UNSWorks DIGITAL PRESERVATION PROCEDURE

Policy Hierarchy link	UNSWorks Digital Preservation Policy		
Responsible Officer	University Librarian		
Contact Officer	Director, Digital Library Services Department		
Superseded Documents	This is a new procedure		
File Number	2013/20317		
Associated Documents			
Version	Authorised by	Approval Date	Effective Date
1.0	Vice-President and Deputy Vice-Chancellor (Academic)	29 November 2013	1 January 2014

Contents

1. Purpose and Scope

This document covers the digital preservation procedure for digital resources held in UNSWorks.

2. Definitions

Key definitions are listed below. Technical terms for preservation actions are listed in Appendix A.

2.1 Preservation copy

A preservation copy is created from the original digital resource deposited in UNSWorks. This is preserved in a preservation package, with preservation metadata and related files. Access copies will be derived from the preservation package. For further detail see 'preservation copy' in Appendix A.

2.2 Level 1 Preservation Support

The Library will provide its highest level of preservation support, making its best effort to maintain the content, structure and functionality of a supported digital resource in the future. The Library will create a preservation copy of the original files (bitstream), preservation metadata and software.

Criteria: Level 1 preservation support is currently provided only for content that represents core business of the University: research output and publications. This includes PhD and Masters by research theses, and copies of research publications eligible for open access.

Level 1 preservation support will only be provided for digital formats in wide use with publicly accessible rendering software and documentation. The main advantage of this approach is that stable or growing user communities typically inspire newer applications

of digital formats, which lead to development of a wide range of supporting software products that increase the longevity of those formats. Additionally, this ensures a sufficient level of existing knowledge about the formats that could be leveraged to develop preservation tools and workflows in the future.

Access: UNSWorks users (see 3.2) will be provided with a preservation copy of the digital resource, preservation metadata and software required for access.

Note: UNSWorks content may be normalized (transformed to another stable format) to facilitate sustained access and re-use.

2.3   Level 2 Preservation Support

The Library will make limited efforts to maintain the usability of a supported digital resource. The Library will create a preservation copy of the original files (bitstream) and preservation metadata, including software metadata required for access.

The encoding file format will be monitored and may be transformed to a more stable format, when significant risk of obsolescence is identified in order to prevent any loss of information content. However, it may not be feasible to control or prevent any changes of structure or functionality of a digital resource resulting from the format migration.

Criteria: Level 2 preservation support is provided for content that supplements core business of the University. This includes research data in any of the supported formats listed in Appendix B.

Access: UNSWorks users will be provided with a copy of the preservation resource, which may not be in the same format as the originally deposited digital resource, and preservation metadata.

Note: The file may be transformed to a more preservable format to ensure that the information content is not lost, even if some structure and functionality are sacrificed.

2.4   Level 3 Preservation Support

The Library will create a preservation copy of the compressed original files (bitstream) and preservation metadata. However, no active effort will be made to monitor the format and associated risks, or to normalize, transform, or migrate the file to a more preservable format. Files may be opened and/or read by future applications, but there is no guarantee that the content, structure, or functionality will be preserved.

Criteria: Level 3 preservation support is provided for digital resources provided in formats that are not supported by Levels 1 and 2. This includes domain-specific formats and/or specialised software applications associated with research materials.

UNSWorks depositors (see 3.1) are required to deposit files for Level 3 preservation in a supported lossless compressed format, such as ZIP or TAR (see Appendix B).

Access: UNSWorks users will be provided with the compressed folder containing the original digital resource(s), along with the corresponding preservation metadata.

Note: Any format not yet reviewed and evaluated by the Library will receive Level 3 support on deposit. A higher level may be assigned after format review takes place.
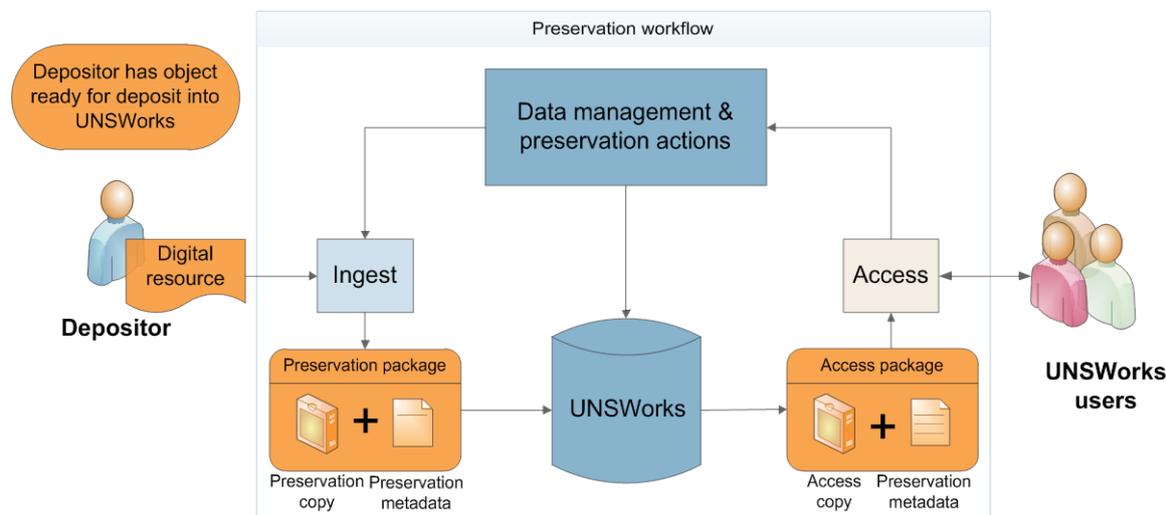
3. Procedure

Figure 1: UNSWorks digital preservation workflow

The UNSWorks digital preservation workflow includes the following main steps:

1. Depositor submits digital resource(s) for deposit in UNSWorks via deposit form;
2. UNSWorks captures preservation metadata about the deposited digital resource(s), creates a preservation package consisting of the preservation copy of deposited digital resource(s) and the preservation metadata, and stores the package for preservation;
3. The resource type determines the preservation level and scheduled preservation actions undertaken by the Library (see Appendix B).
4. Access copies are derived from the preservation copy and made available through the UNSWorks discovery interface.

To realise this workflow, the following roles are defined:

3.1 Depositors

A depositor refers to a member of the UNSW community authorised to deposit supported digital resources, such as PhD or Masters by research theses and scholarly publications, in UNSWorks. Depositors will ensure that objects for deposit are provided in a supported format for the required preservation level (see Appendix B).

3.2 UNSWorks users

An UNSWorks user is a member of the wider research community or the public, who is able to use the UNSWorks discovery interface to find and access the digital resources held in UNSWorks. The user can retrieve an access copy of a digital resource that has been derived from the preservation copy, and use it in accordance with the UNSWorks Rights of use.

3.3 UNSW Library

The UNSW Library will undertake the following responsibilities:
- Provide support and advice to UNSWorks users on preservation policy
- Create a secure preservation copy from deposited resource
- Store standard-compliant preservation metadata (based on PREMIS, an international preservation metadata standard)
- Store copies of necessary software, and software format metadata as appropriate
- Perform maintenance and security operations in line with the Open Archival Information Systems (OAIS) Reference Model, for example:
  - Check file fixity on ingest
  - Virus-check high risk content
  - Check fixity of content at fixed intervals

- o Maintain logs of who performed what actions on files, including deletions and preservation actions
- o Perform format migrations, emulation and similar activities as needed
- o Carry out obsolescence monitoring process for storage systems and media
- o Undertake any actions to detect corrupt data
- o Any other preservation actions as required
- Retain content held within UNSWorks in compliance with minimum standards set out by the General Retention and Disposal Authority – University Records (GDA23) amd the Australian Code for the Responsible Conduct of Research (ACRCR). Refer to Appendix B for UNSWorks preservation levels and retention periods

3.4   Committee on Research

The Academic Board's Committee on Research will provide guidance on interpretation of the UNSWorks Digital Preservation Policy.

The Library will report annually to the Committee on Research the following aspects of digital preservation:
- Levels of preservation support
- Outcome of monitoring and review of content formats for technological obsolescence
- Preservation actions undertaken in the past year
- Review of digital preservation policy, procedure and guidelines.

4. Review & History

Version	Authorised by	Approval Date	Effective Date	Sections modified
1.0	Vice-President and Deputy Vice-Chancellor (Academic)	29 November 2013	1 January 2014	This is a new procedure

5. Acknowledgements

This procedure was created in consultation with the following resources:

Colorado State University, 2013. CSU Digital Repository Preservation Format Support Policy, <http://lib.colostate.edu/repository/csu-digital-repository-preservation-format-support-policy> Accessed 5/8/2013

Consultative Committee for Space Data Systems, 2011. *Audit and certification of trustworthy digital repositories (ISO 16363),* CCSDS 652.0-M-1, <http://public.ccsds.org/publications/archive/652x0m1.pdf> Accessed 5/8/2013

Consultative Committee for Space Data Systems, 2012. *Reference Model for an Open Archival Information System (OAIS)*, Magenta Book, Issue 2, CCSDS 650.0-M-2 <http://public.ccsds.org/publications/archive/650x0m2.pdf> Accessed 7/8/2013

Digital Preservation Coalition, 2012. *Digital Preservation Handbook*, <http://www.dpconline.org/advice/preservationhandbook/media-and-formats/file-format-and-standards> Accessed 28/8/13

Library of Congress (U.S.A.), 2013. *Sustainability of Digital Formats - Planning for Library of Congress Collections,* <http://www.digitalpreservation.gov/formats/content/still.shtml> Accessed 28/8/13

National Digital Stewardship Alliance, 2010. *NDSA Glossary*, <http://www.digitalpreservation.gov/ndsa/ndsa-glossary.html> Accessed 29/7/13

National Digital Stewardship Alliance, 2013. Levels of Digital Preservation Version 1, <http://www.digitalpreservation.gov/ndsa/activities/levels.html> Accessed 9/8/2013

PREMIS 2012, *PREMIS Data Dictionary for Preservation Metadata*, version 2.2, <http://www.loc.gov/standards/premis/v2/premis-2-2.pdf> Accessed 1/8/13

PREMIS Editorial Committee, 2012., *PREMIS Data Dictionary for Preservation Metadata*, version 2.2, <http://www.loc.gov/standards/premis/v2/premis-2-2.pdf> Accessed 1/8/13

UNSW, *Retention periods for records relating to research*, <https://www.recordkeeping.unsw.edu.au/documents/Retentionperiodsforrecordsrelatingtoresearch.pdf > Accessed 1/8/13

Appendix A: Definitions

Note: these definitions are based on the National Digital Stewardship Alliance (U.S.A.) glossary.

Term	Definition
Staff or Academic Staff	Means all employees of UNSW, including continuing, fixed-term and casual employees and persons who are engaged to perform work as contractors of UNSW, or any person who is voluntarily associated with the University.
Researcher	All UNSW staff, conjoint appointments, and visiting appointments undertaking research at UNSW, including staff classified as "professional and technical" and casual staff undertaking research.
Bit Preservation	A baseline preservation approach that ensures the integrity of digital objects and associated metadata over time in their original form, even as the physical storage media which houses them evolves and changes. Also known as "bit preservation."
Checksum	An algorithmically-computed numeric value for a file or a set of files used to validate the state and content of the file for the purpose of detecting accidental errors that may have been introduced during its transmission or storage. The integrity of the data can be checked at any later time by recomputing the checksum and comparing it with the stored one. If the checksums match, the data was almost certainly not altered. See also "Fixity Check."
Derivative	A transformed version of an original source file, often called a "service," "access," "delivery," "viewing" or "output" file, used to facilitate access to or additional use of the content.
Digital content	Any arbitrary item created, published or distributed in a digital form, including, but not limited to, text, data, sound recordings, photographs and images, motion pictures and software. Used interchangeably with "Digital materials" and "Digital resource".
Digital materials	Any arbitrary item created, published or distributed in a digital form, including, but not limited to, text, data, sound recordings, photographs and images, motion pictures and software. Used interchangeably with "Digital content" and "Digital resource".
Digital resource	Any arbitrary item created, published or distributed in a digital form, including, but not limited to, text, data, sound recordings, photographs and images, motion pictures and software. Used interchangeably with "Digital content" and "Digital materials".
Digital preservation	The series of managed activities, policies, strategies and actions to ensure the accurate rendering of digital content for as long as necessary, regardless of the challenges of media failure and technological change.
Fixity check	A mechanism to verify that a digital object has not been altered in an undocumented manner. Checksums, message digests and digital signatures are examples of tools to run fixity checks. Fixity information, the information created by these fixity checks, provides evidence for the integrity and authenticity of the digital objects and are essential to enabling trust. See also "Checksum" and "Digital Signature."
Format migration	A means of overcoming technical obsolescence by preserving digital content in a succession of current formats or in the original format that is transformed into the current format for presentation. The purpose of format migration is to preserve the digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in the face of constantly changing technology.
OAIS Framework	
Permissions	The access available to system users attached to specific roles in a computing environment, as well as the mechanism for administering access to a specific object on a computer system. Depending on the system or application, permissions can be defined for a specific user, specific groups of users, or all users; or for a role, or groups of roles; or based on one or more user attributes.
Preservation copy	Digital content targeted for preservation that is considered the master version of the intellectual content of any arbitrary digital resource. Preservation master files may capture additional information about the original beyond the content itself. Because they are created to

	high capture standards, preservation master files could take the place of the original record if the original was destroyed, damaged, or not retained. Preservation masters generally do not undergo significant processing or editing. Preservation masters are often used to make other copies including reproduction and distribution copies.
Provenance	Information on the origin of a digital object and also on any changes that may have occurred over the course of its life cycle.
PREMIS	"Preservation Metadata Implementation Strategies" an international standard for preservation metadata.
Validation	The process of making sure that data is correct and useful when checked against a set of data validation rules. These might include rules for package or file structure or specific file format profiles.

Appendix B: UNSWorks preservation levels and retention periods

UNSWorks digital resource type	Preservation Level	Retention period	Supported digital formats
PhD and Masters by research theses	Level 1	Retain permanently	• **PDF/A** ISO standardised version of PDF for long-term digital preservation • **PDF** Portable Document Format
Supporting files for Masters by research and PhD theses	Level 1	Retain permanently	• **TIFF** (uncompressed) • **PDF** Portable Document Format • **PDF/A** (for vector graphics) • **Motion JPEG2000** (lossless) • **WAVE-LPCM-BWF** (encoding: uncompressed) • **WAVE-LPCM** (encoding: uncompressed) • **AIFF-LPCM** (encoding: uncompressed) • **CSV strict** (comma-separated values) • **WARC** Web archive format • **XML** Extensible Markup Language
Final research reports	Level 1	Retain permanently (GDA23 – retain as State Archives (GDA23 – 23.4.1)	• **PDF/A** ISO standardised version of PDF for long-term digital preservation • **PDF** Portable Document Format
Open Access research publications (journal articles, conference papers/posters, book/ chapters)	Level 1	Retain permanently	• **PDF/A** ISO standardised version of PDF for long-term digital preservation • **PDF** Portable Document Format
Rights information pertaining to intellectual property (IP)	Level 1	Retain permanently (GDA23 – retain as State Archives (GDA23 – 4.2.4)	• **American Standard Code for Information Interchange (ASCII) Text** • **PDF/A** ISO standardised version of PDF for long-term digital preservation • **PDF** Portable Document Format • **WARC** Web archive format • **XML** Extensible Markup Language

UNSWorks digital resource type	Preservation Level	Retention period	Supported digital formats
Research data (both standalone and associated with a PhD or Masters thesis, or a research report)	Level 2	Retain for maximum of 25 years	• **TIFF** (uncompressed) • **PDF** Portable Document Format • **PDF/A** (for vector graphics) • **Motion JPEG2000** (lossless) • **WAVE-LPCM-BWF** (encoding: uncompressed) • **WAVE-LPCM** (encoding: uncompressed) • **AIFF-LPCM** (encoding: uncompressed) • **CSV strict** (comma-separated values) • **WARC** Web archive format • **XML** Extensible Markup Language
Domain specific files and software (as supporting files for theses or research data)	Level 3	Retain for 25 years	• **ZIP or TAR** (compressed in lossless data compression format)
Working papers	Level 3	10 years (ACRCR – 7 years)	• **PDF/A** ISO standardised version of PDF for long-term digital preservation • **PDF** Portable Document Format