The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

UNIVERSITY OF NOTTINGHAM

LIBRARIES, RESEARCH AND LEARNING RESOURCES

Digital Preservation and Access Policy

2015

Contents

1.0 Document Control

Version	Agent	Date	Notes
0.1	Benjamin Veasey	24-03-2015	First draft
0.2	Benjamin Veasey	10-04-2015	Removed reference to University of Warwick from 7.3
0.3	Benjamin Veasey	10-04-2015	Minor change to Table of Contents
0.4	Benjamin Veasey	16-04-2015	Changes incorporated as a result of feedback from Mark Dorrington Changes incorporated after speaking to Kirsty Lee, Digital Preservation Curator, University of Edinburgh Added section 3.3 - Data Seal of Approval, section 7.5.1 – Staff skills, section 8.0 – audit and certification and section 11.0 - glossary
0.5	Benjamin Veasey	24-04-2015	Changes incorporated as a result of feedback from Christine Middleton Added section 3.6 – ISO 26324:2012
0.6	Benjamin Veasey	06-05-2015	Changes incorporated as a result of feedback from Joanne Chalmers Changes incorporated as a result of a group discussion involving Mark Dorrington, Christine Middleton, Louise Savage, Lydia Johnson, Paul Kennedy, Vineet Chugh and Benjamin Veasey

			Altered the wording of section 2.2
0.7	Benjamin Veasey	26-05-2015	The wording of section 3.4 altered at the suggestion of Paul Kennedy
1.0	LRLR Senior Management Team	01-06-2015	Sign-off from LRLR SMT
1.1	Benjamin Veasey	17-07-2015	Addition to section 7.3
1.2	Benjamin Veasey	26-10-2015	Replaced responsible entities from the generic ("The University") to more specific.
1.3	Mark Dorrington	27-01-2015	Reformatted

Approved by Libraries, Research and Learning Resources Senior Management Team,

1 June 2015

2.0 Aim

2.1 Purpose

The purpose of this policy is to communicate the principles that shall guide the activities to deliver LRLR's digital preservation vision and secure the preservation of its digital information assets. The policy shall set out the University's approach to digital preservation with reference to the following key areas:

- Industry standards and best practice
- Content coverage
- Overview of preservation strategy
- Methods and levels of preservation
- Implementation
- Audit and certification
- Policy review
- Sustainability

2.2 Digital Preservation Vision

To safeguard the University's digital information assets; to ensure their authenticity, reliability and integrity, and to monitor their usability over time, taking appropriate action to protect usability according to best practice and technology available.

2.3 Context

Today, many of the University of Nottingham's digital information assets are born digital. In addition many paper-based and analogue materials are being digitised to reduce physical storage, to safeguard the information they contain and to enable wider dissemination and access.

This valuable content spans academic, administrative and other departments and consists of very different types of materials; e.g. library and archive collections, research project records and datasets, teaching and learning resources, corporate records. Each may have unique characteristics such as access conditions, the need for specific tools to access them, and retention and disposal requirements.

The management of this material, particularly over time, presents new challenges. Although digitisation can be used to preserve the information on vulnerable physical media, the digital medium itself is inherently fragile, dependent on a range of technical processes for us to access and understand the meaningful content encoded within; this meaningful content is at risk. Just as archivists have applied techniques to preserve paper-based materials, digital information will not survive and remain accessible by accident; it requires ongoing active management, both in terms of the data itself as well as the environments required to render the data so that it can be used.

2.4 Scope

This policy applies to the University of Nottingham's records and information assets held in digital form and under the responsibility of Libraries, Research and Learning Resources (LRLR). These include:

- 'Born-digital' acquired by the University according to relevant collecting policies.
- 'Digitised' resources; surrogates scanned or copied from non-digital formats (analogue tape, paper, etc).

3.0 Standards

The first aim of this policy is to ensure that LRLR works towards industry standards and best practice, with regard to the storage and management of digital information assets under its control. The following standards will be used as the basis for meeting the digital preservation vision and a consensus of approach will be developed amongst University functions to ensure that it adopts the aspects suitable for the University environment.

3.1 ISO 15489-1:2001 Records Management Principles

LRLR acknowledges that the primary goal of digital preservation is to maintain access to safe, trustworthy and authoritative information assets held by the archive. ISO 15489-1:2001, a standard for Records Management, summarises this goal through the following four principles:

3.1.1 Authenticity - Adapted from ISO 15489-1:2001

An authentic digital information asset is one that can be proven to satisfy the following characteristics:

1. That it is what it purports to be.
2. That it was created or sent by the agent purported to have created or sent it.
3. That it was created or sent at the time and date it is purported to have been.

3.1.2 Reliability - Adapted from ISO 15489-1:2001

The reliability of a digital information asset is determined by its ability to demonstrate that it has trusted and dependable contents.

3.1.3 Integrity - Adapted from ISO 15489-1:2001

The integrity of a digital information asset is based on proving that its meaningful content is complete and unaltered.

3.1.4 Usability - Adapted from ISO 15489-1:2001

The usability of a digital information asset refers to its ability to be located, retrieved, presented and interpreted.

3.2 ISO 14721:2012 - Open Archival Information Systems reference model

The Open Archival Information Systems (OAIS) model is an international standard that identifies common terms and concepts along with a framework for entities and relationships between entities in a digital preservation archive environment. The OAIS is a conceptual framework and does not act as an implementable system.

In order to achieve its mission the LRLR will reference and follow the broad guidance provided by the OAIS, where appropriate, so that it adheres to best practice.

3.3 ISO 26324:2012 – Digital Object Identifier System

ISO 26324:2012 specifies the syntax, description and resolution functional components of the digital object identifier (DOI) system, and the general principles for the creation, registration and administration of DOI names.

LRLR will adhere to ISO 26324:2012, where appropriate, in order to support its Research Data Management function.

3.4 Information Security Management

Information security management standards will be met by reference to the University's information security policy. This policy is based on the ISO 27001:2013 information security framework and adopts those parts of the standard relevant for a University environment.

3.5 European Framework for Audit and Certification of Digital Repositories

LRLR takes the management of its data seriously. LRLR will therefore aim to acquire the Data Seal of Approval; basic certification within the European Framework for Audit and Certification of Digital Repositories.

4.0 Content Coverage

LRLR will need to preserve an eclectic mix of digital information assets in order to meet its digital preservation vision. We shall develop operating procedures that will link to this policy in order to comply with the following statements.

Digital information assets considered for ingest into the archive will be appraised and retained according to agreed collection policies. It can also be reasonably expected that information assets for ingest may be subject to external policy assessments prior to a request to ingest with the University of Nottingham is made.

The University of Nottingham's digital archive may hold all types of information assets, including, but not limited to:

- - Text documents (plain and marked-up)
- - Still Image collections
- - Datasets (data designed for use in spreadsheets, databases and statistical packages)
- - Digital audio recordings
- - Digital moving image recordings
- - CAD
- - GIS
- - Virtual reality objects
- - Websites
- - Social media snapshots

Where a suitable external discipline-specific archive repository exists the University of Nottingham will accept ingest into that archive assuming that it can meet the minimum requirements set out in this policy. In such a case LRLR will require a persistent link to the digital information assets held externally to be associated with a digital object identifier and a descriptive record maintained by LRLR.

5.0 Overview of Preservation Strategy

In order to meet its digital preservation vision, LRLR recognises that it must adopt a suitable preservation strategy based on the risks associated with its digital information assets.

5.1 Preservation Requirements: Message and the Medium

At the core, there are two risks to digital information assets held by the University of Nottingham; loss of the medium (i.e. loss or corruption of the 0s and 1s that make up the actual bitstreams) and loss of the message (i.e. loss of the ability to correctly interpret the bitstreams as understandable information).

Digital preservation requirements can therefore be sub-divided into two levels of maturity:

5.1.1 Bitstream Preservation

A bitstream preservation function supports the authenticity and integrity of digital information assets stored within the archival storage environment. Essentially bitstream preservation is predicated by the principle of 'what I put in will come back out, intact'.

5.1.2 Content Preservation

A content preservation function seeks to support the usability of digital information assets over time in the face of technological change. Content preservation requires an institution to understand and document what it has got and what is required to correctly interpret the bitstreams so that the content may be rendered; often referred to as characterisation. Furthermore, preservation planning is required which assesses the risks to digital information assets over time based on the information captured through characterisation. Active intervention to maintain the usability of the digital information assets may be required at some point in the future through the implementation of a preservation plan.

5.2 Preservation Strategy: Parsimonious Preservation

LRLR subscribes to the principle of *Parsimonious Preservation,* first put forward by the UK National Archives (Gollins, 2009). This approach places emphasis on the capture of digital information assets and argues that minimal intervention to digital information assets is preferable because this entails 'minimal alteration, which brings the benefits of maximum integrity and authenticity'.

In order to adhere to the principle of *Parsimonious Preservation*, thus supporting its digital preservation vision, those responsible for preserving data within LRLR will adopt a stewardship role by implementing and maintaining:

- A framework for capturing digital information assets and any relevant contextual information;
- a full bitstream preservation function focussing on ensuring that the authenticity and integrity of digital information assets held within the archive is maintained through a secure and sustainable infrastructure, appropriate storage management, access security, integrity verification and administrative metadata management;
- a content preservation function that focusses on the characterisation of digital information assets and preservation planning around their usability.

6.0 Methods and Levels of Preservation

LRLR acknowledges that while the overall aim may be preservation in perpetuity, the best we can do in our (or any) generation is to take a stewardship role. This role focuses on ensuring the survival of material for the next generation of technology.

In order to support its digital preservation vision, LRLR's strategy of a parsimonious approach to preservation will apply to all information assets ingested into the archive; that is LRLR will understand what is being ingested and keep the bitstreams safe. The usability of these assets will be monitored through a preservation planning function.

7.0 Implementation

An important aim of this policy is to communicate the implementation of the digital preservation strategy adopted in support of the digital preservation vision. To be clear, this section is concerned with the functions required to implement the strategy and not the resources or technology to be used.

Those responsible for preserving data within LRLR shall develop a consensus of approach alongside practical operating procedures in order to ensure compliance.

7.1 Procedures for Preservation

7.1.1 Capture

7.1.1.1 Appraisal and Retention

Digital information assets considered for ingest into the archive will be appraised and retained according to agreed collection policies. It can also be reasonably expected that information assets for ingest may be subject to external policy assessments prior to a request to ingest with the University of Nottingham is made.

7.1.1.2 Metadata

Those responsible for preserving data within LRLR will attain suitable metadata for information assets ingested into the archive:

- administrative metadata, including provenance information - used to manage the digital information assets; including rights and permissions;
- technical metadata - describing the technical properties of information assets ingested and the hardware and software needed to maintain usability;
- structural metadata - identifying the relationships between information assets;
- descriptive metadata - meaningful descriptions and labels used to enhance discovery and retrievability

In addition, those responsible for preserving data within LRLR will encourage the ingest of any contextual information relating to the information assets that is likely to aid with supporting their authenticity, reliability, integrity and usability.

Those responsible for preserving data within LRLR shall ensure that the relationship between a digital information asset and its metadata be maintained persistently. This will be achieved by assigning a persistent, unique identifier to every digital information asset at the point of ingest and recording this within the associated metadata to provide a persistent link. This policy is neutral as to the manner in which this is achieved.

7.1.1.3 Rights Management

Those responsible for preserving data within LRLR shall attain a legally-binding agreement from any organisation, department or individual wishing to ingest collections into the archive. This agreement shall confirm the rights and obligations of both parties and provide an opportunity for those providing information assets to specify the conditions under which access may be given.

LRLR expects those who provide information assets to undertake a review of ethical issues relating to any potential ingest to ensure that there is no risk associated with sharing the content with other parties. An assertion of copyright and intellectual property rights shall be sought to ensure the provider of the information assets has cleared any necessary permissions.

LRLR will not ingest information assets into the archive where ownership is unclear or in dispute.

Those responsible for preserving data within LRLR shall seek a license agreement to allow it to perform all necessary actions required for the preservation of the ingested content and allow it to fulfil any legal or regulatory obligations. This license agreement shall also allow the University to make available for access the ingested material, subject to any limitations placed, to its user community.

7.1.2 Bitstream Preservation

A sustainable infrastructure for managing and preserving digital information assets will be maintained. Information Services or a suitable third party shall be responsible for providing this environment.

7.1.2.1 Archival Storage

Archival storage shall be delivered by Information Services or a suitable third party that will include:

- media selection –suitable media for archival storage will be used;
- media refreshment – media will be monitored and either refreshed or replaced periodically based upon the relationship between the longevity of the medium, and that of its supporting technology. Every media refreshment action will be verified at the bit level, to ensure that the content has been copied without corruption or loss. Should corruption or loss occur then these copies will be replaced using redundant copies;
- redundancy – Information Services, or a suitable third party, will maintain multiple redundant copies, stored in at least two different media types in at least two different geographically separated locations. Redundant copies will be periodically verified to ensure that corruption or loss has not occurred.

7.1.2.2 Security

Information Services or a suitable third party shall be responsible for the security of the information assets ingested into the archive, informed by the University of Nottingham’s Information Security Policy. This will include:

- physical security - the physical infrastructure required to store and manage archival collections shall be protected from accidental or deliberate damage. This shall be achieved by way of restricted access to the physical machines and backup power supplies to those machines in the event of a failure;
- systems security - measures to ensure that external attacks from unauthorised users, malicious code or other software attacks against the IT systems deployed for digital preservation shall be enforced. Password protected permissions, firewalls and anti-virus software shall be used in order to achieve this;
- CRUD permissions – access permissions will be managed so that users and other systems have appropriate create, read, update, and delete (CRUD) permissions that comply with the legal and policy conditions placed upon each information asset. This shall be achieved by way of appropriate authentication services.

7.1.2.3 Integrity Monitoring

The integrity of each information asset ingested shall be periodically verified using industry standard methods. In the event of a failed integrity check those responsible for preserving data within LRLR shall replace the copy with an alternative redundant copy.

Information Services or a suitable third party will be responsible for providing the means to verify the integrity of an information asset.

7.1.2.4 Administrative Metadata

Those responsible for preserving data within LRLR shall will maintain administrative metadata relating to the preservation of ingested information assets such as logs illustrating any actions taken on information assets held within the archive. These will help to prove the provenance, authenticity and integrity of the contents of the archive.

7.1.3 Content Preservation

7.1.3.1 Characterisation

Those responsible for preserving data within LRLR shall will characterise the structure and technical properties of information assets submitted for ingest into the archive so that it is understood what is being preserved. File format identification will be recorded as part of this process.

7.1.3.2 Preservation Planning

Those responsible for preserving data within LRLR shall will maintain a preservation planning function that will identify and monitor technological changes and their potential impacts on the usability of preserved digital information assets. This will consist of:

- Technology watch – an assessment of technological changes that may affect the usability of information assets held by the archive;
- Risk assessment – a set of criteria used to identify the level of potential risk identified by the technology watch function;
- Impact and cost assessment – an assessment of the potential implications and costs involved in mitigating the risks identified.

The preservation planning function shall provide this capability.

7.2 Versioning and Withdrawal

Those responsible for preserving data within LRLR shall will not permit changes to ingested items but updated versions may be submitted as new versions of the ingested items. These will be linked within the archive and the persistent URL will point to the latest version.

Withdrawal of information assets from the archive will not be permitted unless there is a legal requirement to do so. Acceptable reasons for withdrawal may include, but are not limited to:

- proven copyright violation or plagiarism;
- national security
- falsified research

Deletion of digital objects shall only be permitted under strictly controlled and authorised circumstances.

7.3 Presentation and Access

Those responsible for preserving data within LRLR shall make available for access information assets held within the archive to its designated community, subject to limitations governed by:

- a suitable user license
- the Copyright, Design and Patents Act, 1988 and amendments to this Act including the Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations, 2014
- Enterprise and Regulatory Reform Act, 2013
- the Data Protection Act, 1998
- the Freedom of Information Act, 2000
- Environmental Information Regulations, 2005

Appropriate access controls shall be applied to all information assets held within the archive. Restrictions of access shall include access to content that has undergone redaction, access to content by specified groups of individuals only or time embargoed access (closed for x years for example).

Access copies of information assets may be provided in a different file format and have different technical properties to that of the original; where it is considered that there is a user need in order to facilitate speed and ease of access.

7.4 Skills

7.4.1 Internal Staff Skills and Training

LRLR will ensure that its digital preservation activities are carried out by sufficient staff with the appropriate skills in order to fulfil its objectives. LRLR will support training to develop, maintain or enhance their digital preservation expertise.

7.4.2 External Contractors

External contractors and service providers may be used to fulfil certain functions within the archive infrastructure. In such a case arrangements shall be made to ensure:

- ownership and control of the archive collection is not jeopardised;
- information assurance standards (including redaction or closure) are fully met;
- digital information assets can be transferred to a nominated party, at a future date, and in a manner that is interoperable.

8.0 Audit and Certification

Those responsible for preserving data within LRLR shall will monitor compliance with this policy by undertaking periodic audits. These audits will be used to measure the effectiveness of its implementation, identify future priorities, and inform future reviews of the Policy.

LRLR shall aim to acquire the Data Seal of Approval; basic certification within the European Framework for Audit and Certification of Digital Repositories in order to support its digital preservation vision.

9.0 Policy Review

This policy will be reviewed every two years to take account of changing circumstances. Reviews will be conducted by LRLR, in conjunction with relevant stakeholders.

10.0 Sustainability

In order to support its digital preservation vision, LRLR is committed to supporting the funding of digital preservation for information assets submitted to the archive. In the event of the archive being shut down a succession plan will be drawn up and every effort to transfer the contents to another suitable archive will be made.

11.0 Glossary

Access	"The OAIS entity that contains the services and functions which make the archival information holdings and related services visible to Consumers." (ISO 14721:2012 Space data and information transfer systems – Open archival information systems – Reference model)
Appraisal	"Appraisal is the process of distinguishing records of continuing value from those of no further value so that the latter may be eliminated" (The National Archives, 2013)
Archive	"An organization that intends to preserve information for access and use by a Designated Community." (ISO 14721:2012 Space data and information transfer systems – Open archival information systems – Reference model)
Authenticity	An authentic digital object is one that can be proven to satisfy the following characteristics: • That it is what it purports to be. • That it was created or sent by the agent purported to have created or sent it. • That it was created or sent at the time and date it is purported to have been. (ISO 15489-1:2001 Information and documentation - Records management - Part 1: General)
Born Digital	The term born-digital refers to materials that originate in a digital form.
Bitstream	A set of bits embedded within a file.
Bitstream Preservation	The aspect of preservation which seeks to ensure the continuing authenticity and integrity of existing manifestations of a digital information asset.
Content Preservation	The aspect of preservation which seeks to ensure the continued usability of digital information assets over time, in the face of technological change.
Designated Community	"An identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities." (ISO 14721:2012 Space data and information transfer systems – Open archival information systems – Reference model)
Digital Information Asset	"The contents of all databases, electronic mailboxes, word processing documents, spreadsheets, web pages, data files, configurations files and other *information systems* created [or managed] by University members in the course of their duties are *information assets* of the University or its members." (University of Nottingham, 2014)
Digital Object	"A physical component of a digital resource. This may be represented as a bitstream, a part of a bitstream, or set of bitstreams within a computer file system." (Parliamentary Archives, 2009)
Digital Preservation	"[A] series of actions and interventions throughout the lifecycle to ensure continued & reliable access to authentic

	digital objects for as long as they are deemed valuable.” (Jisc, 2006)
File Format	A predefined structure for organising a digital object that is managed by a computer file system as a single, named entity.
Ingest	"The OAIS entity that contains the services and functions that accept Submission Information Packages from Producers, prepares Archival Information Packages for storage, and ensures that Archival Information Packages and their supporting Descriptive Information become established within the OAIS." (ISO 14721:2012 Space data and information transfer systems – Open archival information systems – Reference model)
Integrity	The integrity of a digital object is based on proving that its meaningful content is complete and unaltered. (ISO 15489-1:2001 Information and documentation - Records management - Part 1: General)
Meaningful Content	The information encoded within a digital object that is the target of preservation and archiving.
Media Refreshment	"A Digital Migration where the effect is to replace a media instance with a copy that is sufficiently exact that all Archival Storage hardware and software continues to run as before." (ISO 14721:2012 Space data and information transfer systems – Open archival information systems – Reference model)
Metadata	"Data about other data" (ISO 14721:2012 Space data and information transfer systems – Open archival information systems – Reference model). Metadata can usually be classified as the following: • descriptive metadata - meaningful descriptions and labels used to enhance discovery and retrievability; • technical metadata - describing the technical properties of information assets ingested and the hardware and software needed to maintain usability; • structural metadata - identifying the relationships between information assets; • administrative metadata - used to manage the digital information assets; including rights and permissions.
Preservation Planning	An aspect of preservation that is concerned with identifying threats to the continued usability of authentic digital objects. If such threats are identified then appropriate countermeasures should be determined. It incorporates the process of technology watch.
Provenance	"The information that documents the history of the Content Information. This information tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. Examples of Provenance Information are the principal investigator who recorded the data, and the information concerning its storage, handling, and migration." (ISO 14721:2012 Space data and information transfer systems – Open archival information systems – Reference model)
Redundancy	The provision of duplicate copies of data that function if the primary data fails.
Reliability	The reliability of a digital object is determined by its ability to demonstrate that it has trusted and dependable contents.

	(ISO 15489-1:2001 Information and documentation - Records management - Part 1: General)
Technology Watch	An assessment of technological changes that may affect the usability of information assets held by the archive.
Usability	The usability of a digital object refers to its ability to be located, retrieved, presented and interpreted. (ISO 15489-1:2001 Information and documentation - Records management - Part 1: General)

12.0 References

Gollins, T. (2009). *Parsimonious preservation: preventing pointless processes!* Retrieved 2015, from http://www.nationalarchives.gov.uk/documents/information-management/parsimonious-preservation.pdf

(n.d.). *ISO 14721:2012 Space data and information transfer systems – Open archival information systems – Reference model.* International Standards Office.

(n.d.). *ISO 15489-1:2001 Information and documentation - Records management - Part 1: General [adapted].* International Standards Office.

Jisc. (2006). *Digital preservation: contrinued access to authentic digital assets.* Retrieved 2015, from http://www.webarchive.org.uk/wayback/archive/20140615231719/http://www.jisc.ac.uk/media/documents/publications/digitalpreservationbp.pdf

Parliamentary Archives. (2009). *A digital preservation policy for Parliament.* Retrieved 2015, from http://www.parliament.uk/documents/upload/DigitalPreservationPolicy1.0.pdf

The National Archives. (2013). *What is appraisal?* Retrieved 2015, from http://www.nationalarchives.gov.uk/documents/information-management/what-is-appraisal.pdf

University of Nottingham. (2014). *Information security policy 2014/2015.* Retrieved 2015, from http://workspace.nottingham.ac.uk/download/attachments/62358464/IS+Security+Policy.pdf