GVSU Libraries Digital Preservation Policy

Version 1.0, November 15th, 2016

TABLE OF CONTENTS

Section 1. Purpose

Grand Valley State University is committed to the long-term preservation of digital materials collected by the University Libraries (hereafter GVSU Libraries). This policy is being provided to help guide the Libraries' curatorial faculty/staff and inform our **designated community** of stakeholders (see Section 5. Designated Community).

The guidance set forth here applies primarily to acquired or created materials that are deemed of value and importance for continuity and to furthering the teaching and learning mission of the University. As such this policy aligns closely with the collection policies of GVSU's Special Collections & University Archives (hereafter SCUA) and of ScholarWorks@GVSU, the Libraries' institutional repository.

GVSU herein defines **digital preservation** as *a series of managed activities necessary to ensure continued access to digital materials for as long as necessary* (See Section 15. Glossary for further definitions). Because of the rapidly changing nature of digital preservation technologies this policy does not assume, describe, or recommend any specific software/hardware implementations or methodologies.

Section 2. Mandate

The work of digital preservation is an outgrowth of GVSU's dedication to providing students, faculty, staff, and scholars at-large with access to resources to enrich their teaching, learning, scholarly, creative, and administrative activities.

GVSU Libraries considers digital preservation an essential investment in its broader work of collection building and preservation of its institutional history. The Libraries recognize that successful on-going digital preservation requires a commitment to sustainability and thus endeavor to provide dedicated budgetary resources, management, and skilled faculty/staff to this work. In addition the Libraries commit to reviewing needs and adjusting support/priorities as appropriate on a routine and annual basis.

Section 3. Objectives

- Endeavor to preserve access to University-acquired digital materials deemed of value and importance for continuity and for furthering the teaching and learning mission of the University (see Section 4. Scope)
- Support the collecting missions and policies of the SCUA and ScholarWorks@GVSU
- Make use of reliable and proven digital preservation technologies
- Align with best practices and adopt standards in sensible, incremental, and sustainable ways
- Effectively allocate and scale resources/capacity in the area of digital preservation and manage change responsibly over time

- Support faculty/staff professional development and contribute to community developments in the area of digital preservation

- Communicate and advocate responsibly, effectively, and in a timely fashion with University administration and IT for any needed support and resources

- Learn from GVSU's designated community of stakeholders and adapt skills, support, and technologies over time to meet their changing user needs for access to preserved digital materials

- Leverage partnerships and collaborations to advance and enrich digital preservation

Section 4. Scope

This digital preservation policy applies primarily to digital materials (both **born-digital** and digitized) deemed of value and importance for continuity and to furthering the teaching and learning mission of the University, and for which GVSU Libraries has been granted permission to provide access.

4.1. In-Scope Materials

Priority for digital preservation services will be given to materials acquired or created by SCUA and ScholarWorks@GVSU. As resources permit and University or Libraries administration mandate, digital preservation services may be extended to support other partners, projects, or initiatives. Such extensions of service will typically be of a defined nature and subject to signed **memoranda of understanding** (hereafter MOUs) and project plans. Any new projects—or significant changes to existing projects—involving digital preservation work outside the normal scope and capacity of SCUA and ScholarWorks@GVSU will be reviewed by the Digital Collections Working Group (hereafter DCWG), and further by the GVSU Libraries Leadership Team and Dean of Libraries as appropriate.

4.2. Out-of-Scope Materials

This digital preservation policy does not currently apply to the management of digital materials that *are not expressly acquired by agreement* with, and deemed appropriate for long-term preservation by SCUA or ScholarWorks@GVSU—in coordination with the DCWG.

Materials which would be out of scope for the Libraries' digital preservation activities currently include: executable software, research data or digital research materials not shared via ScholarWorks@GVSU, third-party resources (databases, journals, e-books, etc.) licensed by GVSU Libraries, and University records which have not been explicitly deposited for archival preservation in collaboration with SCUA.

4.3. Supported Preservation Levels

GVSU Libraries endeavors at minimum to preserve access to any donated, acquired or created digital materials in their original file formats (this is known as **bit preservation**). The Libraries will then make a **"good-faith effort"** (but cannot guarantee) to provide on-going systematic access to those same digital materials in more preferred or widely supported file formats (this is known as **full preservation**). Accomplishing full preservation is always dependent on available resources, unpredictable changes in technologies, their availability, and policies on **disposition**.

See Section 7. Preservation Strategies below for more specific information on file format policies and preservation levels.

Section 5. Designated Community

GVSU Libraries strives to accomplish digital preservation on behalf of a well-defined set of stakeholders.

5.1. GVSU Faculty, Staff & Students

First and foremost are the faculty, staff, and students who interact with our preserved digital materials in the course of their teaching, learning, and scholarly/creative activity. Their immediate classroom and scholarly needs inform our selection criteria and drive our access solutions.

5.2. University Administration

Another important stakeholder group are the various administrative departments, offices, and units on campus that have established agreements and schedules for delivery of permanent electronic records to the University Archives for safekeeping and controlled access. Their business and institutional history needs also inform our selection criteria and access solutions.

5.3. External Partners & Researchers

Our external partners and scholars/researchers at-large may have unique use cases that require specially agreed upon levels of preservation and access support that are negotiated but ultimately set in accordance with this policy's scope, selection criteria, as well as available resources.

5.4. General Public

Finally, GVSU Libraries engages in digital preservation of cultural heritage and other University-created materials for the on-going benefit of both our regional community and the world. Through our supported access solutions we strive to make preserved digital materials available to the broader public.

Section 6. Selection Criteria

Digital materials will be selected for digital preservation with reference to the following criteria:

- ☐ Materials acquired or created by SCUA and ScholarWorks@GVSU in keeping with their collection and submission policies, which emphasize unique, exceptional, and GVSU-created materials
- ☐ Materials acquired by, or created in collaboration with, external stakeholders through signed MOUs and project plans
- ☐ Materials for which GVSU Libraries has been granted permission to perpetually store and provide access
- ☐ Materials that are able to be supported by available acquisition, preservation, and access technologies, as well as by faculty/staff curatorial expertise
- ☐ Materials that can be reasonably sustained by current and projected budgetary resources for digital preservation

NOTE: GVSU Libraries, primarily through the efforts and direction of the DCWG, reserve the right to weigh criteria and set limits, priorities, and timelines for applying preservation actions and strategies. All digital materials that meet the above criteria will not be given preservation treatment all at once. GVSU Libraries are committed to working collaboratively and constructively with our stakeholders and project partners to accomplish digital preservation over time.

Section 7. Preservation Strategies

For all digital materials that meet the proper scope and selection criteria for preservation, the following strategies will be applied:

7.1. Bit Preservation

GVSU Libraries endeavors to preserve digital materials in their original file formats by aiming to ensure that the individual bits and bytes that make up each acquired or created file remain stable and free of malicious software. Files are given digital signatures known as **checksums**, which are unique alphanumeric strings that enable curatorial faculty/staff to monitor the integrity of files over time. Any loss of quality over time will result in mismatched alphanumeric strings indicating the need for a repair. Files are grouped together with other related files/metadata, checked for known viruses (quarantining and investigating any infected files as necessary), and stored and backed up locally on GVSU network storage servers. GVSU Libraries also aim to store copies of all files and metadata at one or more geographically distant locations. These copies can serve as the basis for repairs. Care is taken to organize digital materials in logical ways and apply good file naming practices. Curatorial faculty/staff will also use tools to attempt to identify specific file formats and the applications used to render them, and record this information as technical metadata (see Section 9). This information can be potentially helpful for transforming files into different formats as needs warrant and time and technology permits (see Section 7.2 below).

7.2. Full Preservation

GVSU Libraries will make a "good-faith effort" to ensure that any bit-preserved files (see Section 7.1 above) can be made accessible in more preferred or widely supported file formats. This is known as **format migration** and can prove helpful when software previously used to render files becomes obsolete or is unavailable for our designated community. Where migration is possible for any given file format it is important to know that the transformation may still result in some loss of look, feel, and functionality of the original content.

To accomplish format migration curatorial faculty/staff actively:

- ☐ Test tools and technologies to assist with file format identification and migration
- ☐ Make use of registry databases of file format information
- ☐ Consult technology watch publications and services
- ☐ Engage with the broader digital preservation community to acquire viable strategies and solutions

Where any rendering software is unavailable due to obsolescence of the underlying computing environment (physical computer hardware and operating system), **emulators** may be available to help recreate the original environment and permit ongoing access. GVSU Libraries curatorial faculty/staff will make a "good faith effort" to support **emulation** as time, resources, technologies, and expertise permit.

7.3. Digital Forensics

Digital forensics involves using specialized software to copy the entire file system (contents and metadata) of a computer or related storage media with the objective of preserving, and in some cases analyzing, some or all of its contents—even content that may have been deleted at one time. GVSU Libraries is in the early stages of testing digital forensics approaches and developing proper procedures. All digital materials acquired through any digital forensic approaches would be subject to the currently supported bit preservation and any/all "good-faith effort" full preservation services.

7.4. Digital Media Preservation

GVSU Libraries in conjunction with SCUA will inventory and store any donated ageing digital media (Floppy Disks, Zip Disks, DigiCams, CDs, DVDs, USB, HDs, etc.) and make a "good-faith effort" to extract and preserve contents prior to reaching **digital obsolescence**. All digital materials acquired through any media extraction would be subject to the currently supported bit preservation and any/all "good-faith effort" full preservation services.

Section 8. Access & Use

GVSU Libraries are committed to digital preservation in an effort to ensure that all designated community stakeholders can reliably access resources of interest over time.

GVSU Libraries will make every effort to ensure that preserved digital collections are at a minimum discoverable. Access is also important but not all preserved resources will necessarily be available immediately or available online. There may be cases where some or all information related to digital material may be restricted/redacted, embargoed, or have other special rights-driven requirements for access. These will be managed through the available access technologies, collection and access policies, and metadata.

Finally, the work of preserving any given digital content at the present moment does not necessarily guarantee on-going future access—something altogether dependent upon sudden unanticipated changes in technologies, their availability, as well as policies on disposition.

GVSU Libraries strives to constantly evaluate its access systems to ensure ongoing functionality and a high degree of user satisfaction for our designated community.

Section 9. Metadata

Metadata is important for digital preservation because it can assist with the long-term curation of material. In addition to gathering together and storing any/all descriptive metadata (e.g., Dublin Core), OCR, administrative documentation, and readmes related to selected digital materials, GVSU Libraries also records a variety of technical metadata.

Types of information collected include checksums, virus reports, date created and last modified information, size and type of file, resolution, frame rate, color space, compression scheme, and creating application (where known). This information assists with current renderings and future transformations.

All metadata is stored in open and widely supported file formats, including XML, comma separated value (CSV), plain text or tab-delimited (TXT), as well as in document formats (e.g., DOC, DOCX, ODF & PDF).

Section 10. Roles & Responsibilities

10.1. Digital Collections Working Group (DCWG)

SCUA and the ScholarWorks@GVSU administrators work within and through the DCWG to coordinate on scope and selection criteria review for all digital materials, and collaborate on the development, review, and update of policies and project work plans.

10.2. Leadership Team & Dean of Libraries

The DCWG works closely with the GVSU Libraries Leadership Team and the Dean of Libraries to approve policies and strategies, and sign-off on preservation for any digital materials that may require MOUs and project plans.

10.3. Library Faculty & Staff

The Metadata & Digital Curation Librarian takes primary responsibility for applying all preservation strategies with support from the Library Specialist for Metadata & Digital Curation; and researches, tests, and implements technologies with support from the Libraries' Technology & Information Services (hereafter TIS).

Section 11. Collaborations

Digital preservation at GVSU Libraries enlists an array of collaborating partners, both across the University and beyond. These partnerships are essential to the program's success.

11.1. Digital Collections Working Group (DCWG)

The DCWG is first and foremost a collaboration between SCUA, TIS, and the Libraries' Research & Instruction division (particularly the ScholarWorks@GVSU administrators). This working group provides direction and oversight for the digital preservation program.

11.2. Library TIS & GVSU IT

Digital preservation within GVSU Libraries could also not be accomplished without the close collaborations between TIS and the University's Information Technology (hereafter IT) Services. GVSU's IT Services provides digital storage for onsite backup and other services to enable partnerships with faculty and projects on campus.

11.3. Third-Party Vendors

Digital preservation at GVSU Libraries is also dependent on a variety of third party technology vendors to provide both computing resources and distributed off-site dark storage.

11.4. Open Source Community

Finally, where possible GVSU Libraries leverages the benefits of **open source software** and considers collaboration with the broader development community of high value and importance.

Section 12. Sustainability

Sustainability of GVSU Libraries' digital preservation program requires fiscal responsibility and on-going support for technology improvement, professional development for curatorial faculty/staff, as well as preservation planning.

12.1. Technical & Fiscal Sustainability

GVSU Libraries has a dedicated budget item for digital collections that includes support for digital preservation—particularly for needed software.

12.2. Professional Development

All Library faculty/staff are generously supported in their professional development—primarily through attendance at conferences and webinars, as well as on-campus continuing education, and other distance learning. This helps to boost capacity for improving digital preservation infrastructure, technologies, and practices over time.

12.3. Preservation Planning

Preservation planning is handled through regular meetings of the DCWG and benefits from that group's proactive assessment of the needs of the University and Libraries' designated community of stakeholders.

Section 13. Security

GVSU Libraries performs bit-level integrity checking of all digital materials that meet scope and selection criteria. This is to detect file corruption and perform repairs from good copies. Checking is currently performed using md5 and sha256 hash functions. Algorithms and tools used to perform integrity checks may change over time or as specific content types warrant. Schedules for auditing will vary somewhat as collections grow and storage solutions and audit tools change. GVSU will generally aim to follow prevailing practice to audit collections on a quarterly and annual basis.

Virus checks of all stored digital materials are also performed, with quarantined content receiving follow-up investigation.

GVSU Libraries pursues geographical distribution and redundancy. Currently this includes one local on-site backup (administered by IT) of all digital materials packaged for long-term preservation, and one remote geographically distant cloud-hosted copy of all digital materials. Secure access is managed via access controls and policies set by GVSU IT and those of any selected third-party vendor(s) that may be in use.

GVSU Libraries curatorial faculty/staff play a role in granting access to other support staff and follow best practices and University policy for account and password security.

Section 14. Challenges

Digital preservation is never a finished project and there is always room to expand and grow services and expertise. The broader library and archives community is also always addressing shared challenges through the advancement of standards, technologies, and new methodologies—on both practical and

organizational levels. GVSU Libraries seeks to engage these common efforts and overcome its own unique challenges over time.

Examples of near-term challenges under active research and development within the Libraries include:

- Balancing adoption and use of standards in-line with available resources, technologies, and expertise
- Scaling and streamlining of computing resources, storage, and services to preserve larger aggregations of digital material
- Outreach and awareness raising amongst donors/creators regarding obsolescence and preferred file formats
- Improved strategies for audit of remote copies of stored digital materials
- Support for replication and redundancy of preservation strategies and services both on campus and with any contracted third party services
- Robust, granular, and flexible integrity reporting mechanisms
- **Versioning** and **de-duplication**
- Appropriate use of compression
- Syncing and repair (healing) of geographically distributed copies of digital materials
- Acquisition of born-digital materials on ageing media
- Redaction of **personally identifiable information (PII)**
- Discontinuation of preservation support (i.e., disposition policies)

Section 15. Glossary of Terms

Bit Preservation: A term used to denote a very basic level of preservation of digital resource as it was submitted (literally preservation of the **bits** forming a digital resource). It may include maintaining onsite and offsite backup copies, virus checking, fixity-checking, and periodic refreshment to new storage media. Bit preservation is not digital preservation but it does provide one building block for the more complete set of digital preservation practices and processes that ensure the survival of digital content and also its usability, display, context and interpretation over time.

Born-Digital: Digital materials which are not intended to have an analogue equivalent, either as the originating source or as a result of conversion to analogue form.

Checksum: A unique alphanumerical signature derived from a file. Used to compare copies.

De-Duplication: Refers to techniques for eliminating duplicate copies of repeating data.
Designated Community: An identified group of potential consumers who should be able to understand a particular set of information from an archive. These consumers may consist of multiple communities, are designated by the archive, and may change over time.

Digital Forensics: Digital forensics is the process of uncovering and interpreting electronic data. The goal of the process is to preserve any evidence in its most original form while performing a structured investigation by collecting, identifying and validating the digital information for the purpose of reconstructing past events.

Digital Obsolescence: A situation where a digital resource is no longer readable because of its archaic format: the physical media, the reader (required to read the media), the hardware, or the software that runs on it is no longer available.

Digital Preservation: A series of managed activities necessary to ensure continued access to digital materials for as long as necessary.

Disposition: The final action that puts into effect the results of an appraisal decision for a series of records. Transfer to an archival institution, transfer to a records center, and destruction are among possible dispositions.

Emulation: A means of overcoming technological obsolescence of hardware and software by developing techniques for imitating obsolete systems on future generations of computers.

Emulator: In computing, an emulator is hardware or software that enables one computer system (called the host) to behave like another computer system (called the guest).

Format Migration: A means of overcoming technological obsolescence by transferring digital resources 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 clients 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 resource with the new generation of technology.

Full Preservation: Defined here as the use of format migration, emulation, digital forensics, and other strategies to ensure that the content of digital materials, rather than just the original bits and bytes, remain protected and accessible over time despite technological obsolescence and the need for refreshed storage media.

Good Faith Effort: Though this policy does not represent a contract with any parties "good faith effort" is defined for the purposes of this policy, as "what a reasonable person would determine is a diligent and honest effort under the same set of facts or circumstances."

Memoranda of Understanding (MOU): A memorandum of understanding (MOU or MoU) is a formal agreement between two or more parties. Companies and organizations can use MOUs to establish official partnerships. MOUs are not legally binding but they carry a degree of seriousness and mutual respect.

Open Source Software: Open-source software (OSS) is computer software with its source code made available with a license in which the copyright holder provides the rights to study, change, and distribute the software to anyone and for any purpose.

Personally Identifiable Information (PII): Information that can be used on its own or with other information to identify, contact, or locate a single person, or to identify an individual in context.

Versioning: Refers to saving and tracking in systematic ways new copies of your files when you make changes so that you can go back and retrieve specific versions of your files later and distinguish authoritative copies.

Section 16. Related Policies & Documentation

- ScholarWorks@GVSU Collection Policy
- GVSU Special Collections & University Archives Collections Policy
- Grand Valley State University Information Technology Policies & Procedures

Section 17. References

Digital Preservation Coalition, *Digital Preservation Handbook* (revised 2nd edition), available at: http://handbook.dpconline.org/

Library of Congress, Analysis of Current Digital Preservation Policies: Archives, Libraries, and Museums (2013), available at: http://blogs.loc.gov/thesignal/2013/08/analysis-of-current-digital-preservation-policies-archives-libraries-and-museums/

Scalable Preservation Environments (SCAPE) Project, *SCAPE Catalogue of Preservation Policy Elements*, available at: http://wiki.opf-labs.org/display/SP/Catalogue+of+Preservation+Policy+Elements

Section 18. Authors & Reviewers

Author

Matt Schultz, Metadata & Digital Curation Librarian

Reviewers

Lee Van Orsdel, Dean of Libraries
University Libraries Leadership Team
Kyle Felker, Digital Initiatives Librarian
Annie Benefiel, Archivist for Collection Management
Matt Ruen, Scholarly Communications Outreach Coordinator

Section 19. Date/Revisions

Version 1.0

Reviewed and approved by GVSU Libraries Dean of Libraries and Library Leadership Team on November 15, 2016