



UNIVERSITY OF CAMBRIDGE

Cambridge University Libraries

Cambridge University Libraries Digital Preservation Policy

Filename: POL_CUL_digitalpreservationpolicy_v1_20181119.pdf

Location: <https://doi.org/10.17863/CAM.32927>

Version: 1.0

Last Amendment: N/A

Policy Owner/Sponsor: Deputy Director, Digital Initiatives

Policy Contact: Director, Digital Transformation Programme

Prepared By: Digital Preservation Specialist – Policy and Planning (Polonsky Fellow)

Approved By: Library Syndicate

Date Approved: 30 October 2018

Next Review: 31 October 2019

Availability: Public

1. Summary

The Digital Preservation Policy (the Policy) sets out the duties and responsibilities for Cambridge University Libraries' (CUL) stewardship of digital content. This Policy applies to the main University Library as well as affiliate and dependent libraries. The Policy provides the overarching framework for supporting the management of digital content (encompassing born-digital and digitised content). CUL is committed to collecting and caring for digital content to support the goals and vision of CUL, and the mission and core values of the University of Cambridge (the University). This contributes to the University's ability to maintain its global reputation as a world-leading university while meeting current and future education, learning and research needs, as well as those of the international research community and wider society. Digital content that is not actively preserved is a significant risk for CUL. This Policy guides CUL's approach to the mitigation of risks as the custodian of digital content and sits within CUL's wider governance framework supporting digital operations, that is currently being established.

While CUL's digital preservation capabilities are at an early stage, creating a sustainable digital environment is vital to supporting the endeavours of CUL and its staff, students, researchers and users. It is therefore expected that this Policy – along with the development of associated governance, resourcing, infrastructure and processes – will be developed iteratively, and will be reviewed and revised on a frequent basis.

2. Target audience

This Policy applies to all CUL and University staff (including, but not limited to, permanent, fixed-term, part-time, casual etc.), contractors, consultants, volunteers, service providers and vendors that acquire, create, manage, preserve, store, make available and/or use any digital content from CUL's collections, and other digital content CUL has responsibility for.

The Policy is also relevant to funding bodies, philanthropic organisations, donors, researchers and all users engaged with CUL's digital content.

3. Policy statement

The Policy establishes CUL's commitment to the management and preservation of digital content it acquires or creates for which it has contractual or legal obligations, determines by policy or chooses to include in its collections, in order to safeguard it for the future. Long-term preservation of digital content is essential to the University's mission of contributing to society through the pursuit of education, learning, and research.

CUL recognises that the management and preservation of digital content is of equal importance to the management and preservation of original physical collection items. The ability to acquire, create and provide sustained long-term access to digital content now and in the future, is underpinned by continuous digital preservation efforts.

The purpose of this Policy is to outline and define:

- CUL’s approach to digital preservation practices;
- The principles used to guide the management and preservation of digital content in CUL’s custody.

The broad scope of CUL’s digital content and the wide-ranging digital preservation efforts required now and in the future, demand that CUL addresses organisational infrastructure (including business, cultural and organisational change) and resources framework concerns (including appropriate levels of staffing, essential skills plus funding), as well as technical infrastructure improvements, as necessary changes to support the acquisition, creation, management and preservation of digital content. CUL’s digital preservation practices are gradually developing to be in line with international good practice, principles and standards, where available. Where appropriate, CUL leads in the development of new and/or improved standards where known gaps currently exist or current standards are not fit-for-purpose.

This Policy aligns with key CUL policy frameworks and strategies including:

- *Cambridge University Libraries Strategy 2019 – 2023 (in draft)*
- *Digital Preservation Strategy 2019 – 2023 (in draft)*
- *Collection Development Policy Framework*

3.1. Scope

This Policy applies to all preservation master, co-master and ‘born-digital original’ digital content, associated materials and associated metadata, in CUL’s custody.

3.1.1. In scope

3.1.1.1. Digital content

CUL’s digital content falls broadly into six classes, all of which are considered in scope of this Policy. Some digital content may fit in or span across several classes; if this is the case, collaborative and cross-departmental approaches to acquiring, creating, managing and preserving digital content are taken.

Class	Type	Description
1	Born-digital personal and corporate records	Digital archives of significant individuals or institutions
2	Born-digital university records	Selected records of the University (including CUL)
3	Research outputs	Research data, research publications, electronic and digitised theses, scholarly digital editions, supplementary research relating to digitised content and associated materials
4	Published born-digital content	Web archives, eBooks, born-digital maps, born-digital music, ephemera, published born-digital content on carriers and copies of electronic subscription materials (archival and/or access copies, as permitted by agreements) etc.
5	Digitised content	Digitised image content: Two-dimensional (2D) photography and three-dimensional (3D) imaging etc. Digitised audiovisual content: Moving image (film and video) and sound recordings etc.
6	In-house created content	Photography and videography of events and lectures, photos of conservation treatments etc.

3.1.1.2. Non-Print Legal Deposit content managed by Cambridge University Libraries

CUL is one of six Legal Deposit Libraries (LDLs) in the United Kingdom (UK) and Ireland. CUL has a legal mandate to acquire, manage, preserve and provide access to digital content that is governed by the *Legal Deposit Libraries (Non-Print Works) Regulations 2013*. Through a Memorandum of

Understanding (MOU) and/or other arrangements, the British Library (BL) acquires, manages, preserves, stores and/or provides access to Non-Print Legal Deposit (NPLD) content on behalf of UK LDLs. The BL's policies cover CUL's NPLD content once it is in the BL's custody, and these policies are agreed in partnership with all LDLs. Responsibility for CUL's NPLD content lies with both CUL and the BL, including ensuring preservation activities meet CUL's requirements. NPLD content collected by CUL that is waiting to be transferred into the BL's custody, falls within the scope of this Policy.

3.1.1.3. Digital content hosted by Cambridge University Libraries

CUL hosts and makes available third-party digital content. CUL endeavours to care for third-party digital content and metadata in the same manner as CUL's own digital content and metadata. Digital content hosted by CUL has in place at a minimum, an agreement, contract or MOU that provides information on hosting arrangements, management, preservation of and access to digital content with third parties. Agreements, contracts or MOUs indicate service level agreements (and definitions) regarding hosting, preservation and access arrangements including time limits and details regarding custodial transfer of digital content and metadata, when hosting arrangements cease.

3.1.1.4. Carriers

Digital content contained on carriers of any format is considered in scope, taking into account CUL's acquisition policies and the University's *Statement of Records Management Practice and Master Records Retention Schedule*. Audiovisual content contained on both digital and analogue carriers is also covered by this Policy. For the care of carriers, this Policy should be read in conjunction with CUL's *Collection Care and Conservation Policy*.

3.1.1.5. Associated materials

Associated materials, such as additional data or files that support, provide context or assist in the preservation and/or provision of access to the digital content in CUL's collections are considered in scope.

3.1.1.6. Metadata

The preservation of metadata supporting digital content, including but not limited to descriptive, administrative, technical, structural, rights and/or preservation metadata is in scope of this Policy. This metadata may be found in associated materials, accession records, catalogue records and/or embedded in files.

3.1.2. Out of scope

3.1.2.1. Corporate and organisational data and current records

Digital content created as part of CUL's operational processes (and required for ongoing business purposes) – unless explicitly related to digital content in CUL's collections (for example, associated materials) – is not considered in scope. The responsibility for the management, maintenance and backup of corporate and organisational data and records, lies with the individual staff member and their department in conjunction with CUL's Digital Initiatives and Strategy (DIS) and the main University Information Services (UIS).

3.1.2.2. Active research data

CUL holds research outputs including research publications, research data (whether or not associated with a publication) and associated materials, that is made available via the University's Apollo service. Active research data is not considered in scope of this Policy. The management, maintenance and backup of active research data is the responsibility of the data creators, their relevant University departments and/or UIS, where UIS services are used for storage and backup of research data.

3.1.2.3. Non-Print Legal Deposit content managed by the British Library

For NPLD content acquired, managed, preserved, stored and/or where access is provided by the BL under the *Legal Deposit Libraries (Non-Print Works) Regulations 2013* on CUL's or other LDLs behalf, CUL works collaboratively with all LDLs to guide and influence the BL's policies for NPLD content. CUL also works collaboratively with the other LDLs to ensure NPLD content – including emerging and innovative published formats – is acquired, preserved and made available for the long-term. The BL is responsible for carrying out the majority of digital preservation activities on NPLD content, including preservation actions. Agreed upon by all LDLs, these preservation activities meet CUL's digital preservation requirements.

3.1.2.4. Access copies

Access copies of digital content are considered out of scope of this Policy. The exception is where an access copy is the only version available, due to a preservation master and/or co-master no longer existing or being available, or because it was never created in the first place. In this instance, the access copy or a version of the access copy is deemed the preservation master or co-master, and is managed and preserved as such.

3.2. Principles

CUL takes a proactive and risk-managed approach to digital preservation. CUL recognises that long-term management and preservation of digital content requires organisational commitment, a skilled workforce, ongoing planning and active management from the point of acquisition or creation, through the entire lifecycle of the digital content, and for the whole time this digital content is in CUL's custody. CUL is committed to developing its digital collections and incrementally improving the ways in which digital content is acquired, created, managed and preserved.

Prioritisation of digital preservation activities is driven by digital preservation strategy, guided by digital preservation plans and supported by preservation needs assessments. Inaccessibility, obsolescence, rarity, significance, urgency, value to staff, students, researchers, users and/or stakeholders and demand for use of CUL's digital content are driving factors that influence prioritisation decisions.

3.2.1. Advocacy and governance

For digital preservation efforts to be successful, a holistic, balanced and continuous commitment to organisational change, resourcing and technical infrastructure improvements are required. In order to support this, CUL acknowledges that ongoing advocacy at all levels across the University, in the UK, and internationally, is necessary. As an identified area of significant risk for CUL, appropriate governance is dedicated to digital preservation efforts and development within the organisation, in order to meet CUL's key strategic goals and vision, and the University's mission and core values.

3.2.2. Organisational engagement

CUL staff across all levels of the organisation that are acquiring, creating, managing, preserving, providing access to and/or working with digital content, ensure digital preservation is factored into organisational and team strategy planning, work practices, current and future project development and funding applications.

3.2.3. Organisational maturity and planning

CUL is committed to continuous improvement of digital preservation capability. CUL acknowledges that there is much work to be done to establish and embed Business-As-Usual (BAU) digital preservation practices, and is yet to transition from project-based digital preservation efforts. This requires significant prioritisation, organisational change and workforce planning while mitigating risks and managing expectations. CUL commits to undertaking annual digital preservation capability maturity assessment in order to work towards developing trustworthy processes and systems, to inform forward planning. In addition, CUL commits to undertake annual workforce assessment to ensure the incremental development of digital preservation capabilities to meet operational needs. During the transition to BAU, any project-based work will have sufficient project management leadership and support.

CUL commits to acting on recommendations produced through organisational, resourcing and digital preservation capability maturity assessments. CUL ensures planning is in place to prioritise and best address these areas, in order to efficiently utilise limited resources. Over the coming five years, CUL is shifting towards a 'digital first' approach where suitable, when acquiring and/or providing access to its collections.

3.2.4. Workforce development

CUL ensures that it maintains appropriate staffing levels, including dedicated staff with expert skills, to support the preservation of digital content in its custody. In line with CUL's *Learning and Development Policy* and *Interim Learning and Development Plan 2018 – 2020*, staff responsible for acquiring, creating, managing, preserving and making available digital content are provided with professional development opportunities to ensure they can fulfil the requirements of their job roles. Appropriate training, guidance and/or supporting resources are provided to all other staff, researchers and users

working with or using digital content to ensure that they possess the necessary skills for safe handling of carriers and digital content.

3.2.5. Resourcing

CUL will establish and enact a robust funding model for digital preservation. To ensure sustainability of digital preservation practices, and therefore the long-term sustainability of digital content in CUL's custody, CUL allocates, and actively seeks out and applies for, adequate and strategically aligned funding. This is necessary in order to mitigate against economic challenges of digital preservation efforts.

3.2.6. Standards

CUL implements existing standards and/or good practices commonly used by the international digital preservation community. CUL also employs good practice in data management procedures and processes. CUL creates internal standards to formalise current and future work practices. These are reviewed and approved via a cross-departmental standards group.

Where the creation of digital content can be influenced, creation of digital content is standards-based or follows good practice, to ensure that preservation-quality digital content and metadata is produced.

3.2.7. Metadata

CUL acknowledges that capturing and/or creating a minimum amount of metadata is essential for the long-term preservation of and continued access to digital content. Minimum metadata is captured and/or created at the first point of contact with digital content. This may include but is not limited to, descriptive, administrative, technical, structural, rights and preservation metadata (as associated materials, in accession records, catalogue records and/or embedded in files). Metadata is created and/or transformed to meet relevant standards. Digital content created by CUL always has accompanying standards-based metadata created. In order for digital content to be acquired by CUL, it must be accompanied by a minimum amount of metadata.

Metadata complies with existing standards and/or uses good practice approaches. Metadata is maintained, indexed and backed up. Any recovery efforts of digital content also include the recovery of all associated metadata.

3.2.8. Maintaining fixity, integrity, authenticity and provenance

In order to be a trustworthy custodian, CUL recognises that maintaining the integrity, authenticity and provenance of digital content is essential. While digital processes alone cannot prove the integrity, authenticity and/or provenance of digital content, they can aid in ensuring knowledge about the digital content is retained over time. 'Fixity' is the measure used by the digital preservation community to ensure that no unauthorised change to digital content occurs. CUL ensures that adequate technical processes are in place for all digital content and/or metadata in its custody, in order to measure the fixity of CUL's digital content. Fixity is regularly monitored – annually for all digital content and more frequently for some classes of digital content, as required. Where unauthorised changes or corruption has occurred, this is immediately reported on and recovery efforts are enacted.

For all digital content, fixity is established immediately after preservation masters and/or co-masters are created by CUL, or transferred into CUL's custody. For digital content on offer, appraisal only takes place after fixity is established. Digital content and/or metadata moved within CUL, has in place appropriate digital processes to evidence that no unauthorised changes or corruption has occurred. Digital content acquired by CUL has persistent identifiers applied, with 'original' filenames retained in metadata.

3.2.9. Associated materials

In order to render, interpret, provide context and/or make available some digital content, additional data or files may be required. These are known as associated materials. In order to render and/or make available digital content that is dependent on associated materials, associated materials are also acquired and preserved. Associated materials are retained, managed and preserved alongside digital content, in order to ensure that access to, and/or enhanced experiences of digital content can be provided to staff, students, researchers and users.

3.2.10. Acquisition and creation

The acquisition of digital content is a developing area for CUL. Dedicated policies, strategies and guidelines for born-digital acquisition and digitisation need developing. Where digital content is on offer, and it meets CUL's *Collection Development Policy Framework* and/or *Cambridge University Libraries Strategy*, CUL undertakes proactive and early acquisition. In the same manner that CUL builds its physical collections, CUL ensures that due diligence checks are undertaken in accordance with relevant legislation and professional ethics regarding the acquisition of cultural property in digital form.

Whether by choice, legislative or contractual obligation, agreement, terms and conditions, MOU or determined by policy, CUL accepts any type of digital content, regardless format, quality, carrier or rights. CUL prefers and selects the highest quality of digital content available, with a preference for open rather than proprietary file formats. Where the creation of content can be controlled or influenced, CUL encourages the creation of digital content in standards-based, open source and/or widely adopted file formats that are considered suitable for long-term preservation, and to the highest quality appropriate for the content type. For certain types of digital content, CUL may accept more than one copy, in different formats and/or qualities.

Digital content transferred to CUL for appraisal purposes will have fixity established, be secured, managed and backed up – with processes in place to ensure CUL's existing digital collections are not placed at risk – until such times that appraisal can take place. Acquisition of digital content must be accompanied by a minimum amount of metadata. For acquisition of orphaned works, risk assessments are undertaken with possible exceptions made.

Digitisation of or transfer from carriers operates on the principle that the data or signal may only be able to be read once. Where it is economically or technically infeasible to acquire, digitise or transfer digital content – or where CUL does not have specialist expertise – alternative arrangements may be pursued, which may influence acquisition decisions. Carriers deemed an immediate risk (by way of a preservation needs assessment and/or technical analysis) or deemed high-priority, have appropriate actions taken in order to digitise or transfer content as soon as is practicable.

3.2.11. Preservation

CUL recognises that preservation is an active process, and commits to undertaking continuous preservation activities across the broad spectrum of work needed, in order to mitigate the risks associated with CUL's digital content. Digital content is preserved in perpetuity, unless subject to agreements, contracts, MOUs or retention schedules. Minimum preservation activities ensure that preservation master, co-master and born-digital original digital content, associated materials, associated metadata and fixity information for each file are retained.

Preservation actions are undertaken within the legal framework and exceptions that CUL is granted. Preservation activities are planned, documented and implemented, and are adequately resourced. Preservation actions are based on industry-based good practice, thoroughly tested, evidence-based and recorded, ensuring there is no unacceptable loss of data.

For at risk carriers, where no preservation solutions are available to immediately stabilise carriers (until such times that content can be digitised or transferred), CUL prioritises transfer or digitisation of content, complying with industry standards, guidelines or good practice. As per international guidelines and good practice, original carriers are retained and preserved, once digital content has been transferred from the carrier, with the exception of temporary transfer devices.

3.2.12. Documentation

Documentation is created at all stages of digital content's lifecycle, and selected documentation is recorded and retained. Where documentation is available from creators, donors or producers of digital content, this is also acquired alongside the digital content. In order to better support research outputs, researchers are encouraged to submit Data Management Plans (DMPs) that provide information about the long-term preservation and access needs of the digital content.

Agreements, contracts and MOUs are created, archived, managed and made available to authorised CUL staff as required, always complying with recordkeeping standards. Acquisition, deaccessioning and disposal decisions are documented and retained.

Policies, guidelines, procedures, standards, processes and workflows, developed to support acquisition and preservation activities for digital content are documented. Configuration and customisation of, and integration between, systems and tools used for acquisition, creation and preservation of digital content are also documented.

3.2.13. Technical infrastructure

CUL develops, implements and maintains technical infrastructure, working with UIS to provide the best possible solutions to secure and safeguard digital content. To ensure the security of the systems as a whole, infrastructure is replaced and refreshed in line with UIS policies.

3.2.13.1. Systems and tools

Systems and tools are developed, implemented and configured to support end-to-end workflow processes for monitoring and reporting on digital content. Where fit for purpose, open source systems and tools are preferred. A standards-compliant digital preservation system and metadata management systems will be implemented and maintained. Only those that ensure adequate migration capabilities or facilitate appropriate exit strategies for both digital content and metadata are considered. Systems and tools implemented must be able to manage the digital content as well as have the capability to report on how they function.

3.2.13.2. Monitor and report

Monitoring and reporting is an essential aspect of digital preservation activities. This is necessary for digital content management as well as organisational, technical and preservation planning activities. Continuous monitoring and reporting on the systems and tools used for maintaining digital content occurs. At minimum, annual fixity reports detailing any unauthorised changes or corruption to digital content are produced and circulated to authorised CUL staff. Digital content risk reports are produced annually, and are circulated to authorised CUL staff.

3.2.13.3. Information security

CUL commits to securing digital content at all times, and to protect against loss, misuse, unauthorised change and/or corruption. Preservation masters, co-masters, born-digital originals and associated materials are located in secure and managed network locations, separate from front-facing services. At all times – including prior to ingest into a digital preservation system – preservation masters, co-masters, born-digital originals and associated materials are located in secure, managed networked locations and are ingested into a digital preservation system as soon as is practicable. The management of digital content, both prior to ingest into and while managed in a digital preservation system, meets UIS policy and data privacy regulations.

3.2.13.4. Storage and storage locations

CUL employs good practice approaches and methods for managing and monitoring digital content. Preservation master, co-master, born-digital original, associated materials and associated metadata are stored in no less than three geographically separate locations and utilise a diversity of storage media. Recovery of all preservation master, co-master, born-digital original, associated materials and associated metadata is possible from all locations and storage media, and is successfully tested on an annual basis.

CUL commits to ensuring that adequate and appropriate infrastructure – including hardware, software, storage media and any other equipment – used in the storage of preservation master, co-master, born-digital original, associated materials or associated metadata is kept up-to-date and subject to regular maintenance and equipment refresh schedules. Maintenance and refreshes are in line with CUL DIS and UIS documented policies, strategies and plans.

3.2.13.5. Shared storage solutions

Where necessary, CUL engages in shared storage solutions for preservation purposes if it meets CUL's preservation requirements, and is economically viable to do so.

3.2.14. Access conditions, copyright, deaccessioning, disposal and rights

CUL ensures that digital content acquired or created has rights metadata generated and/or provided (covering both current and future uses). Where digital content is found to have no access conditions or rights metadata available, risk-based assessments are undertaken, taking into account relevant legislation and/or other factors and metadata is created, prior to access being provided. Access conditions, copyright and/or rights are adhered to when managing, preserving and providing access to digital content. Appropriate practices are established to manage personal and sensitive data, and compliance with data privacy regulations.

Deaccessioning and disposal of digital content, associated materials and/or metadata is undertaken in line with the University's *Statement of Records Management Practice and Master Records Retention Schedule*, and the University's policies, legislation, agreements and/or contracts.

3.2.15. Access and use

The fundamental purpose of digital preservation is to provide staff, students, researchers and users with access to digital content now and in the future. CUL aims for continuous access to digital content, developing mechanisms to provide meaningful and contextualised access, based on the needs of staff, students, researchers and users. CUL ensures that digital content made available is authentic, and that the integrity and provenance is known or can be proven. Access to and use of digital content complies with existing agreements, contracts, MOUs and/or relevant legislation. Take-down requests are assessed, risk-managed and adhered to as required.

Where only one instance of the digital content exists (such as digital content on carriers) and another copy cannot be reasonably obtained, a preservation copy of the digital content is made, prior to use.

3.2.16. Certification

CUL plans for and is working towards CoreTrustSeal certification, as a minimum achievement of Trusted Digital Repository (TDR) status. This requires a certain level of maturity in regards to policies, procedures, skills and associated infrastructure. CUL is working towards the procurement and implementation of a digital preservation system that is OAIS compliant.

3.2.17. Collaboration

CUL seeks to establish strategic collaborations and partnerships across the University, other organisations in the UK and internationally. Collaborations are fundamental to furthering the University's mission and core values, improving digital preservation capabilities, sharing experiences and achieving economies-of-scale. CUL engages with industry leaders and institutions at the forefront of digital preservation, benefitting from the value of a shared community of knowledge. CUL openly shares knowledge and its experiences in order to advance and improve digital preservation information and practices across the University, the UK and internationally. Staff are encouraged to share skills, knowledge and expertise across the University, and through formal and informal digital preservation community and industry channels.

3.2.18. Technology watch

CUL enables staff to undertake appropriate research and technology watch efforts, so as to facilitate digital preservation strategy, planning and BAU activities. CUL works with the international digital preservation community to maintain an awareness of current and emerging trends and issues.

3.3. Standards

To support the management and preservation of digital content, CUL aims to comply with the following British and international standards relevant to digital preservation:

3.3.1. Metadata standards and specifications

- DublinCore Metadata Initiative (DCMI) Specifications
- METS Metadata Encoding and Transmission Standard
- PREMIS Data Dictionary for Preservation Metadata, Version 3.0

3.3.2. Operational standards

- CoreTrustSeal Data Repository Certification
- ISO 14721:2012 Space data and information transfer systems – Open archival information system (OAIS) – Reference model

- ISO 31000: 2009 Risk management – principles and guidelines
- ISO/IEC 27002:2013 – Code of practice for information security controls
- PAS 197: 2009 Code of practice for cultural collections management

3.3.3. Technical standards

- BS ISO 19005-1:2005 Document management. Electronic document file format for long-term preservation
- BS PAS 198:2012 Specification for managing environmental conditions for cultural collections
- ISO 16175:2011 Information and documentation – Principles and functional requirements for records in electronic office environments
- ISO 19262:2015 Photography – Archiving Systems – Vocabulary
- ISO/IEC 15408 Information technology – Security techniques – Evaluation criteria for IT security
- ISO/TR 18492:2005 Long-term preservation of electronic document-based information
- ISO/TR 19263-1:2017 Photography – Archiving systems – Part 1: Best practices for digital image capture of cultural heritage material
- ISO/TS 19264-1:2017 Photography – Archiving systems – Image quality analysis – Part 1: Reflective originals

4. Related legislation

All acts and regulations as amended:

- *Computer Misuse Act (1990)*
- *Copyright, Design and Patent Act (1998)*
- *Copyright and Rights in Performances (Disability) Regulations 2014*
- *Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014*
- *Copyright (Public Administration) Regulations 2014*
- *Data Protection Act 2018*
- *Digital Economy Act (2010)*
- *Environmental Information Regulations (2004)*
- *Equality Act (2010)*
- *EU Copyright Directive (2001)*
- *EU General Data Protection Regulation (GDPR) (2016/679)*
- *Freedom of Information Act (2000)*
- *Health and Safety at Work Act 1974*
- *Human Rights Act (1998)*
- *Legal Deposit Libraries (Non-print works) Regulations (2013)*
- *Limitation Act (1980)*
- *Malicious Communications Act (1988)*
- *Official Secrets Act (1989)*
- *Privacy and Electronic Communications (EC Directive) (Amendment) Regulations (2011)*
- *Privacy and Electronic Communications Regulations (2003)*

4.1. Related Cambridge University Libraries policy frameworks and policies

- *Access Policy: Archives and Modern Manuscripts*
- *Collection Care and Conservation Policy*
- *Collection Development Policy Framework*
- *Collections Development Policy for Cambridge University Library: Special Collections (in draft)*
- *Cambridge University Archives Collection Policy*
- *Ethical Policy*
- *Health and Safety Policy*
- *Memorandum of Understanding between the UK Legal Deposit Libraries for the implementation of non-print legal deposit regulations*
- *Privacy Policy (Notice)*
- *Web Privacy Policy*

4.2. Related University of Cambridge policies, strategies and guidelines

- *Apollo Repository Terms of Use*
- *Business Systems and Data Integration Strategy (in draft)*
- *Cambridge Open Access Policy Framework*
- *Digital Teaching and Learning Strategy (in draft)*
- *DOI policy at the University of Cambridge Research Repository Apollo*
- *Estate Management Records Retention and Disposal Schedule*
- *Information Compliance*
- *Learning and Development Policy*
- *Privacy Policy*
- *Statement of Records Management Practice and Master Records Retention Schedule*
- *University of Cambridge Digital Strategy for Education 2016-2020*
- *University of Cambridge Research Data Management Policy Framework*
- *University Software Policy Terms and Conditions*
- *Web Accessibility Guidelines*
- *Web Strategy (in draft)*

4.3. Other relevant frameworks and guidelines

- *Federal Agencies Digital Guidelines Initiative (FADGI) – Digital File Formats for Videotape Reformatting*
- *Federal Agencies Digital Guidelines Initiative (FADGI) – Digital Imaging Framework*
- *Federal Agencies Digital Guidelines Initiative (FADGI) – Digitizing Motion Picture Film*
- *Federal Agencies Digital Guidelines Initiative (FADGI) – Embedding Metadata in Broadcast WAVE Files, Version 2*
- *Federal Agencies Digital Guidelines Initiative (FADGI) – Minimal Descriptive Embedded Metadata in Digital Still Images*
- *Federal Agencies Digital Guidelines Initiative (FADGI) – Technical Guidelines for the Still Image Digitization of Cultural Heritage Materials*
- *Federal Agencies Digital Guidelines Initiative (FADGI) – TIFF Image Metadata*
- *International Association of Sound and Audiovisual Archives (IASA) – Handling and Storage of Audio and Video Carriers (IASA-TC 05)*
- *International Association of Sound and Audiovisual Archives (IASA) – Guidelines on the Production and Preservation of Digital Audio Objects (IASA-TC 04)*
- *International Association of Sound and Audiovisual Archives (IASA) – Guidelines for the Preservation of Video Recordings (IASA-TC 06)*
- *International Federation of Library Associations and Institutions (IFLA) – Guidelines for Audiovisual and Multimedia Collection Management in IFLA Audiovisual and Multimedia Section Libraries (draft)*
- *Library of Congress Recommended Formats Statement*
- *National Information Standards Organization (NISO) – Framework of Guidance for Building Good Digital Collections*

4.4. Related charters, codes and recommendations

- *Archives and Records Association UK and Ireland (ARA) Code of Ethics*
- *Chartered Institute of Library and Information Professionals (CILIP) Ethical Principles and Code of Professional Practice for Library and Information Professionals*
- *Digital Preservation Coalition – Digital Preservation Community Charter (in draft)*
- *European Commission – Recommendation on the digitisation and online accessibility of cultural material and digital preservation*
- *International Association of Sound and Audiovisual Archives (IASA) Ethical Principles for Sound and Audiovisual Archives*
- *International Association of Sound and Audiovisual Archives (IASA) Safeguarding the Audio Heritage: Ethics, Principles and Preservation Strategy (IASA-TC 03)*
- *United Nations Educational, Scientific and Cultural Organization (UNESCO) – Charter on the Preservation of Digital Heritage*
- *United Nations Educational, Scientific and Cultural Organization (UNESCO) – Recommendation Concerning the Preservation of, and Access to, Documentary Heritage Including in Digital Form*

- *University of Cambridge – University Information Services – Charter for System and Network Administrators*

5. Roles and responsibilities

5.1. Library Syndicate

The Library Syndicate is responsible for:

- Authorising and overseeing the Policy.

5.2. University Librarian

The University Librarian is responsible for:

- Ensuring that digital preservation policy and strategy is effectively led across CUL;
- Driving digital preservation advocacy across the University;
- Developing strategic digital preservation collaborations and partnerships across the University, the UK and internationally.

5.3. Senior Leadership Team

The Senior Leadership Team (SLT) is responsible for:

- Overseeing the Policy's implementation;
- Ensuring the Policy conforms to the University's mission and core values, legislative, funding compliance and other compliance requirements;
- Strategic oversight and actively leading required change for organisational, resourcing and infrastructure arrangements across CUL, so as to support all aspects of acquisition, creation, management and preservation of digital content and processes, including digital preservation efforts;
- Ensuring adequate staffing levels are in place to support digital acquisition, creation and preservation efforts;
- Ensuring professional development opportunities are provided to staff responsible for the acquisition, creation, management and preservation of digital content;
- Securing funding and allocating resources in order to deliver a robust and long-term funding model for digital preservation;
- Overseeing the integration of digital preservation into research projects, including the digital humanities;
- Supporting relevant activities regarding specific policy implementation.

5.4. Senior Management Group

The Senior Management Group (SMG) is responsible for:

- Communicating this Policy effectively to all CUL managers and supervisors;
- Adequately planning for and facilitating access to professional development opportunities for staff responsible for acquiring, creating, managing and preserving digital content, in order for staff to fulfil the requirements of their job roles;
- Actively encouraging digital preservation efforts, in order to collect, preserve and provide access to quality digital content and metadata for all purposes, including education, learning and research.
- Supporting relevant activities regarding specific policy implementation.

5.5. Chief Operating Officer

The Chief Operating Officer is responsible for:

- Overseeing governance of CUL's commitment to digital preservation.

5.6. Deputy Director, Digital Initiatives

The Deputy Director, Digital Initiatives is responsible for:

- Overseeing the creation and maintenance of adequate infrastructure to support digital preservation efforts;
- Allocating appropriate technical resources to support digital preservation infrastructure;
- Overseeing the development and maintenance of policies, strategies, guidelines, procedures, processes and workflows to support digital content and digital preservation efforts.

5.7. Deputy Director, Research Collections

The Deputy Director, Research Collections is responsible for:

- Ensuring the integration of digital preservation into research projects, including the digital humanities;
- Ensuring funding proposals for projects involving current or new CUL digital content and/or metadata factor in and adequately resource the necessary associated digital preservation activities, either in funding applications or by other means;
- Leading the active encouragement of digital preservation efforts, to improve CUL's ability to collect, preserve and provide access to quality digital content and metadata for education, learning and research purposes.

5.8. Director, Digital Transformation Programme

The Director of the Digital Transformation Programme is responsible for:

- Leading digital preservation efforts at CUL;
- Leading the implementation of this Policy;
- Working together with CUL teams (including Development), to secure funding in order to deliver a robust and long-term funding model for digital preservation at CUL;
- Working collaboratively with managers and supervisors across all areas of CUL to define job roles, recruit and allocate resources to digital preservation efforts;
- Coordinating and facilitating the ongoing review and update of this Policy and associated guidelines, processes and workflows in consultation and collaboration with relevant CUL and University stakeholders.

5.9. CUL officers

CUL officers, line managers and managers of teams are responsible for:

- Communicating this Policy effectively to CUL staff;
- Managing processes and workflows relevant to this Policy;
- Providing support and systems to support processes associated with acquiring, creating, managing, preserving and providing access to digital content;
- Ensuring that staff responsible for acquiring, creating, managing and preserving digital content are provided with professional development opportunities so they are able to fulfil the requirements of their job roles;
- Ensuring other staff who use digital content are provided with adequate training and/or guidance as appropriate;
- Ensuring work practices comply with this Policy.

5.10. Digital Initiatives and Strategy

DIS are responsible for:

- Maintaining the infrastructure and systems that support digital preservation efforts.

5.11. CUL staff

CUL staff, contractors, consultants, volunteers, service providers and vendors are responsible for:

- Understanding and complying with this Policy;
- Seeking out professional development opportunities in order to meet the demands of their job roles.

5.12. Donors and researchers

Donors and researchers are aware of how digital content transferred to CUL's custody is managed and preserved. Donors and researchers who provide digital content to CUL endeavour to deliver the highest-possible quality of digital content that is available or can be created.

6. Definitions

Please refer to Appendix A.

7. Implementation framework

CUL is establishing and implementing digital preservation practices, policies, strategy, guidelines, procedures, standards and skills to support the acquisition, creation, management and preservation of its digital content. CUL is also commencing a digital transformation programme. Part of this portfolio

of work involves the procurement and implementation of infrastructure, systems and workflows to support digital preservation activities.

8. Approval

This Policy was approved by the Library Syndicate on 30 October 2018 for immediate implementation.

9. Policy implementation

This Policy is implemented on 22 November 2018.

10. History

This is a new policy.

11. Prepared by

Digital Preservation Specialist – Policy and Planning (Polonsky Fellow)
03 October 2018

12. Document history and version control

Version	Date approved	Approved by	Brief description
1.0	30 October 2018	Library Syndicate	First release

13. Appendix A

Definitions in use by the international digital preservation and digital curation communities are sourced from the Digital Preservation Coalition's [Digital Preservation Handbook Glossary](#), the Digital Curation Centre's [Glossary](#), the National Digital Stewardship Alliance (NDSA) [Glossary](#) and Wikipedia.

Access copy – The version of a file or data that is made available to users. Access copies are distinct from preservation masters or co-masters, however in some cases an access copy file is identical to a preservation master or co-master. Access copies are not intended for long-term preservation. Where preservation masters and/or co-masters have been lost, or can no longer be located, a duplicate of an access copy may be made, and deemed the preservation master.

Active research data – Research data files that are in the process of continuous change and/or development. Files containing this data are accessed, amended and/or updated as new data is gathered and/or processed. Some datasets may never be 'finished'. A 'snapshot' of active research data can be archived to create a version that is fixed and can be cited.

Archival master – See preservation master.

Associated data – See associated materials.

Associated materials – Data or files that provide context, assist in the interpretation of, or are essential in the process of rendering and/or making available digital content. Associated materials may include a wide variety of files such as algorithms, code, sidecar files, scripts, transcripts, Optical Character Recognition (OCR) files etc. They may also be referred to as associated data or supporting documents. Also see sidecar file.

At risk – Digital content that due to obsolescence, fragility of carriers and/or formats, or for another reason, may become corrupted or lost without immediate or near-immediate action.

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 that 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 to differentiate them from digital materials that have been created as a result of converting analogue originals, and digital materials that may have originated from a digital source but have been printed to paper, for example some electronic records.

Born-digital original – The first-known instance of a file that is transferred from a donor or researcher to a collecting institution or other custodian. Due to the nature of digital content, there may be multiple prior versions of a file, however this is the known and documented version of a file, without any alterations made to the content format or other properties of the file.

Carrier – A physical item on which content is recorded, encoded or fixed. This can be stored as analogue or digital information. Selected examples of physical carriers include magnetic tape (carrying analogue or digital audio and/or video content), motion picture film (carrying optical moving image and/or audio content), disks (zip disks, 3½ inch and 5¼ inch floppy disks, carrying digital data), optical media (such as compact discs, DVDs and Blu-ray discs, carrying digital data, audiovisual or multimedia content) portable hard disk drives or USB flash drives (carrying digital data). May also be referred to as a physical carrier, physical format carrier, physical format digital carrier, portable media or transfer device.

Checksum – A unique alphanumeric signature derived from a file. Used to compare copies. A checksum can be conceptualised as the ‘fingerprint’ of a digital file.

Class – CUL digital content is broadly grouped into a number of different ‘types’ of digital content. These types are somewhat associated with the CUL’s collecting streams. How digital content is managed in the long-term, may differ slightly, depending on the class to which it belongs.

CoreTrustSeal – A harmonisation of the Data Seal of Approval (DSA) and the ICSU World Data System (WDS). The CoreTrustSeal is a core-level certification. The requirements reflect the core characteristics of a Trusted Digital Repository.

Co-master – A co-master file is derived from a preservation master file. The co-master may have cropping, filtering and/or other similar actions applied. It is typically the source from which access copies are generated, and may itself be considered a high-quality access file.

Data migration – Moving data, such as metadata, from one location to another. Typically, this term refers to moving data (or metadata) held in one system to a new system (such as from one database to another).

Delivery copy – See access copy.

Derivative – See access copy.

Digital asset – See digital content.

Digital content – Any arbitrary item created, published or distributed in digital form, including, but not limited to text, data, sound recordings, photographs and images, motion pictures and software. May also be referred to as digital assets, digital materials, digital objects or digital resources. This encompasses both born-digital and digitised content.

Digital imaging – See digitisation.

Digital materials – See digital content.

Digital object – See digital content.

Digital Object Identifier (DOI) – A technical and organisational infrastructure for the registration and use of persistent identifiers widely used in digital publications and for research data. The DOI system

was created by the International DOI Foundation and was adopted as International Standard ISO 26324 in 2012.

Digital resource – See digital content.

Digital Rights Management (DRM) – A general term used to describe the control of copying, viewing, altering, printing and other modifications to digital content. DRM is most commonly found to be implemented with copyrighted published digital content. For a more specific implementation, see Technological Protection Measure (TPM).

Digital surrogate – Manifestation of a physical collection item in digital form, for example, a digital image file of a manuscript item. These can be made more widely available than the original physical collection item they represent and allow for new modes of discovery and use that may not be possible with the original physical collection item. A digital surrogate does not replace the original physical collection item. This is typically referred to as digitised content.

Digitisation – The process of creating digital files by scanning or otherwise converting analogue materials. The resulting digital copy, or digital surrogate, can then be classed as digital content and is subject to the same broad challenges as born-digital content, involved in preserving and maintaining access.

Digitised content – The term used to refer to digital content that has been created by way of a digitisation process of an original physical collection item. This may also be referred to as a digital surrogate.

Distribution copy – See access copy.

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 particular file format is often indicated by a filename extension containing three or four letters that identify the format.

Filename – A name used to uniquely identify a computer file stored in a file system.

File path – A specified unique location that is formed by the list of directories (or folders) and sub-directories, that forms a path to where a file is located in a file system.

File system – Controls how data is stored and retrieved.

Fixity – Checksum(s), filename and file path that are generated and/or recorded for a specific file. The combination of these properties can be used to identify and verify whether any changes have taken place to a file between two points in time. For information on this process see fixity check.

Fixity check – A method for ensuring the integrity of a file and verifying it has not been altered or corrupted. During transfer, an archive may run a fixity check to ensure a transmitted file has not been altered en route. Within the archive, fixity checking is used to ensure that digital files have not been altered or corrupted. It is most often accomplished by computing checksums such as MD5, SHA1 or SHA256 for a file and comparing them to a stored value.

Format migration – A means of overcoming technological obsolescence by transferring digital content from one hardware/software generation to the next. The purpose of migration is to preserve the intellectual content of 'digital objects' and to retain the ability for users to retrieve, display, and otherwise use them in the face of constantly changing technology. Migration differs from the refreshing of storage media in that it is not always possible to make an exact digital copy or replicate original features and appearance and still maintain the compatibility of the digital content with the new generation of technology. May also be referred to as migration.

Hosting – Providing a service by which digital content is temporarily in CUL’s custody and that CUL is responsible for managing and/or preserving this digital content.

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.

Legal Deposit Libraries (LDLs) – Cambridge University Libraries is one of six Legal Deposit Libraries in the UK and Ireland. The other Legal Deposit Libraries are the British Library, the Bodleian Libraries, Oxford, the National Library of Scotland, the National Library of Wales and Trinity College, Dublin.

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, 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.

Migration – The process of transferring digital content from one system to another. This may involve export from one system and ingest into the new system. For migration from one file format to another, see Format Migration. For metadata migration, see data migration.

Open source – A philosophy and methodology (often used by software developers) where the source code is made freely available, so that others may continue to develop the software. Open source is the philosophical opposite of proprietary.

Original – A collection item that has physical properties, for example a book, manuscript, painting, sculpture or carrier. May also be referred to as original physical item, physical item, physical original.

Persistent Identifier (PID) – A unique set of characters applied as the filename to a file that remains the same in perpetuity. For information on how persistent identifiers are used, see Digital Object Identifier (DOI).

Physical carrier – See carrier.

Physical format carrier – See carrier.

Physical original – See original.

Portable media – See carrier.

Preservation action – A specific, defined and measurable task undertaken on a digital file or files for the purposes of stabilising the file(s) and/or making them accessible. Preservation actions can be reversible or irreversible. Each preservation action should be documented.

Preservation master – Digital content targeted for preservation that is considered the ‘master’ version of the intellectual content of any ‘digital object’. Preservation master files are created to high capture standards and may also capture additional information about the original, beyond the content itself. In extremely rare cases, preservation master files may take the place of the original file if the original is destroyed, damaged, or lost. Preservation masters generally do not undergo significant processing or editing (this takes place on co-masters). Preservation masters are often used to make other copies including reproduction and distribution copies. May also be referred to as an archival master.

Proprietary – In reference to hardware technology, software applications and/or file formats, the state of being privately owned and controlled. A proprietary design or technique implies that the company has not divulged specifications that would allow other companies to duplicate the product.

Provenance – The origin, or the source of something, or the history of the ownership or location of an item or object. The term is used in a wide range of professional fields including, art collection, archival management, librarianship, computing and law. In most fields, the primary purpose of provenance is to confirm or gather evidence as to the time, place, and – when appropriate – the person responsible for the creation, production, or discovery of the item or object. Also referred to as chain of custody.

Reformatting – Copying information content from one storage medium to a different storage medium (media reformatting) or converting from one file format to a different file format (file re-formatting). This term may be replaced by the terms transfer or format migration, depending on the task needing to be undertaken and the context in which the term is used.

Reproduction copy – See access copy.

Service Level Agreement (SLA) – a contract between a service provider and an organisation (or end user), defining the expected level of service from the service provider.

Sidecar file – Computer files that store data (often metadata), which is not supported by the format of a source file. There may be one or more sidecar files for each source file. They may also take the form of ‘metadata databases’, where one database contains metadata for several source files. In most cases the relationship between the source file and the sidecar file is based on the filename; sidecar files typically have the same base name as the source file, but with a different extension (as they are different file formats). Most computing systems including Operating Systems, file systems, digital preservation systems and tools have no knowledge of these relationships. If the filename or file path of a sidecar file is modified, or if the sidecar file is removed, the relationship to the file is lost. The worst-case scenario is that the file dependent on the sidecar file can no longer be rendered. They are also known as buddy files or connected files.

Supporting documents – See associated materials.

Technological Protection Measure (TPM) – A form of ‘technical lock’, such as encryption, that enables security of digital content, generally disabling the user from copying data and typically preventing data from being separating from the carrier on which it is contained. TPMs are most commonly found to be implemented with copyrighted published digital content. For a more general description of controlling digital access, see Digital Rights Management (DRM).

Technology watch – An organised and structured system for the searching, detection and analysis of the environment, aimed at disseminating and transmitting information and knowledge in a continuous manner, at the required time, so that the target group can be made aware of the main activities in their sector, within the technology field.

Transfer device – A temporary carrier for digital content. Transfer devices are typically external hard disk drives or USB flash drives. Also see carrier.

Trusted Digital Repository (TDR) – A trusted digital repository has been defined as having ‘a mission to provide reliable, long-term access to managed ‘digital resources’ to its designated community, now and into the future’. The TDR must include the following seven attributes: compliance with the reference model for an Open Archival Information System (OAIS), administrative responsibility, organisational viability, financial sustainability, technological and procedural suitability, system security, and procedural accountability. The concept has been an important one particularly in relation to certification of digital repositories.