Te Rua Mahara o te Kāwanatanga
ARCHIVES
NEW ZEALAND

Archives New Zealand
Te Rua Mahara o te Kāwanatanga

Digital Storage Policy

June 2021

CONTENT

1 Introduction

How and where digital information and records are stored affects their viability over time. This Policy outlines principles or requirements for three different types of digital storage required by Archives New Zealand, Te Rua Mahara o Te Kāwanatanga.

Type of storage	Description
Digital preservation storage	Digital storage for long term preservation
Access storage	Digital storage where access copies and access derivatives are kept
Business storage	Digital storage for business purposes

2 Scope

The policy is divided into three parts and covers:

- digital preservation storage
- digital access storage, and
- digital business storage.

The policy does not cover storage hardware procurement, installation or normal hardware maintenance activities.

3 Digital preservation storage

This part of the Digital Storage Policy deals specifically with the storage of born-digital information, records and data of archival value (i.e., digital public archives) that are transferred or deposited to the control of the Chief Archivist.

In association with the Digital Preservation Statement, this is a high-level acknowledgement of Archives New Zealand's responsibilities over time to digital public archives regardless of format, storage cost and type. It affirms Archives New Zealand's commitment to ensure that the digital public archives in its care are actively managed and stored in a way that reflects their status as an asset and taonga for present and future generations.

Archives New Zealand has a statutory responsibility to acquire, preserve and manage public archives and ensure accessibility to their content now and into the future.[1] This obligation should be supported by any contractual, jurisdictional or funding arrangements that may be part of any current and future digital preservation storage decisions.

[1] *Public Records Act 2005* sections 3(c)(ii), 3(f) and 11(c).

Storage and storage management is a crucial pre-requisite for active digital preservation, which is not only about storing born-digital information and records but also about undertaking preservation measures to maintain access to content in obsolete formats in the future.

Digital preservation is defined by the Digital Preservation Coalition as "the series of managed activities necessary to ensure continued access to digital materials for as long as necessary"[2]. The principles for digital preservation storage are based on international best practice.[3] See Appendix A for a list of these principles and whether they are mandatory or desirable for digital preservation storage.

4 Digital access storage

This part of the Digital Storage Policy deals specifically with the storage of digitised copies of physical archives held by Archives New Zealand, created in-house or by a third-party provider.

The use of the term 'access' does not define who may or may not access the digital content. Such restrictions are governed by legislation and existing Archives New Zealand policies. Storage for access prioritises the user experience through speed and capacity. See Appendix A for a list of principles and whether they are mandatory or desirable for access storage.

5 Digital business storage

This part of the Policy deals specifically with the storage of digital information and records created and maintained for Archives New Zealand's internal business purposes.

Archives New Zealand's digital information and records must be created and maintained according to the requirements of the Public Records Act 2005 and the principles of the Information and Records Management Standard (16/S1). See Appendix A for a list of these principles and whether they are mandatory or desirable for business storage.

[2] Digital Preservation Coalition, "What is Digital Preservation?", Digital Preservation Topical Note 1, n.d. Retrieved from https://www.dpconline.org/docs/knowledge-base/1862-dp-note-1-what-is-digital-preservation/file.

[3] A. Goethals, N. McGovern, S. Schaefer, G. Truman, and E. Zierau. "An overview of the Digital Preservation Storage Criteria and Usage Guide", Proceedings of the 16th International Conference on Preservation of Digital Objects (iPres 2019). Retrieved from https://osf.io/sjc6u/ DOI 10.17605/OSF.IO/SJC6U; ISO 16363:2012. "Space data and information transfer systems — Audit and certification of trustworthy digital repositories", 2012; CoreTrustSeal. "Core Trustworthy Data Repositories Extended Guidance v.1.1", 2018. Retrieved from https://www.coretrustseal.org/why-certification/requirements/.

6 Outcome statements

Archives New Zealand:

- uses international standards and best practice to meet its digital preservation, access and business storage responsibilities;
- acknowledges information security, privacy, Māori data and cultural considerations;
- recognises Te Tiriti o Waitangi (Treaty of Waitangi) and acts in accordance with its principles and obligations and with Aratohu Ahurea Tikanga cultural protocols;
- maintains an exit strategy from any technical storage solutions/systems;
- keeps any technical solutions/systems updated;
- ensures that technology and resources have minimal constraint on decisions about the storage of digital public archives; and
- commits to the full range of digital preservation storage requirements over time, including working for a suitable and sustainable legal and economic environment.

This means:

- Digital public archives and their intellectual content are easily available for the people of New Zealand and the rest of the world to consult, subject to any lawful restrictions.
- There is no loss of, or damage to, digital public archives once accepted into the custody of Archives New Zealand.
- There is a clear commitment to, and funding for, digital preservation storage at Archives New Zealand.

7 Appendix A: List of principles/requirements

Reference	Criteria	Category	Description	Related Criteria and References	Digital preservation storage	Digital access storage	Digital business storage	Remarks
DigiPres 1	Integrity checking	Content integrity	Performs verifiable and/or auditable checks to detect changes or loss in or across copies (e.g. checksum recalculation, fixity checking, identifying missing files)		Mandatory	Mandatory	Mandatory	
DigiPres 2	Independent integrity checking	Content integrity	Supports fixity checking by other parties, for example the content-owning institution		Mandatory	Optional	Mandatory	
DigiPres 3	Cost-efficient	Cost considerations	Costs relatively less overall than other comparable solutions, by being designed with cost efficiencies, for example, has resource pooling and sharing, multi-tenancy (multiple users share the same applications)		Mandatory	Mandatory	Mandatory	
DigiPres 4	Energy-efficient	Cost considerations	Takes advantage of energy conservation principles and techniques in full or in part. For example, requires less cooling, consumes less power, uses less rack space, as in green computing initiatives		Mandatory	Mandatory	Mandatory	
DigiPres 5	Storage weight	Cost considerations	Meets relevant requirements for physical weight as documented in SLA, for example, weight may need to be under a certain amount required for a particular floor.		Optional	Optional	Optional	
DigiPres 6	Adapts to requirements	Flexibility	Able to adjust storage infrastructure in response to changing local requirements, for example legal requirements or audit results		Mandatory	Mandatory	Optional	
DigiPres 7	Constrain location	Flexibility	Enables the specification of the location, e.g. by geographic region or geopolitical characteristics		Mandatory	Optional	Mandatory	If not preservation master, still subject to data sovereignty.
DigiPres 8	Customizable replication	Flexibility	Supports user-defined replication rules, for example fewer copies of a particular stream of content		Mandatory	Optional	Optional	

Reference	Criteria	Category	Description	Related Criteria and References	Digital preservation storage	Digital access storage	Digital business storage	Remarks
DigiPres 9	Interoperability	Flexibility	Includes storage components that can be easily integrated with other systems and applications (i.e. plug and play), for example uses standard file access protocols and file system semantics such as Network File System (NFS), SMB, Rest APIs		Mandatory	Mandatory	Mandatory	
DigiPres 10	Open source	Flexibility	Includes storage components that can be integrated with open source tools, systems, and services in accordance with the organization's preferences		Mandatory	Optional	Mandatory	
DigiPres 11	Replaceability	Flexibility	Separates storage layer from other systems in the digital preservation environment so that it could be independently refreshed or replaced without affecting the entire infrastructure	(DP Storage WG, 2018, Independence section)	Mandatory	Optional	Optional	
DigiPres 12	Serviceability	Flexibility	Allows for storage maintenance and changes over time without disruption to availability		Mandatory	Mandatory	Mandatory	
DigiPres 13	Access controls	Information security	Provides role-based, access controls for storage infrastructure, e.g. user, staff, admin, to ensure only the appropriate people have the appropriate levels of access		Mandatory	Mandatory	Mandatory	
DigiPres 14	At-rest server-side encryption with managed keys	Information security	Provides encryption, if required, at the storage layer, with no keys for customers to manage		Optional[4]	Optional	Optional	
DigiPres 15	At-rest server-side encryption with self-managing keys	Information security	Provides encryption, if required, at the storage layer, but customers manage encryption keys		Optional[4]	Optional	Optional	
DigiPres 16	Authentication integration	Information security	Integrates relevant organizational authentication systems to authenticate internal and external users of the system.		Mandatory	Optional	Mandatory	

[4] While this feature is optional we strongly advice against applying encryption (or compression) on files in digital archive in general. As the Digital Preservation Coalition Handbook says: "Information security methods such as encryption add to the complexity of the preservation process and should be avoided if possible for archival copies. Other security approaches may therefore need to be more rigorously applied for sensitive unencrypted files; these might include restricting access to locked-down terminals in controlled locations (secure rooms), or strong user authentication requirements for remote access". https://www.dpconline.org/handbook/technical-solutions-and-tools/information-security

Reference	Criteria	Category	Description	Related Criteria and References	Digital preservation storage	Digital access storage	Digital business storage	Remarks
DigiPres 17	Encrypted transfer	Information security	Uses an appropriate transport layer encryption at all times when moving content		Mandatory	Mandatory	Mandatory	
DigiPres 18	Geographical independence	Information security	Stores multiple redundant copies in geographically-separate locations, at sufficient distances apart, that are not prone to the same natural and human-made disasters and risks	(DP Storage WG, 2018, Bit Integrity section), ISO 27001 and other standards ...	Mandatory	Optional	Mandatory	
DigiPres 19	Multi-tenancy	Information security	Supports separate roles/rules/access controls for separate agencies/departments/colleges/faculties etc	(SNIA, 2017)	Mandatory	Optional	Optional	
DigiPres 20	Organizational independence	Information security	Manages copies under different organizations, preventing any single organization or individual from causing risk to all copies of the content		Mandatory	Optional	Mandatory	
DigiPres 21	Permanent deletion	Information security	Supports requisite deletion by authorized users, in accordance with local policies and rules, in a way that prevents deleted files from being recovered	(SNIA, 2017)	Mandatory	Mandatory	Mandatory	
DigiPres 22	Replication	Information security	Has documented ability to create redundant, distributed copies of content in reasonable timeframes		Mandatory	Optional	Mandatory	
DigiPres 23	Security protocols	Information security	Includes protective measures, controls, and documented procedures to prevent security incidents related to hardware, software, personnel, and physical structures, areas and devices.		Mandatory	Mandatory	Mandatory	
DigiPres 24	System error reporting	Information security	Provides immutable logs and/or reports that show all system errors, failures and other critical system activities		Mandatory	Optional	Optional	
DigiPres 25	Technical independence	Information security	Stores individual copies in different technical solutions (platforms, software including operating systems, hardware, configurations) to prevent all copies from being harmed for example by malware, bugs, or other weaknesses associated with a particular technology.		Mandatory	Optional	Optional	
DigiPres 26	Virus/malware detection	Information security	Includes software that regularly runs virus checks and malware detection.		Mandatory	Mandatory	Mandatory	

Reference	Criteria	Category	Description	Related Criteria and References	Digital preservation storage	Digital access storage	Digital business storage	Remarks
DigiPres 27	Virus/malware remediation	Information security	Provides remediation actions for content with viruses and/or malware, e.g. quarantine, notification, etc.		Mandatory	Mandatory	Mandatory	
DigiPres 28	Diverse storage media types	Resilience	Uses different storage media types / configurations / providers together so that desired levels of independence can be achieved	(DP Storage WG, 2018, Independence section)	Mandatory	Optional	Optional	
DigiPres 29	Durable media	Resilience	Provides documented and acceptable longevity, failure rates, and technical characteristics of the storage media components		Mandatory	Mandatory	Mandatory	
DigiPres 30	Error control	Resilience	Performs error detection and correction 24/7/365 (e.g. using RAID, Erasure coding, ZFS, triple copies/rebuild)		Mandatory	Optional	Optional	
DigiPres 31	High availability	Resilience	Has a high percentage of uptime, i.e. operational for a long length of time, due to techniques such as eliminating single points of failure by having redundant equipment, load-balanced systems and effective monitoring to detect software or hardware failures	(SNIA, 2017)	Mandatory	Mandatory	Mandatory	
DigiPres 32	High resilience	Resilience	Adapts under stress or faults (e.g. resilient to equipment failures, power outages, attacks, surges in user demand)		Mandatory	Mandatory	Mandatory	
DigiPres 33	Recovery and repair	Resilience	Reviews and replaces or repairs missing or corrupt files in acceptable time frames, in a manner that does not propagate errors; or provides ability and tools to perform these actions independently, e.g. by the content-owning institution	Data error notification criteria, Self-healing transparency criteria	Mandatory	Optional	Mandatory	
DigiPres 34	Complete exports	Scalability & performance	Supports the bulk exporting of content and metadata for any reason, at an acceptable rate, for example, as part of an exit strategy		Mandatory	Mandatory	Mandatory	
DigiPres 35	Compute power	Scalability & performance	Meets specified/negotiated computing power for the system or service as documented in the SLA		Mandatory	Mandatory	Mandatory	

Reference	Criteria	Category	Description	Related Criteria and References	Digital preservation storage	Digital access storage	Digital business storage	Remarks
DigiPres 36	Delivery	Scalability & performance	Meets expectations for delivery from the storage layer, e.g. at a reasonable/negotiated rate and supporting concurrent users		Mandatory	Mandatory	Mandatory	
DigiPres 37	File system limits	Scalability & performance	Able to support long file, path or directory names; large amount of files in a directory, and diverse character encodings		Mandatory	Mandatory	Mandatory	
DigiPres 38	I/O performance	Scalability & performance	Meets specified/negotiated input/output performance levels for the system or service as documented in the SLA		Mandatory	Mandatory	Mandatory	
DigiPres 39	Multiple storage tiers	Scalability & performance	Supports use of multiple storage tiers with different availability levels, e.g. on-line, near-line, off-line		Mandatory	Optional	Optional	
DigiPres 40	Scalable to large data sizes	Scalability & performance	Able to support very large amounts of content, in terms of number and size of files, and overall volume		Mandatory	Mandatory	Mandatory	
DigiPres 41	Supports expansion	Scalability & performance	Can increase storage capacity over time as needed in accordance with any SLAs		Mandatory	Mandatory	Mandatory	
DigiPres 42	Supports reduction	Scalability & performance	Can decrease storage over time to support deaccessions, transfer of ownership, etc.		Mandatory	Mandatory	Mandatory	
DigiPres 43	Tiered performance	Scalability & performance	Meets specified/negotiated performance levels appropriate to material being stored, e.g. Tier1 storage for metadata indexing and searching, Tier2 for caching, Tier3 or lower for bulk storage.		Mandatory	Optional	Optional	
DigiPres 44	Accessibility	Support	Ensures people with disabilities equivalent access to reports, documentation and other content		Mandatory	Mandatory	Mandatory	This is not storage
DigiPres 45	Independent preservation services	Support	Supports digital preservation services (e.g. migration and transformations with auditable results) by other parties or external tools		Mandatory	Optional	Optional	This is not storage
DigiPres 46	Support commitment	Support	Documents commitment to support storage infrastructure, e.g. through SLAs (addressing for example responsibilities, data assurance, response times, end-of-service exit provisions, etc.)		Mandatory	Mandatory	Mandatory	This is not storage

Reference	Criteria	Category	Description	Related Criteria and References	Digital preservation storage	Digital access storage	Digital business storage	Remarks
DigiPres 47	Training	Support	Provides requisite training to appropriate staff across all relevant operational and maintenance tasks		Mandatory	Optional	Optional	This is a non-functional requirement
DigiPres 48	Activity monitoring	Transparency	Supports ability to observe or check activity in the storage infrastructure (e.g. see activity in real-time, examine logs, observe the performance status, determine the overall status or drill-down into activities)		Mandatory	Optional	Optional	
DigiPres 49	Activity reporting	Transparency	Provides reports about activity in the storage infrastructure (e.g. fixity or virus results, corruption, replacement with good copies)		Mandatory	Optional	Optional	
DigiPres 50	Allow audits	Transparency	Support independent audits of storage infrastructure and practices in accordance with the SLA		Mandatory	Optional	Optional	
DigiPres 51	Assessment information	Transparency	Provides information needed to support assessments, certifications, audits, and other business activities through for example, documentation, reports, or walkthroughs		Mandatory	Optional	Optional	
DigiPres 52	Content reporting	Transparency	Provides reports about content in the storage infrastructure (e.g. number of objects/files/formats, average file size, types of objects, size of storage in use)		Mandatory	Mandatory	Mandatory	
DigiPres 53	Custom reporting	Transparency	Supports custom (for example configurable and/or on-demand) reporting of content or activity in the storage infrastructure		Mandatory	Mandatory	Mandatory	
DigiPres 54	Data error notification	Transparency	Notifies content-owners of all data errors, remediation actions and issues in reasonable/expected/negotiated timeframes		Mandatory	Mandatory	Mandatory	
DigiPres 55	Documented access	Transparency	Provides immutable logs and/or reports that show all system access		Mandatory	Optional	Mandatory	
DigiPres 56	Documented infrastructure	Transparency	Provides full, complete, current, and available documentation of key processes, services, systems, procedures, known limitations and functions		Mandatory	Optional	Optional	
DigiPres 57	Documented provenance	Transparency	Documents audit/provenance information about all changes, for example about integrity check failures, deletions, modifications, additions, preservation actions; and who or what performed the actions		Mandatory	Optional	Optional	

Reference	Criteria	Category	Description	Related Criteria and References	Digital preservation storage	Digital access storage	Digital business storage	Remarks
DigiPres 58	Expose location	Transparency	Exposes the specific storage location of data to meet SLA requirements		Mandatory	Optional	Optional	
DigiPres 59	Management across storage tiers	Transparency	Supports management and monitoring across multiple storage availability levels, e.g. on-line, near-line, off-line		Mandatory	Optional	Optional	
DigiPres 60	Open storage formats	Transparency	Supports open, standard, non-proprietary storage formats, e.g. TAR, archive eXchange format (AXF), LTFS		Mandatory	Optional	Optional	
DigiPres 61	Self-healing transparency	Transparency	Provides content owners with documentation or notification about any automatic correction or change of data to meet SLA requirements		Mandatory	Optional	Optional	
Access 1a	Speed	Prioritise use experience			Optional	Mandatory	Optional	
ANZ 16/S1: 2.3	Information and records management must be design components of all systems and service environments where high risk/high value business is undertaken.	Principle 2: Information and records management supports business	An organisation must consider at the start how to make system maintenance, migrations and decommissioning easier. In taking this "by design approach", an organisation must ensure:- systems specifications for information and records that are high-risk, high-value, or both, include requirements for managing them- systems specifications include requirements for minimum metadata needed to support information and records identification, usability, accessibility and context- it keeps documents about systems design, configuration and any changes made over time.		Optional	Optional	Mandatory	
ANZ 16/S1: 2.4	Information and records must be managed across all operating environments.	Principle 2: Information and records management supports business	Identify and document where information and records are created and held, across all system environments and physical locations. By identifying where information and records are held, an organisation can better manage them in diverse system environments, storage environments and physical locations, and give appropriate access.		Optional	Optional	Mandatory	

Reference	Criteria	Category	Description	Related Criteria and References	Digital preservation storage	Digital access storage	Digital business storage	Remarks
ANZ 16/S1: 2.6	Information and records must be maintained through systems and service transitions by strategies and processes specifically designed to support business continuity and accountability.	Principle 2: Information and records management supports business	Implement and review a migration strategy. Migrate information, records and metadata from one system to another using a managed process that results in records that people can access easily and that have trustworthy information. Ensure the portability of information and records is addressed in outsourced or service arrangements. Maintain the systems documentation.		Optional	Optional	Mandatory	
ANZ 16/S1: 3.3	Information and records must be identifiable, retrievable, accessible and usable for as long as they are required.	Principle 3: Information and records are well managed	Information and records must be identifiable, retrievable from storage (physical or digital), and accessible, usable and reusable for as long as required. To maintain the accessibility and usability of physical information and records, an organisation must store them in appropriate storage areas and conditions.		Optional	Optional	Mandatory	
ANZ 16/S1: 3.4	Information and records must be protected from unauthorised or unlawful access, alteration, loss, deletion and/or destruction.	Principle 3: Information and records are well managed	Ensure information security and protection mechanisms are in place. Ensure that assessments or audits can test that access controls are implemented and maintained. Security measures must include: - access and use permissions in systems - processes to protect information and records no matter where they are located, including in transit and outside the workplace - secure physical storage facilities.		Optional	Optional	Mandatory	

Reference	Criteria	Category	Description	Related Criteria and References	Digital preservation storage	Digital access storage	Digital business storage	Remarks
ANZ 16/S1: 3.6	Information and records must be kept for as long as needed for business, legal and accountability requirements.	Principle 3: Information and records are well managed	Information and records must be sentenced and disposed of regularly in line with the practices of authorised disposal authorities. This includes information and records located in business systems, in outsourced or service arrangements, or in physical storage. Disposing of digital information and records may be part of a planned migration process or the decommissioning of systems.		Optional	Optional	Mandatory	
ANZ 16/S1: 3.7	Information and records must be systematically disposed of when authorised and legally appropriate to do so.	Principle 3: Information and records are well managed	An organisation must be able to account for their disposal of information and records in business systems, outsourced arrangements, and physical storage. This includes providing evidence that the disposal of information and records is permitted and authorised under disposal authorities' and legal obligations, including the Public Records Act 2005.		Optional	Optional	Mandatory	