Office of the
State Archivist

Digitisation Projects: Plan Before you Scan

IMPLEMENTATION TOOL

INTRODUCTION

To achieve the best outcome for your large-scale digitisation projects, plan the approach.

This implementation tool covers:

- selecting records
- assessing risk
- developing a business case
- preparing a Digitisation Plan.

This tool, and other publications in our *Digitisation Toolkit* will help you manage digitisation projects.

Tasmanian Government

Table of Contents

DIGITISATION DRIVERS

INCREASE BUSINESS EFFICIENCY AND INTEGRATE SYSTEMS

Scanning supports a seamless digital work process combining current and older records. It is easier and quicker to access digital rather than paper records. This can reduce client response times.

IMPROVE ACCESS AND USE

Digitisation can help organisations take advantage of new technologies and support staff flexibility by allowing them to access records anywhere. Digital records can be quickly accessed and viewed by many users at once, including on mobile devices. Digitisation allows continued access to information even if the source record is destroyed.

SAVE PHYSICAL STORAGE SPACE

Disposal of the source records after quality assurance processes have taken place may reduce costs. Only source records covered by an authorised disposal schedule can be disposed of using the *Disposal Schedule for Source Records* (DA2159).

> **Example** Your organisation has a storage room full of records from a recent restructure. Digitisation could reduce storage needs and provide easy access to the records, and expensive office space can now house staff.

DIGITISED RECORDS CAN BE USED AS EVIDENCE IN LEGAL PROCEEDINGS

Organisations can use the digitised record in place of the original. This may reduce time spent on finding records for legal discovery. For more information about the requirements, see *Legal Admissibility of Records*.

IMPROVE SECURITY AND AUDITING OF HIGH-RISK OR SENSITIVE RECORDS

By capturing the digitised copies in your recordkeeping system, audit logs will show who has viewed, edited, distributed, or exported the documents.

DIGITAL TRANSFORMATION

Digital transformation initiatives preference digital formats over other formats. Already, most information is created, stored, and managed digitally. Redesigning business processes and practices often involves digitisation.

TO PRESERVE INFORMATION ASSETS

Some organisations hold paper records that have long-term business use. If used often, they deteriorate. Digitisation may protect the records by limiting use. The digitised version is viewed instead of original paper records, saving them from further damage.

> **Example** You have maps and plans dating back to the earliest European settlement. These are still used for checking boundaries and property rights. However, they are old and fragile and are damaged by physical handling. You plan to digitise the maps and plans and restrict access to the physical records.

DIGITISATION FOR PRESERVATION

Preservation digitisation produces a high-quality reproduction optimised for longevity. Typically, preservation quality records are stored off-line and accessed infrequently to support end-user access and delivery. When digitising to preserve fragile and valuable originals still in regular use, you need to factor in the extra cost of producing access copies at the same time.

> **Example** Prints made from high quality digital reproductions of early Tasmanian land survey plans can be purchased, and lower quality copies are published on the TasMap website. This ensures the fragile original plans can be preserved and protected as State archives.

SELECTING THE RIGHT RECORDS

It can be complex and expensive to digitise records. Your choices need to be linked closely to your aims. These prompts may help you decide if the records are good candidates for digitisation:

Question 1: Are the records often **accessed** and **used**?

- If yes, go to Question 2
- If no, stop. Off-site storage may cost less than digitisation.

Question 2: Are the records in good **condition**?

- If yes, go to Question 3
- If no, stop. Fragile records or unusual formats are harder to digitise. Contact us.

Question 3: Do you have an appropriate **digital storage** solution?

- If yes, go to Question 4
- If no, stop. You may need a better storage solution.

Question 4: Do you know what **metadata** you need to collect?

- If yes, go to Question 5
- If no, stop. Metadata adds vital context, to help understand and access digital records. Find out what metadata standards to use and collect enough metadata to meet your needs and future users' needs.

Question 5: Can you use the *Disposal Schedule for Source Records* (DA2159)?

- If yes, the records are good candidates for digitisation.
- If no, stop. Disposal schedule coverage may be needed to destroy source records. Seek advice from the Office of the State Archivist.

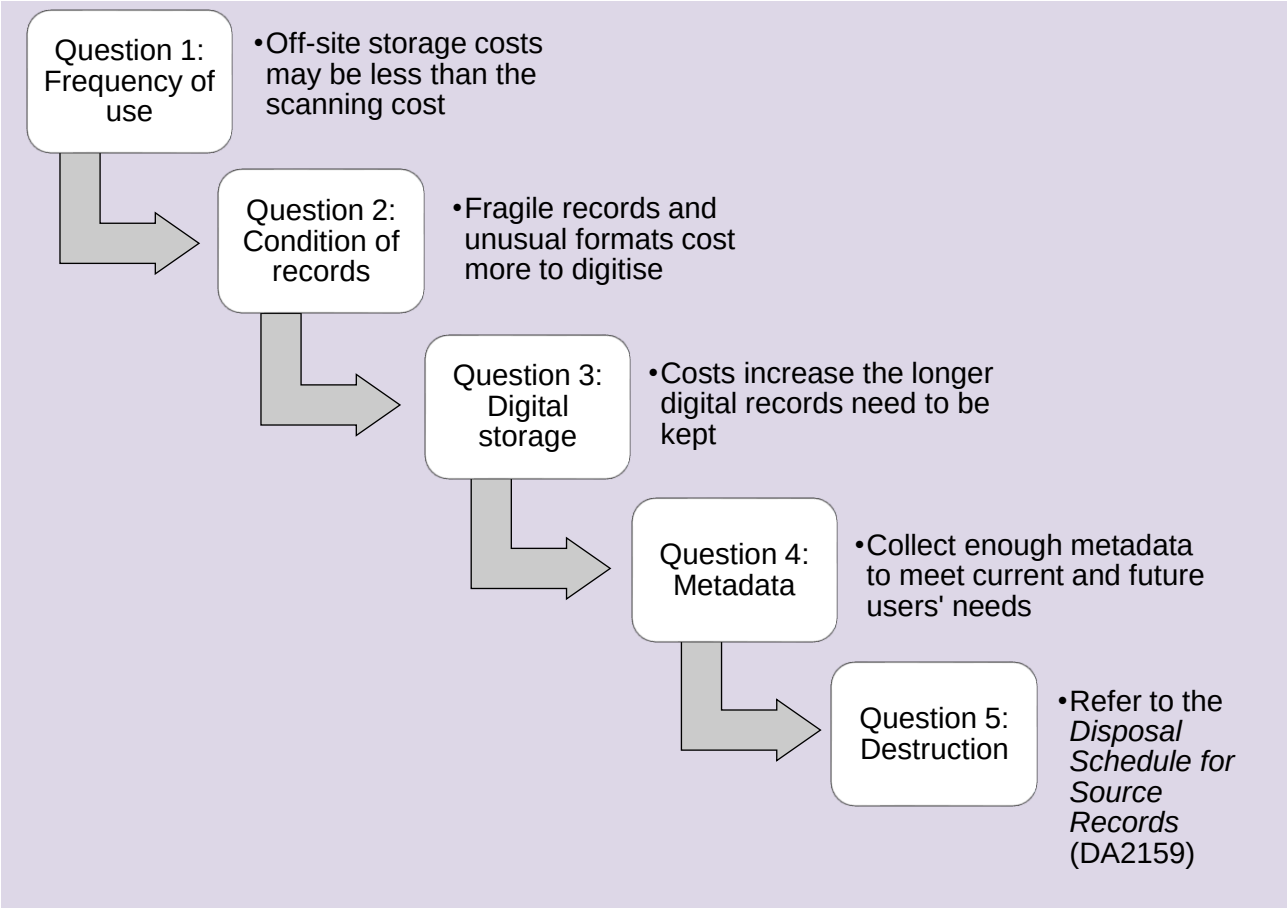

Diagram 1: Selecting the right records

Other questions to consider:

- If your organisation no longer has a business need for them, and the records have a permanent retention, can the paper records be transferred to the Tasmanian Archives immediately rather than continuing to store them?
- It may be harder to protect personally identifiable or sensitive information in digital systems. A benefit of paper records is they cannot be as easily shared. Do you have an appropriate digital system to safely store and manage access to sensitive information?
- Do you know if the source records are covered by an authorised Disposal Schedule? If they are, they can be destroyed after digitisation. Disposal Schedules are published on our website.
- Do any of the exclusions outlined in the *Disposal Schedule for Source Records* (DA2159) apply? If not, source records can be destroyed after digitisation.
- Can you manage the digitised records for the periods specified in Disposal Schedules?
- Will the costs of managing the digitised records for the required retention period be less than the costs of managing paper records?

ARGUING THE BUSINESS CASE

A good business case should discuss the business drivers, options, costs, and risks, and provide a recommendation. The business case should also provide a sound argument for why the recommendation is the best option to address the situation. Your business case can be persuasive if you have considered all aspects of the project. This includes:

- timeframes
- roles and responsibilities of those involved
- resources (including financial and human resources), equipment and training needs
- details of the vendor selection process (if outsourced)
- consideration of tests or pilots before roll-out
- risk assessments
- monitoring, evaluation, and reporting requirements.

You may need to include key performance indicators or success measures so project progress can be tracked. Measures should be tangible and relevant to the project.

DIGITISATION RISK MANAGEMENT

A large-scale digitisation project presents the biggest risk of not delivering on project aims. Before you start, complete a risk assessment, and document your risk mitigation decisions. This should inform your business case and planning.

Below are common risk mitigation strategies.

- Select records carefully to support business aims and achieve efficiencies. Don't digitise everything.
- Implement digital literacy programs to improve uptake and to avoid staff implementing workarounds and establishing parallel paper and digital systems (hybrid systems).
- Establish a breach reporting process to help protect personally identifiable and sensitive information from accidental release or other information security compromises due to the ease with which digital records can be copied and shared.
- Apply appropriate technical standards and implementing quality control processes so the source records can be safely destroyed at the end of your project.
- Capture records in machine-readable format for searchability and potential re-use of the content.
- Manage digitised records in systems with appropriate audit and access controls. Records stored in an approved recordkeeping system are accessed only by authorised users, are secure from alteration or deletion and meet legal and evidential standards. They will also be kept with related records, protected from disaster, and managed for their required retention periods.
- Capture the metadata needed to preserve links between the source record, the digitised record, and any access copies so they are retained for the right length of time. Disposal Schedules are often only applied at aggregated or folder level and may not be automatically applied in all systems.
- Maintain a preservation copy. Choosing open source or sustainable file formats also helps to protect valuable records.
- Avoid use of removable media for storage, where records are at increased risk of loss or damage.
- Allocate funds for ongoing management of the records for accessibility and integrity.
- Actively monitor for technical obsolescence such as after upgrading software platforms.
- Include digitised copies in your organisation's disaster recovery and backup procedures so they aren't lost or inaccessible because of hardware failure or other disaster.
- If outsourcing, set clear benchmarks and standards for contractors. Build your requirements into contracts and monitor the performance of contractors.
- Seek formal endorsement from Senior Management to overcome organisational resistance to accepting the digitised version as the official record.
- Document your reasoning for disposal of the source records in your Digitisation Plan to avoid your organisation keeping them as a just-in-case approach.

IMPLEMENTATION

STAFF & TRAINING

If scanning in-house, you may need to address skills gaps. Depending on the size of your team and programme, some staff may take on multiple tasks. You need to consider if existing workloads will be affected.

If you are outsourcing, you still need a staff member to be responsible for digitisation initiatives.

Arrange training if a skills gap exists. Options include:

- 'train the trainer' where one person is trained and then trains others within an organisation
- tailored training provided by external consultants
- hardware or software-specific training provided by vendors. Some vendors offer training as part of contracts.
- some sectors of projects may require specific skillsets, for example, audio-visual preservation.

GAINING SUPPORT

Leadership support

Secure Senior Management support for the lifetime of your project. It is not enough to have them sign-off on your business case. Ideally, they are sponsors who advocate for the change, and explain its value to your organisation. Ways you can achieve this are by demonstrating the software and hardware, reporting early 'wins' and communicating regularly with them.

Lasting organisational change often depends on the support of business unit managers. You may individually target these managers to discuss and demonstrate potential outcomes such as business efficiencies from your project. Staff champions can see what you are trying to achieve and will support it. They may find innovative ways to improve the digitisation process.

Another means of achieving project aims is by involving staff in discussions on how their work processes will change with the introduction of digitisation. Staff know best about how they might work better. Consider their views. If they feel they can influence change, they may be more likely to accept it.

Communication strategies

Develop education and awareness strategies. Support or reward good practice.

Some staff may revert to former practices or find workarounds if they are not experiencing benefits from digitisation. Offer plenty of training and support to staff who are unfamiliar with the technology and processes.

Develop communication strategies for all affected staff in the organisation. Consider each group and communicate to them in terms of 'what's in it for me?' Simple misconceptions can be removed easily with communication and may prevent some staff members from sabotaging your project.

RESOURCING AND COST CONSIDERATIONS

OUTSOURCING DIGITISATION

Outsourcing can be cost-effective. Some in-house considerations also apply to outsourcing, including migration to new systems, training your users, long-term digital storage, etc.

If outsourcing, you need a thorough understanding of technical standards. Ask for work samples and referees you can speak to. You should also view their set-up.

Don't forget to factor in costs and resources to:

- plan, establish and document the project with suitable parameters and benchmarks
- select contractors that can meet project needs, especially if the records contain sensitive information and be ready to negotiate terms
- prepare (and possibly transport) records to and from the service provider
- liaise with contractors throughout the project
- perform checks and monitor and evaluate services provided. Remember the onus is on you to check the quality of their work and adherence to specified standards before the whole job is done
- train contractors in how to safely handle your records, particularly if they are Permanent records required as State archives.

Other things to keep in mind:

- The cheapest quote will not necessarily deliver the quality you need. For example, have they included OCR costs in the quote?
- Clearly communicate with potential vendors so they understand exactly what is to be digitised, otherwise the final cost could be much higher than the original estimate.
- Clearly specify scanning format requirements. Outputs that require more preparation will have higher costs. For example, multi-page PDF requires more processing than TIFF, separating into unique documents will take longer than scanning as a single file, etc.
- Don't forget about metadata! The more you need them to collect and capture, the higher the cost.

IN-HOUSE DIGITISATION

In-house digitisation may require extra resourcing and equipment. For example:

- the purchase, support and upgrades of digitisation software and hardware
- training and support for staff involved in digitisation work
- costs associated with physical space for digitisation (fit-out, rental, etc)
- staff time to test and develop suitable parameters and benchmarks
- staff time to find and prepare records, digitise, apply metadata, do quality control checks, etc.
- staff time to manage variables like non-uniform or poor-quality originals
- technical infrastructure and server space for storing digitised records temporarily.

Efficient and well-documented procedures will save staff time and therefore costs.

PREPARING CONTRACTS AND SERVICE AGREEMENTS

If you outsource, clearly set out in contracts and service agreements the:

- range, quantity, and type of records to be digitised
- digital formats required, including for preservation and access versions.
- timeframes, costs, and expectations
- roles and responsibilities (including monitoring and compliance processes)
- transport and handling arrangements
- technical, metadata and documentation requirements
- quality assurance measures (including remediation if quality benchmarks are not met)
- backup and disaster recovery arrangements
- written confirmation by the provider that their copies of digitised imagery and metadata have been appropriately destroyed at project completion.

Your organisation always retains responsibility for the records. Ensure this is clearly documented in contracts and service agreements.

CALCULATING COSTS

Digitisation costs can vary widely, according to the scope and aims of your digitisation project and the quality of the digital images required.

When scoping and preparing your business case compare digitisation costs with:

- costs of inaction
- cost savings if destroying original paper records and so reducing off-site storage costs or reducing your office space
- cost savings of improving business practices through providing better access to the records.

Data storage costs may be reduced by choosing smaller file sizes and sacrificing quality.

You may not be able to reduce those data storage costs if you are digitising permanent records, or high-value high-risk records.

IS OPTICAL CHARACTER RECOGNITION (OCR) WORTH IT?

Records that are scanned or photographed (using photocopiers, cameras, flat-bed scanners, etc) create digital images. The text within the digital image cannot be searched or used unless the digitised file is further processed by OCR software, or the text manually keyed in (transcribed).

OCR software converts the record into machine-readable text, which can be searched or manipulated. OCR works best for typed documents and forms with simple, consistent layouts.

For large or outsourced scanning projects, some might consider the cost of OCR too high. However, having machine-readable data may over-ride cost considerations. Postponing this process may increase the risk that images may degrade, the files become inaccessible and knowledge about the content of the records will be lost.

Manual transcription may be better for hand-written or fragile documents and historical registers.

Digitisation Plans

A Digitisation Plan captures decisions and supports large legacy scanning projects, and digitisation of Permanent and long-term Temporary records.

A Digitisation Plan is strongly recommended for digitisation of records containing personally identifiable or sensitive information, such as personnel records and client case files. In 75 or 110 years, when the records are due to be destroyed, the Digitisation Plan will be a key evidential document.

Digitisation plans should cover digital image specifications, metadata, workflow, and quality processes for every large digitisation project you undertake.

The plan typically has six sections:

SECTION 1

Digitisation Activity: What is to be digitised, why it is being digitised and any impact on workflows and users.

SECTION 2

Digitisation Image Specification: The minimum requirements for the digital images.

SECTION 3

Digitisation Processing: The process of converting the source records into digital records. The process starts with retrieving the source records, digitising, capture of metadata, and creation of digital records.

SECTION 4

Management Plan for the Source Records: How the source records are managed after the digitisation activity. Often, the source records will be disposed of after conversion, so this document describes the process of disposal, including any requirements for authorisation from the State Archivist, and approved destruction methods.

SECTION 5

Management Plan for the Digitised Records: How the digitised copies are managed after digitisation.

SECTION 6

Quality Control and Assurance: How the digitisation process will meet quality requirements set by the organisation.

SECTION 1: DIGITISATION ACTIVITY

This section describes the digitisation process in detail, including records selection, assessment, and prioritisation, project planning, and management.

IN YOUR PLAN

- Describe the records to be digitised including the quantity, type, and format of the records.
- Document any appraisal decisions, including value, significance, and risk. Identify relevant Disposal Schedules, whether the records are Temporary or Permanent, and the retention period for the records. NOTE: The source records may not be described in a current disposal schedule, or they may be subject to a freeze or legal action. They may also have value as artefacts and cannot be destroyed. You may need to invest resources in indexing, organising, and describing the records before you scan them, so factor that in.
- Include a statement about the business drivers and the expected benefits, for example: ‘the records are being digitised because we need to relocate’ or ‘they are stored in poor conditions and not easily accessed’.
- Document your stakeholder requirements, including changes to business processes and workflows. You may need to consider future users, such as clients and researchers, not just staff or business users; as well as information security classification, and copyright or other policies governing use and access that may impact on stakeholders.
- Include a risk analysis. Previous risk assessments, for example from your business case, can be used.
- Describe hardware, software, and any technical specifications. Include a statement about the location, equipment used.
- Identify and document any specialist training required by staff or by contractors.

SECTION 2: DIGITISATION IMAGE SPECIFICATION

The Digitisation Image Specification sets out the image requirements for each type of record. You must use our *Digitisation Standard: Technical Standard for Permanent Paper Records* for scanning Permanent records if you plan to dispose of the source records.

Sustainable Digital File Formats for Creating and Using Records in the *Digitisation Toolkit* has information about preferred digital file formats.

IN YOUR PLAN

- Decisions about technical specifications and how they align to business requirements. For example, characteristics of the source records may need to be preserved for evidence purposes, such as notations in pencil. This may mean scanning at higher dots per inch (dpi) and more storage.
- A statement about compliance with file format standards is crucial when digitising Permanent records or if the digitised records will be kept for a long time. For example, personnel records or case management records that may be retained for a person's lifetime.
- If the technical quality of the digitised records needs to be extremely high, focus on documenting in detail any quality control and assurance practices. For example, certain records may need to be captured in colour such as maps, graphs or plans or documents with low contrast. Or you may need to manipulate the scanning process or undertake post-processing activities to capture faded text, browning paper, coloured paper, or coloured backgrounds.

SECTION 3: DIGITISATION PROCESSING

This section describes the workflow process to generate full, accurate and complete records from the source documents. It includes scanner configurations, record retrieval and return processes, scanning procedures, metadata that will be captured, registration into approved recordkeeping systems, etc.

This section contains vital information for future users and external stakeholders, including the Tasmanian Archives.

IN YOUR PLAN

- Document and specify metadata capture that is consistent with your organisational records controls (for example, file plan or managed metadata schemas).
- Include metadata requirements through statements such as:
  - "Titling will be according to organisational naming conventions which are attached to this plan…"
  - "The scanning metadata is uploaded and automatically captured in the recordkeeping system in these fields…"
  - "Extra metadata required to be entered at the point of registration is…"
- Document workflow processes if:
  - the digitised records need to be managed for long retention periods (for example, 20 years or longer) through system migrations and upgrades
  - only a preservation copy is made from which access copies will be generated as needed.
- Handling and transportation requirements may need to be included. For example, if you are:
  - removing staples or paper clips before scanning
  - redacting sensitive information
  - tracking delivery to minimise risk of unauthorised access, loss, or damage to the records.
- Document any special measures to protect personal, sensitive or security classified information in the plan.

SECTION 4: MANAGEMENT PLAN FOR THE SOURCE RECORDS

This section of the Plan addresses:

- disposal status of the source records and how long they are retained after digitisation
- record management processes - this should cover the systems used to manage the source records until disposal, and the relationship (linking metadata) between the digitised records and the source records
- the disposal process and how it needs to be documented.

Use this section of the plan to describe the processes for storage and disposal of the source records after digitisation has been carried out.

This will support legal admissibility and will show that disposal is authorised under the *Disposal Schedule for Source Records* (DA2159).

IN YOUR PLAN

- Describe any requirements for source records. Standard practice is to store source records for a minimum of six months, subject to risk assessments and quality checks.
- If the scanning vendor stores the source records and undertakes disposal, this should be documented in your plan, and in the contract or service agreement with the vendor.
- Insert a disposal statement, for example, “the source records are temporary [cite disposal schedule reference] and will be retained for six months after registration into the recordkeeping system to ensure quality control processes are met. The source records will then be securely destroyed, following organisational procedures”.

SECTION 5: MANAGEMENT PLAN FOR THE DIGITISED RECORDS

This section covers your records management regime, security and access controls, storage, back-up, and the eventual export of the digitised records from your system. This will ensure the digitised records are appropriately managed for the required retention periods. The Management Plan for the Digitised Records should address the following:

- Record management processes for the digitised records. It should cover:
  - The recordkeeping system used to manage the digitised records.
  - Identification of the digitised record (document type, title, unique ID, etc).
  - Indexing and classification.
  - Security and access control.
- Preservation and disposal processes.
- Storage systems.
- Back-up and restoration (daily, weekly, monthly, etc).
- Disaster recovery processes, such as making copies of the records and contextual metadata for the purpose of recovery of information lost due to a catastrophic failure. If included in organisation-wide disaster recovery plans, this should be documented.
- The process for exporting records and associated metadata from the recordkeeping system for migration, decommissioning, and other purposes.

IN YOUR PLAN

- Document decisions about migration and digital preservation strategies so the records are kept for as long as required.
- Describe the recordkeeping system used to manage the digitised records. Document the process for indexing, classifying, and applying security and access controls.
- Document security and access controls, covering system security and physical security of the media and servers. The plan should cover all copies including those held for back-up and disaster recovery regimes.
- Disposal metadata to be applied. All systems holding records must have disposal actions linked to each record folder or group of records with like retention. This needs to be identified and documented.
- Details of data storage and backup procedures, including:
  - back-up software and process (including frequency)
  - quality assurance procedures to ensure that data is restored correctly.

SECTION 6: QUALITY CONTROL AND ASSURANCE

A large back-scanning project requires strict quality assurance. The Quality Control and Assurance Plan for the digitised records addresses:

- image accuracy
- record accuracy
- storage reliability
- quality failure processes
- logging and analysis.

This section describes how the digitisation process produces full, complete, and accurate digital records. Quality assurance processes should include both routine activities as part of the digitisation workflow and irregular activities, such as audits.

The *Digitisation Standard: Technical Standard for Permanent Paper Records* includes guidance about quality checks.

IN YOUR PLAN

- Insert a basic benchmark statement for stakeholders, such as: “To meet accuracy standards the digital image must capture all of the content of the source records and be legible.”
- Examples of image accuracy processes:
  - “The operator will sample 100% of digitised images in the quality control process”
  - “View images to check the scanned image is an accurate and legible reproduction of the original”
  - “Acceptance quality is achieved when the image is verified as a true and accurate representation of the original document”
  - “The vendor verifies the Upload File before delivery”
  - “Verify image accuracy and check the metadata is linked to the correct scanned image”.
- The storage system needs to reliably hold records for as long as they are required. Consider and document tasks such as back-up and restore processes, periodic checks, and long-term preservation plans for storage media so records can be accessed and exported when required.
- If quality failures such as inaccuracies or poor-quality images are identified, describe the processes in place to check if this is a systematic problem (for example operator error, hardware/software error, storage error).
- Describe processes for monitoring trends and detecting systematic problems. If outsourcing, the onus is on you to ensure the scanning vendor sends regular status reports. Include this requirement in outsourcing agreements.

ACKNOWLEDGEMENTS

Tasmanian Archive and Heritage Office 2015, *Information Management Advice 21: Plan Before You Scan,* TAHO, 12 May 2015, viewed 21 May 2024, <https://www.osa.tas.gov.au/information-management-framework/older-guidelines-and-advice>.

MORE INFORMATION

The Digitisation Toolkit helps you digitise and preserve your high-value, high-risk information. Using this toolkit will help you meet reproduction standards so you can destroy original or source records. The Toolkit includes:

- *Application to Dispose of Permanent Paper Source Records* – a checklist to help you prepare your application to our Office.
- *Digitisation Projects: Plan Before you Scan* – will help you manage large volume digitisation or back-scanning projects.
- *Digitisation Standard: Technical Standard for Permanent Paper Records* – describes minimum requirements for preservation quality digital copies. Use this Standard for permanent records and to meet the conditions in our *Disposal Schedule for Source Records* (DA2159).
- *Digitisation Toolkit FAQs* – your guide to the Toolkit.
- *Disposal Schedule for Source Records* (DA2159) – the legal instrument you use to dispose of the source records.
- *Sustainable Digital File Formats for Creating and Using Records* – this CAARA publication recommends and lists digital formats.
- *Three Steps for Better Scans* – a staff educational poster.

Other relevant resources include:

- Office of the State Archivist 2020, *Legal Admissibility of Records*, OSA, viewed 21 May 2024, <https://www.osa.tas.gov.au/information-management-framework>.

CONTACT US

Office of the State Archivist | www.osa.tas.gov.au | osa@libraries.tas.gov.au | 03 6165 5581

Digitisation Projects: Plan Before you Scan is part of the *Tasmanian Government Information Management Framework.* It supports the *Information and Records Management Standard* and the *Digitisation Standard.* This is a living document, and we will make minor changes as needed. If you notice anything that needs updating, please let us know.

Document Development History

Version	Date	Comments
2.0	21/05/2024	Major rewrite, new template
1.0	12/05/2015	Initial release