Digital Preservation

Purpose

This Records Advice has been prepared to define what digital preservation is and give an overview outlining some of the broad issues and tasks involved in preserving digital resources. It synthesises current thinking on digital preservation and aims to provide a focus for further discussions within ACT Government directorates.

Digitisation vs digital preservation

Digitisation is a process to convert paper or analogue items (e.g. audio cassettes, photographs) into electronic formats for better accessibility, usage, storage, and security. Digital preservation stores digital data for as long as necessary.

While ***digitisation*** converts analogue information into digital form, ***format shifting*** transforms processes using digital technology formats. So, format shifting can be defined as the conversion of media files into different file format or data compression, for example converting music on a CD to MP4. Format shifting is central to preservation and archiving, particularly for audio recordings and film. It usually takes place when a digital hardware or software platform is in danger of becoming obsolete, so that the digital content can be retained on a new, ongoing platform.

It is important to conduct an analysis and be clear about the purpose of any digitisation or format shifting work, to ensure that the format chosen is fit for that purpose. The chosen format becomes more critical the longer a record must be kept.

For recordkeeping requirements when digitising and converting records, information and data refer to: https://www.territoryrecords.act.gov.au/data/assets/pdf_file/0009/1315827/Retain-Digitisation-Technical-Specifications.pd

What is digital preservation?

Preservation is just one part of managing digital information to ensure that it can be used in the way that is required for as long as is needed. Digital preservation is a vital aspect of digital continuity: some records have retention periods that outstrip the life of the systems they are captured in.

The United Nations Archives and Records Management Advice defines digital preservation ‘as the active management and maintenance of digital objects so they can be accessed and used

by future generations.'[1] It is a set of procedures and activities to keep information safe and accessible in the long-term.

Digital preservation involves planning, the allocation of resources and applying preservation methods and technologies to ensure ongoing access of both born digital and reformatted content. Its goal is 'the accurate rendering of authenticated content over time to ensure its authenticity, accessibility and usability'. [2]

Why is digital preservation important?

Digital preservation is essential to the process of preserving content for future use. With it, we can achieve long-term retention and archiving. For example, the decisions we make today about environmental management have are-reaching implications. Our community is likely to expect that they can access records about those decisions far into the future, which means we need to make sure that these digital assets are preserved and accessible for as long as they are required, which could mean for many future generations.

Backup is not the same as archiving. Backup retains the data as long as it's being actively used. Archives retain the data indefinitely or for a defined period. Duplicate copies can be overwritten in backups, but archive data must not be deleted or altered.

The first step to preserving digital assets is to have a digital preservation plan as part of a governance framework. One of the steps should include doing an appraisal to determine what digital records should be preserved. This involves determining what should not be lost, but also taking the opportunity to dispose of digital records where appropriate e.g. before migrating to a new system.

Threats to digital records

Digital records face various threats, particularly in relation to obsolescence. These include:

- **File format obsolescence**. Creators/vendors can update their file formats with new versions or open-source software communities might withdraw support for older formats. File format obsolescence is a risk that needs to be considered. But many established file formats are still supported, and usable.
- **Software obsolescence**. The software required to render file formats may change. Sometimes the vendor withdraws support for software. This risk can also apply to complete operating systems.
- **Specialist software**. Formats that are optimal for long-term preservation and access tend to be open, well established, and not dependent on only one software application, hardware, or operating system. Software without these characteristics can be at risk in the longer term. E.g. Files such as GIS/CAD are proprietary formats and at greatest risk of obsolescence.

[1] Digital Preservation | Archives and Records Management Section (un.org). Accessed 23 February 2023.
[2] Ibid.

- **Fragility of storage media**. Data fragility and loss risk have come under increasing scrutiny with the rise in digital storage devices like USB sticks, external hard drives, and cloud storage. The storage of digital data faces ongoing challenges including easily corruptible and deleted files.
- **Hardware obsolescence**. Computer storage and hardware can change quickly, and storage devices and media can become obsolete within a few years. Recovering digital records stored on obsolete hardware can be expensive and difficult.

Components for ensuring sustainable digital preservation

Digital records are data and information that is evidence of business activity stored digitally. There are several components to create a sustainable digital preservation workflow.

- **Data protection**. Preserving digital data allows you to maintain access for the future and helps mitigate risk of inadvertent access to sensitive material.
- **Data storage locations.** These are central to the concept of digital archiving. You can make sure your data is safe by choosing a location that is physically secure while also being accessible to its users.
- **Data backup**. Organisations should prioritise data security. Backups are one way to ensure business continuity and that your work is safe and secure.
- **Snapshots.** Snapshots of continuously updating files and data sets can be an important part of digital preservation. Snapshots should be timed with the frequency that content is updated on a server or a website**.**

Principles for digital preservation

The following principles are standard for digital preservation strategies at institutions including the National Archives of Australia and the State Library of NSW.

- **Recommended, sustainable file formats**. It is important to create and acquire content that is kept in recommended file formats. Sustainable formats are those which comply with standards, support metadata and interoperability, and have a high rate of user acceptance. The recommended list of sustainable formats provided by CAARA and ADRI is a good example of this.
- **Standardised metadata**. Good quality preservation metadata is vital to ensure long-term accessibility. It should be Open Archive Information System (OAIS) compliant to allow interoperability, sharing and harvesting by other systems and organisations.
- **Migration**. Transferring data to newer hardware and software. This can include conversion from one file format to another or from one business system to another, so the data remains fully accessible. For more information please reference https://www.territoryrecords.act.gov.au/data/assets/pdf_file/0008/1315826/Retain-Decommissioning-business-systems.pdf
- **Authentication**. Ensure that the archival master files match the attributes of recommended file formats.
- **Normalisation**. The process of transforming a digital file/record from one data format to another.

Digital Preservation Planning

The TRO recommends an active program to identify at-risk records and take the necessary steps to ensure their ongoing viability. This is essential to setting up an effective digital preservation plan. This approach ties into our Records By Design Advice.

A digital preservation plan is the organisation's public statement regarding its commitment to preserve its digital collections through the development of a comprehensive digital preservation program. The plan will provide the mission, specific goals and objectives, and policies and procedures.

Challenges

Aside from technical issues with digital preservation, there are also challenges and decisions to be made by the agency/organisation. This includes how the workflow for digital preservation is set up and delivered, how those tasks and responsibilities change over time and the lifecycle of the digital materials. While there are common digital preservation challenges, every operating context is different with a unique set of problems. It is important to identify risks and work out practical solutions to fit different organisational needs.

Roles and responsibilities

The creation and implementation of a successful digital preservation strategy will:

- require working across the organisation or ACT Government to use the full range of expertise and skills needed
- include Content creators, Recordkeeping professionals and ICT specialists.
- involve identifying and engaging various stakeholders
- call for working collaboratively with colleagues
- need stakeholders to understand the digital records and the context they were created in – their format, the medium used to create them, the retention requirements
- require appropriate organisational and technical infrastructure.

Many of the activities will converge, including decisions about creation/acquisition and preservation, which should be made at the same time. It is important to have the right people with the right knowledge and skills involved in planning and decision-making.

Staffing skills

There are several issues to consider when considering human resources for digital preservation, as it involves a diverse range of skills. This ranges from special knowledge of metadata standards and auditing through to more general experience in risk management and project planning. No single individual or team is likely to have all of the skills required so collaboration will be important to successful digital preservation work.

There are several international initiatives which have investigated and identified the skills required for digital preservation work, as well as the staff roles involved at different organisational levels. These include:

Digital Preservation Outreach & Education Network (dpoe.network)

Initiated by the Library of Congress Digital Preservation and Outreach Program (DPOE), this model defines job levels or career stages to support their model of training.

DigCurV - Skills in Digital Curation Curriculum Framework (gla.ac.uk)

DigCurV is a project funded by the European Commission under the Lifelong Learning Programme to develop a curriculum framework for training in Digital Curation. It examined training issues involved with preservation training and digital curation and developed core competencies against DPOE's job levels framework. The work undertaken by DigCurV provides a useful toolkit to assist in identifying the range of competencies and skills needed in this space and to identify skill gaps within an organisation.

Stakeholder responsibilities

- Developing an information strategy plan integrating ICT training within the overall mission
- With key staff, develop a skills audit to work out what competencies are necessary to meet organisational objectives
- Monitor the professional landscape to identify new and emerging skills, activities, responsibilities
- In collaboration with staff, design training and development programs
- Facilitate timely and effective multidisciplinary collaboration. This work requires networking across ICT, records management, and policy/project teams.

Access and Security

Security and access are always strongly linked in digital preservation planning, and a balance needs to be maintained between the need to protect and the need to share. Different types of digital records selected for long-term preservation can contain sensitive and confidential information that needs to be protected from non-authorised users. There might also be regulatory and legal requirements affecting access to digital records in the repository.

Storage

Storage hardware is very important to digital preservation facilities and has changed rapidly in recent years. The developments in magnetic media have made handheld media such as CDs or DVDs redundant. Enterprise storage systems provide large storage space at reasonable costs. They have a typical lifespan of 4-8 years and are usually easy to monitor and replace. Organisations also have the option of cloud services to rent preservation infrastructure.

Digital repositories

Dedicated digital preservation systems provide many of the vital requirements for preserving digital records. A digital preservation repository will manage the storage of objects, and their metadata, and facilitate access. This is essential to managing digital records effectively over time.

Further information

What is digital preservation? (nsw.gov.au)

Digital preservation briefing - Digital Preservation Handbook (dpconline.org)

Australasia Preserves

Home - Digital Preservation Coalition (dpconline.org)

DigitalPreservationEurope - Digital Preservation (Library of Congress)

Born-digital file format standards | naa.gov.au

The global skills and competency framework for a digital world — English (sfia-online.org)

University of Melbourne Digital Preservation Policy 2015-2025

CONTACT US

Territory Records Office | www.territoryrecords.act.gov.au| TRO@act.gov.au