding the organisation's goals and missions; and
- knowing a little about the digital materials in question.

Get to Know Your Organisation and Your Data

Creating a Digital Asset Register

As part of a rapid assessment: it is vital to understand the nature and extent of your digital collections. A digital asset register will be incredibly useful for assessing the extent and significance of the collection, identifying priorities and planning digital preservation actions. A high-level assessment of the collection will help with more detailed mapping later: a comprehensive and detailed audit could be time consuming. So the advice in early stages is to keep the asset register simple. Ask the following questions:

- What is the subject of the collection?

- Where does it come from and what is its function?
- Where is it stored and what kinds of media are used?
- Why is it being retained?
- Who is responsible for it; who are the users; who are the subjects of the data?
- How is the data accessed?
- How is the data likely to change and grow in the near future?

Assessing Your Organisation's Readiness

Organisational maturity is another factor to consider. The National Digital Stewardship Alliance ([NDSA, 2013](#)) in the USA has recommended a simple 4 level model to help organisations understand and improve their technical capacity in digital preservation. The four levels are:

- Level 1 - protect your data
- Level 2 - know your data
- Level 3 - monitor your data
- Level 4 - repair your data

These 'Levels of Preservation' are intended to be progressive, and are used to measure maturity against four components: storage, file fixity, information security, metadata, and file formats. An organisation's capacity to undertake digital preservation is indicated by its maturity level across these five components. More comprehensive maturity models are available, such as the Digital Preservation Capability Maturity Model ([Dollar and Ashley, 2014](#)), if a more well-rounded exploration of organisational maturity is required.

First Steps to Securing Your Data

This section provides an overview of initial actions to secure your data once you have assessed your organisation's readiness and compiled basic information about your data. The following steps are essential to ensuring a minimum level of preservation when a new collection of digital material is received. This is typically referred to as bit preservation. Quite literally, preserving the streams of binary digits, or bits, that make up your digital files (without preserving the means to decode the bits into meaningful information).

Prompt check in on receipt

When a new collection of digital material is received from a supplier it is essential to ensure that what has been received is what is expected. Depending on the source of the material, it may be possible to request new copies of missing or poor quality files. These checks are made and any replacement requests submitted, the greater the likelihood of successful resolution.

Key tasks include:

- Scan for viruses and malware to make sure there are no unwanted surprises in the collection. Perhaps keep the collection 'in quarantine' until you have checked it.
- Check all expected files are present. If the material is accompanied by a manifest, check the files against it

- Open a random selection of files to verify their integrity and/or expected quality levels
- Promptly request replacements for any damaged or missing files, where possible

Create a Verifiable File List

In order to check over time that your digital files are being preserved, it is first necessary to record exactly what files are in your possession. It is, therefore, important to create a verifiable list of files in each collection. These lists should likely contain information such as file names, locations and sizes, format types and checksums. A checksum is a short alphanumeric string that represents the contents of a file which acts as a 'digital fingerprint' allowing comparison over time. Once the list has been created, it is a simple process to verify that all files are present and undamaged, at any point in the future (see fixity checking below).

Various software [tools](#) can be used to automatically generate this data; these are commonly referred to as characterisation tools. For example, you may wish to use The National Archives [PRONOM](#) (a register of file formats and their behaviours) and [DROID](#) (a tool that uses PRONOM to analyse the files on a system). Having a list of the file formats, versions and quantities in your collections will help you make a case to senior management for the support and resources that will be needed to do the job properly and sustainably. This information can also be used to update and enrich your digital asset register. The range of formats in use should be consolidated to minimize duplication and eliminate problem formats. This process is known as normalization.

Key tasks:

- Generate a verifiable file list
- Update digital asset register

Stabilise your files: make copies

No matter how good your digital storage, your digital material will always be at risk of damage, decay or accidental deletion. Making more than one copy of your digital materials and utilising more than one type of storage solution mitigates a variety of digital preservation risks.

Key tasks include:

- Keep (at least) one copy easily accessible on non-removable disk. You will need to regularly revisit your material to ensure its fixity, so keeping it accessible will make this easier
- Make (at least) one additional copy, if necessary on a less accessible, but cheaper storage medium such as tape
- Keep one copy in a different geographical location to the others to mitigate against disaster

Revisit and inspect: Fixity checking

By revisiting your digital materials on a regular basis (e.g. every 6 months) you can ensure that no damage or accidental loss has occurred. If it has, you can recover problematic files from the copies or backups you have made previously. Future fixity checks will generate new digital fingerprints (or checksums) for the files in your collections. If they do not match the ones originally created, bit loss or damage has occurred.

Key tasks include:

- Revisit your collection on a frequent basis, recalculate the checksums, identify files that have become damaged

- Retrieve copies of damaged files and repair as necessary
- Perform test recoveries of data backed up by third party services, to ensure backups are being performed as agreed

Document your processes

From the outset of creating a digital collection, it is important to document as much as possible about a collection's assets, the tools and workflows. This documentation is an important component of [technical and descriptive Metadata](#). It is necessary to retain this information for the purposes of longevity. As, with any project, staff retention can be an issue. If staff leave they often take essential knowledge and skillsets.

Where next?

Having taken the first steps in digital preservation, where do you go next? This will obviously depend on your own requirements and priorities, but this table provides a number of suggestions and other sections of the Handbook will help you move forward with them:

Next steps
Develop advocacy and outreach, an understanding of risk , the business case, costs, benefits and impact
Establish an organisational preservation strategy and policies . As well as ensuring a consistent approach to preservation it can be a useful tool to achieve buy in across an organisation and in particular with senior management
Establish a digital repository. Technical solutions and tools either on local IT infrastructure or offered as a cloud service will help you understand, manage and preserve your digital material for the long term
Establish your long-term storage , preservation planning and action
Revisit and expand your collection audits: <ul style="list-style-type: none"> • Characterise priority collections in more detail • Periodically update collection audits as required
Establish a digital preservation working party. Effective digital preservation often requires buy in across many departments within an organisation. A representative working party can be vital in making coordinated steps forward
Build the necessary staff training and development and skill sets
Establish a professional network and collaborations . Join a digital preservation membership organisation such as the Digital Preservation Coalition

Keep up to date with new developments:

- Email lists for digital preservation include the [digital preservation](#) announcement list on JiscMail, and the USA-focussed [digipres](#) list
- A weekly [DP News](#) blog selects recent tweets and news links on digital preservation
- Journals with a digital preservation focus include: [International Journal of Digital Curation](#), and [D-Lib](#)
- Events with a digital preservation focus include [iPRES](#), and [PASIG](#)
- The [Digital Preservation Coalition organises briefing days on particular digital preservation topics](#)

Resources



A Preservation Primer

<http://knconsultants.org/a-preservation-primer/>

This clear practical short primer on preservation for beginners was written by staff at Portico. It summarizes the issues and outlines various short and long-term preservation options that an organization might take to begin planning for long-term digital preservation of its content, beginning with near-term protection and concluding with full preservation and long-term protection. (83 pages).

Don't Panic: The Archivist's Guide to Digital Preservation

http://www.wyjs.org.uk/documents/archives/dont_panic_digital_preservation_first_steps_guide.pdf

A practical and concise guide produced in 2011 by Stefanie Davidson at the West Yorkshire Archive Service. It is intended to act as a signpost to assist in taking the first steps in understanding some of the issues involved rather than a comprehensive guide and an introduction to the topic to help you find your feet. (8 pages).

Putting Parsimonious Preservation into Practice

<http://www.nationalarchives.gov.uk/documents/information-management/parsimonious-preservation-in-practice.pdf>

The principle of Parsimonious Preservation was originally developed in 2009 at The National Archives in the UK as an approach for small or medium sized institutions to permit them to begin work on digital preservation but is also practical for large scale institutions. It now underpins advice and guidance given to the UK archive sector on digital preservation. (11 pages).



Community Owned digital Preservation Tool Registry COPTR

http://coptr.digipres.org/Main_Page

COPTR describes tools useful for long term digital preservation and acts primarily as a finding and evaluation tool to help practitioners find the tools they need to preserve digital data. COPTR aims to collate the knowledge of the digital preservation community on preservation tools in one place. It was initially populated with data from registries run by the COPTR partner organisations, including those maintained by the Digital Curation Centre, the Digital Curation Exchange, National Digital Stewardship Alliance, the Open Preservation Foundation, and Preserving digital Objects With Restricted Resources project (POWRR). COPTR captures basic, factual details about a tool, what it does, how to find more information (relevant URLs) and references to user experiences with the tool. The scope is a broad interpretation of the term "digital preservation". In other words, if a tool is useful in performing a digital preservation function such as those described in the OAIS model or the DCC lifecycle model, then it's within scope of this registry



DPC Getting Started in Digital Preservation Workshops

<http://www.dpconline.org/events>

The DPC Getting Started in Digital Preservation workshops are events designed to raise awareness of digital preservation issues, increase involvement with digital preservation activities and sign-post the support and resources available to help you on your way. They provide an introduction to digital preservation, build an understanding of the risks to digital materials, include practical sessions to help you apply digital preservation planning and tools, and feature speakers sharing their own experience of putting digital preservation into practice. You can find details of forthcoming workshops and the programmes and speaker presentations at previous workshops on the DPC events page.

Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Solutions

<http://www.dpworkshop.org/>

An excellent free online tutorial that introduces you to the basic tenets of digital preservation. It is particularly geared toward librarians, archivists, curators, managers, and technical specialists. It includes definitions, key concepts, practical advice, exercises, and up-to-date references. The tutorial is available in English, French, and Italian.

Canadian Heritage Information Network (CHIN) Digital Preservation Toolkit

http://www.rcip-chin.gc.ca/carrefour-du-savoir-knowledge-exchange/outils_preservation_numerique-digital_preservation_toolkit-eng.jsp

CHIN has released a suite of documents to identify digital material found in museums, the potential risk and impact of lost material, and how to get started in the development of Preservation Policies, Plans and Procedures. The toolkit includes a Digital Preservation Inventory Template, Digital Preservation Policy Framework Development Guideline, Decision Trees, and a Digital Preservation Plan Framework.

Digital Preservation 101, or, How to Keep Bits for Centuries

http://scholar.harvard.edu/jcs/presentations/dhttp://handbook.dpconline.org/administrator/index.php?option=com_content&view=article&layout=edit&id=86igital-preservation-101-or-how-keep-bits-centuries

This 2015 presentation by Julie Swierczek Digital Asset Manager and Digital Archivist at Harvard University Art Museums is a good advocacy for and explanation of, digital preservation to other non-specialist institutional colleagues including "why archivists cry themselves to sleep at night when the general public conflates archives with backup copies of data" (142 slides but many are images with good slide notes that make this easily understandable).

The National Archives Digital Continuity Guidance

<http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/step-by-step-guidance/step-2/>

This guidance takes you through the process of creating an information asset register, and includes a template in Excel spreadsheet format. The register can be useful for Records Managers/Information Managers as a model which they can demonstrate aligns with business risk management.



Risk Management for Digital Preservation

<https://vimeo.com/171082277>

From a series of video covering topics from the 'Getting Started in Digital Preservation' roadshows, this video provides a brief introduction to the use of risk management for Digital Preservation.

Case studies



Bishopsgate library case study

http://wiki.dpconline.org/index.php?title=Bishopsgate_library_case_study

A collections audit and business case focused on taking the first steps of digital preservation at the Bishopsgate Institute Library. (28 pages).

Starting Small: Practical First Steps in Digital Preservation

<http://www.slideshare.net/hakbailey/starting-small-practical-first-steps-in-digital-preservation-13385434>

One example of how digital preservation principles can be added to the collections management activities of a small institution (Dartmouth College USA from 2010–2012), without needing a lot of additional resources. (26 slides).

DPC case note: West Yorkshire Archive Service accepts a digital collection

http://www.dpconline.org/component/docman/doc_download/511-casenotemlawyas.pdf

In this Jisc-funded case study staff from West Yorkshire Archives Service report on their experience in taking their first large digital archive. This made them confront new problems and new ways of working, they conclude that "If we try we may fail; if we don't try we will certainly fail". October 2010 (4 pages).

DPC case note: Glasgow Museum takes first steps in turning an oral history headache into an opportunity

http://www.dpconline.org/component/docman/doc_download/502-casenoteglasgowmuseums.pdf

This Jisc-funded case study examines how Glasgow Museums' took some simple steps in addressing digital preservation and created short and long term opportunities. Activities such as creating an inventory, assessing significance and promoting access provide the basis for building confidence to manage the wider challenges, and can bring early rewards if properly embedded within the mission of an organization. September 2010 (4 pages).

Digital Preservation Planning Case Study

http://www.dpconline.org/component/docman/doc_download/863-2013-may-getting-started-london-planning-case-study-ed-fay

A set of DPC Getting Started in Digital Preservation workshop presentation slides by Ed Fay from May 2013. An excellent concise overview of planning for digital preservation and how to approach it . (20 slides).

References

Dollar, C.M. and Ashley, L.J., 2014. *Assessing Digital Preservation Capability Using a Maturity Model Process Improvement Approach*. Available:

<http://static1.squarespace.com/static/52ebbb45e4b06f07f8bb62bd/t/53559340e4b058b6b2212d98/1398117184845/DPCMM+White+Paper+Revised+April+2014.pdf>

NDSA , 2013. *The NDSA Levels of Digital Preservation: An Explanation and Uses, version 1 2013*.

National Digital Stewardship Alliance. Available:

http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf