>DIGITISATION STANDARD

>ARCHIVES NEW ZEALAND

>GOVERNMENT RECORDKEEPING GROUP

> ISSUED UNDER PUBLIC RECORDS ACT 2005, SECTION 27
> SCOPE: ALL PUBLIC OFFICES AND LOCAL AUTHORITIES
> STATUS: DISCRETIONARY STANDARD

S-6 ARCHIVES NEW ZEALAND'S STANDARD FOR DIGITISING NON-ELECTRONIC RECORDS FOR RECORDKEEPING PURPOSES AND RETENTION OF NON-ELECTRONIC RECORDS IN ELECTRONIC FORM ONLY.

THIS STANDARD SETS OUT REQUIREMENTS FOR DIGITISATION PROCESSES FOR PUBLIC OFFICES AND LOCAL AUTHORITIES TO SUPPORT THE REQUIREMENTS OF THE ELECTRONIC TRANSACTIONS ACT 2002 AND PUBLIC RECORDS ACT 2005.

> CHIEF EXECUTIVE AND CHIEF ARCHIVIST

> SIGNED: *D. M. Macaskill*

> DATE: 29 January 2007

> REVIEW DATE: 2010

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

› ACKNOWLEDGEMENTS

This Standard was developed for the Government Recordkeeping Group, Archives New Zealand by Barbara Reed, Recordkeeping Innovation Pty Ltd., and Stephen Clarke, Archives Analyst, Archives New Zealand, with advice and guidance from an advisory group comprising:

> › *Adam Stapleton,* Project Manager, State Services Commission;

> › *Alison Midwinter,* Senior Advisor, Appraisal, Archives New Zealand;

> › *Alison Nevill,* Manager, Knowledge Management Unit, Ministry of Economic Development;

> › *Andy Fenton,* Managing Director, NZ Micrographic Services;

> › *Janine Delaney*, Manager, Digital Preservation, Archives New Zealand;

> › *Joanna Newman,* City Archivist, Wellington City Council;

> › *Liz Kolster,* Senior Analyst, e-GIF, State Services Commission;

> › *Mike Chapman,* Information Administrator and Democratic Services Manager, Porirua City Council; and

> › *Shane Beverley,* Information Manager, Department of the Prime Minister and Cabinet.

The Chief Archivist gratefully acknowledges the contributions of the advisory group.

The development of this Standard was informed by international professional literature and draws upon various institutional standards and policies. In particular the ISO standards on digital imaging and information documentation[1]; and two publications issued by the Queensland State Archives: 'Digitisation Disposal Authority' (April 2006) and 'Guidelines for the Digitisation of Paper Records' (May 2005). Please refer to *Appendix 9: Further Resources* for further information.

---

1. ISO 12653 – 1: 2000, *Electronic imaging – Test targets for the black and white scanning of office documents, Part 1 – characteristics;*

ISO 12653-2: 2000, *Electronic imaging – Test targets for black and white scanning of office documents, Part 2 – methods of use;*

ISO/TR 15801: 2004, *Electronic imaging – Information stored electronically – Recommendations for trustworthiness and reliability;*

ISO 23081 – 1: 2006, *Information and documentation – Records management processes – Metadata for records, Part 1 – Principles; and*

ISO 23081 – 1: 2006, *Information and documentation – Records management processes – Metadata for records, Part 2 – Implementation issues.*

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

› CONTENTS

‹ DIGITISATION STANDARD

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

DIGITISATION STANDARD

> INTRODUCTION

This Standard is for use in the design and conduct of responsible digitisation by all organisations which are covered by the Public Records Act 2005.

Meeting the mandatory requirements of this Standard will enable public offices and local authorities to meet their legislative obligations and dispose of original source records after digitisation. This applies to records where the business action takes place on the digitised record, rather than on the non-digital original source record. Public offices may then implement the *General Disposal Authority: Digitised Original Source Records.* The Standard sets out the criteria that will allow local authorities to be confident that they have met the requirements of the Electronic Transactions Act 2002 and dispose of the original source protected records.

The Standard provides a framework for digitisation back-capture projects although not the disposal authority; these records must have a specific disposal authority and are subject to an appraisal process prior to any consideration of disposal.

With the shift to electronic recordkeeping systems, many public offices and local authority organisations in New Zealand are digitising original paper, or other non-digital, records. The official record of activity, regardless of the format, is the one on which the business action was taken; this is increasingly becoming the digitised copy of a non-digital original. Public offices and local authorities are legally responsible for ensuring that they create full and accurate records of their activities and that these records are maintained over time for subsequent reference. This responsibility applies regardless of the records' storage media.

Digitisation, also known as imaging or scanning, is the means of converting hard-copy, or non-digital, records into digital format. It may also involve taking digital photographs of the source records. Once converted to digital objects, they may be:

> captured as a static picture (raster image) represented by pixels;

> processed by optical character recognition technology which converts the pixels into digital representations which are searchable, editable and manipulable; or

> captured into both formats.

For the purposes of this Standard, digitisation can broadly be categorised into two types, further defined in the Glossary and Scope sections:

> **Business-process digitisation:** ongoing, routine digitisation for daily business use. For which the Standard's compliance certification process applies and the *General Disposal Authority: Digitised Original Source Records* MAY apply; and

> **Digitisation projects:** which are subject to an appraisal process prior to any consideration of disposal.

Digitisation offers the following potential benefits to organisations:

> Capacity of more than one person to access the images concurrently;

> Distribution over networks enabling access from multiple locations at any time;

> Greater integration with business systems;

> Capacity to distribute images in a structured workflow, thus assisting work processing;

> Elimination of hybrid (both paper and electronic) systems which can cause confusion to users who require access to the whole history of a matter;

DIGITISATION STANDARD
CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

5

> Capacity to re-use existing resources limited in their re-use by their format;

> Application of consistent classification and indexing for document retrieval;

> Integration with existing organisational disaster recovery and back-up regimes;

> Provision of a protected and secured master image; and

> Potential to reduce physical storage space occupied by hard-copy records.

There are a number of risks associated with implementing a digitisation process:

> Short-term cost savings in space may be fleeting when balanced with longer-term costs in maintaining the accessibility of digital images over time;

> Technical standards used to create digital images will significantly affect longevity and capacity to re-use the images in the future;

> Requirements to maintain authentic and reliable representations of original records may limit the capacity to deploy commonly offered digitisation features (such as image manipulation, etc.);

> Processing requirements may be complex and require additional resources;

> It may not be permissible to dispose of the original source records after the digitisation process, especially where the physical records may have intrinsic value, e.g. records with importance for national identity or other intrinsic spiritual or cultural significance, including items of high value to Māori;

> It may not be permissible to dispose of the original source records after the digitisation process for legislative reasons (by law, some specific classes of records have to be retained in their original format).

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

6

❯ GLOSSARY

Further explanation of technical terminology is included in the glossary in Appendix 5: Recommended Technical Specifications.

Obligation Terms:

In this Standard, obligation is indicated by the following terms, when they are capitalised within the text of the requirements:

MUST	Means that the item is a mandatory requirement of the Standard.
SHOULD	Means that valid reasons may exist in particular circumstances to ignore a particular item, but the full implications have to be understood and carefully weighed before choosing a different course.
MAY	Means that an item is truly optional. Some implementers may choose to include the item because it serves a particular local requirement or because it enhances the digitisation process, while others may omit the same item.

General Terms:

In this Standard, the following terms are used as defined below:

Born digital
Digital materials which are created and retained only in digital form, without ever having had a non-digital equivalent. This term is used to differentiate them from: 1. digital materials which have been created as a result of converting non-digital originals; and   2. digital materials which may have originated from a digital source but have been printed to paper.   [Derived from *Digital Preservation Coalition, Handbook*. http://www.dpconline.org/graphics/intro/definitions.html]

Business system
An organised collection of hardware, software, supplies, policies, procedures and people, which stores, processes and provides access to an organisation's business resources [Source: National Archives of Australia, *Recordkeeping Glossary*. http://www.naa.gov.au/recordkeeping/rkpubs/recordkeeping_glossary.html].   A business system will create records, but may or may not manage them according to records management requirements. A recordkeeping system is a specific type of business system with the dedicated functionality of managing an organisation's record resources.

DIGITISATION STANDARD

7

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

Business-process digitisation

Digitisation of records, often on receipt, and incorporation of the digital record into a business system supporting ongoing organisational activities. Such digitisation may take place in conjunction with the operation of an electronic document and records management system. In business-process digitisation, the business action takes place on the digital record and, therefore, the official record of the action is the digital record.

Derivative

An image that has been created from another image, such as the master image, through some kind of image editing process to create a user or working copy. The process usually involves a loss of information to reduce the size by sampling it to a lower resolution, using lossy compression techniques, or altering an image using image processing techniques. Typically, such derivatives are made for purposes such as web access, including "thumbnail" images that might be only 100 pixels square, or as "reference" or "service" images that should fit completely within an average monitor. Images created for this purpose commonly have smaller file sizes and, therefore, do not require a fast network connection and are in web-viewable formats.

Digitisation

Also known as imaging or scanning, digitisation is the means of converting hard-copy or non-digital records into digital format. Hard-copy or non-digital records include audio, visual, image or text. Digitisation may also be undertaken by taking digital photographs of the source records, where appropriate.

Digitisation projects

Retrospective or project-based digitisation of existing sets of non-digital records to enhance accessibility, maximise re-use or for preservation purposes. In digitisation projects, the business action is already completed and was undertaken on the non-digital record, which remains the official record regardless of the existence of a digital copy.

Indexing

The process of establishing access points to facilitate retrieval of records and/or information. [Source: ISO 15489-2001]

Master

A faithful digital reproduction of a document, optimised for longevity and for production of a range of delivery versions (derivatives). Masters are captured at the highest practicable quality or resolution and stored for long-term usage. Typically, masters are stored in an off-line mode on tape or CD and are accessed only for the production of derivative images.

Metadata

Data describing context, content and structure of records and their management through time. [Source: ISO 15489-2001]

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA

JANUARY 2007

8

Protected record (local government)

A local authority record declared by the Chief Archivist to be a protected record by notice in the *Gazette*. This list of local authority protected records is known as the 'Local Government Schedule'. Local authority records are the classes of local government records that are created or received by a local authority in the conduct of its affairs.

Public record

Any record created or received by a public office in the conduct of its affairs. See section 4 of the Public Records Act 2005 for a full definition.

Re-use

A use of the record that is different from and separate to the original process that resulted in the creation of the record.

Recordkeeping System

"[I]nformation systems which capture, maintain and provide *access* to *records* over time." (AS 4390-1996, Part 1, 4.20). These systems must have records management functionality embedded; refer to Archives New Zealand's *Continuum* Resource Kit for further guidance. http://www.archives.govt.nz/continuum/rkpublications.php

Redaction

The process of removing or masking unwanted or sensitive areas of a document prior to showing it to others.

Source record

A document or record that has been copied, converted or migrated or will be the input for such a process. A source record may be an original record or it may be a reproduction that was generated by an earlier copying, conversion or migration process. [Source: National Archives of Australia, *Recordkeeping Glossary*. http://www.naa.gov.au/recordkeeping/rkpubs/recordkeeping_glossary.html]

> PURPOSE

This Standard establishes a framework for maintaining public records and local government protected records in digital format only, where the original paper, or other non-digital source record, has been copied electronically by digitising or other means. It sets out the mandatory requirements for digitisation processes and for the implementation of the accompanying

Disposal Authority: Digitised Original Source Records, and guidance for local authorities to meet the requirements of the Electronic Transactions Act 2002. The Standard outlines best practice recommendations, particularly addressing the following broad areas:

> The reliability and authenticity, hence the trustworthiness and legal admissibility, of digitised records;

DIGITISATION STANDARD

9

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

> The accessibility of digitised records for as long as they are required;

> Strategies to assist in creating digitised records fit for long-term retention; and

> The management of original source records following digitisation.

> APPLICABILITY

This Standard applies to public offices and local authorities. Where public offices and local authorities contract third parties to carry out (local) government functions, this Standard applies to the records documenting those (local) government functions.

> MANDATE

This is a discretionary standard issued under section 27 of the Public Records Act 2005.

However, where public offices wish to use the accompanying *General Disposal Authority: Digitised Original Source Records* to destroy original source records, the mandatory requirements of the Standard, as set out in Appendix 1, must be met, and certified as outlined in Appendix 2. Compliance with the Standard also means that the Chief Archivist and the public office can be confident that the requirements of s 25(1) of the Electronic Transactions Act 2002 are met by the digitised record.

Local authorities which meet the mandatory provisions of the Standard can be confident they meet the requirements of s 25(1) of the Electronic Transactions Act 2002. They can then dispose of the original source records.

> LEGISLATIVE FRAMEWORK

Legislation that has an impact on digitisation of public and local authority protected records includes:

> The Public Records Act 2005, which establishes a framework for recordkeeping in the wider state sector and local government and ensures the preservation and accessibility of records required for long-term retention:

> Section 17 requires that public offices and local government organisations create and maintain full and accurate records of their activities and ensure the records' ongoing accessibility over time;

> Section 18 states that no-one may dispose of public records and local authority protected records without authorisation from the Chief Archivist.

> The Electronic Transactions Act 2002, which provides legal validity to transactions carried out electronically:

> Section 25 permits the recording and retention of information and documents in an electronic form and destruction of non-digital source records, providing that the electronic form provides reliable maintenance of the integrity of the information and that the information is readily accessible and usable for subsequent reference;

> Section 25(2) requires the Chief Archivist's authorisation for retention of public records in electronic form, before destruction of the original records.

Public offices need to satisfy the mandatory requirements established in this Standard as a pre-condition to implementing the *General Disposal Authority: Digitised Original Source Records,* issued under the Public Records Act 2005.

By satisfying the requirements of this Standard, a local authority can be confident that it meets the requirements of s 25(1) of the Electronic

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

10

Transactions Act 2002. It may then legally dispose of the non-digital form of the protected record and retain the record in electronic form only.

A checklist to assist agencies to assess whether they meet the requirements of this Standard is included as Appendix 1.

> Evidence Act 2006 (as amended), governs the admissibility of documents as evidence in legal proceedings.

Section 138 of the Evidence Act 2006 provides that a copy which is certified to be a copy of the public document held by a local authority is admissible in court to prove the truth of its contents.

Being able to show that the organisation's digitisation processes met the Standard would be relevant in the event of a challenge to the authenticity of a public office or local authority's digitised form of the original source record.

> SCOPE

This Standard applies to the processes of converting non-digital original records (including audio, visual, image, text, or microform) into digital images. This Standard also covers the subsequent management of both the original records and their digitised counterparts.

It applies to:

> **Business-process digitisation:** Digitisation of records, often on receipt, and incorporation of the digital record into a business system supporting ongoing organisational activities. Such digitisation may take place in conjunction with the operation of an electronic document and records management system. In business-process digitisation, the business action takes place on the digital record and therefore the official record of the action is the digital record (i.e. routine digitisation for daily business use for which the Standard's compliance certification process applies and the *General Disposal Authority: Digitised Original Source Records* may apply) and;

> **Digitisation projects:** Retrospective or project-based digitisation of existing sets of non-digital records to enhance accessibility, maximise re-use or for preservation purposes. In digitisation projects, the business action is already completed and was undertaken on the non-digital record, which remains the official record regardless of the existence of a digital copy (i.e. records that must have a specific disposal authority and which are subject to an appraisal process prior to any consideration of disposal).

Once non-digital records are converted to digital images, many of the management and preservation issues associated with the digital image are the same as those for born-digital records.

Out of scope:

This Standard does not cover:

> The capture and management of born-digital records;

> Records covered by other current Archives New Zealand Disposal Authorities; or

> Long-term management and preservation of digital records[2].

2. For guidance on these and other issues, please see Archives New Zealand's other *Continuum* publications: Fact Sheets, General Disposal Authorities, Guides and Standards.

DIGITISATION STANDARD

11

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

> REQUIREMENTS

> SUMMARY OF MANDATORY REQUIREMENTS

1.0 > ALL DIGITISATION AND DIGITISATION PROCESSES MUST BE PLANNED, SCOPED AND DOCUMENTED:

1.1 > An appropriate digitisation approach **MUST** be selected, documented and implemented;

1.2 > Technical specifications aligned to the digitisation requirements **MUST** be selected, documented and implemented; and

1.3 > Equipment and software aligned to the digitisation requirements **MUST** be implemented.

2.0 > SYSTEMS TO SUPPORT MANAGEMENT OF THE DIGITAL OUTPUT OF DIGITISATION AS RECORDS MUST BE IN PLACE:

2.1 > Source record preparation guidelines **MUST** be documented and implemented;

2.2 > All digitised images **MUST** be assigned metadata to document digitising processes and to support ongoing business processes;

2.3 > Quality assurance procedures **MUST** be defined, documented and implemented; and

2.4 > Storage media and back-up procedures **MUST** be defined, documented and implemented.

3.0 > DISPOSAL OF ALL RECORDS MUST BE AUTHORISED AND DOCUMENTED:

3.1 > Disposal of source records **MUST** be authorised in accordance with relevant legislation and documented; and

3.2 > Disposal of the digitised records incorporated into business systems **MUST** be authorised and documented.

4.0 > LONG-TERM MANAGEMENT SYSTEMS, WHERE REQUIRED, FOR BOTH SOURCE AND DIGITISED RECORDS, MUST BE DOCUMENTED AND IMPLEMENTED:

4.1 > Sound management systems **MUST** be in place for source records until their authorised disposal within a recordkeeping framework; and

4.2 > Migration and/or preservation strategies and processes **MUST** be defined, documented and implemented.

(See Appendix 1 for a full checklist of mandatory requirements.)

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

12

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

> PRELIMINARY CONSIDERATIONS

The rationale for digitising SHOULD be carefully aligned to a costed business case geared at improving the organisation's ability to carry out its functions. The business case SHOULD clearly outline the benefits and anticipated business or cost efficiencies. The business case SHOULD involve appropriate project budgets, resource commitments and be realistically costed.

Digitisation can involve extensive document preparation and requirements for indexing, which combined can comprise the majority of a digitisation budget.

Agencies should not be misled by casual advice that digitisation is a cheap option. Digitisation undertaken as a space saving device is rarely justified, especially when the costs of future migration projects are factored in.

Appendix 4 provides a set of questions to assist in assessing the viability of digitisation for non-digital records.

Master Copies and Derivatives

Master copies of digitised records are those maintained as a separate and inviolable record in a safe storage environment, executed to the highest technical specifications available.

Derivative versions, where required, SHOULD be made during the digitisation process. Master copies SHOULD be made available for the creation of subsequent derivative images, where necessary.

The maintenance of a master record separated from the business process MAY not be necessary for business-process digitisation, where the digitised record:

> Is the version used for making business decisions;

> Is evidence of a business action; or

> Acts as supporting reference material.

Official Record

The official record of action is the record on which the business action takes place, identified as the record which is used in the transaction of business, regardless of whether this is the master or a derivative. Where a digitised copy of an original source record is the official record for business actions then the original is no longer the official record, it can then be assessed for disposal. If the original non-digital record is the record on which the business action still takes place, then it is still the official record and CANNOT be considered for disposal.

Mandatory Requirement

Preliminary Requirement: *The official record of action* **MUST** *be identified as the one which is used in the transaction of business.*

NB *If the original source record is still the record of business action it cannot be disposed of. If the subsequent business action is taken on the digitised copy the original source record can be considered for disposal.*

1.0 > ALL DIGITISATION AND DIGITISATION PROCESSES <u>MUST</u> BE PLANNED, SCOPED AND DOCUMENTED

The project documentation SHOULD include:

> Scope definition: with clear identification of business drivers, objectives, scale, size and constraints of the project;

> Statement of the purpose and expected uses of the digitised records, illustrated if necessary with examples;

> Statement of benefits: clear identification of the benefits anticipated from the digitisation;

> Statement of user needs and impacts: for example, how the digitised records will be used, accessed and impact on user;

DIGITISATION STANDARD
13
CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

> Statement of technical standards adopted: including format, compression and metadata;

> Equipment and resources to support the digitisation;

> Processes for the planning, control and execution of the digitisation, including those undertaken prior to, during and after digitisation;

> Quality assurance processes;

> Strategies for integrating the digitised image into work processes to support the business action taking place; and

> Strategies for the ongoing management of the digitised records for as long as they are required to be maintained.

Mandatory Requirement

1.0   All digitisation and digitisation processes **MUST** be planned, scoped and documented.

1.1 > AN APPROPRIATE DIGITISATION APPROACH <u>MUST</u> BE SELECTED, DOCUMENTED AND IMPLEMENTED

A number of approaches, or a combination of approaches, to digitisation can be adopted. The approach selected needs to be documented.

Decisions on each of the following areas will contribute to the overall approach employed:

> In-house or outsourced digitisation:

> In-house digitisation will allow an organisation to develop and acquire all the equipment and expertise required to digitise and process documents into their own systems. The alternative to this is to outsource the digitisation to a third party contracted to perform this service on behalf of the organisation.

> An ongoing process or a batch process:

> Batch processing is the collecting of source documents into sequences prior to digitising until documents have been accumulated in numbers sufficient to provide efficiencies of scale in undertaking the digitising process. The alternatives to this approach are on-demand imaging or digitising individual documents as they arrive in the digitising facility.

> Centralised or decentralised:

> Centralised digitisation involves the establishment of a single site for digitising in which all records to be processed are accumulated prior to digitising. The alternative to this is decentralised digitisation which involves placing multiple digitising stations in different locations throughout the organisation.

The way these decisions are approached will be different for business-process digitisation and for digitisation projects and will vary across organisations.

Regardless of which approach to digitisation is adopted, the following will apply:

> Quality control and assurance processes **MUST** be implemented regardless of the digitisation approach adopted; and

> The digitisation approach SHOULD be regularly reviewed for continuing relevance and cost effectiveness.

Mandatory Requirements

1.1   An appropriate digitisation approach **MUST** be selected, documented and implemented.

1.1.1   Quality control and assurance processes **MUST** be implemented regardless of the digitisation approach adopted.

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

14

1.2 > TECHNICAL SPECIFICATIONS ALIGNED TO THE DIGITISATION REQUIREMENTS <u>MUST</u> BE SELECTED, DOCUMENTED AND IMPLEMENTED

A large volume of technical standards associated with digitisation are available. Such standards include recommendations on:

> File formats;

> Resolution;

> Colour resolution or bit depth;

> Compression; and

> Colour Management.

Specific recommendations on technical specifications for digitisation are detailed in Appendix 5.

These standards are rapidly evolving, especially in the area of technical capacity of equipment to accommodate such standards. The primary consideration in adopting technical specifications is to ensure the legibility of the digitised image. The following basic criteria SHOULD be adhered to when selecting technical standards:

> The highest technical specifications that can be realistically supported SHOULD be incorporated into the digitisation process;

> Formats SHOULD be open source (that is, non-proprietary), have published technical specifications available in the public domain, or be widely deployed within the relevant sector;

> Formats SHOULD not contain embedded objects, or link out to external objects beyond the specific version of the format;

> Formats SHOULD be supported by many software applications and operating systems;

> Formats SHOULD be able to be read by utilising

a readily available viewing plug-in if the specific production software is not available to all users;

> A body of accessible and product-independent technical expertise SHOULD be available to support the decision;

> Adequate technical support SHOULD exist to enable ongoing maintenance and assurance of migration capability when necessary;

> Master copies SHOULD be created to the highest technical standards achievable;

> Master copies SHOULD be retained inviolable in secure storage;

> Derivative copies MAY be made in formats most convenient to the business requirements (e.g. thumbnails for distribution over the internet, etc.).

Mandatory Requirement

1.2   *Technical specifications aligned to the digitisation requirements* ***MUST*** *be selected, documented and implemented.*

1.3 > EQUIPMENT AND SOFTWARE ALIGNED TO THE DIGITISATION REQUIREMENTS <u>MUST</u> BE IMPLEMENTED

The quality of equipment and software used in digitising significantly affects the capability to support appropriate technical standards and, therefore, to ensure longevity of the digital image produced. Where destruction of the original source records is contemplated, agencies **MUST** be able to assert confidence in the long-term viability of those digitised images requiring on-going retention.

The following general rules apply:

> During the digitisation process, the use of techniques that enhance the digitised image to make the image have a more exact resemblance

DIGITISATION STANDARD

15

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

to the original, or reproduce the source records' appearance more faithfully, **MUST** be documented. Such procedures may, if not undertaken in routine and documented ways, enable the challenge that the image is not an authentic copy of the original. Such techniques include 'sharpening' and/or 'clipping' of highlights or shadows, 'blurring' to eliminate scratches, 'spotting' or 'de-speckling' to touch up specific areas of a digital image; and

⟩ Where software that is employed to manage digital images after capture enables additions of annotations to images, such as highlighting, stamps, redaction or addition of sticky notes, these annotations SHOULD be managed as overlays that do not change the actual image. Printing of the image SHOULD be possible with or without the annotations.

Readers available to users SHOULD support the display of the digital image in a manner, and to a quality, acceptable for the business being conducted. This can involve reviewing equipment specifications. For example, if the quality of the colour on a map is critical, the quality of the equipment used to render the image needs to support the capacity to retrieve and analyse this quality. If, on the other hand, it is only essential to be able to read the contents to gain the sense of the text, the quality of display could be appropriately less critical.

Response times required by end users can affect the storage media used for images. For example, if DVDs are used as an offline storage media, the retrieval or response time for a user to receive a copy of the requested image will be at least partly dependent on the hardware loading the relevant DVD onto the system. This timeframe might be too long for the purpose of the query.

Mandatory Requirements

1.3   *Equipment and software aligned to the digitisation requirements **MUST** be implemented.*

1.3.1   *Where destruction of the source record is contemplated, agencies **MUST** be able to assert confidence in the long-term viability of the images requiring on-going retention.*

1.3.2   *During the digitisation process, the use of techniques that enhance the digitised image to make the image have a more exact resemblance to the original, or reproduce the source records' appearance more faithfully, **MUST** be documented.*

2.0 ⟩ SYSTEMS TO SUPPORT MANAGEMENT OF THE DIGITAL OUTPUT OF DIGITISATION AS RECORDS <u>MUST</u> BE IN PLACE

The purpose of the digitisation will determine the functionality and system requirements necessary to manage the digital outputs.

Prior to digitisation, consideration of third party copyright or other constraints inherent in the record SHOULD be resolved.

Before the final version is reached, the digitising process could create a number of versions of a digitised record – for example, a raw file, one which has been enhanced to maximise resemblance to the original, or one to which quality assurance processes have been applied. Each digitisation process will specify the parameters of acceptable enhancements (see Requirement 2.1). The final output of the digitisation process **MUST** be regarded as the record for incorporation into the organisation's recordkeeping framework.

Many public offices and local authorities use electronic document and records management systems as their image control systems. Such systems SHOULD be assessed for functionality against Archives New Zealand's *Electronic Recordkeeping Systems Standard* (see the *Continuum* Resource Kit).

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

16

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

Where the digital image is to be used as a record in current or continuing business, the system governing the business process which will use the image SHOULD be integrated with records control systems. This ensures that the digital image will inherit the business controls and metadata associated with the business process, will be lodged within its business context and will have its authenticity enhanced by integration with the normal business systems.

Digitisation projects are often undertaken as a preservation measure to extend the life of fragile original records or to increase access by users to the informational content of records. Where the results of such projects are not immediately linked to a pre-existing business system, consideration of the software to be used to manage the images is necessary. Where the digital images represent a considerable investment of time and resources, serious consideration SHOULD be given to purchasing a management system to ensure appropriate control of processes such as identification, indexing, classification, security and access controls, rights management and preservation.

Mandatory Requirements

2.0   *Systems to support management of the digital output of digitisation as records **MUST** be in place.*

2.0.1   *The final output of the digitisation process **MUST** be regarded as the record for incorporation into the organisation's recordkeeping framework.*

2.1 › SOURCE RECORD PREPARATION GUIDELINES **MUST** BE DOCUMENTED AND IMPLEMENTED

The aim of producing digital images is to reproduce the original as faithfully as possible so that the digital image can act in place of the original, where it is required to act as evidence of business activities.

Source record preparation guidelines are not restricted to, but SHOULD include:

› An assessment of the source record's capability to sustain a digitisation process (e.g. paper quality, creasing, stapling, condition of microform jackets, attributes of the informational content – e.g. graphics);

› Methods of digitising non-standard source records (e.g. by digitising a photocopy of records on fragile or thin paper; creating a standard-sized document by using photocopy enlargement or reduction; enclosing fragile originals in plastic wallets or using specialty devices such as overhead scanners) and quality checks to ensure against data loss in any such processes;

› Methods for dealing with source records containing handwritten annotations, marginalia, white opaque paint or highlighted areas;

› Methods of distinguishing between original source records and photocopies;

› Guidance on what types of material need not be digitised as they are of only ephemeral or short-term value (guidance on this is available from Archives New Zealand's *General Disposal Authority 3: General Housekeeping Records*);

› Physical preparation for digitisation (e.g. careful staple removal, alignment of single pages, batching of like documents – size, technical settings, shared indexing fields);

> Processes for assigning links between associated documents to be regarded as a single item, so that the digitised image can faithfully represent the source record (e.g. a document and a 'post it' note attached; a document enclosing an attachment, etc.);

> Processes for assigning links between the source record and the digitised copy. Such links will usually be documented using identification protocols. In some applications barcode technology could be used to link paper and digitised versions;

> Procedures to enable checking and verification that all target source records have been included in the digitisation process;

> Principles governing the assembly of batches or groups of source records suitable for digitisation at the same time (e.g. size, colour, date order, document formats, orientation – portrait or landscape, single or double sided, etc.).

Mandatory Requirement

2.1   *Source record preparation guidelines* **MUST** *be documented and implemented.*

2.2 > ALL DIGITISED IMAGES <u>MUST</u> BE ASSIGNED METADATA TO DOCUMENT DIGITISING PROCESSES AND TO SUPPORT ONGOING BUSINESS PROCESSES[3]

Metadata attributed to, or associated with, images is an essential component in the management and retrieval of images. Wherever possible metadata SHOULD be inherited from existing systems.

Two types of metadata **MUST** be captured:

> Metadata specific to the particular image and the imaging process; and

> Metadata about the record, the business being transacted and the agents associated with the business.

The majority of this metadata can be automatically sourced from the software and hardware used to manage the digitising process. The manual attribution or application of metadata SHOULD be minimised.

Metadata can be embedded with the resource in header information, or can be managed in a separate system, or both, but in either case there has to be a direct relationship or association between them. That is, while metadata may reside in a separate system, it has to link directly to the resources. Metadata can also be encapsulated within the image format. Choices for storage of metadata SHOULD be made according to the principles outlined in ISO 23081 – 1: 2006, *Information and documentation – Records management processes – Metadata for records.*

Image-level metadata

Image-level metadata SHOULD be generated automatically at the point of digital capture direct from the digitisation equipment and SHOULD avoid manually-assigned data entry wherever possible.

In addition to the metadata inherited from recordkeeping capture and processes, or from indexing and searching metadata, image-level metadata **MUST** include:

> Unique digital image identifier;

> Date and time of digitisation;

> The name of the agent associated with the digitisation process (e.g. name of the outsourced bureau or name of the in-house operator);

---

3. Please note: This section has been designed not to be prescriptive, but to identify areas of critical importance for maintaining the authenticity and integrity of the digitised record. This enables organisations to adapt specific elements as required and to maximise the inheritance of data values from existing systems and equipment. It is designed to maximise automatic capture of metadata, not manual attribution.

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

18

> Capture device (hardware and software);

> Calibration settings (resolution, colour, dimensions, etc.); and

> Date of last calibration.

Additional image-level metadata MAY be assigned at the discretion of the public office or local authority.

Recommendations on naming protocols for digital images and directories are included in Appendix 6: File-naming Metadata Recommendations.

Business-process digitisation

Wherever possible the metadata controlling the business process and the recordkeeping functions associated with the business process SHOULD govern, and be inherited by, the specific digital image. This metadata **MUST** be incorporated into the organisation's electronic recordkeeping framework and be consistent with *ISO 23081 – 1: 2006, Information and documentation – Records management processes – Metadata for records*. Additional metadata describing the process of digitisation and specific characteristics of the digitised image **MUST** be included, as outlined above.

Digitisation projects

Where access to content is the primary driver, more intensive attention to providing additional indexing and search entry points is appropriate. The images can be managed as individual items, rather than as context-linked records, particularly if intended for web access by external users. Relevant elements from the *New Zealand Government Locator Service* (NZGLS) metadata SHOULD be associated with each image.

Mandatory Requirements

2.2   All digitised images **MUST** be assigned metadata to document digitising processes and to support ongoing business processes.

2.2.1   Two types of metadata **MUST** be captured:

• Metadata specific to the particular image and the imaging process; and

• Metadata about the record, the business being transacted and the agents associated with the business.

2.2.2   In addition to the metadata inherited from recordkeeping capture and processes, or from indexing and searching metadata, image-level metadata **MUST** include:

• Unique digital image identifier;

• Date and time of digitisation;

• The name of the agent associated with the digitisation process;

• Capture device;

• Calibration settings; and

• Date of last calibration.

2.2.3   In business-process digitisation, records metadata **MUST** be incorporated into the organisation's electronic recordkeeping framework and be consistent with ISO 23081 Metadata for records, Parts 1 and 2.

2.3 > QUALITY ASSURANCE PROCEDURES <u>MUST</u> BE DEFINED, DOCUMENTED AND IMPLEMENTED

Quality assurance to ensure the digital copy of the source record is a true and accurate copy is critical to being able to assert that the records possess integrity and are authentic.

Such quality assurance procedures SHOULD be documented and built into the ongoing operation of the digitising process, not considered simply a check on output only.

Quality assurance procedures SHOULD, at minimum, address the following issues:

> Any acceptable variations from normal procedures;

> Scanner operation quality control;

DIGITISATION STANDARD

19

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

- Verification that digital output matches the quantity of original record input;

- Extent and frequency of sampling of digitised images;

- Criteria for checking image quality;

- Frequency and criteria for checks on metadata;

- Processes for re-digitising; and

- Operator training.

Quality checking **MUST** be completed before the digitised images are accepted into a business process, or as a master copy in the case of digitisation projects. Quality checking **MUST** be complete before the destruction of the source records is considered.

The results of quality assurance processes and quality checks **MUST** be documented.

A review of quality procedures for digitising SHOULD be undertaken regularly to ensure that the procedures continue to meet the business requirements.

Appropriate training SHOULD be provided to all staff who create, manage or work with digitised records. Documentation on the level and the frequency of training provided to those staff involved with digitisation SHOULD be created and maintained.

For further information on quality assurance, see Appendix 7: Quality Assurance Recommendations, and Appendix 8: Recommended Staff Skill Sets.

Mandatory Requirements

2.3   Quality assurance procedures **MUST** be defined, documented and implemented.

2.3.1   Quality checking **MUST** be completed before the digitised images are accepted into a business process, or as a master copy in the case of digitisation projects.

2.3.2   Quality checking **MUST** be complete before the destruction of the source records is considered.

2.3.3   The results of quality assurance processes and quality checks **MUST** be documented.

2.4 > STORAGE MEDIA AND BACK-UP PROCEDURES <u>MUST</u> BE DEFINED, DOCUMENTED AND IMPLEMENTED

Maintaining technology-dependent records in networked storage is currently the storage strategy most likely to ensure the continuing accessibility of such records over time. However, digitised records can occupy significant storage space depending on the quality, resolution and compression ratios employed. Strategies for storage MAY include:

- A dedicated server or other digitised record storage solution;

- Writing the digitised records to magnetic tape;

- Writing the digitised records to WORM (write once, read many) storage media (e.g. a CD or DVD).

The requirements for storage include:

- That the digitised records **MUST** be unalterable in all storage media;

- That security and access controls for storage media **MUST** be capable of detecting and logging unauthorised attempts at access;

- That the retrieval times implicit in off-line storage SHOULD be acceptable for the business being conducted;

- Wherever possible digitised records sharing similar retention periods SHOULD be co-located to enable execution of destruction processes as required.

All digitised records, and their associated metadata, **MUST** be included in the organisation's back-up regime. Back-up procedures are designed to provide sufficient up-to-date copies of business records to be used in the event of loss or corruption of all or part of the data.

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

20

Back-up regimes SHOULD be documented and back-up copies maintained to a level of security that will ensure the authenticity of the records used in recovery situations.

All system failures **MUST** be documented, and use of the back-up copies for restoration purposes **MUST** be accompanied by verification testing to ensure the integrity of the restored records.

Information technology professionals often use the term 'archiving' to describe back-up regimes. For recordkeeping purposes conducting back-ups does not constitute an archiving or preservation strategy, it is a business continuity or disaster recovery precaution.

Mandatory Requirements

2.4   *Storage media and back-up procedures* **MUST** *be defined, documented and implemented.*

2.4.1   *Digitised records* **MUST** *be unalterable in all storage media.*

2.4.2   *Security and access controls for controlling storage media* **MUST** *be capable of detecting and logging unauthorised attempts at access.*

2.4.3   *All digitised records, and their associated metadata* **MUST** *be included in the organisation's back-up regime.*

2.4.4   *All system failures* **MUST** *be documented.*

2.4.5   *The use of the back-up copies for restoration purposes* **MUST** *be accompanied by verification testing to ensure the integrity of the restored records.*

3.0 > DISPOSAL OF ALL RECORDS MUST BE AUTHORISED AND DOCUMENTED

3.1 > DISPOSAL OF SOURCE RECORDS MUST BE AUTHORISED IN ACCORDANCE WITH RELEVANT LEGISLATION AND DOCUMENTED

Original source records cannot be disposed of without first meeting the requirements of any relevant statutes. This Standard sets out mandatory requirements that will satisfy the Chief Archivist that the requirements of the Public Records Act 2005 and Electronic Transactions Act 2002 have been met.

The full list of mandatory requirements that will ensure that a digitised record provides a reliable means of assuring the maintenance of the integrity of the source record are set out in Appendix 1.

Once the mandatory requirements are met and compliance has been signed-off by the Chief Executive using the certification form in Appendix 2, then public offices can implement disposal actions under the *General Disposal Authority: Digitised Original Source Records*. By meeting the mandatory requirements established in this standard, a local authority organisation can be confident that it meets the requirements of s 25(1) of the Electronic Transactions Act 2002, and may then legally dispose of the non-digital form of the protected record and retain the record in electronic form only.

All decisions to dispose of source records and disposal actions **MUST** be documented and this information **MUST** be accessible and be produced on request. Authorisation for destruction and the instance of destruction of the original source record SHOULD be documented in the metadata associated with the digitised record. Disposal actions SHOULD be documented and authorised by the relevant authority in an organisation

Digitised records **MUST** be (re)producible in their original format on request. Any person who requests information from an organisation is not obliged to accept it in electronic form without their consent, in accordance with section 16 of the Electronic Transactions Act 2002.

General Disposal Authority: Digitised Original Source Records

Public offices should note that the *General Disposal Authority: Digitised Original Source Records* includes exclusions and appraisal considerations before records can be routinely destroyed. (See the accompanying *General*

DIGITISATION STANDARD

21

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

Disposal Authority: Digitised Original Source Records document for full details of relevant records classes).

Summary of Exclusions

After being digitised, the following categories of records will **NOT** be authorised for disposal under the *General Disposal Authority: Digitised Original Source Records* (see accompanying document for full details):

> Records predating 1946;

> Source records digitised prior to the adoption of the *Digitisation Standard*, which do not comply with its requirements;

> Source records digitised in specific digitisation projects, where the source records are maintained in existing paper systems; or

> Items with significant intrinsic or cultural value, including items of high value to Māori. This category only applies where value is significant enough to warrant the extra expenditure of resources. It includes items with value due to their rarity, uniqueness or scarcity, artistic or aesthetic qualities, or items where notions of national or cultural identity require the original object as the true and correct representation.

Statement of criteria for disposing of digitised original source records

The accompanying *General Disposal Authority: Digitised Original Source Records* outlines the following criteria that public offices must consider when disposing of any digitised original source records:

> That the digitised record is an accurate and complete rendition of the original it replaces (including colour reproduction, where applicable);

> That appropriate quality assurance procedures and certifications for the digitising processes are defined, implemented and monitored routinely;

> That the digitised record is individually identified and linked to the context of its creation and use;

> That an appropriate management system is in place to ensure the ongoing business use of the digitised record;

> That the digitised record is relied upon in the normal conduct of business;

> That the necessary metadata about the digitised record is created and maintained;

> That the disposal program in place within the public office encompasses the business system(s) incorporating digitised records;

> That a migration and/or preservation strategy is defined, documented and implemented for electronic records, including digitised records;

> That legislative or regulatory requirements to maintain the records in a specific form are not breached;

> That no known or anticipated legal action will be imperilled by the destruction of the original; and

> That the risk of challenge to the authenticity and integrity of the digitised record has been assessed and considered and is acceptable to the organisation.

Mandatory Requirements

3.0   *Disposal of all records **MUST** be authorised and documented.*

3.1   *Disposal of source records **MUST** be authorised in accordance with relevant legislation, and documented (this information must be accessible and be produced on request).*

3.1.1   *Certification of compliance against the requirements of this Standard **MUST** be authorised by the Chief Executive, before the disposal of original source records can be considered.*

3.1.2   *Digitised records **MUST** be (re)producible in their original format on request.*

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

22

CONTINUUM
create & maintain
tāhuhu *te hanga me te tiaki*

3.2 > DISPOSAL OF THE DIGITISED RECORDS INCORPORATED INTO BUSINESS SYSTEMS <u>MUST</u> BE AUTHORISED AND DOCUMENTED

Where digitised records are incorporated into business systems as the official business record, they are regarded as the authentic 'official record' for evidential purposes. Digitised business records are subject to normal processes for disposal authorisations, established by the Chief Archivist under the Public Records Act 2005, in the same way as the original source records. A digitised record acting in place of an original source record **MUST** be retained for the authorised minimum period as established in the relevant disposal authority for the non-digital original source record. Once disposal is authorised, all extant copies of the digitised record SHOULD be destroyed (e.g. back-up copies). Authorisation of disposal of the digitised copy **MUST** be documented.

Disposal actions **MUST** be documented in the metadata associated with the record.

Mandatory Requirements

3.2   *Disposal of the digitised record incorporated into the business system **MUST** be authorised and documented.*

3.2.1   *Digitised records acting in place of an original source record **MUST** be retained for the authorised minimum period as established in the relevant disposal authority for the original source record.*

3.2.2   *Disposal actions **MUST** be documented in the metadata associated with the record.*

4.0 > LONG-TERM MANAGEMENT SYSTEMS, WHERE REQUIRED, FOR BOTH SOURCE AND DIGITISED RECORDS, <u>MUST</u> BE DOCUMENTED AND IMPLEMENTED

4.1 > SOUND MANAGEMENT SYSTEMS <u>MUST</u> BE IN PLACE FOR ORIGINAL SOURCE RECORDS UNTIL THEIR AUTHORISED DISPOSAL WITHIN A RECORDKEEPING FRAMEWORK

Source records **MUST** be managed appropriately until their authorised disposal within a recordkeeping framework. Where source records are retained for whatever reason other than quality control, or are not authorised for destruction, systematic controls **MUST** be applied. The digital image and the source record SHOULD be linked.

Where digitised records are incorporated into business systems and the source records are retained for reasons other than quality control checks, controls **MUST** be applied to their maintenance. The source records SHOULD be organised to maximise retrieval and to enable the efficient application of management and disposal processes. Day boxing, the process of accumulating source records in chronological order, or in their digitisation sequence, is not recommended. Day boxing is rarely suited to efficient management and disposal processes as it eliminates contextual linkages and mixes records intended for short-term retention with material subject to longer retention periods.

Digitisation projects rarely have the principle aim of enabling the destruction of the source record. After the digitisation process, source records SHOULD be returned to their original context and order, reflecting the processes of their creation and management in their original format. This enables the existing finding aids to continue to function as a retrieval tool for the records.

4.2 › MIGRATION AND/OR PRESERVATION STRATEGIES AND PROCESSES MUST BE DEFINED, DOCUMENTED AND IMPLEMENTED

Technology-dependent electronic records are inherently vulnerable to hardware, software and media obsolescence.

Digitised records SHOULD be included in the record-keeping framework adopted by the organisation to support the continuing existence of records for as long as they are required.

A migration strategy aims to transfer the record (and its associated metadata) into subsequent generations of software, hardware and media, in ways that preserve the authenticity and integrity of the record, and enables it to continue to be used as an authorised record of business.

All migration strategies **MUST** identify which record objects and associated metadata are required to enable agencies to continue to access the authentic digitised record. The original metadata **MUST** be migrated into any system replacing the one in which the digital image was originally managed.

Any decision not to migrate a digitised record into subsequent generations of software and hardware is a disposal action, and as such **MUST** be supported by either:

› Authorisation of disposal of the record and its associated metadata in accordance with the Public Records Act 2005; or

› Transfer of the record and its associated metadata into a dedicated preservation environment.

A dedicated preservation environment, either within an organisation or within Archives New Zealand, is one which identifies and documents the content and context of the record, including its linkages to other records and an event history relevant to processes applied to and uses made of the record. Typically, such an environment applies strict procedural controls on the format of both the record and its associated metadata. The preservation environment aims to extend the existence of the record once it is no longer required for active conduct of the business which created or managed it. The preservation environment **MUST** support the retrieval of the record for as long as required.

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

24

› MANDATORY APPENDICES

› APPENDIX 1: CHECKLIST OF MANDATORY REQUIREMENTS

This checklist is a statement of mandatory requirements to meet the Archives New Zealand Digitisation Standard. Compliance with these requirements will satisfy the Chief Archivist that the requirements of the Public Records Act and Electronic Transactions Act have been met by public offices and local authorities. Organisations **MUST** also certify that these criteria have been met (see Appendix 2) before disposing of original source records. In order for public offices to use the General Disposal Authority: Digitised Original Source Records meeting and certifying that these requirements have been met is a mandatory pre-condition.

Standard Section Reference	Requirement	No	Yes
Preliminary Requirement	The official record of action **MUST** be identified as the one which is used in the transaction of business. **NB** If the original source record is still the record of business action it cannot be disposed of. If the subsequent business action is taken on the digitised copy of the original source record it can be considered for disposal.	☐	☐
1.0	All digitisation and digitisation processes **MUST** be planned, scoped and documented.	☐	☐
1.1	An appropriate digitisation approach **MUST** be selected, documented and implemented.	☐	☐
1.1.1	Quality control and assurance processes **MUST** be implemented regardless of the digitisation approach adopted.	☐	☐
1.2	Technical specifications aligned to the digitisation requirements **MUST** be selected, documented and implemented.	☐	☐
1.3	Equipment and software aligned to the digitisation requirements **MUST** be implemented.	☐	☐
1.3.1	Where destruction of the source record is contemplated, agencies **MUST** be able to assert confidence in the long-term viability of the images requiring on-going retention.	☐	☐
1.3.2	During the digitisation process, the use of techniques that enhance the digitised image to make the image have a more exact resemblance to the original, or reproduce the source records' appearance more faithfully, **MUST** be documented.	☐	☐
2.0	Systems to support management of the digital output of digitisation as records **MUST** be in place.	☐	☐
2.0.1	The final output of the digitisation process **MUST** be regarded as the record for incorporation into the organisation's recordkeeping framework.	☐	☐
2.1	Source record preparation guidelines **MUST** be documented and implemented.	☐	☐
2.2	All digitised images **MUST** be assigned metadata to document digitising processes and to support ongoing business processes.	☐	☐

Standard Section Reference	Requirement	No	Yes
2.2.1	Two types of metadata **MUST** be captured: • Metadata specific to the particular image and the imaging process; and • Metadata about the record, the business being transacted and the agents associated with the business.	☐	☐
2.2.2	In addition to the metadata inherited from recordkeeping capture and processes, or from indexing and searching metadata, image-level metadata **MUST** include: • Unique digital image identifier; • Date and time of digitisation; • The name of the agent associated with the digitisation process; • Capture device; • Calibration settings; and • Date of last calibration.	☐	☐
2.2.3	In business-process digitisation, records metadata **MUST** be incorporated into the organisation's electronic recordkeeping framework and be consistent with *ISO 23081 Metadata for records, Parts 1 and 2*.	☐	☐
2.3	Quality assurance procedures **MUST** be defined, documented and implemented.	☐	☐
2.3.1	Quality checking **MUST** be completed before the digitised images are accepted into a business process, or as a master copy in the case of digitisation projects.	☐	☐
2.3.2	Quality checking **MUST** be complete before the destruction of the source records is considered.	☐	☐
2.3.3	The results of quality assurance processes and quality checks **MUST** be documented.	☐	☐
2.4	Storage media and back-up procedures **MUST** be defined, documented and implemented.	☐	☐
2.4.1	Digitised records **MUST** be unalterable in all storage media.	☐	☐
2.4.2	Security and access controls for controlling storage media **MUST** be capable of detecting and logging unauthorised attempts at access.	☐	☐
2.4.3	All digitised records, and their associated metadata **MUST** be included in the organisation's back-up regime.	☐	☐
2.4.4	All system failures **MUST** be documented.	☐	☐
2.4.5	The use of the back-up copies for restoration purposes **MUST** be accompanied by verification testing to ensure the integrity of the restored records.	☐	☐
3.0	Disposal of all records **MUST** be authorised and documented.	☐	☐
3.1	Disposal of source records **MUST** be authorised in accordance with the relevant legislation, and documented (this information must be accessible and be produced on request).	☐	☐
3.1.1	Certification of compliance against the requirements of this Standard **MUST** be authorised (See Appendix 2 for Compliance Certification). Public offices may then apply the *General Disposal Authority: Digitised Original Source Records*. Local authorities may dispose of original source records having demonstrated compliance with the Electronic Transactions Act, 2002.	☐	☐

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA

JANUARY 2007

26

Standard Section Reference	Requirement	No	Yes
	NB If all the mandatory requirements of this standard are not met an *ad hoc* disposal authority for public records **MUST** be sought from the Chief Archivist before any disposal action.		
3.1.2	Digitised records **MUST** be (re)producible in their original format on request.	☐	☐
3.2	Disposal of the digitised record incorporated into the business system **MUST** be authorised and documented.	☐	☐
3.2.1	Digitised records acting in place of an original source record **MUST** be retained for the authorised minimum period as established in the relevant disposal authority for the original source record.	☐	☐
3.2.2	Disposal actions **MUST** be documented in the metadata associated with the record.	☐	☐
4.0	Long-term management systems, where required, for both source and digitised records, **MUST** be documented and implemented.	☐	☐
4.1	Sound management systems **MUST** be in place for original source records until their authorised disposal within a recordkeeping framework.	☐	☐
4.1.1	Where digitised records are incorporated into business systems and the source record retained for reasons other than quality control checks, systematic controls **MUST** be applied to their maintenance.	☐	☐
4.2	Migration and/or preservation strategies and processes **MUST** be defined, documented and implemented.	☐	☐
4.2.1	All migration strategies **MUST** identify which record objects and associated metadata are required to enable organisations to continue to access the authentic digitised record.	☐	☐
4.2.2	The original metadata **MUST** be migrated into any system replacing the one in which the digital image was originally managed.	☐	☐
4.2.3	Any decision not to migrate a digitised record into subsequent generations of software/hardware is a disposal action and **MUST** be supported by either: • Authorisation of disposal of the record and its associated metadata in accordance with the Public Records Act 2005; or • Transfer of the record and its associated metadata into a dedicated preservation environment.	☐	☐
4.2.4	The preservation environment **MUST** support the retrieval of the record for as long as required.	☐	☐

DIGITISATION STANDARD

27

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA

JANUARY 2007

⟩ APPENDIX 2: COMPLIANCE CERTIFICATION

*This certificate of compliance **MUST** be signed by the Chief Executive of a public office prior to the disposal of original source records under the* General Disposal Authority: Digitised Original Source Records *issued by the Chief Archivist. The purpose of this certification is to enable public offices to qualify to implement the* General Disposal Authority: Digitised Original Source Records*.*

For local authorities it enables them to show they have met the requirements of s 25(1) of the Electronic Transactions Act 2002. Local authorities may then dispose of their original source records.

It will also act as clear evidence of senior management responsibility for digitised records, to ensure that there is organisational confidence in the business systems incorporating digitised images. It is an acknowledgment of the complex management requirements necessary to ensure that public records or local authority protected records in digitised forms are maintained with integrity and authenticity for as long as required.

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

28

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

Certificate of Compliance with Archives New Zealand's Digitisation Standard

I .... Chief Executive of

(Name)

....

(Name of organisation)

confirm that the organisation has demonstrated to my satisfaction that the mandatory requirements of Archives New Zealand's *Digitisation Standard* as set out in Appendix 1 (summarised below) have been and can continue to be met.

Public office Chief Executives – I understand that meeting this Standard is a pre-condition for the organisation implementing the **General Disposal Authority: Digitised Original Source Records** issued by the Chief Archivist under s 20(1) of the Public Records Act 2005.

Local authority Chief Executives – I understand that meeting this Standard provides assurance that the organisation's digitisation procedures meet the requirements of s 25(1) of the Electronic Transactions Act 2002 and that the original paper or non-electronic format records may now be destroyed.

Summary of disposal requirements

> *That the digitised record is an accurate and complete rendition of the original it replaces (including colour reproduction, where applicable);*

> *That appropriate quality assurance procedures and certifications for the digitising processes are defined, implemented and monitored routinely;*

> *That the digitised record is individually identified and linked to the context of its creation and use;*

> *That an appropriate management system is in place to ensure the ongoing business use of the digitised record;*

> *That the digitised record is relied upon in the normal conduct of business;*

> *That the necessary metadata about the digitised record is created and maintained;*

> *That the disposal program in place within the public office or local authority encompasses the business system(s) incorporating digitised records;*

> *That a migration and/or preservation strategy is defined, documented and implemented for electronic records, including digitised records;*

> *That legislative or regulatory requirements to maintain the records in a specific form are not breached;*

> *That no known or anticipated legal action will be imperilled by the destruction of the original; and*

> *That the risk of challenge to the authenticity and integrity of the digitised record has been assessed and considered and is acceptable to the organisation.*

(The full list of mandatory requirements to ensure that a digitised record provides a reliable means of assuring the maintenance of the integrity of the source record is set out in Archives New Zealand's Digitisation Standard: Appendix 1.)

DIGITISATION STANDARD

29

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

GUIDANCE APPENDICES

APPENDIX 3: CHECKLIST OF BEST PRACTICE RECOMMENDATIONS

This checklist is a statement of the best practice recommendations outlined in the Standard.

Standard Section Reference	Best Practice Guidance	No	Partly	Yes
Preliminary Considerations	• The rationale for digitising SHOULD be carefully aligned to a costed business case geared at improving the organisation's ability to carry out its functions.	☐	☐	☐
	• The business case SHOULD clearly outline the benefits and anticipated business or cost efficiencies.	☐	☐	☐
	• The business case SHOULD involve appropriate project budgets, resource commitments and be realistically costed.	☐	☐	☐
	• Derivative versions, where required, SHOULD be made during the digitisation process.	☐	☐	☐
	• Master copies SHOULD be made available for the creation of subsequent derivative images, where necessary.	☐	☐	☐

Standard Section Reference	Best Practice Guidance	No	Partly	Yes
1. Planning and Processes	The project documentation SHOULD include: • Scope definition: with clear identification of business drivers, objectives, scale, size and constraints of the project;	☐	☐	☐
	• Statement of the purpose and expected uses of the digitised records, illustrated if necessary with examples;	☐	☐	☐
	• Statement of benefits: clear identification of the benefits anticipated from the digitisation;	☐	☐	☐
	• Statement of user needs and impacts;	☐	☐	☐
	• Statement of technical standards adopted: including format, compression and metadata;	☐	☐	☐
	• Equipment and resources to support the digitisation;	☐	☐	☐
	• Processes for the planning, control and execution of the digitisation, including those undertaken prior to, during and after digitisation;	☐	☐	☐

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA

JANUARY 2007

30

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

Standard Section Reference	Best Practice Guidance	No	Partly	Yes
1. Planning and Processes continued	• Quality assurance processes;	☐	☐	☐
	• Strategies for integrating the digitised image into work processes to support the business action taking place; and	☐	☐	☐
	• Strategies for the ongoing management of the digitised records for as long as they are required to be maintained.	☐	☐	☐
	The digitisation approach SHOULD be regularly reviewed for continuing relevance and cost effectiveness.	☐	☐	☐
	• The highest technical specifications that can be realistically supported SHOULD be incorporated into the digitisation process.	☐	☐	☐
	• Formats SHOULD be open source, have published technical specifications available in the public domain, or be widely deployed within the relevant sector.	☐	☐	☐
	• Formats SHOULD not contain embedded objects, or link out to external objects beyond the specific version of the format.	☐	☐	☐
	• Formats SHOULD be supported by many software applications and operating systems.	☐	☐	☐
	• Formats SHOULD be able to be read by utilising a readily-available viewing plug-in if the specific production software is not available to all users.	☐	☐	☐
	• A body of accessible and product-independent technical expertise SHOULD be available to support the decision.	☐	☐	☐
	• Adequate technical support SHOULD exist to enable ongoing maintenance and assurance of migration capability when necessary.	☐	☐	☐
	• Master copies SHOULD be created to the highest technical standards achievable.	☐	☐	☐
	• Master copies SHOULD be retained inviolable in secure storage.	☐	☐	☐
	Where software that is in place to manage digital images after capture enables additions of annotations to images, these annotations SHOULD be managed as overlays that do not change the actual image.	☐	☐	☐
	Printing of the image SHOULD be possible with or without the annotations.	☐	☐	☐
	Readers available to users SHOULD support the display of the digital image in a manner, and to a quality, acceptable for the business being conducted.	☐	☐	☐
2. Management Systems	Prior to digitisation, consideration of third party copyright or other constraints inherent in the record SHOULD be resolved.	☐	☐	☐
	Where an electronic document and records management system is used as the image control system, it SHOULD be assessed for functionality against Archives New Zealand's Electronic Recordkeeping Systems Standard.	☐	☐	☐
	Where the digital image is to be used as a record in current or continuing business, the system governing the business process which will use the image SHOULD be integrated with records control systems.	☐	☐	☐

DIGITISATION STANDARD
31

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

Standard Section Reference	Best Practice Guidance	No	Partly	Yes
2. Management Systems continued	Where digitisation is undertaken for preservation purposes or primarily to enhance external access to information, care is exercised to ensure that the appropriate systematic controls are in place to manage the digital image. Consideration SHOULD be given to purchasing a management system to ensure appropriate control of processes such as identification, indexing, classification, security and access controls, rights management and preservation.	☐	☐	☐
	Source record preparation guidelines SHOULD include:			
	• An assessment of the source records' capability to sustain a digitisation process;	☐	☐	☐
	• Methods of digitising non-standard source records;	☐	☐	☐
	• Quality checks to ensure against data loss in digitisation;	☐	☐	☐
	• Methods for dealing with source records containing handwritten annotations, marginalia, white opaque paint, or highlighted areas;	☐	☐	☐
	• Methods of distinguishing between original source records and photocopies;	☐	☐	☐
	• Guidance on what types of material need not be digitised as they are only of ephemeral or short-term value;	☐	☐	☐
	• Physical preparation for digitisation;	☐	☐	☐
	• Processes for assigning links between associated documents to be regarded as a single item, so that the digitised image can faithfully represent the source record;	☐	☐	☐
	• Processes for assigning links between the source record and the digitised copy;	☐	☐	☐
	• Procedures to enable checking and verification that all target source records have been included in the digitisation process; an authorisation for destruction and the instance of destruction of the original source record SHOULD be documented in the metadata associated with the digitised record.	☐	☐	☐
	Where possible, metadata attributed to, or associated with, images SHOULD be inherited from existing systems.	☐	☐	☐
	The manual attribution or application of metadata SHOULD be minimised.	☐	☐	☐
	Choices for storage of metadata SHOULD be made according to the principles outlined in ISO 23081 Metadata for records.	☐	☐	☐
	Image-level metadata SHOULD be generated automatically at the point of digital capture direct from the digitisation equipment and SHOULD avoid manually-assigned data entry wherever possible.	☐	☐	☐
	In business-process digitisation, wherever possible the metadata controlling the business process and the recordkeeping functions associated with the business process SHOULD govern, and be inherited by, the specific digital image.	☐	☐	☐
	Particularly if intended for web access by external users, relevant elements from NZGLS metadata SHOULD be associated with each image	☐	☐	☐

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA

JANUARY 2007

32

Standard Section Reference	Best Practice Guidance	No	Partly	Yes
2. Management Systems continued	Quality assurance procedures SHOULD be documented and built into the ongoing operation of the digitisation process.	☐	☐	☐
	Quality assurance procedures SHOULD, at minimum, address the following issues:			
	• Any acceptable variations from normal procedures;	☐	☐	☐
	• Scanner operation quality control;	☐	☐	☐
	• Verification that digital output matches the quantity of original record input;	☐	☐	☐
	• Extent and frequency of sampling of digitised images;	☐	☐	☐
	• Criteria for checking image quality;	☐	☐	☐
	• Frequency and criteria for checks on metadata;	☐	☐	☐
	• Processes for re-digitising; and	☐	☐	☐
	• Operator training.	☐	☐	☐
	A review of quality procedures for digitising SHOULD be undertaken regularly to ensure that the procedures continue to meet business requirements.	☐	☐	☐
	Appropriate training SHOULD be provided to all staff who create, manage or work with digitised records.	☐	☐	☐
	Documentation on the level and the frequency of training provided to all levels of staff involved with digitisation SHOULD be created and maintained.	☐	☐	☐
	Retrieval times implicit in offline storage SHOULD be acceptable for the business being conducted.	☐	☐	☐
	Wherever possible digitised records sharing similar retention periods SHOULD be co-located to enable execution of destruction processes as required.	☐	☐	☐
	Back-up regimes SHOULD be documented.	☐	☐	☐
	Back-up copies SHOULD be maintained to a level of security that will ensure the authenticity of the records used in recovery situations.	☐	☐	☐
3. Disposal Processes	Authorisation for destruction and the instance of destruction of the original source record SHOULD be documented in the metadata associated with the digitised record	☐	☐	☐
	Disposal actions SHOULD be documented and authorised by the relevant authority in an organisation.	☐	☐	☐
	Once authorised for destruction, all extant copies of the digitised record SHOULD be destroyed (e.g. back-ups).	☐	☐	☐

DIGITISATION STANDARD

33

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

Standard Section Reference	Best Practice Guidance	No	Partly	Yes
4. Long-term Management	The digital image and the source record SHOULD be linked.	☐	☐	☐
	Where digitised records are incorporated into business systems and the source records are retained for reasons other than quality control checks, the source records SHOULD be organised to maximise retrieval and to enable efficient management of retention and disposal processes.	☐	☐	☐
	In digitisation projects, source records SHOULD be returned to their original context and order after the digitisation process.	☐	☐	☐
	Digitised records SHOULD be included in the framework adopted by the organisation to support the continuing existence of records for as long as they are required.	☐	☐	☐

See the following appendices for specific guidance on:

> *Recommended Technical Specifications;*

> *File-naming Metadata Recommendations;*

> *Quality Assurance Recommendations; and*

> *Recommended Staff Skill Sets.*

DIGITISATION STANDARD
CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

34

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

> ## APPENDIX 4: ISSUES TO CONSIDER WHEN ASSESSING THE VIABILITY OF DIGITISATION

This appendix is for guidance. It provides a set of questions to assist organisations in assessing the viability of digitisation for original source records. It relates to the 'Preliminary Considerations' section of the 'Requirements' element of the Standard.

Business-process digitisation

> Will digitising enable the organisation to manage all incoming and outgoing transactions electronically, thus eliminating a hybrid paper/electronic system?

> Are the records required for transmission across internal or external networks?

> Do the records need to be available to all staff at centralised and remote locations?

> Are all incoming paper documents to be digitised – that is, is the process to be consistently applied, or is a selection process required?

> Is the digitising taking place prior to action being initiated on the item?

> Is there guidance for staff on what not to capture (following the guidelines in the Archives New Zealand's General Disposal Authority 3: General Housekeeping Records)?

> Will the digital images be incorporated into an established business or recordkeeping system application (e.g. EDRMS)?

> Will all further action arising from the digital record be documented in the business system or recordkeeping system (e.g. EDRMS) application?

> Can adequate controls be established to ensure the reliability of the digital image?

> Have the records to be digitised been appraised for retention and disposal scheduling?

Digitisation projects

> Do the records document a process or function that continues to be done?

> Is the equivalent information in the records now recorded electronically?

> Do the records contain a rich information source, required by current business processes?

> Are the records a coherent and complete set?

> Are the records required for constant or continuing access?

> Are the records constantly or continuously requested for use by people either internal or external to the organisation?

> Is the information in the records enhanced by being available in digital form (e.g. additional indexing, sorting capability)?

> Are the records in a form which is suited to digitisation?

> Are the items in a standard format which will enable single equipment settings to be applied?

> Will this act as a precedent for future projects?

> Have the records been appraised as being required for long-term retention?

> Is there legislation that prevents these records being held only in electronic form?

> Are the records fragile or subject to damage through repeated physical handling which will be minimised after digitisation?

35

DIGITISATION STANDARD
CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

⟩ APPENDIX 5: RECOMMENDED TECHNICAL SPECIFICATIONS

This appendix contains recommended technical specifications for digitisation. The recommendations are the least prescriptive commonly-used industry specifications − that is, public offices and local authorities can exceed these recommendations.

*These recommendations are not mandatory. That is, where justifiable, public offices and local authorities can adopt other technical standards. The primary consideration in adopting technical standards is to ensure the legibility of the digitised image. This appendix relates to Requirement 1.2: Technical specifications aligned to the digitisation requirements **MUST** be selected, documented and implemented.*

Document Type	Resolution	Bit Depth	File Format	Compression
Text only, black and white	Minimum 300ppi	1 bit (bi-tonal)	TIFF PDF/A[4] containing TIFF or JPEG 2000[5]	Lossless compression
Documents with watermarks, grey shading, grey graphics	Minimum 600 ppi	8 bit greyscale	TIFF JPEG 2000 PDF/A containing TIFF or JPEG 2000	Lossless compression
Documents with discrete colour used in text or diagrams	Minimum 600 ppi	Minimum: 8 bit colour	TIFF JPEG 2000 PDF/A containing TIFF or JPEG 2000	Lossless compression
Black and white photographs	Sufficient to provide > 3000 pixels across long dimensions	8 bit greyscale	TIFF JPEG 2000 PDF/A containing TIFF or JPEG 2000	Lossless compression
Colour photographs	Sufficient to provide > 3000 pixels across long dimensions	24 bit colour	TIFF JPEG 2000 PDF/A containing TIFF or JPEG 2000	Lossless compression
Black and white negatives	Sufficient to provide > 3000 pixels across long dimensions	8 bit greyscale or 24 bit colour	TIFF JPEG 2000 PDF/A containing TIFF or JPEG 2000	Lossless compression
Colour negatives and transparencies	Sufficient to provide > 3000 pixels across long dimensions	24 bit colour	TIFF JPEG 2000 PDF/A containing TIFF or JPEG 2000	Lossless Compression
Microforms	See note below			

Note: When scanning microforms, the approach should be to emulate the methods detailed above consistent with the source document on the (typically greyscale) microform − that is, to produce a minimum resolution of 600 ppi (in relation to the original document). However, this may vary for textual records to focus more on creating digital images with reasonable or good legibility. JPEG 2000 and PDF/A are recommended formats.

4. PDF/A is a constrained version of PDF version 1.4 with various proprietary fonts and formats removed, issued as ISO 19005-1:2004.

5. JPEG 2000 is defined in ISO 15444-1:2000.

DIGITISATION STANDARD
CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

36

CONTINUUM

create & maintain
tāhuhu te hanga me te tiaki

Recommended Technical Specifications Glossary

Area	Issues
File formats	**Categories of file formats are:** **Raster:** also known as bit-mapped formats, where images take the form of a grid or matrix with each picture element (pixel) having a unique location and independent colour value. Examples are TIFF, JPG/JPEG, GIF, PNG; **Vector:** also known as object oriented, based on a set of mathematical instructions typically used by drawing programs to construct an image – not of relevance to digitisation which will use raster formats; and **Encoding:** or metafiles which may contain either vector or raster images. Such formats enable the contents to be consistently displayed and used across different computer programs and operating systems. Typically, they support internal metadata, support multi-page images and enable security management. Examples include Adobe PDF and TIFF.
Resolution	A measure of the ability to capture detail in the original work, often quantified in pixels per inch (ppi). The optimum resolution depends on the nature of the documents being scanned. Photographs, for example, require much greater resolution than text-based documents. **ppi:** (pixels per inch) is a measurement of resolution for computer display. **dpi:** (dots per inch) is often used interchangeably with ppi, but actually refers specifically to measurement of the resolution for computer printers.
Colour resolution or bit depth	A measure of the number of colours (or degree of brightness, in greyscale images) available to represent the colours (or shades of grey) in the original document. For example: **1 Bit, Black and White or line art:** only black and white pixels; **Greyscale:** black and white in addition to a range of intermediate greys, requiring 8 bits to describe each pixel; **8 Bit Colour:** uses a palette of 256 colours; **24 Bit Colour:** a resolution that enables storage of 8 bits of information describing the red, green and blue components of every pixel, thus enabling a much greater palette of colours; and **36-48 Bit RBG Colour:** uses an extended colour space, creating a much larger file, and requiring storage in formats that explicitly support this colour depth (TIFF or PNG).
Compression	Algorithms designed to reduce the size of the image for storage or transmission. Multiple options exist but decisions should be made on the characteristics of the document to be imaged. The two major categories of compression are: **Lossy:** where information is removed from the stored information during the compression process; and **Lossless:** where no information is irretrievably lost and where the decompressed object will always appear exactly the same as the original. Examples include LZW or ZIP lossless compression with TIFF files. Newer forms of compression emerging are fractals and wavelets.
Colour management	Means of attempting to ensure that the image looks the same across a range of different output devices. Monitors and printers typically use different colour spectrum. The standard for colour representation is the ICC colour management system, which uses a standardised and known 'colour space' based on the human eye and then compares all devices to the known standard.

DIGITISATION STANDARD

37

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

APPENDIX 6: FILE-NAMING METADATA RECOMMENDATIONS [6]

*This appendix is for guidance. It outlines a number of best practices in relation to determining a file naming protocol, particularly for digital images. It relates to Requirement 2.2: All digitised images **MUST** be assigned metadata to document digitising processes and to support ongoing business processes.*

A file naming scheme SHOULD be established prior to capture. The development of a file naming scheme should take into account whether the identifier requires machine or human-indexing (or both — in which case, the image may have multiple identifiers). File names can either be meaningful (such as the adoption of an existing identification scheme which correlates the digital file with the source material), or non-descriptive (such as a sequential numerical string). Meaningful file names contain metadata that is self-referencing; non-descriptive file names are associated with metadata stored elsewhere that serves to identify the file. In general, smaller-scale projects may design descriptive file names that facilitate browsing and retrieval; large-scale projects may use machine-generated names and rely on a database for sophisticated searching and retrieval of associated metadata.

In general, file names SHOULD:

- Be unique;

- Be consistently structured;

- Take into account the maximum number of items to be digitised and reflect that in the number of digits used (if following a numerical scheme);

- Use leading zeros to facilitate sorting in numerical order (if following a numerical scheme);

- Avoid using spaces within the file name, using underline character as an alternative;

- Avoid an overly complex or lengthy naming scheme that is susceptible to human error during manual input;

- Restrict the length of file names to under 30 characters to avoid potential problems with migration between different systems;

- Use lowercase characters and file extensions;

- Use numbers and/or letters but not characters such as symbols or spaces that could cause complications across operating platforms; and

- Record metadata embedded in file names (such as scan date, page number, etc.) in another location in addition to the file name. This provides a safety net for moving files across systems in the future, in the event that they have to be renamed. In particular, sequencing information and major structural divisions of multi-part objects SHOULD be explicitly recorded in the structural metadata and not only embedded in filenames.

Directory structure

Regardless of file name, files will likely be organised in some kind of file directory system that will link to metadata stored elsewhere in a database. Production master files might be stored separately from derivative files, or directories MAY have their own organisation independent of the image files, such as folders arranged by date or classification structure, or they MAY replicate the physical or logical organisation of the originals being scanned.

The files themselves can also be organised solely by directory structure and folders rather than embedding meaning in the file name. This approach generally works

---

6. This section is sourced from: US National Archives and Records Administration, *'Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files — Raster Images'*, June 2004.

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

38

well for multi-page items. Images are uniquely identified and aggregated at the level of the logical object (i.e. a document, a record, a file/folder, etc.), which requires that the folders or directories be named descriptively. The file names of the individual images themselves are unique only within each directory, but not across directories. For example, book 0001 contains image files 001.tif, 002.tif, and 003.tif. Book 0002 contains image files 001.tif, 002.tif, and 003.tif. The danger with this approach is that if individual images are separated from their parent directory, they will be indistinguishable from images in a different directory.

Versions

For various reasons, a single scanned object may have multiple but differing versions associated with it (for example, the same image prepared for different output intents; versions with additional edits; layers, or alpha channels that are worth saving; versions scanned on different scanners, scanned from different original media, or scanned at different times by different scanner operators). Ideally, the description and intent of different versions should be reflected in the metadata; but, if the naming convention is consistent, distinguishing versions in the file name will allow for quick identification of a particular image. Like derivative files, this often implies the application of a qualifier to part of the file name. The reason to use qualifiers rather than entirely new names is to keep all versions associated with a logical object under the same identifier. An approach to naming versions should be well thought out; adding 001, 002 to the base file name to indicate different versions is an option; however, if 001 and 002 already denote page numbers, a different approach will be required.

Naming derivative files

The file naming scheme should also take into account the creation of derivative image files made from the production master files. In general, derivative file names are inherited from the production masters, usually with a qualifier added on to distinguish the role of the derivative from other files (i.e. "p" for published version, "t" for thumbnail). Derived files usually imply a change in image dimensions, image resolution, and/or file format from the production master. Derivative file names do not have to be descriptive as long as they can be linked back to the production master file.

For derivative files intended primarily for web display, one consideration for naming is that images may need to be cited by users in order to retrieve other higher-quality versions. If so, the derivative file name should contain enough descriptive or numerical meaning to allow for easy retrieval of the original or other digital versions.

DIGITISATION STANDARD

39

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

> APPENDIX 7: QUALITY ASSURANCE RECOMMENDATIONS [7]

This appendix is for guidance. It covers a number of areas relevant to quality assurance, including guidance on tests, standards and frequency of checking. It is intended as an aid to assist public offices and local authorities in devising their own quality assurance processes. It relates to Requirement 2.3: Quality assurance procedures **MUST** be defined, documented and implemented.

Scanner Operation Quality Control

Scanners SHOULD be tested periodically to monitor their operational performance and check that operating performances are within agreed tolerances as set out by existing standards[8] and benchmarks. Results of previous tests SHOULD be used as benchmarks for system performance over time.

Simple preventive measures such as ensuring the cleanliness and routine servicing of equipment SHOULD be implemented.

Validation of Output

The equipment SHOULD routinely record the number of discrete documents and the number of documents comprising a record (more than one page bundled) that were scanned during a session.

Sampling

The frequency of sampling should be determined according to the system usage and expected or anticipated deterioration periods. Advice from the system vendor may assist in determining the frequency period. Initially, it MAY be appropriate to scan a test target every few thousand pages. However, once benchmarks have been established

and equipment and processes stabilised, this MAY be reduced to a random sampling of between 5 and 10%.

Sample sets

Assemble a sample set of source documents for the purposes of evaluating scanner results against agreed quality criteria. These documents SHOULD be representative of the records to be scanned and SHOULD include examples of source documents whose quality is poor relative to the majority of the sample documents.

Test target sheets and detailed tests included in ISO 12653, Parts 1 and 2, MAY be used for black and white scanning.

Where digitisation involves colour as an intrinsic part of the record, implementers should consider including a standard colour sheet with the image (such target sheets are commercially available from manufacturers such as Fuji or Kodak, and comply with the ISO 12641 standard). Colour calibration, matching and profiling SHOULD be monitored where colour scanning is used.

Quality Criteria for images

Quality criteria for images SHOULD include consideration of overall legibility:

> smallest detail legibly captured (e.g. smallest type size for text; clarity of punctuation marks, including decimal points);

> completeness of detail (e.g. acceptability of broken characters, missing segments of lines);

> dimensional accuracy compared with the original;

> scanner-generated speckle (i.e. speckle not present on the original);

---

7. Much of the material in this section derives from ISO/TR 15801: 2004, *Electronic imaging – Information stored electronically – Recommendations for trustworthiness and reliability*, and *Technical Advisory Service for Images (TASI), In-depth Report: 'Quality Assurance'*. http://www.tasi.ac.uk/advice/creating/qassurance.html.

8. ISO 12653 – 1: 2000, Electronic imaging – *Test targets for the black and white scanning of office documents, Part 1 – characteristics*;

ISO 12653-2: 2000, *Electronic imaging – Test targets for black and white scanning of office documents, Part 2 – methods of use*; and

ISO/TR 15801: 2004, *Electronic imaging – Information stored electronically – Recommendations for trustworthiness and reliability.*

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

40 >

- completeness of overall image area (i.e. missing information at the edges of the image area);

- density of solid black areas; and

- colour fidelity.

Metadata

Procedures SHOULD specify the checks to be implemented to assess metadata quality assigned to images[9].

Issues which may be considered in the quality checking of metadata are:

- Adherence to standards set by institutional policy or by the requirements of the digitisation project;

- Procedures for accommodating images with incomplete metadata;

- Relevancy and accuracy of metadata;

- Grammar – check for correct grammar, spelling and punctuation, especially for manually-keyed data;

- Consistency in the creation of metadata and in interpretation of metadata;

- Evaluation of the usefulness of the metadata being collected;

- Synchronisation of metadata stored in more than one location – procedures SHOULD be in place to make sure metadata is updated in a synchronised manner across more than one location (e.g. information related to the image might be stored in the TIFF header, the digital asset management system, and other databases); and

- Completeness of metadata - all mandatory fields should be complete.

Of particular importance are:

- Verifying the accuracy of the file identifier. File names SHOULD consistently and uniquely identify both the digital resource and the metadata record (if it exists independently of the image). File identifiers will likely exist for the metadata record itself in addition to identifiers for the digitised resource, which may embed information such as page or piece number, date, project or institution identifier, among others. Information embedded in file identifiers for the resource should parallel metadata stored in a database record or header. Identifiers often serve as the link from the file to information stored in other databases and have to be accurate to bring together distributed metadata about a resource. Verification of identifiers across metadata in disparate locations should be made;

- Verifying the correct sequence and completeness of multi-page items. Pages SHOULD be in the correct order with no missing pages. If significant components of the resource are recorded in the metadata, such as the presence of attachments, documents with identifiable chapters or multi-page records, they should match up with the actual image files. A convention for describing these views SHOULD be followed and SHOULD match with the actual image files.

Documentation

Quality control data (such as logs, reports, decisions) SHOULD be captured in a formal system and SHOULD become an integral part of the image metadata at the file or the project level. This data may have long-term value that could have an impact on future preservation decisions.

---

9. ISO 23081 – 1: 2006, *Information and documentation – Records management processes – Metadata for records, Part 1 – Principles;*

ISO 23081 – 1: 2006, *Information and documentation – Records management processes – Metadata for records, Part 2 – Implementation issues.*

DIGITISATION STANDARD

41

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

Processes for Re-digitisation

If more than 1% of the total number of images and associated metadata examined in a randomly selected sampling are found to be defective for any of the reasons listed above, the entire output since the last quality check SHOULD be re-inspected. Any specific errors found in the random sampling and any additional errors found in the re-inspection SHOULD be corrected. If less than 1% of the batch is found to be defective, then only the specific defective images and metadata that are found should be redone.

Common Faults

Broadly speaking, quality faults can be categorised as 'implementation faults', 'process faults' or 'operator faults'. Implementation faults are those that can be avoided, providing appropriate procedural controls are in place to guide the digitisation. Process faults are normally out of the control of the operator and need to be addressed by a supervisor to the process. Operator faults are the day-to-day faults that are made by the operator as they work.

Implementation Faults

There are a number of faults that can be avoided with appropriate specification of procedures to guide the implementation. These include:

> Dirty originals;

> Incorrect file-size and format, where files are made to wrong size or with wrong choice of file format;

> Compression, where files are made with an inappropriate type or level of compression.

Process Faults

There are a wide variety of process faults that can be caused by many problems within the workflow. These problems can include:

> Incomplete or inaccurate specifications or process documentation;

> Faulty capture hardware (incorrectly calibrated and characterised devices);

> Faulty software (inaccurate image processing or faulty image links within database);

> Incorrectly established colour management systems;

> Low quality original data (either non-digital surrogates or legacy digital image files); and

> Inaccurate source metadata.

Operator Faults

These faults are caused by some form of operator error within the workflow and can include:

> Basic capture faults;

> Cropping that has cut into the image, is too loose, or is uneven;

> Orientation of the image is the wrong way around, or upside down;

> Exposure of the image is too light or too dark;

> Focus, where the image is out of focus;

> Daily calibration, where the capture device has not been calibrated;

> Basic image processing faults;

> File Optimisation Faults, where incorrect adjustments are made to the colour, contrast and brightness of the image during processing;

> Incorrect file-naming, where image files are incorrectly named or use non-unique names;

> Basic metadata attribution faults;

> Incorrect data entry, where data is incorrectly entered into the management control system; and

> Incorrect use of controlled vocabulary, the use of words not established within scope notes.

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

42

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

> APPENDIX 8: RECOMMENDED STAFF SKILL SETS

*This appendix is for guidance. It outlines the types of skill areas and tasks that staff engaged in digitisation should be expected to possess. It is included as an aid to public offices and local authorities in meeting Requirement 2.3: Quality assurance procedures **MUST** be defined, documented and implemented.*

Staff training to support digitisation SHOULD include the following areas, and be provided to all levels of skill expected within the organisation:

Skill area	Tasks
Management	• Assessing the business case for digitisation • Negotiating purchase, ongoing service and maintenance of equipment and supplies
Business Analysis	• Defining the workflow for the digitisation process • Defining the integration of digitised records into existing business systems/workflows • Selecting image format • Determining image enhancement requirements • Identifying information architecture for business-process support
Systems Analysis	• Selection of scanner hardware • Defining storage requirements • Integration of computer hardware, imaging equipment and software • Integration of digitisation requirements into existing organisational IT infrastructure • Compliance with national and organisational IT standards • Testing of configurations • Ongoing support of digitisation equipment (where necessary) • Definition of policies and procedures to ensure authenticity and integrity of digital images
Recordkeeping	• Ensuring legislative compliance • Integration with organisational records and business systems • Integration with existing classification and disposal regimes • Defining file naming conventions • Defining and implementing disposal process • Defining metadata • Monitoring of quality of metadata • Managing the source records after digitisation
Equipment Operations	• Operating scanners • Applying any defined selection criteria • Carrying out quality checking on digital records • Adding metadata to digital records

DIGITISATION STANDARD

43

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA

JANUARY 2007

> APPENDIX 9: FURTHER RESOURCES

This appendix is included for guidance. A plethora of advice and standards exist for further reference; however, much of the information is specific to particular types of digitisation – either project-based digitisation for archival quality, or vendor-driven commercial advice. The context of the advice should always be indicated within the project methodology documentation.

The resources listed here represent those most helpful in the development of this Standard, or assessed as useful for public offices' and local authorities' reference.

ISO 12641: 1997, *Graphic technology – Prepress digital data exchange. Colour targets for input scanner calibration.*

ISO 12653 – 1: 2000, *Electronic imaging – Test targets for the black and white scanning of office documents, Part 1 – characteristics.*

ISO 12653-2: 2000, *Electronic imaging – Test targets for black and white scanning of office documents, Part 2 – methods of use.*

ISO/IEC 15444-1: 2000, *Information technology – JPEG 2000 image coding system – Part 1: Core coding system [using lossy compression].*

ISO 15489-1: 2001, *Information and documentation – Records management – Part 1: General.*

ISO/TR 15489-2: 2001, *Information and documentation – Records management – Part 2: Guidelines.*

ISO/TR 15801: 2004, *Electronic imaging – Information stored electronically – Recommendations for trustworthiness and reliability.*

ISO/DIS 19005-1: 2004, *Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1).*

ISO 23081 – 1: 2006, I*nformation and documentation – Records management processes – Metadata for records, Part 1 – Principles.*

ISO 23081 – 1: 2006, *Information and documentation – Records management processes – Metadata for records, Part 2 – Implementation issues.*

PDF Tools AG, PDF/A – 'The Basics' – Understanding PDF White Papers, v. 1. February 7, 2006. http://www.pdf-tools.com/public/downloads/whitepapers/whitepaper-pdfa.pdf

Queensland State Archives. 'Guidelines for the Digitisation of Paper Records', v. 1, May 2005. http://www.archives.qld.gov.au/publications/digitisation/digit1.pdf

Queensland State Archives, 'Digitisation Disposal Authority' April 2006. http://www.archives.qld.gov.au/publications/digitisation/QSA6814%20DisposalPolicy.pdf

New Zealand State Services Commission, 'Trusted Computing and Digital Rights Management Principles and Policies', September 2006. http://www.e.govt.nz/policy/tc-and-drm/principles-policies-06/tc-drm-0906.pdf

Technical Advisory Service for Images (TASI), In-depth Report: 'Quality Assurance'. http://www.tasi.ac.uk/advice/creating/qassurance.html'

Technical Advisory Service for Images (TASI). http://www.tasi.ac.uk

United States National Archives and Records Administration, 'Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images', June 2004. http://www.archives.gov/research/arc/digitizing-archival-materials.pdf

DIGITISATION STANDARD

CONTINUUM CREATE & MAINTAIN
TĀHUHU TE HANGA ME TE TIAKI
ARCHIVES NEW ZEALAND
TE RUA MAHARA O TE KĀWANATANGA
JANUARY 2007

44

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

NOTES:

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

NOTES:

CONTINUUM
create & maintain
tāhuhu te hanga me te tiaki

NOTES: