

Digital Preservation Policy Manual

Joint Operations Group – Policy



Document control

Date approved	
Review date	
Policy owner	Joint Operations Group - Policy
CMS number	NL_CIMS-#414999
Version	1

Revision history

Revision	Date	Author	Reason for change
1.0	12 February 2010	Euan Cochrane Peter McKinney	
1.1	25 May 2011	Peter McKinney	Access, representations, preconditioning, file identification and deletion policies added.
2	25 November 2011	Peter McKinney	Addition of common actor descriptions
3A	15 December 2011	Peter McKinney	Addition of common term for Archives and Library and DC Team and NDHA. Also addition of description of “content owner” in the general roles.
4	1 March 2011	Peter McKinney	Formatting and additional text in “Introduction” for passing to NDHA Transition Committee and other stakeholders.
5	26 June 2012	Peter McKinney	Addition of signed-off policies (Access and Deletion). Addition of new text on consultation in the introduction.
6	20 September 2012	Peter McKinney	Addition of signed-off policies (TAW and Preconditioning)
7	7 November 2012	Peter McKinney	Addition of draft Preservation Management.

Approved by:

Name _____ Position _____

Signature: _____ Date: _____

Name _____ Position _____

Signature: _____ Date: _____

Name _____ Position _____

Signature: _____ Date: _____

TABLE OF CONTENTS

POLICY MANUAL INTRODUCTION	5
TECHNICAL ASPECTS OF ACCESS TO DIGITAL CONTENT POLICY	10
TECHNICAL ANALYST WORKBENCH POLICY	14
DIGITAL CONTENT PRECONDITIONING POLICY	19
DELETION OF DIGITAL OBJECTS POLICY	23
PRESERVATION MANAGEMENT POLICY	26
ANTI-VIRUS	32
FIXITY ASSURANCE.....	37
ROSETTA DATA FILES BACKUP.....	40
REPRESENTATIONS.....	45
ACCEPTABLE CHANGE	48
FORMAT LIBRARY USAGE	49
RISK MANAGEMENT	50
DISASTER RECOVERY	51
SECURITY	52
REFERENCES	53
GLOSSARY	54

POLICY MANUAL INTRODUCTION

Archives New Zealand Te Rua Mahara o te Kāwanatanga (henceforth 'Archives') and the National Library of New Zealand Te Puna Mātauranga o Aotearoa (henceforth 'the Library') have agreed to jointly preserve, manage and give access to digital objects that come within their legislative mandate. Codified within the Digital Preservation Strategy¹, this undertaking is guided by twelve digital preservation principles. The fulfillment of these principles is supported by this Policy Manual (see Diagram 1). This manual contains all the policies that Archives and the Library require to manage their digital preservation programme.

PURPOSE OF THE MANUAL

Well-described and achievable processes ensure that digital content remains accessible and most importantly, integrity. Small deviations from agreed work practices can have severe and irreversible impacts on the content.

The Policy Manual outlines the processes and operating rules that the Digital Preservation Team at Archives and the Library are required to follow in the course of caring for the digital content under their purview.

CONTEXT OF THE MANUAL

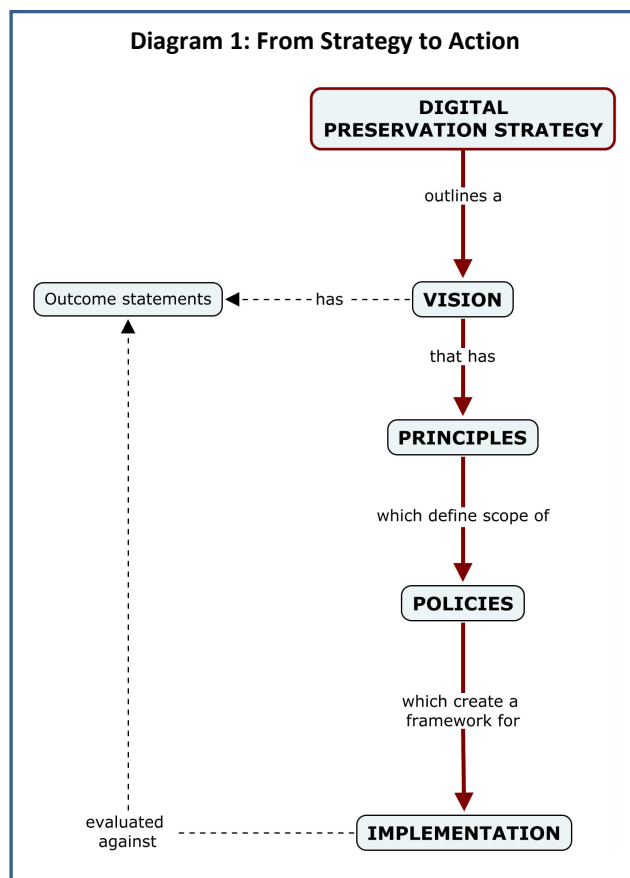
The work of the Digital Preservation Team is undertaken within the parameters outlined by both the National Library of New Zealand Te Puna Mātauranga o Aotearoa Act 2003 and the Public Records Act 2005.

The Digital Preservation Strategy² is the overarching statement by Archives and the Library in relation to the preservation of digital content. It states that a unified and joint approach to digital preservation will best serve the content and those that wish to interact with it. The Policy Manual covers the operational layer below the level of the Strategy.

GOVERNANCE OF THE MANUAL

The policy manual is managed by the Joint Operations Group – Policy. Current members are the Preservation Managers from the Library and Archives, NDHA Policy Analyst, and Senior Analyst of the Digital Continuity Team.

Ownership and final sign-off of the Manual is provided by the Programme Manager of the Government Digital Archive Programme, and the Director of Preservation Consultancy and Research.



¹ Archives New Zealand Te Rua Mahara o te Kāwanatanga & National Library of New Zealand Te Puna Mātauranga o Aotearoa, *Digital Preservation Strategy*, June 2011, http://ndha-wiki.natlib.govt.nz/ndha/attach/News/Digital_Preservation_Strategy.pdf.

² *Ibid.*

SCOPE OF THE POLICY MANUAL

The manual covers all functions of a digital preservation business unit. These functions range from rules guiding error-handling during the ingest of files, to the security requirements of the storage and database layers of the digital preservation system.

STRUCTURE OF THE MANUAL AND POLICIES

The Manual contains all policies required by the Digital Preservation Team. The manual also has a list of references and a glossary. Each policy itself describes:

- The goal of the policy;
- The scope of the policy, defining in particular areas that are out of scope;
- Operating rules that are designed to help the digital preservation team achieve the policy goal;
- The actors that are expected to play a role in implementing the operating rules; and,
- Measures that can be used to monitor the success of the implementation.

CONSULTATION

During the development of all policies a great deal of consultation takes place. Consultation is vital to ensure the policies are grounded in best practice, practicable and understood by all actors. Consultation takes place across the development of a policy. Typically, this includes:

- initial consultation with the digital preservation teams as the scope is developed;
- detailed work with affected actors;
- review with the digital preservation teams as the policy approaches sign-off;
- final review by key actors;
- final review by JOG Committee;
- sign-off by programme directors (and others as required).

It is the responsibility of the policy authors to carry out all consultation and reflect this in the final document. It is the responsibility of the JOG Committee to ensure that all relevant actors have been consulted. Where operating rules must be fully understood and put into operation by another business unit, that business unit will become a signatory at sign-off to demonstrate full and proper understanding of the policy. Finally, the programme directors must satisfy themselves that the consultation has been both broad and detailed enough to make the policy effective and meaningful.

DEVIATIONS/NON-COMPLIANCE

There will be deviations from the operating rules. It is expected that deviations will fall under one of two categories:

1. The deviation is due to a uniqueness in a discrete set of content that cannot conform to the expected operating rules. These should be noted in the Policy Deviations Record.³

³ *Preservation Research & Consultancy Record of Decision Document*, CMS Ref: 429223.

2. The deviation is in place for a short period of time while system functionality or business process are developed to ensure compliance. These should be noted in the Policy Deviations Record.⁴ A plan of work to achieve compliance should also entered into the Policy Compliance Workplan document.⁵

MEASUREMENT

Each policy contained within this manual has a set of measures that when met, will signify that the operating principles and the policy goal are successfully being achieved. Measures are clear indicators of specific outcomes, with the outcomes in this case being the result of following the policy operating principles.

The measures should also be cognisant of the actions that may need to be undertaken to rectify issues that the policies are trying to minimise. For example, the goal of the Fixity Policy is to monitor that no corruption or unauthorised change to the objects has taken place. The measurements will show that checks are taking place, but also that when errors are found, the correct procedures are followed.

The measures do not have to be directly equatable to reports. Indeed, it is more likely that a number of reports will need to be brought together with some manipulation of data coming from reports. Archives New Zealand and the National Library will use the same measures and parameters. In general:

- The policies do not necessarily have one measure per operating principle;
- The responsibility for reviewing the measures within the context of the policy will be with the Policy Analyst, unless otherwise indicated;
- The responsibility for collecting the measures will be defined by business-as-usual reporting procedures; and,
- The methods of collecting the information required for measurement will be defined by business-as-usual reporting procedures.

TERMINOLOGY

The language used in the policy operating rules has specific meaning throughout the manual.

Must: That the operating rule is mandatory in order to be compliant with the policy.

Should: Compliance with the operating rule would be beneficial, but that deviations do not have a critical impact.

Digital Preservation Team: Rather than use the names of the teams that are charged with carrying out the preservation of the digital content of the Archives and Library (Digital Continuity Team and Preservation, Research and Consultancy) the policies use the term Digital Preservation Team.

Preservation System: This is the technical solution implemented by Archives and the Library to preserve digital content.

ACTORS

The policies contained in this manual list a number of actors. These actors have certain responsibilities in relation to the operating rules. The nomenclature used is not a direct reference to roles that exist in either the Digital Continuity Team or the NDHA team. However, the two teams should map the actors to roles within their teams to ensure that both teams understand their responsibilities as outlined in the policies.

⁴ *Ibid.*

⁵ *Digital Preservation Policy Compliance Workplan*, CMS Ref: 450826.

Content Owner: This actor is the role which has ultimate decision making authority over content. Ownership in this case does not denote proprietorship, but could rather be translated as “power of attorney” with regard to the content.

Curator: A role within the Alexander Turnbull Library that is a content owner.

Preservation Analyst: This actor’s main concern is technical oversight of the content within the preservation system. They work they undertake includes; preconditioning, risk analysis, preservation planning, and preservation actions.

Preservation Manager: This is a decision-making position. The position will normally have management of preservation and technical actors.

Preservation Policy Analyst: This actor develops and reviews policy that is directly related to the work of the Digital Continuity and NDHA teams.

System Administrator: This actor is defined as a person who has the ability and mandate to manage change configurations in the preservation system.

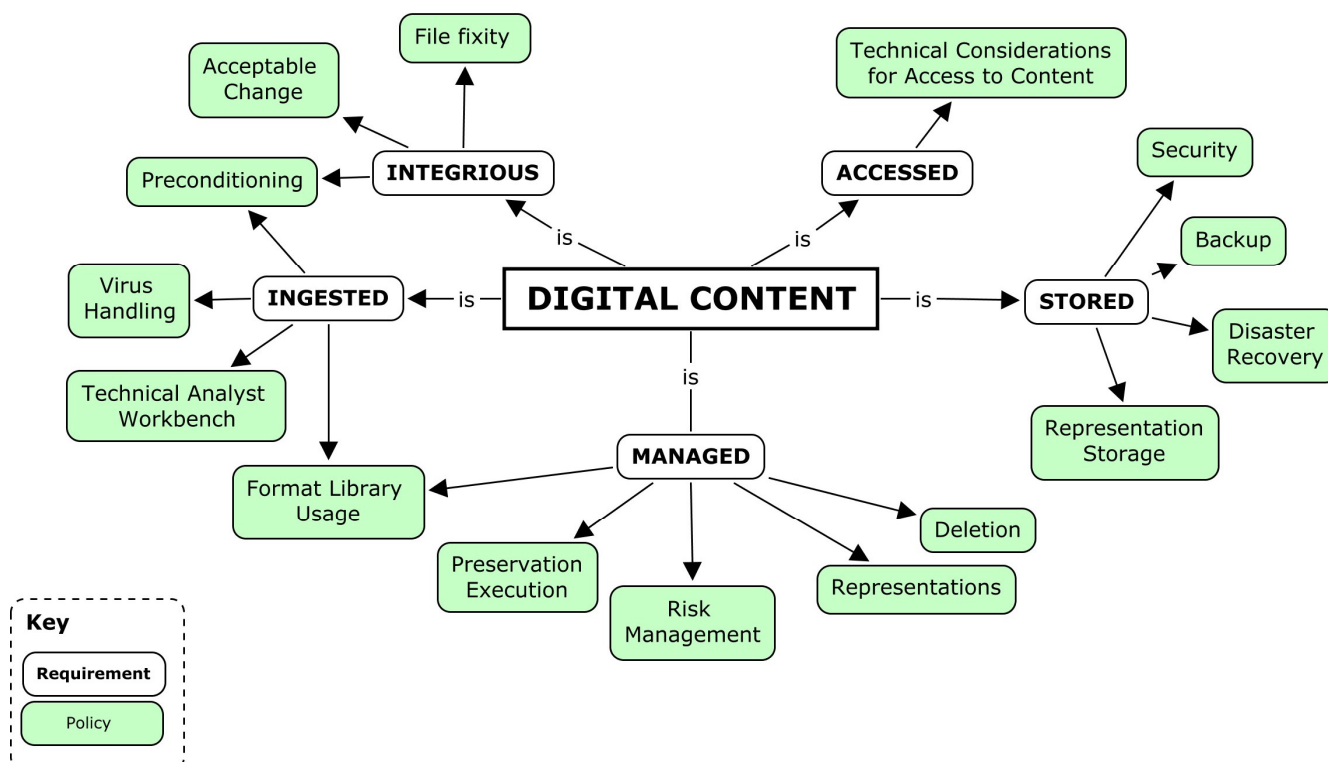
Technical Analyst: The technical analyst is a role specific to trouble shooting errors with content as it is passed through the validation stack. It can be very close in function to the preservation analyst.

REVIEW

The policies must be reviewed regularly to ensure that; a) that the goals are still aligned with best practice in digital preservation, b) the operating rules allow the digital preservation team to fulfill the policy goal.

Individual policies can be reviewed and updated without affecting the status of the manual or other policies. Each policy has its own review date and version number in order to track such reviews.

Diagram 2: Relationships of all policies



TECHNICAL ASPECTS OF ACCESS TO DIGITAL CONTENT POLICY

Date approved	23 May 2012
Approved by	<i>Alison Fleming</i> : Programme Manager Government Digital Archive <i>Steve Knight</i> : Programme Director Preservation Research
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

POLICY OVERVIEW

This policy outlines the operating rules for giving access to digital content held in the preservation repository.

INTRODUCTION

Archives New Zealand and the National Library have a commitment to give access to the objects in their care. This includes born digital and digitised content. The Digital Preservation Team has a role to play in delivering this digital content. This role must be undertaken in accordance with all relevant Department of Internal Affairs (DIA) policies.

For the sake of this policy, ‘access’ means allowing users to obtain or retrieve a copy of the requested content.

‘Users’, within the context of this policy, are individuals, organisations or systems that interact with content.

SCOPE

This policy does not define who may access digital content. Such restrictions are governed by legislation and existing DIA policies.

This policy also does not define restrictions placed on the form of the content that users are given access to. Where appropriate, this is defined by policies created by other parts of the organisation (Alexander Turnbull Library Reproduction of Materials Policy^{*}). Such policies can define the resolution of images, or compression of audio that will be offered for access.

Access for internal staff is not included in this policy.

In addition, this policy does not detail the process or rules for generating derivative copies for access. This is covered in the Representation Policy.

Finally, security restrictions to ensure there is no unauthorised access are not detailed here. These are covered by various DIA policies⁶.

POLICY GOAL

To clearly define the role of the Digital Preservation Team, and the preservation repository that it manages, in giving access to digital content preserved by Archives New Zealand and National Library of New Zealand.

⁶ See the “Reference” section below for more details.

POLICY OPERATING RULES

1. Access from the preservation repository must be made with full reference to all relevant legislation and policies developed by the DIA. These include, but are not limited to, the National Library Access Policy and Archives New Zealand Access Standard.⁷
Rationale: Access to digital content is controlled by a suite of policies. This policy is only one part of that suite and must not impede or impinge upon the other policies.
2. Where the user is not entitled to access the content, they will be provided with information about the denial of access.
Rationale: Some content preserved by Archives and the Library has restrictions placed on it. Where this is the case, the user should be able to understand why they cannot view the material they have requested to view.
3. The organisations will endeavour to give access to reliable copies of the original digital content, irrespective of what version or representation is being accessed.
Rationale: Multiple versions of the digital content can and will exist. Irrespective of whether the copy delivered is a derivative copy or a version of the preservation master, it should always reflect best efforts to convey the intellectual essence of the content.
4. The user should be able to quickly understand what version and representation they are accessing and how it relates to other versions. They should also be given an indication of any known differences.
Rationale: The user should be made aware of what version and representation of the content they are viewing. This should allow them to recognise that, for example, the original was a BWF file, but that they are viewing a MP3 derivative.
5. The version and representation of the content most appropriate to the user's request should always be delivered.
Rationale: There may exist multiple versions and representations of the content. The user should be delivered the version and representation that best fulfils their request. For example, a user's request placed through an online channel for a particular music file may result in a compressed derivative copy being delivered (e.g. MP3 file). However, a subsequent request through a different channel seeking the non-compressed version of the audio may result in the Modified Master being delivered (e.g. BWF).
6. Where practical, content will be delivered in a form that minimises the technical burden on users.
Rationale: Every effort should be made to offer content in a manner that eases the user's ability to interact with it, in terms of technical requirements. This may mean that derivative copies are created for content that is in uncommon formats, or that viewers are developed to present content to users.
7. When applicable, an indication of software and hardware that may be used to interact with the content should be provided to the user, when available.
Rationale: Such information will give users an indication of how both organisations render the content. This information will not be definitive and will only reflect the current status of the Institutional Technology Profile⁸ as it exists in the Format Library.

⁷ See the "Reference" section below for more details.

⁸ The profile is an indication of what application(s) the Department of Internal Affairs uses to render particular formats. See "Key terms" section for more information.

ACTORS

Preservation Analyst:

- Responsible for technical comprehension of the formats in the repository, with particular reference to understanding user needs.

Preservation Manager:

- Is responsible for management of this policy.
- Is responsible for managing liaison with staff to ensure proper development and implementation of DIA policies and services affecting access to digital content.

Preservation Policy Analyst:

- Is responsible for managing the review process of this policy.

System administrator:

- Responsible for supporting the generation of derivative copies by the preservation system.

IMPACTS

The degree and frequency of changes in preferred access mechanisms may have a detrimental impact on the Digital Preservation Team and preservation repository. A representative of the Digital Preservation Team should therefore always be part of requirements capture for new access developments by the DIA.

MEASUREMENT

	Related Operating Rule	Measure	Parameter
1	2	The content delivered to a user is the same as the content requested by the user.	99% of all user requests.
2	2, 5	All content delivered through viewers is rendered correctly	99% of all viewer deliveries.
3	5, 6	The number of user queries about rendering or opening files	Report of count. No limit as will be used as a measure of the efficacy of 5 & 6.

KEY TERMS

Institutional Technology Profile: The profile is an indication of what application(s) the Department of Internal Affairs uses to render particular formats. It is not a fully comprehensive view of what applications *may* be used to render the format. It is a statement of what the DIA currently uses.

RELATED POLICY CHAPTERS

Document	Location
Representation Policy	Chapter

REFERENCES

Alexander Turnbull Library, 'Reproduction of Materials from the Collections Staff Guidelines', CMS ID#77435.

Archives New Zealand, 'Access Standard', http://archives.govt.nz/sites/default/files/s4_0.pdf.

Department of Internal Affairs, 'Security Policy', <http://1840.dia.govt.nz/policies/information-technology/high-level-security-policy>.

National Library of New Zealand, 'Access Policy', <http://www.natlib.govt.nz/about-us/catalogues/library-documents/access-policy>.

National Library of New Zealand (Te Puna Mātauranga o Aotearoa) Act 2003.

Public Records Act 2005.

NOTES OF DIVERGENCE

There are no points of divergence between the Archives and Library in terms of the Operating Rules. Archives specific technical implementation will differ from the Library's. Archives have yet to be fully define their implementation.

TECHNICAL ANALYST WORKBENCH POLICY

Date approved	20th September 2012
Approved by	<i>Alison Fleming</i> : Programme Manager Government Digital Archive <i>Steve Knight</i> : Programme Director Preservation Research
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

POLICY OVERVIEW

This policy states that both Archives and the Library should respond to technical errors in the preservation system in a consistent and timely fashion. It also calls for the institution of a technical committee that has oversight of error-handling.

INTRODUCTION

Technical errors can be discovered both during ingest of content into the preservation system and when processes are carried out on content that is already within the system. These errors can be caused by a number of factors, including content failing validation by tools, tools themselves failing, or processes being interrupted.

The purpose of the policy is not to define what actions should be taken to resolve the errors. The purpose is to detail necessary steps and general guidance for actions that may be taken, define responsibilities, and define timescales within which errors should be resolved.

Functionality within the preservation system does allow for institutions to undertake their own error handling and invoke their own rules. However, it is the aim of Archives and the Library that error handling should be consistently applied to all digital content, irrespective of the institution that is managing it. We acknowledge that this may not be always achievable and any deviations must be managed and documented appropriately.

SCOPE

Within scope for this policy is content that triggers an error during normal ingest. This includes content that is added using the “Add Representation” process.

The policy also includes the resolution of any errors discovered when content from the permanent repository is passed back through the validation stack.

It does not cover content that lands in the preservation workbench. Errors triggered during preservation actions will be covered in the Preservation Action policy.

POLICY GOAL

To ensure that error-handling actions taken in the preservation system are consistent across both institutions whenever possible, completed in a timely fashion, and are well-documented.

POLICY OPERATING RULES

1. Archives New Zealand and the National Library should take consistent actions, where possible, on content that has been identified as containing technical errors.
Rationale: The Joint Digital Preservation Strategy details the intent of both organisations to share principles and processes during the management and preservation of digital content [Joint Strategy, p.5].
2. A technical committee must be established and meet regularly.⁹ It will have oversight of all error-handling activities. It should comprise the responsible Technical Analysts from each institution. Possible areas of difference must be escalated to the Preservation Manager from each institution for sign-off.
Rationale: The committee will ensure that decisions taken are uniform for both organisations unless there is a strong reason to deviate. Where deviation is required, the decision must be approved by the Preservation Manager of each institution.
3. There are at least seven classes of errors that the Technical Analyst may have to respond to. Table 1 describes these levels and details the response type that must be taken and the timescale that should be adhered to.
Rationale: This classification scheme describes levels of actions described by the degree of change introduced by the action. It is a way of categorising errors in order to describe the appropriate responses.

Table 1: Classification of errors and expected responses

Level	Description	Example	Response	Timescale to conclusion*
A	Errors for which there is an agreed precedent.	The content is erroneously truncated due to faulty download or transfer from producer.	These require no recourse to the technical committee. Existing documentation must be robust enough to support ongoing resolution.	One week.
B	Errors caused by a system process failing.	Content failing the loading stage of the ingest process.	These require no recourse to the committee and are the responsibility of the system administrator, who should consult any relevant stakeholders.	Two days.
C	Errors that require ANZ or NLNZ to introduce an element of change to address the error.	A non-standard extension is used in the file naming.	Any actions should be discussed by the technical committee and relevant stakeholders. The Preconditioning Policy should be used as the guiding document in areas of change.	Depends on complexity of work. Stakeholders to be kept informed.
D	Errors that require a new rule to be added to the format auto-correction rules.	Multiple format identifications for the content.	Proposed rules must be discussed by the technical committee.	Two weeks.

⁹ The technical committee is defined in the Technical Committee Terms of Reference [CMS ID#461848].

E	Errors that require a change to the format library.	A new format type is ingested.	The technical committee should agree on a response and submit a change request to the Format Library Working Group as per the FLWG guidance.	Three weeks.
F	Errors for which there is no clear resolution, requiring detailed research.	The content type is new and technically complex, but has errors.	Research undertaken. Full communication with the technical committee and relevant stakeholders.	One month for detailing a resolution.
G	Errors so numerous that a solution will not be immediate.	A deposit containing 1,000 files that all have some sort of error.	Research undertaken. Full communication with the technical committee and relevant stakeholders.	One month for detailing a resolution.
H	Errors for which a system or tool update is required.	Content cannot be processed with current ingest flow.	Technical committee to define and enter change request. Full communication with relevant stakeholders required.	One month for detailing a resolution.

*** NB:** A valid “resolution” is to propose a path for finding a solution, rather than having to find a solution. A valid solution may be that the content is ingested with the errors remaining. Such a solution should detail a future plan for resolving those ignored errors.

4. All actions taken within the workbench must be properly documented. Documentation can include system notes, minutes of the committee, and external papers. The technical committee will determine the form of that documentation and where it should reside.
Rationale: Documentation will allow analysts to understand what has happened to what content at what time. The analyst will also be able to understand patterns of error types.
5. The time that content spends awaiting a resolution should be minimised.
Rationale: Other staff users are dependent on the objects passing cleanly and quickly through the system so they can continue their ingest flow. In addition, the content must be made secure in permanent storage quickly, from where it can be properly managed and offered for access (if permitted).

ACTORS

Preservation/Technical Analyst:

- Responsible for monitoring the preservation system for errors in their institution’s content.
- Responsible for communications with internal content owners about resolution of errors.
- Responsible for taking the required steps to resolve the error.
- Responsible for documenting all actions taken.

Technical Committee:

- Has oversight of responses to content found to have technical errors.
- Responsible for ensuring documentation is fit for purpose.
- Responsible for documenting deviations.

System Administrator:

- Responsible for resolving particular errors caused by system faults.

Preservation Manager:

- Responsible for resolving conflicts or issues arising from the Technical Committee.
- Responsible for management of this policy.
- Responsible for approving deviations.

Preservation Policy Analyst:

- Is responsible for managing the review process of this policy.

IMPACTS

Coming to common agreement on procedures may impact on the time taken to resolve errors. The benefits of taking agreed and consistent action make this potential impact acceptable. The Technical Committee should meet regularly and work efficiently to minimise delays. Dealing with errors consistently between both organisations and the Rosetta user community (where possible) can benefit from user community knowledge and can avoid dependency on one Technical Analyst's knowledge within each organisation.

MEASUREMENT

No.	Related Operating Rule	Measure	Parameter
1	1	The number of rules to deal with technical errors that are institution specific	0
2	4	The number of errors left with no resolution or solutions after the timescale to conclusion.	0
3	3	The number of times that documentation is incomplete for error-handling actions.	0

KEY TERMS

n/a

RELATED POLICY CHAPTERS

Document	Location
Preconditioning Policy	Chapter <i>n</i> , page <i>x</i>

REFERENCES

Document	Location
Technical Committee Terms of Reference	CMS ID#461848
Digital Preservation Strategy: Archives New Zealand and National Library of New Zealand.	http://ndha-wiki.natlib.govt.nz/ndha/attach/News/Digital_Preservation_Strategy.pdf
NDHA/GDAP – System Admin processes	CMS ID#441316
NDHA TA Workbench – Handling Process Guide 2010.	CMS ID#390043

NOTES OF DIVERGENCE

There are no points of divergence between

DIGITAL CONTENT PRECONDITIONING POLICY

Date approved	17 October 2012
Approved by	<i>Alison Fleming</i> : Programme Manager Government Digital Archive <i>Steve Knight</i> : Programme Director Preservation Research
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

POLICY OVERVIEW

This policy is concerned with changes that may be made to digital content before it is ingested into the preservation system.

INTRODUCTION

Preconditioning deals specifically with changes to digital content that have come within the control of the Archives or Library, but have not yet been ingested into the preservation system.

The diverse nature of digital content means that there are times when it is desirable to make changes to it before it is ingested into the preservation system. These changes are classed under the term 'preconditioning'.

Preconditioning can aid in resolving technical issues found in digital content. In particular, preconditioning can ensure that a version of transferred digital content will be able to be ingested into the digital preservation system without any issues or errors being presented by the system. Issues that preconditioning avoids include those presented by validation, metadata extraction and format identification tools. While it is the role of preservation analysts, digital archivists and other domain experts to decide when this type of change is necessary, there are some factors that should be considered before deciding; a) whether a change is desirable, and b) whether that change should be done through preconditioning or as a preservation action. These include, but are not limited to:

- Can it be shown that the preconditioning will not affect, in any way, the intellectual message of the content of any and every object that the preconditioning procedure will be applied to?
- Is the proposed change completely reversible?

SCOPE

This policy covers actions taken on content that has been transferred to or acquired by Archives or the Library but has not yet been ingested into the preservation system. It does not cover the limits of change that can be introduced once the content is within the controlled boundaries of the preservation system.

The policy also only covers changes to content that do not result with the original and a copy being ingested. Preconditioning changes are made on the original. They do not generate a new copy.

POLICY GOAL

The goal of this policy is to describe the limits of change that can be introduced to digital content from the time it is brought within the control of Archives or the Library to its acceptance into the preservation system.

POLICY OPERATING RULES

1. Certain changes may be made to content before it is ingested into the preservation system. Any changes must adhere to the following operating rules. Where the proposed changes do not sit within the boundaries defined, an action within the preservation system should be considered.¹⁰
Rationale: There are situations where it is desirable to change content in order that it can be ingested into the preservation system in the desired way.
2. All changes must be reversible.
Rationale: The lack of full control and audit external to the system means that preconditioning changes must be fully reversible. It may become desirable to return the content to its original state in the future.¹¹
3. All changes must have two critical pieces of information created and stored appropriately. These are:
 - a. documentation detailing the reasons for the preconditioning. It should be stored outside the preservation system but must be able to be associated with the affected content.
Rationale: Documentation will be necessary to understand why the change was made, how it was made, and how it may be reversed. It should also detail who was involved in the decision making process.
 - b. a system-based provenance note that clearly describes the change that has been made to the file. This note should remain as part of the file's preservation metadata throughout its existence.
Rationale: The provenance note allows users to understand exactly what preconditioning has been undertaken on the file. It must give sufficient information for operating rule two to be actioned.
4. A final approval of the acceptability of all proposed changes must be given by the appropriate authority.
Rationale: The preservation team will advise on technical changes but they must be approved by the appropriate authority (defined in the actors section below).

ACTORS

Curator/Archivist:

- A content owner. Is responsible for decision making on permissible change.

¹⁰ The key difference between preconditioning work and preservation actions is the degree of change and reversibility of the work. Preconditioning deals with small changes that are lossless and completely reversible, hence can be made without generating a new copy. Preservation actions are more complex and will often be impossible to reverse. This is why preservation actions result in a new digital object being generated with the original untouched.

¹¹ If necessary it is possible to reverse the change and get exactly the same file as original one. Non-reversible change is considered as preservation action.

Conservators:

- May act as a proxy for the content owner in terms of decision making.
- May carry out changes.
- Is responsible for understanding any intellectual impacts of changes on the file in question.

Digital Archivist:

- May act as a proxy for the content owner in terms of decision making.
- May carry out changes.
- Is responsible for understanding any intellectual impacts of changes on the file in question.

Preservation/Technical Analyst:

- Responsible for advising content owners on impacts.
- May carry out changes.
- Is responsible for understanding any technical impacts of changes on the file in question.

Preservation Manager:

- Is responsible for management of this policy.

Preservation Policy Analyst:

- Is responsible for managing the review process of this policy.

IMPACTS

The process of proposing, understanding and undertaking preconditioning changes will extend the length of time that the content remains outside the preservation system. This must be minimised where possible.

MEASUREMENT

No.	Related Operating Rule	Measure	Expected outcome
1	3	The number of files that have undergone preconditioning that do not have a provenance note attached.	0
2	2	The number preconditioning actions that cannot be reversed.	0

KEY TERMS

Intellectual message: That content which the digital file is trying to convey: the story, the picture, the song, the record. This is defined as required by the Archives and the Library.

Preconditioning: Deliberate and agreed changes made to content that is not in the preservation system. These changes are undertaken with the purpose of preparing the content for ingest into the system.

RELATED POLICY CHAPTERS

Document	Location
Preservation Action Policy	Chapter <i>n</i> , page <i>x</i>
Technical Analyst Workbench Policy	Chapter <i>n</i> , page <i>x</i>

REFERENCES

N/A

NOTES OF DIVERGENCE

There are no points of divergence between Archives and the Library.

DELETION OF DIGITAL OBJECTS POLICY

Date approved	15 May 2012
Approved by	<i>Alison Fleming</i> : Programme Manager Government Digital Archive <i>Steve Knight</i> : Programme Director Preservation Research
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

POLICY OVERVIEW

This policy contains guidance for the deletion of digital content from the preservation system. The precondition for such deletion is that all legislative requirements have been met before an object has been marked for deletion.

INTRODUCTION

This policy is concerned with ensuring that the correct processes are followed to ensure full and proper deletion of digital content.

In general, no digital object once selected for long-term preservation may be deleted from the collections of Archives New Zealand or National Library of New Zealand unless the circumstances meet those determined in the relevant legislation. Both organisations have processes for seeking permission to delete material. There are other conditions under which items within the preservation system may be deleted. These are detailed below in operating rule 2.

SCOPE

This policy covers public records and documentary heritage items that have been created in digital form. It also applies to digital copies of material held by Archives and the Library where there is clear intent to preserve these objects.

The policy does not cover copies of content made to facilitate access. These objects are not classed as content created for long-term preservation

The policy is not concerned with the business processes that relate to the selection of content through legislative reasoning. It is concerned however, with detailing conditions that must be met for material to be deleted outwith this legislative process.

POLICY GOAL

The goal of this policy is to ensure that objects selected for deletion in the preservation system meet the criteria for deletion, and that the correct processes for deletion are followed.

POLICY OPERATING RULES

1. No digital content selected to be public records and digital documentary heritage may be deleted unless the relevant legislation allows it, or operating rule 2 applies.

Rationale: The Digital Preservation Team must always comply with the two Acts that govern the management of documentary heritage and public records.

2. Beyond the parameters outlined in operating rule 1, content may be deleted where the one of the following conditions has been met:
 - a. The item has been ingested in error. Commonly, this means it is a duplicate or it is the wrong item.
 - b. A sanctioned authority has authorised the deletion of the item.
 - c. The content has been deposited for the purpose of testing.

Rationale: Items ingested in error are not covered by the Acts and can be deleted without legislative permissions. Either institution may also have to act on the instruction of a judge to delete material after the outcome of cases that seek to have the content deleted.

3. The process of deletion must ensure that copies of the object are expunged. It is likely that in legal cases, the level of deletion (for example, the deletion of copies on offsite tapes) will be stated. In cases where the deletion is driven by error or testing, offsite copies will not be deleted.

Rationale: For ease of operation, the default should be to delete only from on-site storage.

4. The process of deletion must also ensure that where applicable, all related metadata are deleted (for example descriptive, structural and administrative metadata). Applicability is governed by the reason for deletion. However, rule 5 below must also be adhered to.

Rationale: Different levels of deletion may require more or less aspects to be deleted. Ideally a shell record will remain in order to aid system management.

5. Notwithstanding 4 above; where deletions have taken place, a record must exist of the process that was undertaken.

Rationale: The operational default should be to leave a trace of the deletion.

6. Two different members of staff are required to complete the deletion process.

Rationale: This will ensure a check that the process has been followed and that the correct content is being deleted.

ACTORS

Preservation Manager

- Responsible for ensuring that all relevant processes are complied with before continuing with the deletion process.
- May be the second actor in the deletion process.
- Responsible for management of this policy.

Preservation Analyst

- Responsible for processing the deletion.

Preservation Policy Analyst

- Responsible for the review of this policy.

IMPACTS

Ensuring all copies of the object to be deleted could require off-site tapes to be recalled and the item deleted from them.

MEASUREMENT

No.	Related Operating Rule	Measure	NLNZ Parameter	ANZ Parameter
1	1, 2, 3, 4, 5, 6	The number of cases not complying with correct deletion procedures as detailed in legislation, legislative ruling, or in the process of correcting business errors.	0	0

KEY TERMS

n/a

RELATED POLICY CHAPTERS

Document	Location

REFERENCES

n/a

APPENDIX 1: NOTES OF DIVERGENCE

There are no areas of divergence between Archives and the Library.

PRESERVATION MANAGEMENT POLICY

Date approved	18 December 2012
Approved by	<i>Alison Fleming</i> : Programme Manager Government Digital Archive <i>Steve Knight</i> : Programme Director Preservation Research
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

POLICY OVERVIEW

This policy provides guiding principles and rules for preservation actions undertaken on digital content that has been brought under the control of either Archives or the Library.

INTRODUCTION

Archives and the Library are mandated to collect, preserve and give access to public records and documentary heritage.¹² In order to fulfil these mandates, digital content must be subject to an ongoing programme of digital preservation management.

Preservation management includes the planning and execution of actions taken on content that is in the permanent repository of the preservation system.¹³ These actions must balance a potential change in format with the requirement that the integrity and authenticity of the digital content must not change¹⁴. For the purpose of this policy, preservation actions are either migration (including normalisation) or emulation.¹⁵

Preservation planning is the result of a considered risk analysis of digital content. This risk analysis looks at the ability of Archives and the Library to manage and give access to the digital content. Preservation planning is the process of defining possible methods of mitigating risks that have occurred or may occur in the future. The purpose of a plan is to provide sufficient information for a) decision makers to choose the preferred action to take, and b) maintenance of a fully transparent and auditable process of decision making and action.

The Joint Preservation Strategy clearly details four principles that are the framework for this policy:

1. Preservation actions will always result in new versions of the content and will not directly affect the original item.

¹² See Public Records Act 2005, and the National Library of New Zealand Te Puna Mātauranga o Aotearoa Act 2003.

¹³ The Preconditioning Policy applies when reversible changes are being undertaken on content that is not yet under the control of the preservation system.

¹⁴ As noted in: Archives New Zealand and National Library of New Zealand, *Digital Preservation Strategy 2011*. http://ndha-wiki.natlib.govt.nz/ndha/attach/Welcome/Digital_Preservation_Strategy.pdf.

¹⁵ While emulation does not change the original object, the validity of the emulation environment will be subject to preservation planning.

2. The integrity of digital content (as defined by each organisation) will be retained through all preservation actions.¹⁶
3. The authenticity (as defined by each organisation) of the content that is being preserved will be guarded and assured through all preservation actions.
4. Preservation actions will be transparent and auditable.¹⁷

The operating rules contained in this policy address each of these four principles.

SCOPE

This policy covers content that has been selected for long-term preservation by either Archives or the Library. Specifically, it is concerned with the management of migration and emulation techniques that ensure the content remains manageable, accessible and authentic. These techniques have the potential to alter the integrity and authenticity of the digital content, which we would like to avoid.

Moving the digital content to new hardware is out of scope.

POLICY GOAL

Preservation management is conducted in a way that retains the authenticity and integrity of digital content, and all plans and actions are consistent in their approach, process and outcomes.

POLICY OPERATING RULES

Preservation actions will be transparent and auditable

1. Preservation management will be a shared activity across both Archives New Zealand and the National Library.

Rationale: Sharing one preservation systems requires a collaborative approach to sharing solutions, processes, and knowledge in order to maintain consistency in the system.

2. All preservation actions must be preceded by a robust planning process shared by both institutions. This process must explore all available options, engage with the content owner(s), and document all the information required to make a decision about what action to undertake. The planning process will be an agreed activity coming out of business planning.

Rationale: The decision to undertake an action on content must be well researched and fully auditable. The planning process is therefore critical in supporting the transparency of actions.

3. A preservation committee will manage preservation planning and actions.

Rationale: Planning and action processes must always be controlled at the consortium level.

4. The final decision about what preservation action to undertake will be shared between both institutions and signed off by each institution's Preservation Manager and the content owner.

¹⁶ Further refinement of authenticity and integrity requirements will result in a shared definition between Archives and the Library. The Digital Preservation Strategy will be updated accordingly at its next review.

¹⁷ See *Digital Preservation Strategy 2011*.

Rationale: The preservation manager of each institution must satisfy themselves that the option chosen has been subjected to rigorous planning and testing. The content owner will play a role in this sign-off.

5. Detailed documentation for preservation management (including, but not limited to the information used in the decision-making process, changes introduced by the action, and the processes used) must be maintained either in the institution's EDRMS system or in the preservation system itself.

Rationale: Detailed documentation provides an in-depth description of the whole preservation planning process, including; testing different alternatives and options; changes introduced by the action; criteria for assessment (intellectual and automatic); sign off from both institutions.

6. It must be possible to associate the new item created by the preservation action to the creation process.

Rationale: Metadata about the preservation action should be created and kept with files to provide transparency around what has been done, why, by whom, when and with what outcome. This is an essential part of proving the authenticity of the content in the future.

Preservation actions will always result in new versions of the content and will not directly affect the original item

7. The outcome of a migration preservation action is always a new representation, which creates a new version of the intellectual entity. The original master file will be kept along with the new master.

Rationale: The preservation action is conducted on any master representation. The original will be kept for the sake of being able to refer to it in the future or provide the original (i.e. first archived iteration) file in case of need to the users.

The integrity of digital content (as defined by each organisation) will be retained through all preservation actions¹⁸

The authenticity (as defined by each organisation) of the content that is being preserved will be guarded and assured through all preservation actions¹⁹

8. Preservation actions should be carried out on all files in the permanent repository at consortium level that fit the profile covered by the preservation plan.

Rationale: The decision to run a preservation action on a certain file format is made for serious reasons (obsolescence, no available rendering application or other risk) and is generally applicable to the same files across all institutions in the consortium.

9. As content is ingested that fits the profile of the preservation plan, actions should run on it at regular intervals.

Rationale: As new content is ingested that fits the profile of signed-off plans, it is important to ensure that it undergoes the preservation action as soon as is practical.

¹⁸ See footnote 5 above.

¹⁹ See footnote 5 above.

10. The preservation system must always be used to manage and control preservation actions. The actual work (migration) may be done within or outside of the system.

Rationale: The preservation system is designed to manage and run preservation actions both within and outside the system. Either way, the system keeps track of all changes and actions and generates relevant metadata.

11. The object that is created in the course of the preservation action must be able to stand as an authentic representation of the digital content. This is defined through the planning process and will reference existing institutional statements on the meaning of authentic.

Rationale: The preservation planning process must ensure that the proposed object to be created contains all aspects relating to the intellectual message and any other characteristics of the content deemed to have value (as defined during the planning process). This includes objects created through migration or the rendering of original files in an emulated environment.

12. Preservation plans should be reviewed at regular intervals to ensure that they are still relevant and achieve the best outcome.

Rationale: As technology and knowledge changes, decisions made in the past should be reviewed. Signed-off preservation plans should be reviewed periodically (ideally every time before running relevant preservation action) to understand whether they are still valid and give the best outcome for the at risk content.

ACTORS

Curator/Archivist

- A content owner. Is responsible for decision making on permissible change of the content.

Preservation/Technical Analyst:

- Responsible for proposing technical solution, tool etc. for preservation action.
- Responsible for taking care about metadata creation (ingest, preservation action).
- Responsible for documenting all actions taken.
- Responsible for running the preservation action.
- Responsible for communications with internal content owners during the planning process about possible solutions and outcomes of the preservation action.
- Responsible for periodically reviewing plans.

Preservation Manager:

- Is responsible for decision-making against preservation plans.
- Is responsible for management of this policy.

Preservation Policy Analyst:

- Responsible for managing the review process of this policy.

Preservation Committee

- Will comprise a preservation analyst, preservation manager, and content owner (or their representative) from each institution.
- Is responsible for managing preservation planning and action, including agreeing criteria for making preservation decisions and setting priorities.
- Responsible for ensuring documentation is fit for purpose.

IMPACTS

N/A

MEASUREMENT

No.	Related Operating Rule	Measure	Expected outcome
1	5,6	Each preservation plan and action is documented.	100%
2	4, 5, 6, 10	We are able to prove that the preserved content of the documents is authentic.	To be defined.
3		There are no IEs which we cannot render, manage or understand in permanent repository.	To be defined.

KEY TERMS

Preservation management:

An overarching term that encompasses preservation planning and actions.

Preservation planning:

The process of comparing alternative methods and tools for mitigating an identified risk to content. The purpose of a plan is to create sufficient information on processes and outcomes that a decision can be made about which one to implement.

Preservation action:

Any activity done with aim of keeping the content of the document usable and renderable in the current (or future) technical environment. Preservation actions must be conducted on files which already are in the preservation system. Preservation actions do not necessarily have to be reversible, however for this reason, the original file has to be kept.

A typical preservation action is file format migration. However the creation and use of an emulated original environment for rendering complex objects such databases or websites might be considered a preservation action also. For the purposes of this policy, normalisation is classes as a form of migration.

Preconditioning:

Deliberate and agreed changes made to files that are not in the preservation system. These changes are undertaken with the purpose of preparing the files for ingest into the system and do not result in a new file being created.

RELATED POLICY CHAPTERS

Document	Location
Digital Content Preconditioning Policy	A670592, CLIO#437447
Joint Digital Preservation Strategy	A577312, CLIO#420467

REFERENCES

Digital Content Preconditioning Policy A670592

Archives New Zealand and National Library of New Zealand, *Digital Preservation Strategy 2011*. http://ndha-wiki.natlib.govt.nz/ndha/attach/Welcome/Digital_Preservation_Strategy.pdf

NOTES OF DIVERGENCE

There are no points of divergence between Archives and the Library.

ANTI-VIRUS

Date approved	
Approved by	
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

POLICY OVERVIEW

This policy states that no viruses will be knowingly ingested into the preservation repository. It details the process that should be followed upon discovery of a virus during ingest and also lists the requirements for periodic checking of content stored in the preservation repository.

INTRODUCTION

Archives and the National Library have a responsibility to safeguard the assets they control from viruses or other malicious software. These can potentially alter or delete content. The organisations must also safeguard services delivered by the DIA which may be impacted by investigation of the virus alert.

Neither organisation believes that viruses are inherent pieces of the original content that must be preserved. Moreover, it is not the role of either organisation to document the history of the virus.

There is therefore no reason for viruses or malicious software to exist in the preservation repository.

SCOPE

The policy is limited to the point at which digital objects come under the control of the preservation repository. It does not cover virus checking of digital objects before they are brought within the boundaries of the preservation system. This is covered by existing DIA policies ('Security Policy: Anti-virus' 2011.)

POLICY GOAL

The goal of this policy is that no content with a confirmed virus will be ingested into the archival storage of the preservation repository.

POLICY OPERATING RULES

1. All objects will be checked for viruses during the ingest process using appropriate virus checking tools.
Rationale: Any viruses should be identified before the objects are passed to the permanent repository.
2. A weekly check should be made for definition and security updates for those tools.
Rationale: Anti-virus software is updated regularly to keep pace with the development of viruses. The Preservation System should reflect those updates.
3. An annual review of the tools used must be undertaken to ensure that they are still valid.
Rationale: New Anti-virus tools may bring improvements in virus monitoring, and new content types may be better served by running a different tool over them.

4. No content that has been confirmed as containing a virus shall be passed into the permanent storage of the preservation system.
Rationale: Viruses are not considered part of the collection item or public record being preserved. They are classed as undesirable additions to a file. Effects may range from infecting customer's computers to affecting critical services delivered by DIA.
5. Investigation of possible viruses will be done in a fashion that does not put at risk the preservation system, a DIA computer, or the Department's internal network.
Rationale: The NDHA and Digital Continuity teams have a responsibility to contain any suspected virus in order that it does not affect any other machines or networks.
6. When a virus warning is reported by the preservation system, the *Process for handling virus warnings on ingest* diagram (Diagram 1) should be followed by all staff involved.
Rationale: The Diagram [will have been] agreed upon by all relevant stakeholders including the DIA Security Team. It is the current best practice method for dealing with potential viruses.
7. Virus checks will be run over content held in permanent storage twice a year.
Rationale: This will allow new signature updates in the virus tools to find any positive identifications to be made on content that was ingested before that particular virus was understood and identified by the anti-virus community.
8. If a virus warning is given during the running Rule 6, then the *Virus check on objects in permanent storage* diagram (Diagram 2) must be followed by all staff involved.
Rationale: The Diagram [will have been] agreed upon by all relevant stakeholders including the DIA Security Team. It is the current best practice method for dealing with potential viruses.
9. All stakeholders (NDHA Manager, Producer or Producer Agent, GTS Technical Representative) will be informed of actions being undertaken at every point from the initial identification of the virus to the conclusion of the process.
Rationale: Communication is vital during the process of dealing with potential viruses. Stakeholders must be kept informed in order that they can feed into their own risk management processes.

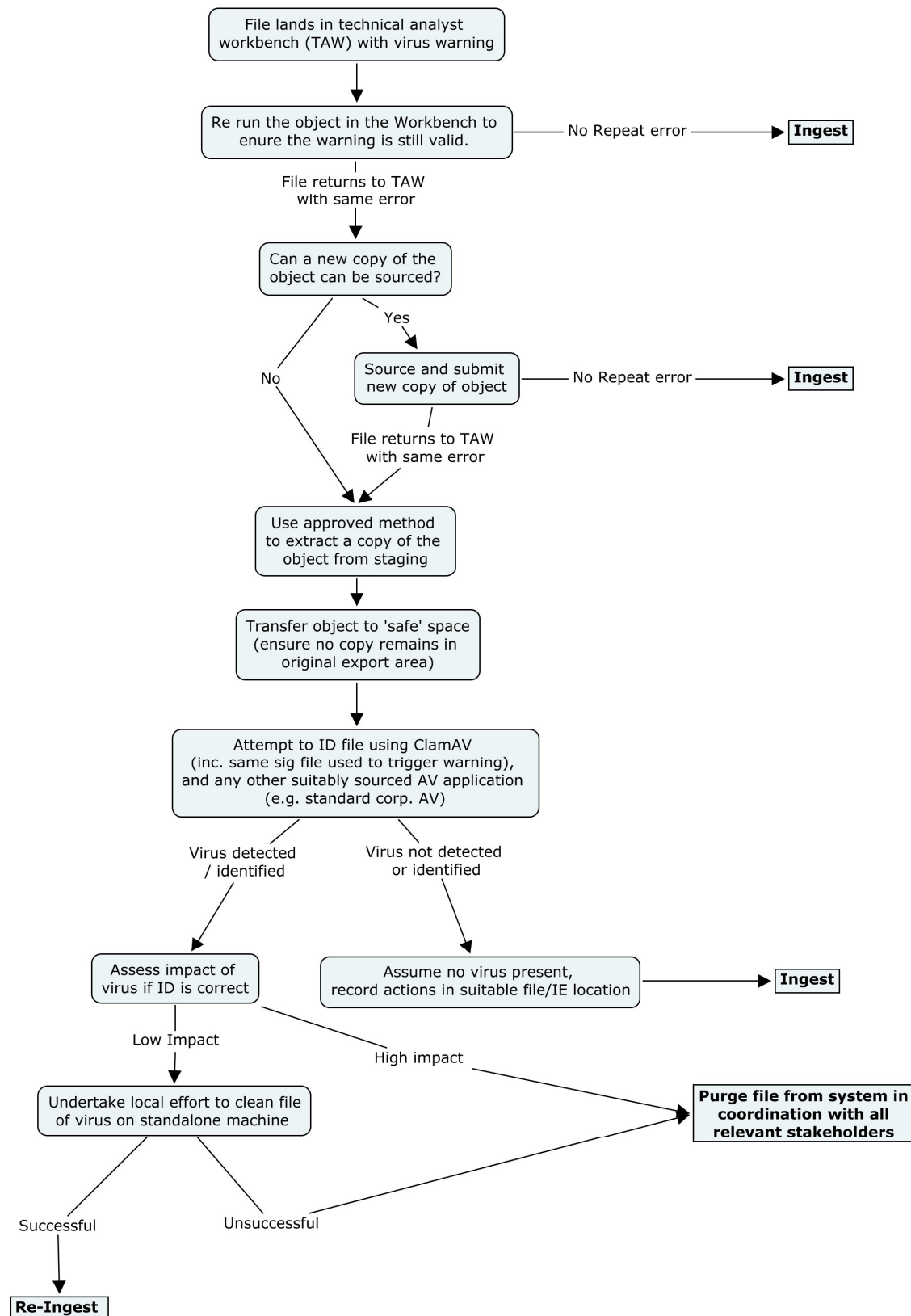


Diagram 1: Process for handling virus warnings on ingest

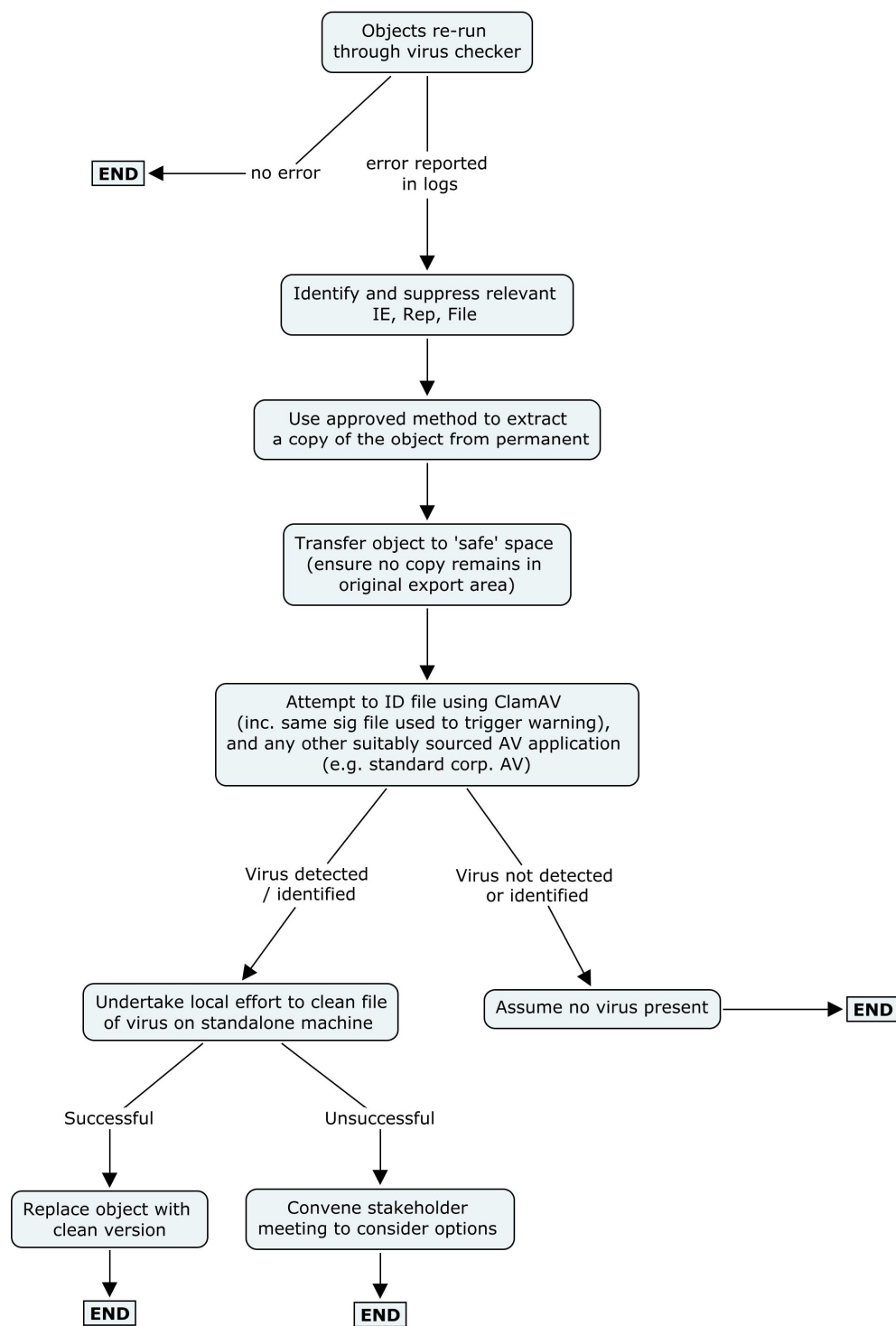


Diagram 2: Virus check on objects in permanent storage.

ACTORS

Technical Analyst:

- Responsible for monitoring the Technical Analyst Workbenches for virus alerts.
- Responsible for moving the infected file to safe storage for testing;
- Responsible for testing the infected file.

GTS Technical Representative

- Responsible for all providing a secure environment into which the infected file will be initially downloaded.
- Responsible for supplying a safe environment within which tests on the infected file can be carried out.
- Responsible for updating definition and security updates.
- Responsible for carrying out the annual review of the anti-virus tools.

System Administrator:

- Responsible for scheduling and running virus checks on content held in permanent storage.
- Responsible for notifying the Technical Analyst of any errors in this process.

Manager/Digital Continuity Manager:

- Responsible for overseeing this policy, and ensuring that no viruses are knowingly ingested into the permanent storage of the preservation system.

Policy Analyst

- Responsible for reviewing this policy annually and making any updates required by that review.

IMPACTS

The impacts of running virus checks on items in permanent should be monitored, with the potential to either increase or decrease the number of annual checks made on the content.

MEASUREMENT

No.	Related Operating Rule	Measure	Parameter
1	4	No viruses are knowingly ingested into the permanent storage of the preservation repository	
2	8, 9	Any infected files found on sweeps of the permanent storage are dealt with within one week of discovery.	

KEY TERMS

Virus: A virus is defined as any code or string of bits that has been identified by anti-virus tools as a risk. It can include, but is not limited to items such as malware, adware and spyware.

RELATED POLICY CHAPTERS

Document	Location

REFERENCES

Security Standards Anti-Virus Policy, <http://1840.dia.govt.nz/policies/information-technology/anti-virus-policy>.

FIXITY ASSURANCE

Date approved	
Approved by	
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

POLICY OVERVIEW

This policy details the requirements for monitoring the integrity of digital objects that have been selected for long-term preservation by Archives New Zealand and National Library of New Zealand.

INTRODUCTION

The National Library of New Zealand Te Puna Mātauranga o Aotearoa (the Library) and Archives New Zealand Te Rua Mahara o te Kāwanatanga (Archives) must be able to prove the integrity of the content they are charged with collecting, preserving and giving access to.

This policy deals with the integrity of digital content. Within this context, integrity of an object is the quality of it remaining “uncorrupted and free of unauthorised and undocumented changes” (UNESCO 2003).

Integrity is foundational to both Archives and the Library. Trust is key to their status as holders of the memory of New Zealand and integrity is one of the blocks upon which this trust is built. Undocumented and unauthorised change to the digital content they preserve is an unacceptable outcome.

SCOPE

This policy applies to digital documentary heritage collected and created by the Library and digital archival items collected and created by Archives that have been selected for long-term preservation.

It applies equally to content that is being held in pre-deposit areas and content held within the permanent repository.

While security is a critical component of protecting the integrity of content, this is dealt with in a separate policy.

It is understood that integrity plays a role in tracing the authenticity of heritage items. However, this policy is concerned only with monitoring integrity from the time that digital objects come under the control of the Library or Archives. Descriptive and other contextual information is not included in this policy.

POLICY GOAL

To ensure that all content under the control of the Archives and Library can be monitored for corruption and unauthorised change.

POLICY OPERATING RULES

1. Any digital file selected for long-term preservation must have information associated with it that can be used to check its integrity.

2. This information will take the form of fixity information. Any proposed form of fixity information should be unique to the file, stable, and accurate enough to aid the monitoring of any possible change to the object. The choice of method must be documented.
3. More than one piece of fixity information should be used for each object to prove its integrity.
4. If deposited with the object, the fixity information should be checked for validity. That is, is the fixity information supplied the same as that generated by the institution? If valid, it should be retained and kept with the object. If not supplied with the object, it should be generated by the institution.
5. The fixity information should be rechecked at the end of every calendar year. A log of this activity must be retained.
6. Where multiple copies are kept of the digital information object, all copies should be checked for their validity.
7. Any movement of the digital object should trigger a recheck of the integrity information.
8. The type of fixity information used should be reviewed annually.
9. Where a new form of fixity information is selected for use, all items in the permanent repository must have this new form associated with it.

ACTORS

Rosetta System Administrator:

- Is responsible for putting in place the mechanism to check the integrity information of objects stored in the permanent repository annually.

Preservation Policy Analyst:

- Is responsible for managing the review process of this policy.

Digital Preservation Manager:

- Is responsible for management of this Fixity Information Policy.

Digital Archivists/staff collecting content:

- This class of users is responsible for a) collecting content from external producers and either a) checking fixity information associated, or b) associating new fixity information.

Internal content creators:

- Responsible for generating initial fixity information for the content they create.

IMPACTS

It is possible that rechecking fixity information could drain system resources. All fixity rechecking should be done as a background task, preferably at the times of lowest use by internal and external users.

MEASUREMENT

No.	Related Operating Rule	Measure
1	Goal	There is no unauthorised and undocumented change to content once it has come under the control of the organisation.
2	1, 2, 3	All digital content has the correct amount and type of fixity information associated at all stages from transfer to the organisation, temporary storage, and ingest to the preservation system.
3	2	Documentation is in place that details current fixity information practice, including its generation and rechecking.

KEY TERMS

Fixity Information: A stored value, derived from the file properties that is used as a basis for assuring physical file integrity.

Integrity: The state of being whole, uncorrupted and free of unauthorised and undocumented changes. (UNESCO)

Control: The point of control for the purpose of this document is taken to be the point that the digital content is submitted or transferred to Archives and the Library or created by internal producers.

RELATED POLICY CHAPTERS

Document	Location

REFERENCES

UNESCO 2003, *Guidelines for the Preservation of Digital Heritage*, Prepared by the National Library of Australia, Paris: UNESCO. <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>.

NOTES OF DIVERGENCE

There are no areas of divergence.

ROSETTA DATA FILES BACKUP

Date approved	
Approved by	
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

POLICY OVERVIEW

INTRODUCTION

BACKGROUND

There are many risks that if they became issues would cause loss to the data files preserved in, and critical to the on-going successful use of, the Rosetta Digital Preservation System.

These risks are documented in the Rosetta File Integrity Policy and include such things as media failure, malicious destruction of data (through hacking or other security breach), natural disaster and random bit-rot.

In order to mitigate against these risks one of the approaches that can be taken is to create multiple “backup” copies of valued files and store these on different media in different locations.

This policy covers the rules surrounding the backup of valued files stored in the Rosetta DPS and its supporting Oracle database.

SCOPE

The Rosetta Digital Preservation System (DPS) installation used by Archives New Zealand and the National Library of New Zealand includes a number of data-files that are either components of digital public archives and documentary heritage objects that have been ingested for long term preservation, or are critical to supporting the management of and access to those digital objects. These data-files can be separated into three main components that may have differing backup requirements. These classes of files and other excluded classes are documented in the policy coverage table below.

#	Class of Data Files	Description	Coverage
1.	Oracle Database Data Files	The Oracle Database Data Files are the files that are used to store metadata about the digital objects stored in Rosetta. This metadata is used to manage the objects in Rosetta and can	Included
2.	Access Derivative Representation Files	The Access Derivative Representation files are the files make up the Access Derivative Representation Digital Objects in Rosetta.	Included
3.	Preservation Master and Modified Master	“The Preservation Master and Modified Master Representation Files and their DNX Formatted	Included

	Representation Files and their DNX Formatted Metadata Files.	<p>Metadata Files” are the files that make up the Preservation Master representations and Modified master representations of the Digital Public Archives and Documentary heritage objects stored in Rosetta. Their DNX-formatted metadata files are also included in this category. These are the files that document the technical aspects of the representation files and include limited descriptive metadata about the contents of those digital objects.</p> <p>These metadata files provide an ultimate fall-back in the event that the rest of the system fails and the Repository has to be repopulated from the backups.</p> <p>This class of files are the most valued and important files in the Rosetta DPS and are the essential asset files that the system is intended to help preserve.</p>	
4.	Rosetta System Files	“Rosetta System Files” are the software application files that make up the Rosetta DPS.	Excluded
5.	Rosetta System Documentation Files	“Rosetta System Documentation Files” are the Files that include documentation about the use and maintenance of the Rosetta DPS	Excluded
6.	Files that are part of tools used to access objects in Rosetta	“Files that are part of tools used to access objects in Rosetta” includes all files that make up the software applications that enable users to open or otherwise access the digital objects in Rosetta.	Excluded
7.	Files that are part of applications that are part of the Institutional Technology Profile of Archives NZ or NLNZ	“Files that are part of applications that are part of the Institutional Technology Profile of Archives NZ or NLNZ” are the files that make up the applications that provide the ability for Archives NZ and NLNZ to render the objects that they preserve in the Rosetta DPS.	Excluded
8.	Files that are part of Archway, Tapui, Voyager or and other documentation/context control system or access catalogue.	“Files that are part of Archway, Tapui, Voyager or and other documentation/context control system or access catalogue” are the files that make up or support the documentation systems that include detailed descriptions of the content or access aids for the content in the files preserved in the Rosetta DPS	Excluded

POLICY GOAL

To ensure that enough copies are made and retained of files in the Rosetta Repository so that the rules and requirements outlined in the Rosetta File Integrity Policy are fulfilled.

ASSUMPTIONS

1. Access Derivative files and Oracle Database files are not as valued as Preservation Master, Modified Master Representation files and their DNX Formatted Metadata files. This is because the access copies can be recreated (at a cost) from the preservation master and modified master files and the Oracle Database files can be recreated by re-ingesting the preservation master and modified master files. This difference in value means that Access Derivative and Oracle Database Files can have a backup strategy applied to them with a greater risk of loss associated with it (e.g. fewer copies can be made of these files or copies can be stored in places with similar risk profiles etc).
2. All files need to be backed up (have multiple copies made of them and those copies potentially stored in different places) in order to reduce the risk of loss to those files.
3. Storing copies of files in the same building increases the risk of loss to all of those copies.
4. Storing copies of files in geographic locations with the same risk profile increases the risk of loss to all of those copies.

POLICY RULES

Rules for Oracle Database Data Files

1. At least two copies of each file (including the primary copy) should be created and stored on different media within one day of the file being created.
Rationale: Data files can be lost due to a number of contributing risk factors including random changes to storage media, failures in storage media, failures in infrastructure that manage storage media, accidental deletion, deliberate malicious deletion, natural disaster etc. Creating and storing two copies of each file on different media within a day is perceived to provide the minimum number of copies required such that there will always be at least one copy of each file available in the event one or more of these risks becoming issues within that time period, while also satisfying the risk appetite and budgetary limitations of Archives NZ and NLNZ.
2. At least two copies of each file should be stored in different buildings within three days of the file being created (i.e. at least one copy must be in one building while at least one other copy is in another building).
Rationale: Many risks that may affect the integrity of data files are related to the properties of the building in which they are stored including fire risks, security risks and accidental damage risks. By storing a copy of a data file in a different building these risks are mitigated.
3. At least three copies of each file should be stored in different buildings within two weeks of the file being created (i.e. at least one copy must be in one building while at least one other copy is in second building and another copy is in a third building).
Rationale: Many risks that may affect the integrity of data files are related to the properties of the building in which they are stored including fire risks, security risks and accidental damage risks. By storing copies of each data file in different buildings these risks are mitigated.
4. At least one copy of each file should be stored in a geographic location with a different risk profile within two weeks of the file being created.
Rationale: Many risks that may affect the integrity of data files are related to the Geographical location in which they are stored including Natural Disaster and security related risks. By storing a copy of a data file in a geographical location with a different risk profile these risks are mitigated. Two weeks is perceived to be the minimum time necessary to achieve this while satisfying the risk appetite and budget of Archives NZ and NLNZ.

5. At least three copies of each file (including the primary copy) should be created and stored on different media within one week of the file being created.
Rationale: Data files can be lost due to a number of contributing risk factors including random changes to storage media, failures in storage media, failures in infrastructure that manage storage media, accidental deletion, deliberate malicious deletion, natural disaster etc. Creating and storing three copies of each file on different media within a week is perceived to provide the minimum number of copies required such that there will always be at least one copy of each file available in the event one or more of these risks becoming issues at that time period, while also satisfying the risk appetite and budgetary limitations of Archives NZ and NLNZ.

Rules for Access Derivative Representation Files

6. At least two copies of each file (including the primary copy) should be created and stored on different media within one day of the file being created.
Rationale: Data files can be lost due to a number of contributing risk factors including random changes to storage media, failures in storage media, failures in infrastructure that manage storage media, accidental deletion, deliberate malicious deletion, natural disaster etc. Creating and storing two copies of each file on different media within a day is perceived to provide the minimum number of copies required such that there will always be at least one copy of each file available in the event one or more of these risks becoming issues within that time period, while also satisfying the risk appetite and budgetary limitations of Archives NZ and NLNZ.
7. At least three copies of each file should be stored in different buildings within two weeks of the file being created (i.e. at least one copy must be in one building while at least one other copy is in second building and another copy is in a third building).
Rationale: Many risks that may affect the integrity of data files are related to the properties of the building in which they are stored including fire risks, security risks and accidental damage risks. By storing copies of each data file in different buildings these risks are mitigated.
8. At least one copy of each file should be stored in a geographic location with a different risk profile within two weeks of the file being created.
Rationale: Many risks that may affect the integrity of data files are related to the Geographical location in which they are stored including Natural Disaster and security related risks. By storing a copy of a data file in a geographical location with a different risk profile these risks are mitigated. Two weeks is perceived to be the minimum time necessary to achieve this while satisfying the risk appetite and budget of Archives NZ and NLNZ.
9. At least three copies of each file (including the primary copy) should be created and stored on different media within one week of the file being created.
Rationale: Data files can be lost due to a number of contributing risk factors including random changes to storage media, failures in storage media, failures in infrastructure that manage storage media, accidental deletion, deliberate malicious deletion, natural disaster etc. Creating and storing three copies of each file on different media within a week is perceived to provide the minimum number of copies required such that there will always be at least one copy of each file available in the event one or more of these risks becoming issues at that time period, while also satisfying the risk appetite and budgetary limitations of Archives NZ and NLNZ.

Rules for Preservation Master Representation Files, Modified Master Representation Files and Their DNX Formatted Metadata Files.

10. At least two copies of each file (including the primary copy) should be created and stored on different media within one day of the file being created.
Rationale: Data files can be lost due to a number of contributing risk factors including random changes to storage media, failures in storage media, failures in infrastructure that manage storage media,

accidental deletion, deliberate malicious deletion, natural disaster etc. Creating and storing two copies of each file within a day is perceived to provide the minimum number of copies required such that there will always be at least one copy of each file available in the event one or more of these risks becoming issues within that time period, while also satisfying the risk appetite and budgetary limitations of Archives NZ and NLNZ.

11. At least three copies of each file should be stored in different buildings within two weeks of the file being created (i.e. at least one copy must be in one building while at least one other copy is in second building and another copy is in a third building).

Rationale: Many risks that may affect the integrity of data files are related to the properties of the building in which they are stored including fire risks, security risks and accidental damage risks. By storing copies of each data file in different buildings these risks are mitigated.

12. At least one copy of each file should be stored in a geographic location with a different risk profile within two weeks of the file being created.

Rationale: Many risks that may affect the integrity of data files are related to the Geographical location in which they are stored including Natural Disaster and security related risks. By storing a copy of a data file in a geographical location with a different risk profile these risks are mitigated. Two weeks is perceived to be the minimum time necessary to achieve this while satisfying the risk appetite and budget of Archives NZ and NLNZ.

13. At least four copies of each file (including the primary copy) should be created and stored on different media within one week of the file being created.

Rationale: Data files can be lost due to a number of contributing risk factors including random changes to storage media, failures in storage media, failures in infrastructure that manage storage media, accidental deletion, deliberate malicious deletion, natural disaster etc. Creating and storing four copies of each file on different media within a week is perceived to provide the minimum number of copies required such that there will always be at least one copy of each file available in the event one or more of these risks becoming issues at that time period, while also satisfying the risk appetite and budgetary limitations of Archives NZ and NLNZ.

ACTORS

IMPACTS

MEASUREMENT

KEY TERMS

RELATED POLICY CHAPTERS

Document	Location

REFERENCES

NOTES OF DIVERGENCE

REPRESENTATIONS

Date approved	
Approved by	
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

POLICY OVERVIEW

This Policy determines that there can be three types of representation within an Intellectual Entity. It also details the points of difference between these types.

INTRODUCTION

The Library's model of preservation works on the basis that the thing to be preserved, the digital content, is contextualised as an Intellectual Entity. That entity can be presented in different forms. These forms are known as representations. For example, a digitised book may be represented in both PDF format and as jpeg image per page. These two representations are portraying the same content, but through a different format.

Currently the Library has defined three varieties of representation, which are described below. This policy is written however during a period where there is reassessment of the number of types and where the technical infrastructure is being opened to allow for the definition of more types.

SCOPE

The definition of the representations should be used with caution with any materials other than those under the control of the preservation system. The terms are used throughout the Library, and while some areas are consistent with NDHA meaning, others are not.

POLICY GOAL

To define clearly the representation types that can exist for digital content within the digital preservation system.

POLICY OPERATING RULES

1. There are three representation types. These are:
 - a. Preservation Master: A representation of intellectual entity that has long-term preservation value. This is the intellectual entity in its most pristine form.
 - b. Modified Master: A representation of the intellectual entity with long-term preservation value. It differs from the preservation master through work undertaken on the files before they are deposited with the preservation system. This work is done for various reasons and can include actions such as de-skewing or inversion. Modified masters are most common for digitised content.

- c. **Derivative Copy:** A copy of the intellectual entity that has been created as the accessible representation. The format and parameters used should be controlled by the policy that has defined the need for the derivative.
2. Preservation and modified masters can have multiple versions. Versions of the representations are generated when a newer version is associated with the IE. The rationale behind this can be two-fold:
 - a. The current representations are of insufficient quality and a new version has been created. This is only applicable for content that has been digitised.
 - b. A preservation action has generated a new version of the representation.
3. The decision to make derivative copies of masters, generally used to lower the barriers to accessing digital content, should take account of relevant access and reproduction policies, network considerations and end user requirements.

Rationale:
4. Processes that create derivative copies must be robust and result in a valid representation of the intellectual entity. The processes must be signed-off by relevant curators, collection managers, content specialists and preservation analysts.

Rationale:
5. Derivative copies are not collection items. That is, they will not be subject to preservation considerations.

Rationale:

ACTORS

Content owners:

- Responsible for provide derivative requirements.
- Responsible for quality assuring the derivative creation process.

Preservation analyst:

- Responsible for testing for advise on derivative formats.
- Responsible for testing the derivative creation process.

System Administrator

- Deploy derivative generation plugins.

IMPACTS

MEASUREMENT

KEY TERMS

RELATED POLICY CHAPTERS

Document	Location

REFERENCES

NOTES OF DIVERGENCE

ACCEPTABLE CHANGE

Date approved	
Approved by	
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

FORMAT LIBRARY USAGE

Date approved	
Approved by	
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

RISK MANAGEMENT

Date approved	
Approved by	
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

DISASTER RECOVERY

Date approved	
Approved by	
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

SECURITY

Date approved	
Approved by	
Review date	One year from approval date
Policy owner	Joint Operations Group - Policy
Version	1

REFERENCES

List of documents referenced in the policies.

GLOSSARY

List of key terms used in the policies.