UNIVERSITY OF
LEADING
THE WAY
WESTMINSTER

University of Westminster Records and Archives

Digital Preservation Policy 2019

Published by: University Records and Archives

Information Category: Public

Version Record

Date	Version	Author	Description
November 2019	Digital Preservation Policy 2019	Elaine Penn	Approved by Information Governance Advisory Group on 23 January 2019.
May 2016	Digital Preservation Strategy	Rebecca Short & Elaine Penn	Approved by Information Management Group on 22 June 2016.

UNIVERSITY OF
LEADING
THE WAY
WESTMINSTER

1.0 Introduction

University Records and Archives (URA) collects, preserves, and provides access to records created by staff and students of the University and its predecessors.

Records are selected for long-term preservation in line with URA's *Collection and Acquisition Policy*. Records are selected for their evidential, historical and cultural values which make them of interest to a variety of users for multiple purposes.

Traditionally, the University has created records in analogue formats such as paper, photographs, slides, drawings, or physical film. Today the University creates a large percentage of its information in a digital format, which will continue to grow in the coming years. As a result, the University Archive is collecting, preserving, and providing access to a variety of digital records.

There are significant challenges associated with maintaining access to digital records over time, which are different to the challenges posed by analogue material. These are broadly as follows:

- Digital records require hardware and software in order to access the information within them. Both the software and the hardware that is necessary to read digital records is subject to rapid changes and developments in technology. If this is not monitored, and appropriate action taken, then digital records cannot be read.
- Digital information is easily edited and/or changed. If this is not mitigated against or managed, then it can impact the integrity and authenticity of digital records.
- Storage media such as CD's, DVD's and flash drives are unstable and degrade over time. This poses a risk to the ability to access digital records effectively.

As a result, digital records require pro-active management and continued maintenance to ensure that they are accessible for as long as they need to be.

2.0 Purpose

This policy outlines how URA intends to manage, preserve and make accessible its digital records selected for long-term preservation in a manner that retains the records' authenticity, integrity, usability and reliability.[1]

3.0 Scope

The University Archive contains both 'digitised' and 'born digital' records:

- 'Digitised' records are digital copies of analogue information where the record was created in a physical, tangible form and has subsequently been recreated, through scanning or photographic techniques, as a digital object. For example, a digital scan of a photographic print.
- 'Born digital' records are records native to the digital environment, where the record was created using software and hardware, and saved in digital format. For example, a database record, a word document or an email account.

For the purposes of this policy, the term 'digital records' refers to both these types. See Appendices 1 and 2 for examples.

3.1 Digital formats and media

[1] Characteristics of trustworthy and authoritative records adapted from ISO 15489-1:2016.

- URA will accept digital records in most media types (such as text, graphic, image, video, audio, database, website and email) and will apply standard archival appraisal criteria, codes of practice and best practice to determine suitability for preservation.[2]
- The policy does not apply to the content or subject matter of digital records. URA's *Collection and Acquisition Policy*, the University's *Records Management Policy* and Records Retention Schedules will be referred to when determining if the digital records complement the University Archive's existing holdings.
- URA will accept digital records held on physical media (such as CD Rom, external Hard Disc Drive, USB flash drives), but, depending on condition and age, cannot guarantee that their contents can be fully extracted.

3.2 Out of scope

This policy does not include research data, which is currently managed by the Research and Scholarly Communications team and/or the Research Office.

4.0 Principles

- URA will provide consistent and relevant guidance to record creators on the short term and long term preservation of the University's digital records.
- URA will work with the wider University to advise on best practice in the creation of digital records, in an effort to minimise the complexity of preservation activities required. This includes procurement processes for any technology (software and hardware) that will manage digital records likely to be transferred to the University Archive in the future.
- Record creators, both internal to the University and external, are encouraged to be mindful of the preservation of digital content at the point of its creation. This is to ensure that records, deemed sufficient in value to be preserved for the long-term, are created in a manner that will facilitate their preservation.
- URA will take all reasonable measures to ensure digital objects managed and preserved by them are, and remain, trustworthy and accessible.
    - Authenticity – URA will carry out regular audits to ensure that digital records held by them have not been subjected to unauthorised or accidental alteration, corruption or loss.
    - Reliability – All archival processes and procedures undertaken to preserve digital records will be fully documented, in line with current international standards and best practice to determine that the records have trusted and dependable contents.
    - Integrity – URA will maintain an audit trail of actions and activities that have been carried out throughout the lifecycle of a digital record in their custody to demonstrate that the meaningful content of the record is complete and unaltered.
    - Usability – URA will preserve digital records in line with best practice and provide sufficient metadata to allow the records to be located, retrieved, presented and interpreted.
- URA will provide public access to its digital collections, unless subject to restrictions imposed by legislation, contractual obligations imposed by a donor/depositor, or technological issues that limit accessibility.
- URA will follow international standards and established best practice in all its digital preservation actions and activities. See section 6 for relevant standards.
- URA will engage with the wider digital preservation community, and, where appropriate, use processes, procedures and tools already developed and in use.
- All preservation processes will be transparent and auditable.
- The University recognises that the preservation of its digital records is an active process that requires sustainable management and resources.

[2] URA's ability to preserve digital records is subject to the necessary resourcing and technical solutions being in place. See Appendix 2 for details of formats at risk.

5.0 Policy requirements[3]

5.1 Selection and appraisal
The selection of digital records to be managed and preserved by URA will be carried out in line with URA's *Collection and Acquisition Policy*. Appraisal of digital records will be carried out through adopting best practice procedures and the use of industry standard applications.

5.2 Accessioning
At the point of accession it is important that digital records are properly screened and documented to ensure that the 'chain of custody' is maintained, the records retain their authenticity and the preservation process begins with good quality data and metadata. To achieve this objective, URA will:

- Quarantine records prior to accession and conduct thorough anti-virus checks to ensure they pose no threat to the integrity of other records.
- Identify, characterise and validate file formats.
- Gather appropriate descriptive, administrative and preservation metadata.
- Conduct fixity checks to ensure the authenticity of accessioned records.
- Generate a 'preservation' and 'access' copy of the original where appropriate.

5.3 Preservation strategy
URA will adopt a suitable preservation strategy based on the risks associated with its digital information assets.

There are two risks to digital records held by the University:

- Loss of the medium (i.e. loss or corruption of the 0s and 1s that make up the actual bitstreams).
- Loss of the message (i.e. loss of the ability to correctly interpret the bitstreams as understandable information).

5.3.1 Bitstream preservation
In all cases URA will preserve the original bitstream as well as any other manifestations created as a bi-product of the preservation process. In order to adopt such strategies, URA will develop appropriate workflows for preservation planning and a technological infrastructure to manage the ingest, preservation process, storage, back up and accessibility of its digital collections.

5.3.2 Content preservation
Content preservation requires an institution to understand and document what it has got and what is required to correctly interpret the bitstreams so that content may be rendered; often referred to as characterisation.

Those responsible for preserving data within URA will characterise the structure and technical properties of digital records 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.

5.4 Preservation Planning
Preservation planning is at the core of content preservation. Its role is to monitor the technological, financial, legislative and institutional environment and mitigate the risks of change to the accessibility of digital records. URA will carry out preservation planning under the following areas:

- Risk assessment – URA will perform regular risk analysis on the digital records it holds to determine the type and level of preservation action required.
- Technology watch – URA will continually monitor the technological landscape both internally and externally to identify where changes or developments may impact upon its digital records, the type and level of impact and recommend appropriate actions.

[3] This section is concerned with the functions required to implement the policy and not the resources or technology to be used.

- Impact assessment – In response to outcomes from the risk assessment and technology watch URA will prioritise actions it needs to take and implement changes accordingly.

5.5 Access and use
The online archive catalogue will be the entry point for access to digital records held by URA. The catalogue will be open to the public.

5.5.1 Open access – where access can be granted fully to digital records the user will be able to view them online, through a browser, via the archive catalogue provided the user either has access to an internet connection or access to the university network.

5.5.2 University staff only – there may be a requirement, in some instances, to restrict access to some digital records to internal users only. In this instance only users with the appropriate level of access will be able to view those records.

5.5.3 Partially closed access – some digital records described within the catalogue may be subject to rights management restrictions and may therefore have limited access. Access to such records will be purely onsite, within the University Archive reading room.

5.5.4 Closed – where digital records have to remain closed for reasons described within the rights management section (5.6) there will be no access (either online or onsite) to both the catalogue record and the accompanying digital record.

5.6 Rights management
It is likely that certain rights and access conditions will apply to digital records held by URA. The University will adopt open metadata standards, such as PREMIS, METS, and Dublin Core, to express the rights status of a record or collection within the catalogue record. This may result in restrictions to the accessibility of some records. Such restrictions typically relate to:

- The presence of personal data which restricts access under Data Protection legislation.
- Where the records are subject to Copyright legislation.
- Contractual obligations made by the donor/depositor of the digital records at the point of acquisition.
- Where an access copy cannot be made due to current technological limitations.

5.7 Storage, duplication and backup
Archival storage shall be delivered by the University's Information Systems and Support (ISS) 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 content has been copied without corruption or loss. Should corruption or loss occur, then these copies will be replaced using redundant copies.
- Redundancy – ISS 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.

Digital records acquired by the University which are stored on removable or other physical media, will be transferred from their physical carrier onto secure, server-based storage by URA staff using industry standard applications at the point of accession. The original physical container will only be retained by URA if it holds archival value in itself.

5.8 Security
ISS or a suitable third party shall be responsible for the security of the digital records ingested into the archival storage. 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 to that users and other systems have appropriate create, read, update and delete (CRUD) permissions that comply with the legal and policy conditions placed upon the digital records. This shall be achieved by way of appropriate authentication services.

6.0 Legal constraints and professional standards

URA will ensure compliance with all relevant legislation and will adopt key professional industry standards in its approach to Digital Preservation. Standards enable URA to define its Digital Preservation requirements, processes and workflows and to thereafter benchmark its success against established best practice. The most relevant industry standards applicable are (but not limited to):

- Space data and information transfer systems – Open Archival Information System (OAIS) reference model (ISO 14721:2012)
- Space data and information transfer systems – Producer-Archive Interface Methodology Abstract standard (ISO 20652: 2006)
- Information and documentation – Records management – Part 1: Concepts and principles (ISO 15489-1:2016)
- Space data and information transfer systems – Audit and certification of trustworthy digital repositories (ISO 16363:2012)
- International Standard for Archival Description (General) (ISAD(G))
- Preservation Metadata Implementation Strategies (PREMIS)
- Metadata Encoding and Transmission Standard (METS)
- Dublin Core Metadata Initiative

7.0 Roles and responsibilities

- The implementation and management of Digital Preservation activities will require expertise from within the University as well as potentially from external sources. The University will endeavour to ensure that sufficient resources are available to enable URA to carry out is Digital Preservation mandate to the highest industry standard.
- URA will ensure its Digital Preservation activities are carried out by trained staff and will provide training opportunities for staff to develop and enhance their Digital Preservation skills.
- URA will actively raise awareness of Digital Preservation issues and approaches across the University and will provide training, where appropriate.

8.0 Audit and certification

URA 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.

The University will pursue appropriate accreditation and certification relevant to its Digital Preservation activities in line with other university collections-based accreditation, worked on or achieved.

9.0 Policy review

This policy will be reviewed on a periodic basis as circumstances within the University and URA changes. This review period will be at least every two years, depending on the rate of technological

changes and how this impacts on the policy, and will be conducted in conjunction with senior management and other stakeholders.

10.0 Glossary

Accessioning
The process of taking custodianship of a digital record or collection of records for the purposes of long-term preservation and access.

Appraisal
The process of distinguishing records of continuing value from those of no further value so that the latter may be eliminated.

Bitstream
A set of bits embedded within a digital file.

Bitstream preservation
A preservation strategy that involves management of the original manifestation of a digital record. It ensures that the original retains its authenticity and is maintained in a secure environment with appropriate security and backup.

Chain of custody
A system of controls that extends over the lifecycle of the digital record to ensure trustworthiness of its provenance.

Content preservation
A preservation strategy that ensures the continued accessibility of digital objects over their lifetime to mitigate the effect of technological obsolescence. It involves active intervention, and format migration, to ensure accessibility and readability of digital records.

Digital Information Asset
The contents of all databases, electronic mailboxes, word processing documents, spreadsheets, web pages, data files, configuration files and other information systems created or managed by University staff in the course of their duties are information assets of the University.

Digital object
An individual digital component that either singly, or collectively with other digital objects, forms a digital record.

Digital record
Information in an electronic format that demonstrates evidence of an action or activity.

File characterisation
The process whereby information about the digital record, such as format and version, is identified and extracted in the form of metadata.

File Validation
The process whereby digital records can be checked to establish if their format conforms to standard specifications.

Ingest
The process of moving digital records from the record creator and into a Digital Preservation system.

Manifestation
A digital derivative or copy of an original bitstream object.

Metadata
The literal definition is ‘data about data’, and is classified as either descriptive, administrative or structural and which in some way will enable the continued management, preservation and access to digital records.

Redundancy
The provision of duplicate copies of data that function if the primary data fails.

Technology watch
The process whereby the technological landscape is monitored to assess the likely impact any changes may have on the preservation and accessibility of digital records.

11.0 References

Digital Preservation Coalition (2015). *Digital Preservation Handbook*, 2nd edition. Retrieved 2019, from
http://handbook.dpconline.org/

Parliamentary Archives (2008). *A digital preservation policy for Parliament.* Retrieved 2019, from http://www.parliament.uk/documents/upload/digital-preservation-strategy-final-public-version.pdf

The National Archives (2015). *What is appraisal?* Retrieved 2019, from https://www.nationalarchives.gov.uk/documents/information-management/what-is-appraisal.pdf

University of Westminster (2019). *University Records and Archives Collection and Acquisition Policy*. Retrieved 2019, from http://recordsandarchives.westminster.ac.uk/home/collection-policy/

University of Westminster (2019). *University Records Management Policy*. Retrieved 2019, from https://www.westminster.ac.uk/sites/default/public-files/general-documents/uow-records-management-policy.pdf

Appendix 1 - Inventory of current digital assets at December 2019

Please note - these are estimates only and are based on the level of detail available in catalogue records and other sources. Assets are currently held in various departments and in different storage media across the University.

Asset	No. of items	Volume (GB)	Further accruals likely?	Format	Location	Current vulnerability	Risk Type
Digitised Polytechnic Magazine	5907	10.4	No	pdf	ISS servers; DVD copies	4	Financial
Oral History recordings	89	4.8	Yes	mp3; wav	DP storage; Shared Drive (Archive); DVD copies	Mitigated	Reputational
BA Fashion Course digital records	29796	111	Yes	jpg; tiff	Shared Drive (Archive)	1, 4	Reputational
BA Film Course digital records	unknown	unknown	Yes	unknown	DCDI, Harrow	1, 3, 4	Operational, Reputational
Ambika P3 and LGW digital records	6394	46.9	Yes	jpg; tiff; RAW	Shared Drive (Archive)	1, 4	Reputational
Senior members of staff digital records	unknown	14.5	Yes	pst; pdf; Microsoft office	DVD copies, Shared Drive (Archive); DP Storage	1	Operational, Financial
Exam papers	1818	0.45	Yes	pdf	DP Storage, Shared Drive (Archive); DVD and CD	1	Operational
Presentation Ceremony DVD's	104	c. 312	Yes	unknown	DVD	3, 4	Reputational
UOW History Project Books	31	0.95	Yes	pdf	Shared Drive (Archive)	4	Reputational
VC videos – addresses to staff and students	40	123	Yes	mp4; mov; wav	Shared Drive (Archive)	1, 4	Reputational
Ceramics course digital records	49	unknown	No	website; jpg	1 box in Archive; archived website in DP storage	1, 3, 4	Reputational
Archigram - http://archigram.westminster.ac.uk/	unknown	unknown	No	website; image files	Managed by ISS; archived website in DP storage	5, 2	Reputational; Financial

Asset	No. of items	Volume (GB)	Further accruals likely?	Format	Location	Current vulnerability	Risk Type
Student Records (SITS)	unknown	900	Yes (av. 40GB per month)	database (inc pdf)	ISS Servers	5	Operational, Financial, Reputational
Arts on Film - http://artsonfilm.wmin.ac.uk/	unknown	unknown	No	website; image files	Managed by ISS; archived website in DP storage	5, 3	Reputational; Financial
Staff Records (SAP)	unknown	260	Yes	database (inc pdf)	ISS Servers	5	Operational, Financial, Reputational
Financial Records (Agresso)	unknown	300	Yes	database (inc pdf)	ISS Servers	5	Operational, Financial, Reputational
UoW senior committee records	unknown	0.019	Yes	pdf; Microsoft Office	DP storage; Shared Drive (Archive)	Mitigated	Operational, Financial, Reputational
Publications from Design team	14500	c.1TB	Yes	Pdf; Adobe InDesign	NextCloud (tempstorage)	1, 4	Operational, Financial, Reputational
General - UoW Business Records (various university departments)	unknown	unknown	Yes	msg; pdf; Microsoft office	Shared Drive (U and Archive); DVD copies	1, 4	Operational, Financial, Reputational
General - Digital images on shared drive	unknown	169	Yes	jpg; tiff	Shared Drive (U)	1, 4	Reputational, Financial, Operational
General - Digital images on removable media	568	unknown	Yes	unknown	DVD; CD	1, 3, 4	Reputational, Financial, Operational
Menswear digital images	unknown	unknown	Yes	Jpeg; tiff; png	USB; OneDrive; Shared Drive (U); Dropbox	1, 3, 4	Reputational, Financial, Operational
General - Moving image	17	unknown	Yes	unknown	DVD; CD	1, 3, 4	Reputational
General - Audio	5	unknown	Yes	unknown	DVD; CD	1, 3, 4	Reputational

Appendix 2 - Inventory of assets requiring digitisation activities in order to gain or maintain access at December 2019
Please note - these are estimates only and are based on the level of detail available in catalogue records and other sources. Assets are currently held in the University Archive, with additional analogue moving image and audio records held in DCDI.

Asset	No of items	Format	Vulnerability Type	Risk
BA Film Course analogue records	872	U-Matic; Mini-DV; VHS	2, 1	Reputational
Known series of VHS recordings (Uncatalogued)	63	VHS	2, 1	Reputational
UoW Analogue Oral History recordings	2	Audio Cassette	2, 1	Reputational
Presentation Ceremony Videos	73	VHS	2, 1	Reputational
GENERAL - Analogue Moving Image	62	VHS; Reels	2, 1	Reputational
GENERAL - Analogue Audio	27	Audio Cassette	2, 1	Reputational
GENERAL - Analogue UoW Business records[1]	c.300	Floppy disc	2, 1	Reputational, Operational
Ceramics course – analogue	3	VHS	2, 1	Reputational

Vulnerability types[2]

1 Content for which there is no current provision for user access

2 Content that is currently inaccessible due to technology obsolescence

3 Content that is currently at risk of corruption or loss, due to storage on vulnerable removable media

4 Content that does not have suitable storage or back up processes in place

5 Content within business systems for which there is not a process in place for longer term management ('archiving' process)

[1] To include a large amount of records from UOW/2/8/COP on floppy disc.
[2] These criteria are based on those found in Adrian Brown, *Practical Digital Preservation: a how to guide for organizations of any size* (2013).