**West Virginia and Regional History Center**

**Digital Preservation Policy**

# **Summary and Purpose**

As part of its mission as an academic institution, the West Virginia & Regional History Center (WVRHC) acquires, makes accessible, and preserves materials of all formats that pertain to our collecting scope and which serve the needs of West Virginia University. To accomplish these activities, a structured and sustainable digital preservation program that adheres to professional standards within the library and archives fields and is supported by an institutional commitment of resources and an institutional mandate is required.

This Digital Preservation Policy defines a framework that governs the operation, management, and scope of the WVRHC digital preservation program, which supports the preservation and access of authentic digital records in our collections throughout their entire lifecycle. Principles – such as robust administrative support and adoption of best practices and accepted preservation standards – guide the program and its stakeholders. These principles call for collaborative efforts directed by the WVRHC that involve partnerships with record creators and multiple departments within West Virginia University.

# **Mandate**

The mandate for digital preservation at the WVRHC is linked to our institutional responsibility as part of a land grant university, legal obligations, scholarly commitment, contractual obligations, and grant obligations. Digital preservation supports our role as the keepers of the University Archives, which is the official repository of certain university documents as noted in [WVU Policy G-1 Record Retention Policy & Schedule](https://generalcounsel.wvu.edu/record-retention-policy-schedule), [Exhibit A: University Retention Schedule](https://generalcounsel.wvu.edu/files/d/0f0df2fb-c424-45f6-8b98-304648de4c7c/exhibit-a-7-28-21-w0111882xd4daa.pdf). Additionally, strong support of digital preservation furthers two key goals within the [2021-2024 WVU Libraries Strategic Roadmap](https://library.wvu.edu/about/strategic-roadmap) that outline that the library should, “be a leader in the creation and dissemination of knowledge,” and “be an exemplary land-grant university library that advances and promotes West Virginia and the Appalachian region.”

# **Objectives**

The digital preservation program is intended to preserve, maintain, and make accessible over the long-term the digital assets of the WVRHC’s collections, including those created and managed digitally, sometimes referred to as “born-digital,” and those transformed to a digital form, referred to as “digitized.”

The primary objectives of the digital preservation program are to:

* Preserve born-digital and digitized assets collected by the WVRHC by ensuring that contextual and intellectual meaning, authenticity, integrity, validity, and long-term accessibility of digital objects in the custody of the WVRHC are maintained.
* Comply with professional standards and best practices for digital preservation and access, particularly the *Open Archival Information System (OAIS) reference model* (ISO 14721) and *Audit and certification of trustworthy digital repositories* (ISO 16363) as applicable.
* Create transparent and accessible policies and procedures for the preservation and access of digital assets.
* Provide appropriate levels of access to digital assets while maintaining and supporting intellectual property and privacy rights in conjunction with WVU Identity and Authentication Management.
* Where applicable or necessary, collaborate with campus, regional, and national partners to make the best use of resources and avoid duplication of effort.

# **Scope and Levels of Commitment**

Digital preservation always entails making choices about what to preserve. Finite resources of time, dedicated staffing, and technology necessarily limit what an institution can actively preserve. As such, this policy addresses preservation of digital assets for which the WVRHC is the primary custodian. Preservation decisions are always made within the context of the WVRHC’s appraisal and collection development activities which balance costs and the limitations of resources, historical and scholarly value of the materials, and the needs of users. The primary focus of preservation activities is on preserving the intellectual content and context of the materials. The WVRHC shall preserve, as much as possible given limitations in available software and technology, the following:

1. Digital materials which have been officially appraised as having permanent archival value in accordance with the WVRHC Collecting Policy, legally transferred to the WVRHC, and accessioned by the WVRHC.
2. Digital surrogates of analog WVRHC collection materials that have been previously digitized.

In recognition that digital preservation is an ongoing, resource-intensive endeavor, the WVRHC has outlined levels of commitment to different categories of materials. Priority within these categories will vary based on the research value, uniqueness, and level of danger of obsolescence or loss:

1. Born-digital materials
   * The highest degree of effort will be made to ensure preservation in perpetuity of material selected for preservation through inclusion in the WVRHC’s collections. This effort may include preservation strategies such as migration, geographic distribution, and redundant bit-level replication. Digitized materials with no available analog versions will be treated with the same care as born-digital materials.
2. Digitized materials (available analog version)
   * Whenever possible, digitized materials will be created using file formats conducive to long-term preservation activities. The cost of re-digitizing materials should be weighed against the cost of long-term preservation. In most cases the analog version of these materials will be considered the preservation format, and preservation activities for the digital materials will be limited to local and non-distributed bit-level replication. In cases where the analog carriers of information are obsolete or at great risk of obsolescence, such as audio/visual materials, the digitized materials will be considered the preservation copy and treated with the same care as born-digital materials.
3. Legacy digital materials
   * Legacy digital materials are materials donated to the WVRHC prior to the development of a digital preservation system. Many of these materials were not officially appraised, accessioned, or evaluated for enduring value. Some of these materials are stored on obsolete media, encoded in obsolete file systems or formats, or are otherwise inaccessible. When possible, the WVRHC will attempt to recover this data and evaluate it for inclusion into the digital preservation system. The WVRHC makes no guarantee that recovery will be successful or that it will be able to provide the resources necessary to attempt recovery.

# **Challenges and Risks**

Without the adoption of strategic measures to preserve and make accessible the digital assets held by the WVRHC, much of the unique and irreplaceable historical and cultural legacy of the university, state, region, and nation will be lost. To mitigate the potential occurrence of such a scenario, risks and challenges must be addressed to create and maintain a successful digital preservation program.

The primary challenge in implementing an effective digital preservation program is the technical and organizational complexity inherent in the design, development, and operation of a digital preservation system. It is worth noting that digital preservation is never “done” and that digital materials require consistent preservation actions and resources due to the constantly evolving nature of technology. Beyond that, recognized challenges and risks include:

* Sustainability: The longevity of a program is reliant upon affordability and a stable cost model. WVU Libraries requires sufficient funding for startup and major improvements, as well as designated library funding to sustain ongoing preservation efforts. The program should reflect reasonable expectations of requisite resources, i.e., WVU Libraries should not promise more than can be delivered. A sustainable and well-paced program will also help prioritize and maintain progress to capture digital assets from their container media before the media degrades to the point that the assets are corrupted or lost.
* Rapid Growth and Evolution of Technology: The variety of digital asset formats and dissemination mechanisms change rapidly, introducing the threats of an ever-expanding list of formats to preserve, technological dependencies, and technical and functional obsolescence. Our program must be flexible enough to respond to evolving technological capabilities and changing user expectations without jeopardizing the ongoing care of the digital content. Additionally, with an increasing proportion of materials in a digital format, adequately describing digital assets so that they are accessible; ensuring their technical characteristics are captured and their archival provenance is maintained; and more is critical to a successful program.
* Infrastructure and Internal Partnerships: Infrastructure challenges include guarding against hardware, software, and network failures that could corrupt or erase digital assets. Internal partnerships, such as between the WVRHC and the Libraries’ Systems Development and Infrastructure departments, will also be crucial.
* Commitment and Management: Well-managed and preserved digital collections require institutional effort, partnership development, and financial commitment. Leadership must maintain a vision that prioritizes preservation or at least balances preservation and access. Management must also commit to continually updating staff expertise as needed.
* External Partnerships: The WVRHC should strive to work with creators and providers of crucial content to employ appropriate maintenance prior to deposit that will facilitate future preservation.
* Privacy, Ethical, and Intellectual Property Rights: There are many intellectual property and other rights-based constraints on providing access that impact digital preservation efforts. These rights change as legislation and legal precedent evolves and impact the actions that can be taken by a digital preservation program as well as options for providing access.
* Political: Over time, political shifts at a variety of levels may put pressure on the program to prioritize preservation of certain materials, deaccession certain materials, or may increase the need to preserve at-risk materials. External pressures to increase certain services or reappropriate resources should be mediated by professional standards and ethics in the library and archives fields.

**Program Principles**

The digital preservation program is based on the following program principles:

* Maintenance of Authenticity
  + The WVRHC strives to ensure the authenticity of digital records because the mutable nature of digital assets opens the possibility for unauthorized and undetectable changes. From the moment that digital resources are acquired, the WVRHC will undertake protective procedures to prevent, discover, and correct loss or corruption of digital assets due to either inadvertent or malicious intent. In addition, WVRHC staff will endeavor to secure supporting evidence, ideally in the form of metadata, from those creating the resources. This will enable those who access the resources to accurately evaluate the authenticity of all preserved digital resources.
* Collection, Creation, and Preservation of Metadata
  + Metadata (descriptive, preservation, technical, and administrative) is fundamental to preserving and providing access to the WVRHC’s digital assets. The preservation process includes maintenance of metadata submitted with the digital asset as well as the creation of additional metadata to manage the long-term and active preservation and access of that asset. WVRHC staff are committed to identifying and collecting the needed metadata to preserve digital assets in its collection and to provide access to those assets.
* Emphasizing Access and Use by Our Designated Communities
  + The purpose of preserving digital assets is to ensure that the assets remain accessible in the future. The WVRHC is committed to providing access where technology, collection restrictions, and ethical, legal, and professional obligations and standards allow. To ensure continued access and understandability of materials, the WVRHC may need to migrate assets to new formats, resulting in some information loss during the migration but maintenance of the fundamental information in the records.
* Collaboration and Consultation with Record Creators
  + As with analog materials, record creators help ensure that digital content is properly identified and collected by the WVRHC. However, because of the immensely greater risk of loss associated with digital formats, there will need to be greater collaboration between producers and the WVRHC to ingest and manage digital assets throughout their entire lifecycle.
* Collaboration and Consultation with Experts
  + No one individual or department can have the needed expertise in all of the areas relevant to digital preservation work. A collaborative approach to realizing a digital preservation program means involving stakeholders at all levels of planning, development, and implementation.
* Adherence to Standards and Best Practices
  + The WVRHC digital preservation program will adhere to national and international standards, such as *The Reference Model for an Open Archival Information System (ISO 14721)* and *Audit and Certification Of Trustworthy Digital Repositories (ISO 16363),* as well as a range of applicable community-based standards and best practices, that inform digital preservation procedures and technology as well as archival requirements such as provenance, chain of custody, intellectual property rights, and authenticity.
* Scalability and Sustainability of Systems, Processes and Infrastructure
  + Digital preservation is a resource-constrained activity. The ability to preserve digital materials will necessarily be scaled to the level of committed institutional support. Choices to implement programs or initiatives must be made in the context of the overall environment at the organization and seek technologically, organizationally, and financially sustainable solutions, including through collaboration and cooperation within the organization and with external organizations and collaborative networks. A sustainable digital preservation infrastructure, framework, and program must have a realistic cost model and be interoperable, scalable, reliable, sustainable, and auditable using the best technology available given institutional contexts and resource constraints, ideally incorporating open-source options whenever feasible. Adopting open-source solutions reduces risks related to vendor reliance and ensures transparency in preservation practices. The WVRHC recognizes that open-source solutions also require continued development and maintenance over time. A sustainable model is responsive to staffing, equipment, software, and infrastructure changes as needed, without under- or overestimating the needs imposed by these changes.
* Adherence to Legal and Ethical Obligations and Requirements
  + The digital preservation program will comply with WVU’s existing policies, procedures, and or relevant applicable legislation, including compliance with the WVU Policy G-1 Record Retention Policy & Schedule as well as intellectual property; regulatory; copyright; privacy; ethical and professional obligations; and ownership rights related to copying, storage, modification, access, and use of digital resources.
* Creation of Robust and Transparent Documentation
  + Documented guidelines and procedures for each stage of the lifecycle of digital objects, i.e., creation, selection/appraisal, acquisition, accession, ingest, preservation actions—including reformatting and producing derivatives—storage, identification/cataloging, access and use, transformation, and disposal, are critical to creating a program that is auditable and accountable to our designated communities.
* Flexibility and Responsiveness
  + The program must be flexible and responsive to organizational, environmental, and technological changes, positive or negative, that impact the program without jeopardizing the ongoing care of digital content. Policies, plans, and procedures will be subject to periodic review to adjust priorities, evaluate effectiveness, and incorporate the capabilities afforded by new and emerging technologies in cost-effective and responsible ways. The program is committed to on-going training and education of staff and will share that knowledge with WVU Libraries and the campus community.

# **Roles and Responsibilities**

The preservation of WVRHC digital assets requires collaboration with WVU, WVU Libraries, and external groups. The following high-level roles exist within our digital preservation ecosystem:

1. Digital Archivist
   * Individual who designs, enables, and carries out the workflows to ensure that preservation occurs. This person will liaise and work with external departments to facilitate these tasks as necessary.
   * Serves in administration and management roles in the OAIS model.
2. Collection Curators and Archivists
   * WVRHC faculty librarians who are responsible for appraising, selecting, and curating content that is managed as part of the digital preservation program.
   * Serves in administration in the OAIS model as well as in the role of consumer when accessing information.
3. Library Systems Infrastructure and Systems Development Personnel
   * Technical services and information technology personnel who work with the Digital Archivist to support the WVRHC with the procurement, configuration, customization, and maintenance of hardware and software solutions, network infrastructure, and preservation storage.
   * Serves in a management role in the OAIS model on these issues. Serves in a limited administration role with regard to maintaining configuration management of system hardware and software.
4. Administrative Personnel
   * Advocates for the program, ensures that it stays in line with the institution’s mission, and allocates resources to help ensure the program’s success. Administrators in WVU Libraries and the WVRHC will provide adequate managerial, personnel, and financial commitment for the digital preservation program.
   * Serves in a management role in the OAIS model on these issues.
5. Legal Counsel
   * WVU General Counsel, as creators and maintainers of the University Retention Schedule, will work with the University Archives curators at the WVRHC on any schedule updates or changes, to help ensure compliance with the schedule and relevant records retention laws. As the processors of FOIA requests for the University, they may act as consumers as they need to access university records in a timely and efficient fashion. As the legal arm of the University, General Counsel will also be available to consult on issues of copyright, privacy and confidentiality of digital assets as needed.
   * Serves in a management role in the OAIS model on these issues as well as in the role of consumer when accessing information.
6. University Information Technology Services and Security
   * WVU Information Technology Services, particularly Identity and Access Management, will provide consultation on security and systems as they relate to University policies on acceptable use, security, authentication, and privacy.
   * Serves in a management role in the OAIS model on these issues.
7. Record Creators
   * Content owners or creators who submit content to the repository. Record creators will be responsible for complying with established submission requirements and working with the management of the digital archive to ensure successful transfer. This is especially true of record creators within WVU.
   * Serve as role of producer in OAIS model.
8. Designated Community of Users

* The designated community is composed of people who interact with the system to find and access archived content. They can be individuals from a variety of communities, including WVU students, staff, and faculty; members of the public; scholars; and more.
* The designated community is composed of consumers as defined in the OAIS model.

# **Financial Commitment**

This policy acknowledges that a viable and successful digital preservation program requires an ongoing commitment of resources—financial, technical, and human—from the WVU Libraries and especially the WVRHC. Enduring preservation of digital assets requires substantial and ongoing resource management over time for such activities as:

1. Creation, acquisition, appraisal, description, and storage of digital assets,
2. Creating and managing access solutions for digital assets,
3. Management of descriptive and technical metadata and data,
4. Technical infrastructure to support the above activities, and
5. Staff training.

# **Policy Review Schedule**

This policy may be revised and updated at any time, in order to reflect technological, infrastructural, and operational developments in the digital preservation program. The policy is intended to be updated at least biennially, with the first review scheduled to occur in 2024.

The current policy was approved and made effective on 2022-05-09 by Karen Diaz, Dean of WVU Libraries.

# **Resources Consulted**

The following is a selection of resources that were consulted in the development of this policy:

Other Policies

* University of Maryland Libraries: Digital Preservation Policy, second revision 2014 (<http://hdl.handle.net/1903/14745>)
* NDSA Levels of Digital Preservation (<http://ndsa.org/activities/levels-of-digital-preservation/>)
* Digital Preservation Policy Framework: A Case Study, 2014 (<https://er.educause.edu/articles/2014/7/digital-preservation-policy-framework-a-case-study>) and The Ohio State University, University Libraries, [Digital Preservation Policy Framework](http://library.osu.edu/documents/SDIWG/Digital_Preservation_Policy_Framework.pdf) (August 2013).
* Cornell University Library Digital Preservation Policy Framework, 2004 (<https://hdl.handle.net/1813/11230>)
* NEDCC Digital Preservation Policy Template (<https://www.nedcc.org/assets/media/documents/SoDAExerciseToolkit.pdf>)
* MetaArchive Cooperative, October 2010 (<https://metaarchive.org/wp-content/uploads/2017/03/ma_dp_policy_template.pdf>)
* ICPSR Digital Preservation Policy Framework, 2018 (<https://www.icpsr.umich.edu/web/pages/datamanagement/preservation/policies/dpp-framework.html>)
* Rockefeller Archive Center Digital Preservation Policy (<https://docs.rockarch.org/digital-preservation-policy/>)
* University of Washington Digital Preservation Policy (<https://www.lib.washington.edu/preservation/preservation_services/digitization-and-digital-preservation/digital-preservation-policy>)
* Northwestern Digital Preservation Policy (<https://www.library.northwestern.edu/about/administration/policies/digital-preservation-policy.html>)
* Dartmouth Digital Preservation Policy, 2014 (https://www.dartmouth.edu/library/preservation/docs/dartmouth\_digital\_preservation\_policy.pdf)

Standards

* Library of Congress Sustainability of Digital Formats: Planning for Library of Congress Collections (<https://www.loc.gov/preservation/digital/formats/index.html)>
* Open Archival Information System (OAIS) reference model (ISO 14721)
* Audit and certification of trustworthy digital repositories (ISO 16363)

# **Glossary**

Acquisition: A group of materials physically and legally transferred to a repository.

Administration (OAIS Model): The entity that contains the services and functions needed to control the operation of the OAIS on a day-to-day basis.

Authenticity: The quality of digital materials being what they purport to be and free from tampering or corruption.

Bit-level preservation: The maintenance of a digital object’s original bitstream, as opposed to file format migration. It is the minimum standard for digital preservation and allows for future preservation actions. This approach may include checksum validation, maintaining onsite and offsite backup copies, and virus checking.

Born-digital: Information created in a digital format.

Checksum: A method of verifying the integrity of digital files and is commonly used for the monitoring of whether a file has changed over time. Checksums are often called a “digital fingerprint” because the checksum algorithm used provides a unique alphanumeric string for a particular manifestation of a file.

Consumer (OAIS Model): The role played by those persons or client systems that interact with OAIS services to find preserved information of interest and to access that information in detail.

Designated community (OAIS Model): A group of primary users that the preserving institution has identified as able to access and understand the preserved information without expert assistance. This means that an appropriate level of contextual information must be preserved alongside the materials themselves. Note that multiple designated communities are possible and that they may change over time.

Digital preservation: Digital preservation combines policies, strategies, and actions to ensure the accurate rendering of authenticated digital content over time, regardless of the challenges of media failure and technological change. Digital preservation applies to both born-digital and reformatted content.

Digital object: An aggregation of one or more individual files and/or bitstreams designated as the subject of digital preservation actions and activities. For example, a journal article composed of XML full text plus a series of PNG figures is a digital object; a single PDF containing the same information would also be a digital object.

Digital preservation system: A system, consisting of people and technologies, for ensuring the long-term usability of digital objects and materials, beyond the limits of media failure or technological change.

Digital surrogate/digitized materials: Materials that were originally created in analog form but were reformatted (such as by scanning) into digital form, usually for preservation or access.

Emulation: The imitation of a computer system in order to allow programs and media designed for a particular system to operate in a different, usually newer, system. A method of overcoming technological obsolescence.

Information security: The practice of protecting the integrity and privacy of data, both in storage and in transit.

Ingest: The process of accepting electronic content or metadata into an electronic repository or database.

Integrity: Establishing that a file remains complete and unaltered over time. One method to do so is using checksums. Part of authenticity.

Management (OAIS Model): The role played by those who set overall OAIS policy as one component in a broader policy domain.

Metadata: Metadata is information describing the significant aspects of a resource that adds contextual information to that resource to aid in its discovery and management. In the context of digital preservation, descriptive metadata enables access to resources stored in preservation systems. Technical metadata such as fixity information via checksums and file format identifications assist in the preservation of a resource over time. Administrative metadata such as rights information enables determinations around access and other preservation events that may be undertaken by the preserving institution.

Migration: The transfer of digital resources from one hardware or software generation to the next, while preserving the essential characteristics of the data. A method of overcoming technological obsolescence. The purpose of migration is to preserve the intellectual content of digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in the face of constantly changing technology.

Open Source: A software development methodology and licensing approach that makes computer source code freely available, open to modification, and redistributable.

Producer (OAIS Model): The role played by those persons or client systems that generate or provide the information to be preserved

Stakeholders:

Internal stakeholders are the individuals and units within an organization tasked with identifying their organization’s past, existing, and future content and then creating a digital preservation strategy and executing its implementation. These internal stakeholders include, for example, content creators, selectors, and curators; developers; platform administrators; and anyone who solicits, maintains, stores, and will ultimately partner in preserving the organization’s content.

External stakeholders are those individuals who are not directly attached to an organization by employment, membership, studentship, or other such connection. These individuals and groups in some cases are impacted by the organization’s content but do not share in the responsibility for collecting it, hosting it, or preserving it. However, these external stakeholders are interested in and impacted by the longevity and future availability of this content. Such external stakeholders include but are not limited to end-users, user communities, readers, authors, librarians, members, students, historians, and any other interested parties who are sustained in any way by the organization’s content.