



University of
St Andrews

Digital Preservation Policy

Document type	Policy
Scope (applies to)	All relevant staff – see Scope
Applicability date	15 March 2021
Expiry date	
Approved date	
Approver	VP (Collections, Music and Digital Content)
Document owner	Sean Rippington
School / unit	University Libraries and Museums
Document status	
Information classification	Public
Equality impact assessment	None
Keywords	
Purpose	The purpose of this policy is to provide a statement of the University's ongoing commitment and approach to preserving its digital content selected for permanent retention, so that materials remain accessible and usable to relevant wide audiences in perpetuity.

Version number	Purpose / changes	Document status	Author of changes, role and school / unit	Date
1.0	Final draft – approved by ICG. For approval by ISB.	For approval	Anna Clements (Head of Digital Research, RIS) and Sean Rippington (Digital Archives and Copyright Manager, Library) incorporating feedback from Ester Ruskuc (DoSP) and Chris Milne (Head of Information Assurance and Governance) following discussion at ICG	21/11/2019
1.1	Response to feedback from Dec 2019 ISB. Strengthened sections on strategic alignment, governance, risks, roles and technology.	For feedback from CIO, VP Collections, Music and Digital Content and Director of Libraries and Museums	Daniel Farrell (Head of Data Transformation, ITS) and Sean Rippington (Digital Archives and Copyright Manager, Library)	1/10/2020

1.2	Response to feedback from VP Collections, Music and Digital Content. Changes made at 1.2, 5.3, 8.5.2 and 11	For feedback and/or approval by Director of Libraries and Museums and CIO	Daniel Farrell (Head of Data Transformation, ITS) and Sean Rippington (Digital Archives and Copyright Manager, Libraries and Museums)	16/10/2020
1.3	Response to feedback from Director of Libraries and Museums, Senior Manager (Metadata, Acquisitions & Scholarly Communications), and Head of Open Research	For feedback and/or approval by Director of Libraries and Museums and CIO	Daniel Farrell (Head of Data Transformation, ITS) and Sean Rippington (Digital Archives and Copyright Manager, Libraries and Museums)	02/11/2020
1.4	Response to final feedback from Head of Information Assurance and Governance- see yellow highlights	For submission to the Governance Zone	Daniel Farrell (Head of Data Transformation, ITS)	04/03/2021

1. Introduction

- 1.1. This policy is structured around principles, activities and processes including roles and responsibilities. This policy aims to ensure pro-active management of digital content through identifying and applying appropriate records management controls at the point of record creation and long-term preservation.
- 1.2. Long-term is defined as 10+ years or in perpetuity.
- 1.3. Long-term preservation of digital content¹ is essential to the University's mission to deliver research and teaching of the highest quality. It underpins the University's commitment to providing robust digital platforms supporting its 'digital first' approach.
- 1.4. The University has kept an excellent record of its more than 600-year history through its physical collections. The University's current and future research, teaching, learning and administrative activities rely on preserving its digital content to the same excellent standard and with the same level of commitment.
- 1.5. The University generates and acquires digital content during its normal business activities. This content is often unique to the University, critical for business continuity and corporate memory, and has high value.
- 1.6. The University operates within a complex digital environment in which digital content, much of it valuable, is not actively preserved. There is an increasing risk that this digital content will become inaccessible, unusable, or unreliable over time, due to rapid changes in technology and infrastructure.

¹ For the purposes of this policy 'digital content' (also known as digital assets, digital materials, digital objects or digital resources) includes MS Office documents, data in databases, digital photographs, website content, digital video, digital museum objects, digital artwork, and digital audio. It encompasses content that has only ever existed digitally ('born-digital content') and physical items that have been converted to digital files ('digitised content').

- 1.7. The objective of this policy is to guide the University's approach to the treatment of institutional risks associated with disruption and/or loss of access to digital content.
- 1.8. The University recognises digital preservation as a necessary capability supporting many aspects of its core organisational functions (research, learning and teaching) and consequently recognises digital preservation is essential to the ongoing reputation of the University and its ability to deliver critical business activities.
- 1.9. Digital preservation can support the integrity and longevity of digital content and safeguards it against loss and corruption. A successful digital preservation capability means that digital content can be reused in a meaningful way by future generations of staff, students, external researchers, and other users.

2. Context

- 2.1. This policy sits within a wider governance framework supporting digital operations and is dependent on the existence or development of associated governance, resourcing, infrastructure and processes.
- 2.2. The selection of digital content for long-term preservation is described in relevant policy and guidance, including:
 - University Records Retention Schedules
 - University Records Management Policy
 - Research Data Management Policy
 - University of St Andrews Library Collections Policy
 - Collections Policy [Museums]
- 2.3. The assignment of value to digital content may change over time. These changes will be recorded in the relevant policy and guidance documents.
- 2.4. Appraisal will be delivered through adoption of best practice procedures and the use of industry standard applications.

3. Purpose

- 3.1. The purpose of this policy is to provide a statement on the University's ongoing commitment and approach to preserving digital content selected for permanent retention, so that materials remain accessible and usable in perpetuity.
- 3.2. This policy outlines how the University intends to manage, preserve and make accessible its digital content selected for long-term preservation in line with its principles.
- 3.3. This policy underpins the University's core organisational functions of research, teaching and learning by directly addressing or supporting:
 - Mitigation of operational risks to digital content, including irretrievable loss of core university digital records required to support business continuity
 - Compliance with national legislation and funder requirements
 - Service delivery improvements, including improved ability to manage digital content
 - Increased staff efficiency due to improved access to, and retrieval of, archived digital content
 - Long-term availability of digital research outputs, e.g. data, e-theses
 - Investment protection in the move to paper-lite or paper-free processes

- Stimulated research and collaboration through improved access to the University's world-leading collections
- Reputational sustainability for the long-term dissemination of knowledge

4. Audience

4.1. The target audience for this policy includes:

- University staff who create, collect, and manage digital content
- University staff tasked with the procurement and implementation of management information systems
- Funding bodies, researchers, and internal and external users of digital content
- Donors and potential donors of digital content to the University
- The wider records management and digital preservation communities

5. Scope

5.1. This policy applies to unique digital content owned by the University as well as digital content for which the University plays a stewardship role (e.g. digital content donated to the University or added to University repositories under licence).

5.2. All digital content, except for the exclusions listed below, is covered by this policy. This includes administrative content to meet the University's obligations, content for historical collection/archival purposes as well as research data.

5.3. Specifically, all born-digital content selected by the relevant University authority for long-term or permanent retention is within scope of this policy including, but not limited to:

- University archives, including all records created by the University during its business, retained after application of the University Retention Schedules;
- Research data deposited in the institutional research data repository (Pure);
- Scholarly communications, including theses, deposited in the St Andrews research repository (DSpace);
- Heritage content, including museum objects, music, and artworks; and
- The core student and core staff record.

5.4. Digital content may take forms including but not limited to: MS Office files, pdfs, xml database exports, image files, audio-visual files, zipped files, proprietary formats, LaTeX files, and web and social media content. In addition, all digitised content, and content that will be digitised, selected by the relevant University authority for permanent retention is within scope of this policy including, but not limited to:

- Digitised University archives (including all digitised physical records created by the University during its business, retained after application of the University Retention Schedules);
- Digitised Museums collections;
- Digitised Special Collections materials (including manuscripts and photographs); and
- Digitised audio-visual materials.

5.5 Digital content not within scope of this policy include:

- Born-digital administrative content generated in the University's business but not intended for long-term preservation (e.g. informal chats and emails, MS Office

documents created by day-to-day administrative tasks) which are governed by other institutional retention policies;

- Digitised administrative content generated in the University's business but not intended for long-term preservation (e.g. digitised reference material, digitised documents created by day-to-day administrative tasks);
- Commercial content not managed by the University, but which the University has access to; and
- Content hosted by the University on behalf of external organisations.

6. Roles and responsibilities

6.1. Planning for preservation activities including the definition of current capacity and the forecasting of future needs is the responsibility of the Digital Preservation Steering Group.

6.2. The Digital Preservation Steering Group is chaired by the Vice Principal (Collections, Music and Digital Content) and comprises colleagues from across the Libraries and Museums, other Professional Services, and academic staff:

- 2 x senior academic staff tbc;
- Head of Information Assurance and Governance;
- Head of Open Research;
- Head of Data Transformation;
- Senior Archivist (Keeper of Manuscripts and Muniments);
- Digital Archives Manager;
- Senior Manager (Metadata, Acquisitions & Scholarly Communications); and
- Museums representative tbc.

6.3. The membership of the Digital Preservation Steering Group represents:

- custodians of digital content selected for permanent preservation
- staff responsible for processes and infrastructures related to digital preservation
- representatives of the academic community

6.4. Membership of the Digital Preservation Steering Group can be changed as appropriate to meet the requirements of 6.1 above.

6.5. Establishing how digital preservation will be achieved is the responsibility of the Information Strategy Board (ISB).

6.6. The University will implement different preservation strategies concurrently and over time as required. The most appropriate strategies will be selected based on the individual requirements of digital content, ongoing risk assessment, evolving best practice, and reviews of the resources required to execute the strategy in accordance with this policy.

6.7. The management and implementation of digital preservation activities requires expertise, and the University is committed to ensuring enough resources are available to enable the delivery of this policy to the highest industry standard.

7. Principles

7.1. Digital content will be managed in accordance with commonly adopted, openly documented, standards and practice within the digital preservation community.

- 7.2. The University will ensure compliance with all relevant legislation (see appendix) and will adopt key professional industry standards in its approach to digital preservation in order to better define its digital preservation requirements, processes and workflows and to benchmark its success against established best practice.
- 7.3. Any digital content selected for permanent preservation will be managed according to three overarching principles:
- Digital content will be integral, i.e. unaltered, or change-controlled through documentation;
 - Digital content will be authentic, i.e. trustworthy, being what it purports to be; and
 - Digital content will remain usable, i.e. available as appropriate, readable, meaningful, and understandable.
- 7.4. Digital content may exist in multiple versions and the University is committed to preserving the authentic version of record, i.e. the fullest highest-quality available authoritative version of the content, whenever possible. In addition, the University will preserve the hierarchical and structural relationships associated with the digital content and all descriptive, structural and administrative metadata that make the content usable by future technologies.
- 7.5. Staff, students, external researchers and other users will be provided with appropriate access to digital content now and in the future.
- 7.6. The University will foster partnerships between University staff and organisations around the world (such as the Digital Preservation Coalition, Ligue des Bibliothèques Européennes de Recherche, Research Libraries UK, Jisc) to benefit mutually from a sustainable shared community of knowledge, helping the University future-proof its processes by keeping up to date with best practice and evidence-based solutions.
- 7.7. All digital preservation processes will be transparent and auditable.

8. Processes and workflows

8.1. The above principles will be underpinned by the development of sustainable processes and workflows along with appropriate technical infrastructure and staff development.

8.2. Sustainable processes

8.2.1. The University will ensure that adequate resourcing is provided to support the long-term sustainability of digital preservation activities, including documenting approaches and solutions and ensuring that this documentation is kept up to date.

8.2.2. Sustainable digital preservation capabilities will be developed through collaboration between the University Libraries and Museums, the Information Assurance and Governance function, ITS colleagues and other stakeholders involved in the creation of digital content at the University.

8.3. Technical infrastructure

8.3.1. The University will procure, develop, maintain and review the technical infrastructure required to carry out preservation of its digital content, including appropriate storage, to ensure future accessibility, usability and integrity of archived records and other digital content.

- 8.3.2. The University recognises that digital content requires management beyond the lifespan of the technical infrastructure and systems currently used to manage them. Systems used for managing digital content, including storage, will be chosen and/or developed using appropriate and tested exit strategies for digital content and metadata.
- 8.3.3. The University will use technical infrastructure which supports reporting and monitoring of systems, as is required to undertake preservation activities and proactive planning.
- 8.3.4. The University will continue to document, monitor and refresh its technical workflows and processes over time.

8.4. Workflows for creation, management and access

- 8.4.1. Digital content will be created and acquired in sustainable digital formats where possible.
- 8.4.2. Physical material (e.g. paper records) selected for digitisation and permanent retention will be digitised and managed according to the principles in section 7 above.
- 8.4.3. Access conditions, copyright and other rights will be adhered to when managing, preserving and providing access to digital content. Appropriate practices are established to manage personal and sensitive data, in compliance with data protection laws.

8.5. Preservation activities

- 8.5.1. The University will actively monitor its digital content for corruption and unauthorised changes.
- 8.5.2. The University may undertake preservation activities, such as migration of digital content from one format to another format, to ensure that digital content can render in current digital environments. For example, it may be necessary to migrate files held in a proprietary format to a more sustainable, widely used format to avoid vendor lock in or obsolescence.
- 8.5.3. Any preservation activities performed on digital content will be tested, evidenced-based, and recorded.

8.6. Metadata

- 8.6.1. The University will collect and generate metadata with creators for managing and describing its digital content in accordance with relevant content policies.
- 8.6.2. The University will permanently retain and actively manage persistent identifiers used to identify and locate digital objects (e.g. a DOI) that are to be preserved in the long-term.
- 8.6.3. All actions undertaken as part of the preservation process (including any new metadata created as part of the process, e.g. preservation activity, custody/ownership) will be fully documented in the preservation metadata associated with each record, to provide an audit trail.

8.7. Staff development

- 8.7.1. The University will ensure its digital preservation activities are fit for purpose and carried out by trained staff. The University will provide training opportunities for staff to develop and enhance their digital preservation skills.
- 8.7.2. In addition, the University will actively raise awareness of digital preservation issues and highlight inconsistencies in approaches across the University and will provide training, where appropriate.

8.8. Communication of progress

8.8.1. Digital preservation activities at the University will be communicated to staff via a dedicated web page and relevant internal communication channels.

8.8.2. Digital preservation activities at the University will be communicated to the wider user community via a dedicated web page, social media, and contributions to relevant professional communication channels.

9. **Guidance and control**

9.1. The University will provide training and guidance in records and data management, including metadata management, to all staff working with digital content.

9.2. The University will provide training and guidance to help digital content creators transfer content to any digital preservation system.

9.3. Creators of digital content will be supported with policies and guidance documents relevant to the creation, management, and retention of specific content types. For example:

<i>Creators of ...</i>	<i>Will be given training/guidance in accordance with ...</i>
• administrative records	the records management policy
• research data	the research data management policy
• web content	digital communications standards
• Open Access scholarly communications	Open Access and thesis policies

10. **Communication and review**

10.1. This policy will be published online via the University Governance Zone and will be communicated to stakeholders publicly via the University website www.st-andrews.ac.uk.

10.2. The University's digital preservation requirements will change over time. Regular review will ensure ongoing dialogue with users in the University and external communities. This policy will be reviewed at least annually in order to keep pace with those conversations and the maturity of experience. If there are periods of rapid change this policy will be modified as needed to reflect current priorities, infrastructure, research or investment.

10.3. This policy applies from the date of publication.

11. **Published related documents and regulatory frameworks**

11.1. University of St Andrews documents

Strategies

- [University Strategy](#)
- [Library Strategy](#)
- [Digital Strategy](#)

Policies

- [Records Management Policy](#)
- [Information Classification Policy](#)

- [Research Data Management Policy](#)
- [Library Collections Policy](#)
- [Collections Policy \[Museums\]](#)
- [Open Access Policy](#)
- [Special Collections Digitisation Policy](#)
- [Data Protection Policy](#)
- Core Student Record Policy
- [Assessment of Postgraduate Research Students Policy](#)

Guidance

- University Records Retention Schedules
- [Library Digital Recovery Plan](#)

11.2. National and international standards related to digital preservation

- [DublinCore Metadata Initiative \(DCMI\) Specifications](#)
- [METS Metadata Encoding and Transmission Standard](#)
- [PREMIS Data Dictionary for Preservation Metadata, Version 3.0](#)
- [CoreTrustSeal Data Repository Certification](#)
- [ISO 14721:2012 Space data and information transfer systems – Open archival information system \(OAIS\) – Reference model](#)

11.3 Related legislation

- [The Data Protection Act 2018 and the UK-General Data Protection Regulation](#)
- [Copyright, Design and Patent Act \(1998\), and amendments to this Act](#)
- [Copyright and Related Rights Regulations 2003](#)
- [Freedom of Information \(Scotland\) Act 2002](#)
- [Environmental Information Regulations 2004](#)
- [Limitation Act 1980](#)
- [The Control of Substances Hazardous to Health Regulations 2002](#)
- [The Control of Asbestos Regulations 2012](#)
- [The Ionising Radiations Regulations 2017](#)

12. Bibliography

- Digital Preservation Coalition, (2015). Digital Preservation Handbook, 2nd Edition. URL: <http://handbook.dpconline.org/>
- Halvarsson, E. and Langley S., (2019). Digital Preservation Policy Toolkit. Dataset. Oxford University Research Archive. DOI: <https://doi.org/10.5287/bodleian:qq9Z7wddQ>
- University of Glasgow, (2017). Digital Preservation Policy. URL: https://www.gla.ac.uk/media/media_598622_en.pdf

13. Glossary

Authenticity – The digital material is what it purports to be. In the case of electronic records, it refers to the trustworthiness of the electronic record as a record. In the case of "born digital" and digitised materials, it refers to the fact that whatever is being cited is the same as it was when it was first created unless the accompanying metadata indicates any changes. Confidence in the authenticity of digital materials over time is particularly crucial owing to the ease with which alterations can be made.

Born digital – Digital materials which are not intended to have an analogue equivalent, either as the originating source or as a result of conversion to analogue form. This term has been used in this policy to differentiate them from 1) digital materials which have been created as a result of converting analogue originals; and 2) digital materials, which may have originated from a digital source but have been printed to paper, e.g. some electronic records.

Digital preservation – Refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. Digital preservation is defined very broadly for the purposes of this policy and refers to all of the actions required to maintain access to digital materials beyond the limits of media failure or technological and organisational change. Those materials may be records created during the day-to-day business of an organisation; "born-digital" materials created for a specific purpose (e.g. teaching resources); or the products of digitisation projects.

Digital content – For the purposes of this policy 'Digital content' (also known as digital assets, digital materials, digital objects or digital resources) includes MS Office documents, data in databases, digital photographs, website content, digital video, and digital audio. It encompasses content that has only ever existed digitally ('born-digital content') and physical items that have been converted to digital files ('digitised content').

File format – A file format is a standard way that information is encoded for storage in a computer file. It tells the computer how to display, print, process, and save the information. It is dictated by the application program which created the file, and the Operating System under which it was created and stored. Some file formats are designed for very particular types of data, others can act as a container for different types. A file format is often indicated by a filename extension containing three or four letters that identify the format.

Integrity – Data that has remained unchanged. For example, data that has undergone a process (such as transmission or storage and retrieval) and is identical after the process to how it was before the process began.

Metadata – Information that describes significant aspects of a 'digital object'. This encompasses many types of information including descriptive, administrative, technical, structural, rights and preservation metadata. Most discussion to date has tended to emphasise metadata for the purposes of resource discovery. For digital preservation purposes, the emphasis is on what metadata are required to successfully manage and preserve digital materials over time and assists in ensuring essential contextual, provenancial, historical, and technical information are preserved along with the digital content. The PREMIS Data Dictionary for Preservation Metadata has become a key de facto standard in digital preservation.

Muniments – Records created by the University during its business, held in the University archive. This would include senate minutes, matriculation and graduation records, titles for University properties, administrative files, records of students and staff and financial records. All other archival material will be found within the manuscripts or photographic collections.

Persistent Identifier (PID) – A long-lasting reference to a digital object. Typically, it has two components: a unique identifier; and a service that locates the resource over time even when its location changes. The first helps to ensure the authenticity of a digital resource (that it is what it purports to be), while the second will ensure that the identifier resolves to the correct current location. A DOI (Digital Object Identifier) is commonly used as a PID.

PREMIS – Preservation Metadata: Implementation Strategies. A de facto standard for digital preservation metadata. See <http://www.loc.gov/standards/premis/>

Version control

Version number	Purpose / changes	Document status	Author of changes, role and school / unit	Date
0.1	First draft	Draft	Sean Rippington, Digital Archives Officer, University Library	11/04/2019
0.2	Second draft	Draft	Sean Rippington, Digital Archives Officer, University Library [incorporating feedback from Gabriel Sewell (Assistant Director, University Library), Rachel Hart (Deputy Head of Special Collections), Anna Clements (Head of Digital Research, RIS), Eva Borger (Data Scientist, RIS)]	14/05/2019
0.3	Third draft	Draft	Sean Rippington, Digital Archives Officer, University Library [incorporating feedback from Gabriel Sewell (Assistant Director, University Library), Rachel Hart (Deputy Head of Special Collections), Anna Clements (Head of Digital Research, RIS), Eva Borger (Data Scientist, RIS)]	17/05/2019
0.4	Fourth draft	Draft	Sean Rippington, Digital Archives Officer, University Library [incorporating feedback from Daniel Farrell (Head of Data Transformation, ITS)]	01/10/2020
0.5	Fifth draft	Draft	Sean Rippington, Digital Archives Officer, University Library [incorporating feedback from VP Collections, Music and Digital Content and Daniel Farrell (Head of Data Transformation, ITS)]	16/10/2020
0.6	Sixth draft	Draft	Sean Rippington, Digital Archives Officer, University Library [incorporating feedback from Director of Libraries and Museums, Senior Manager (Metadata, Acquisitions & Scholarly Communications), and Head of Open Research]	30/10/2020