Digital Preservation Handbook

Hillary Ellis, Director of Conservation

Beth Merkle, Director of Library & Archives

Julia Novakovic, Archivist

Andrew Borman, Digital Games Curator

Victoria Gray, Collections Manager

Edited by Chris Bensch, JP Dyson, Racquel Gonzales, and Jeremy Saucier

digitalcollections@museumofplay.org

Rev. 1/2020

Contents

Preface

In 2016, the Collections Team formed a Digital Preservation Working Group to discuss the challenges of access and preservation of media in the Strong museum's collections.

In 2017, The Strong undertook a systematic collections and archives assessment to identify, quantify, and make plans to care for endangered media, with the goal of eventually being able to perform digital conversion and preservation of content at the point of acquisition. Digital conversion and preservation are crucial to ensure future access.

The Brian Sutton-Smith Library and Archives of Play contains more than 5,200 individual components that the collections team (conservator, library staff, and curators) identified as endangered media formats. These include floppy disks, CD-ROMs, DVDs, audiocassettes, audio CDs, VHS tapes, Betacam tapes, U-matic tapes, and other magnetic data and audiovisual formats.

The International Center for the History of Electronic Games (ICHEG) collection holds more than 45,000 individual components. These include floppy disks, data cassettes, ROM cartridges, solid-state cartridges, CD-ROMs, DVDs, Blu-Ray disks, and other unique card formats. Within the video games collections, The Strong has identified 25,000 artifacts that are imperiled, including all games created before 2005, which reside on endangered media formats considered to be at-risk of immediate loss of data, requiring data migration to more stable formats.

As the collection of video games and related archival materials continues to grow in numbers at The Strong, the museum must adapt to collect, display, and preserve video game media and digital assets with the same expertise and confidence as toys, games, art, housewares, and ephemera related to play. For this reason, it is important to understand the unique challenges of digital preservation. This document provides strategies, digitization methodology, and decision-making processes involved in digital preservation. The Director of Conservation, Director of Library and Archives, Archivist, Digital Games Curator, Collections Manager, and Arcade Game Conservation Technician have written the Digital Preservation Handbook to communicate the preservation actions each of these team members performs. Given that digital preservation requires additional technology and digital asset management, the Director of Information Technology provides the Team additional digital storage and IT support.

The Digital Preservation Handbook is a working document, intended to respond to changes in technologies available and adapt to current best practices in the library, archives, museum, and game communities. The authors welcome comments to digitalcollections@museumofplay.org.

Measuring Institutional Progress in Digital Preservation

Project-Based Key Performance Indicators (KPI)

The model approach to digital preservation projects is listed in six steps:

1. Schedule and hold a start-of-project meeting with the Digital Preservation Working Group and leadership to share Objective Key Results (OKR).
2. Identify total number of digital media or electronic games in the project.
3. Identify rough amount of digital storage required for these objects.
4. Evaluate the objects by RAVE standards of Rare, At-Risk, Valuable, and Engaging. Prioritize a quantity of objects to digitize.
5. Estimate the digital storage required. Estimate a five-year total cost of ownership, including backup equipment and conservation needs.
6. Schedule and hold an end-of-project meeting with leadership to present your findings and Objective Key Results (OKR).

Examples of Project-Based KPI

The Digital Preservation Working Group will report project statistics in a variety of ways. Each example below is a measure of the success of a digitization project in the library, archives, video and digital game, or general museum collection.

Library Example

Percentage of catalogs scanned and saved, percentage of scanned materials made searchable by Optical Character Recognition (OCR), and where and how images are stored according to National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation.

Archives Example

Percentage of a media collection digitized or preserved in a pilot project, and where and how it is stored according to NDSA Levels of Digital Preservation.

ICHEG Example

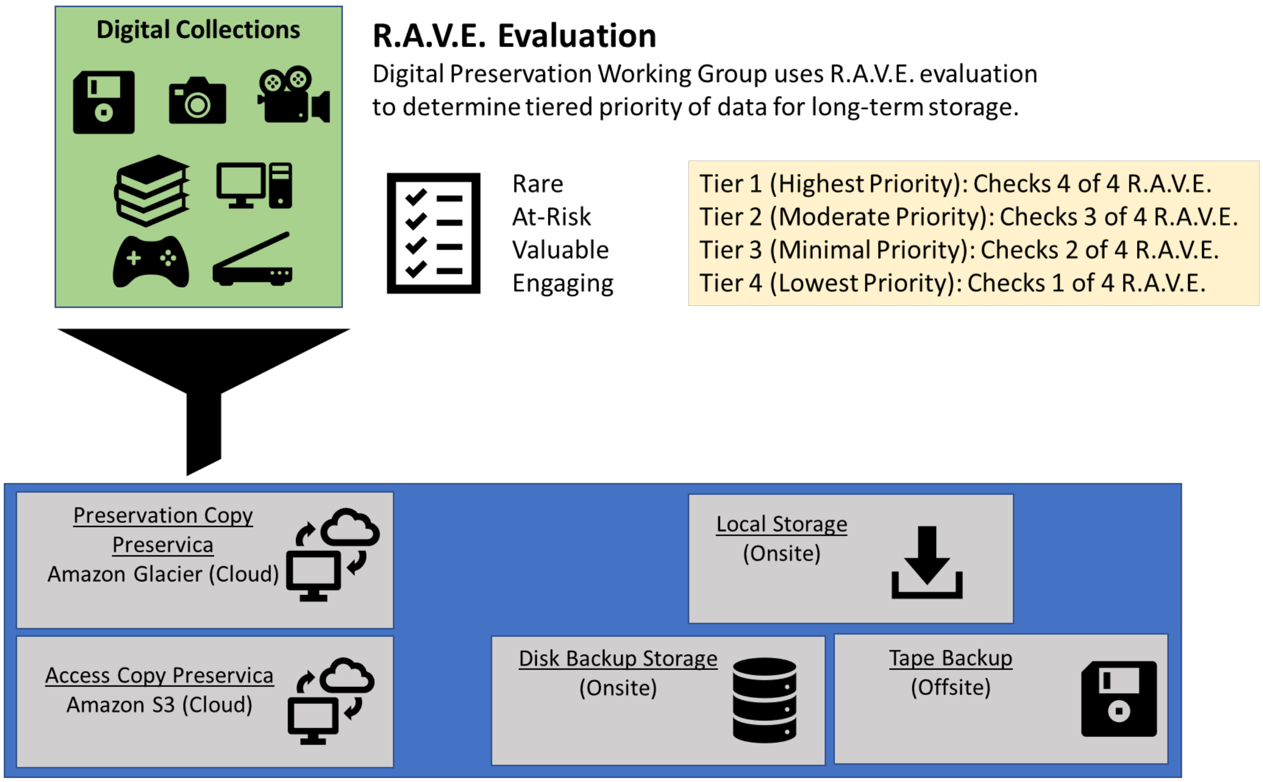
A semi-annual game census quantifying how many games are functional and supported by existing equipment, and review acquisition of new equipment to enable read or play access for remaining games. Evaluation of where and how game saves are stored according to NDSA Levels of Digital Preservation.

Museum Objects Example

Number of photos for Online Collections as a result of acquisitions processing and how many are saved in Argus collections database, Google Images, or other searchable databases.

Workflow of Digitization and Storage

Whether it is a demand created by the acquisitions process, preservation and access grants, or image requests for marketing, publication, exhibits, or researcher use, digitization and storage of digital files is a necessity. Each of these demands on the collection will lead to the creation of new digital files. The files created are evaluated and stored. The flow chart below is a graphical representation of the path of digital collections through evaluation and selection of storage.



Evaluation by RAVE Digitization Criteria

The Strong collects and preserves historic objects, primary source documents, monograph and serial publications, and other materials that reflect and document the importance of play in cultural history. As resources allow, The Strong implements selected best practices from the communities of archivists, conservators, and other museum-allied professions to preserve digital and virtual collections. These preservation efforts make both the digital collections and traditional archival and artifact collections accessible digitally in accordance with rapidly changing user needs. In addition to the objects and digitized materials themselves, The Strong supplies descriptions and other interpretive metadata that places materials in context and gives them meaning.

Rare: One-of-a-kind items or items not owned by other institutions. (Examples: trade catalogs and trade sheets, design documents, source code, game builds, iconic documents and images from the archives)

At-risk: Items threatened by deterioration or other preservation issues. (Examples: source code, born-digital files created or saved in proprietary formats from the late 20th century, items that should not be handled due to condition, scrapbooks, other items on acidic paper or coated with adhesives)

Valuable: Items that document important historical events, social trends, or are otherwise culturally significant. (Examples: *Playthings* magazine, electronic gaming magazines such as *Nintendo Power)*

Engaging: Items that are visually compelling, fun, recognizable, or iconic. (Examples: electronic gaming magazines, archival collections, and videogame play capture)

Levels of Digital Preservation

The storage of digital assets is evaluated according to the metric developed in the 2019 Levels of Digital Preservation established by the National Digital Stewardship Alliance. The overall structure of the chart is progressive across the preservation areas of location, integrity, security, metadata, and formats—the Level 1 actions themselves are the most pressing activities to accomplish first or can function as necessary prerequisites for those in Level 2, Level 3, or Level 4. Broadly speaking, as a digital object's preservation storage moves up from Level 1 to Level 4 in each area of preservation, the preservation level goes from the basic need to ensure bit preservation towards broader requirements for keeping track of digital content to ensure that it can be made available over longer periods of time.

Levels of Digital Preservation Chart

	Level One	Level Two	Level Three	Level Four
Storage and Geographic Location	Two complete copies that are not collocated. For data on heterogeneous media (optical discs, hard drives, etc.), get the content off the medium and into your storage system.	At least three complete copies. At least one copy in a different geographic location. Document your storage system(s) and what media you need to use them.	At least one copy in a geographic location with a different disaster threat. Obsolescence monitoring process for your storage system(s) and media.	At least three geographic locations with different disaster threats. Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems.
File Fixity and Data Integrity	Check file fixity on ingest if it has been provided with the content. Create file fixity into if it wasn't provided with the content.	Check file fixity on all ingests. Use write-blockers when working with original media. Virus-check high-risk content.	Check fixity of content at fixed intervals. Maintain logs of fixity info; supply audit on demand. Ability to detect corrupt data. Virus-check all content.	Check fixity of all content in response to specific events or activities. Ability to replace/repair corrupted data. Ensure no one person has write access to all copies.
Information Security	Identify who has read, write, move, and delete authorization for individual files. Restrict who has those authorizations to individual files.	Document access restrictions for content.	Maintain logs of who performed what actions on files, including deletions and preservation actions.	Perform audit of logs.
Metadata	Inventory of content and its storage location. Ensure backup and non-collocation of inventory.	Store administrative metadata. Store transformative metadata and log events.	Store standard technical and descriptive metadata	Store standard preservation metadata
File Formats	When you can give input into the creation of digital files, encourage use of a limited set of known open formats and codecs	Inventory of file formats in use	Monitor file format obsolescence issues	Perform format migrations, emulation, and similar activities as needed

(Source: https://ndsa.org/activities/levels-of-digital-preservation/)

Overview of Data Storage

DAMP

The Digital Asset Management Team was formed in 2015 in order to develop a Digital Asset Management Plan (DAMP). See the Digital Asset Management Plan for more information. The Digital Asset Management Team is comprised of members of the Collections and Information Technology Teams. The Strong's Digital Asset Management Plan centralizes information regarding the acquisition, production, storage, description, access, and preservation of the museum's digital assets. Digital asset management and preservation are rapidly changing fields; the DAMP is updated annually to account for evolving technology standards and best practices. In 2017, the Digital Asset Management Team selected Preservica as the museum's Digital Asset Management System (DAMS). Preservica performs all operations on the POWRR Tool Grid checklist https://digitalpowrr.niu.edu/digital-preservation-101/tool-grid/ and stores digital assets in both Amazon S3 and Glacier cloud-based storage. The addition of Preservica gave the digital asset management three tiers of data storage for digital collections and a fourth tier for low-priority digital files.

Digital POWRR Tool Grid Checklist: Digital Asset Management

	digitalpowrr.niu.edu	**Preservica Amazon Glacier**	**Preservica Amazon S3**	**Local Storage with Tape Backup**
Ingest	Copy	x	x	x
	Fixity Check	x	x	
	Virus Scan	x	x	x
	File Dedupe	x	x	
	Auto Unique ID	x	x	
Processing	Auto Metadata Creation	x	x	
	Auto Metadata Harvest	x	x	
	Manual Metadata	x	x	x
	Rights Management	x	x	
	Package Metadata	x	x	
	Auto SIP Creation	x	x	
Access	Public Interface	x	x	
	Auto DIP Creation	x	x	
Storage	Auto AIP Creation	x	x	
	Reliable, Long-Term Bit Preservation	x	x	
	Redundancy	x	x	x
	Geographically Dispersed Data Storage Model	x	x	
	Exit Strategy	x	x	x
Maintenance	Migration	x	x	
	Monitoring	x	x	x
	Auto Recovery	x	x	
Other	Open Source			x
	Clear Documentation	x	x	
	Cost	Varies	Varies	Varies

Library Collections

Principal Statement

The library collects and preserves published materials in print, audiovisual, or digital formats. Decisions to digitize physical materials are made according to RAVE standards.

Physical to Digital: Library Materials

In 2018, the library was awarded an IMLS Museums for America Collections Stewardship grant to catalog 5,000 catalogs and digitize 2,500 trade catalogs. Catalogs published prior to 1960 and the most frequently used catalogs were digitized. Digitization began in June 2019, using the Bookeye A1 scanner, and was completed in November 2019.

Oversize Materials, including serials and trade catalogs, can be scanned using the Widetek or Bookeye A1 scanners.

Preservation of Born-Digital Library Materials

The library preserves born-digital annual reports, press kits, press releases, dissertations, and trade catalogs in a designated folder on the Library drive. File formats include PDF, JPG, Excel, and Word.

Limitations of Digital Preservation

Scanning physical materials takes time and effort from a staff member. The creation of metadata for scanned materials should also occur at the same time as digitization, which also involves staff time to analyze and produce accurate and adequate metadata.

Digital storage is still costly. As we determine tiers of digital storage, we are required to maintain copies that may seem redundant but are part of best practices.

Chronology of Technology used for Library Preservation

Dates	Technology Capability	Purpose
2009 – present	Nikon Slide Scanner	Creates TIFF files from standard sized slides
01/2014 – present	Plustek OpticBook A300 Scanner	Flatbed scanner with book edge, scans up to 11” x 17” (general scanning use)
01/2018 – present	Widetek Scanner	Up to 36” wide scans, unlimited length (oversized paper/flat materials)
11/2017 – present	ImageBox MF 4-in-1 Converter	Converts film negatives to positives via scan, also digitizes slides and photos
02/2018 – present	Kryoflux and floppy disk conversion station	Digitize and read information from 3.5” and 5.25” floppy disks in the archives.
02/2019 – present	U-matic player and conversion station	Digital capture of analog U-matic videocassettes in the archives.
03/2019 – present	Bookeye V1A Professional Scanner	High-speed, hi-res, books and bound volumes, flat materials up to 23” x 33”
06/2019 – present	Beta player	Play and convert Beta tapes

Archival Collections

Introduction to archival collections

The archival collections in the Brian Sutton-Smith Library and Archives of Play at The Strong fall into three distinct categories: collections related to the study of play, collections related to artifacts of play, and collections related to the history of electronic and video games. These collections contain both physical and digital materials. Our goal is to make these documents, photographs, audiovisual files, and digital files available to researchers, scholars, museum staff, and the public.

Digitization criteria

Physical materials in the archives are digitized for preservation and access. These include paper, photographs, audiovisual media, and magnetic media. Selection of archival materials for digitization falls under our RAVE criteria:

Rare: One-of-a-kind items or items not owned by other institutions. (Examples from the archives: game design documents, source code, game builds, photographs, non-published audiovisual media)

At-risk: Items threatened by deterioration or other preservation issues. (Examples from the archives: source code, born-digital files created or saved in proprietary formats from the later 20th century, items that should not be handled due to condition, scrapbooks, other items on acidic paper or coated with adhesives)

Valuable: Items that document important historical events, social trends, or are otherwise culturally significant. (Examples from the archives: arcade fliers, game design documents, marketing materials, photographs, audiovisual media)

Engaging: Items that are visually compelling, fun, recognizable, or iconic. (Examples from the archives: toy and game design documentation, photographs, audiovisual media)

As museum staff developed these values, the need to address our “at-risk” materials in the archives became most urgent. The archival collections, by definition, are already considered rare and valuable based on their significance to the museum collection and mission, and a large quantity of the collections’ contents (especially from graphic or audiovisual formats) are publicly engaging.

Key Digitization Projects

New York Heritage

Upload facilitated by Rochester Regional Library Council (RRLC), 2015-2016

Selected items from two archival collections are available on New York Heritage: the Atari Coin-Op Division Corporate Records, 1969-2002, and the Margaret Woodbury Strong Papers, 1897-1969. These online collections – the images and metadata – were created through the work of interns, under the supervision of the Director of Libraries and the Archivist. The Rochester Regional Library Council (RRLC) facilitated the upload to New York Heritage site.

Endangered Media Assessment

Performed in 2016

For the past several years, the Archivist and Director of Conservation have focused on what we call "endangered media formats" – time-based storage that requires specific hardware or players to access the information. These formats include floppy disks, CD-ROMs, DVDs, audiocassettes, VHS tapes, Betacam tapes, U-Matic tapes, and other magnetic data and audiovisual formats. In 2016, the Archivist conducted a survey of the endangered media formats within existing archive collections. At the time, 41% of the museum's fully processed archival collections contained at least one component considered to be endangered. Without the proper hardware and/or software to read the media, the information contained on such formats remained inaccessible and unknowable.

"Preservation of Endangered Media Pilot Project" (Floppy disks)

Project funded by Rochester Regional Library Council (RRLC) Technology Grant, 2017-2018

The Strong developed a pilot project to read a sampling of disks across several different archival collections: our goal was to test 100 floppy disks out of all the 3.5" and 5.25" floppy disks identified as endangered media. By performing this task, we aimed to digitize a diversity of material with varying condition issues. This would provide a better representation of the percentage of viability in the assumed obsolete media collection. It also served as a pilot project for treatment and an approach that could be applied to the other unlabeled disks in archival collections, as we moved beyond the pilot project to digitize more of the endangered archival media. This grant supplied the necessary funds to purchase equipment (such as two Kryoflux floppy controllers and associated licenses, 3.5" floppy disk drives, 5.25" floppy disk drives, a standalone workstation, and software) and supported a part-time Digital Conversion Technician. Within a three-month window, we imaged 1,756 floppy disks with a 99% success rate of reading the disks. Museum staff presented our findings at several area conferences and provided museum researchers with rare information imaged from these floppy disks.

"Preservation of Endangered Media: U-Matic Tape Pilot Project" (U-matic tapes)

Project funded by Rochester Regional Library Council (RRLC) Technology Grant, 2018-2019

Based on our previous success with our floppy disk project, we developed a similar model to address U-matic tapes in the archival collections. During the application process, we determined that the data generated during this project should be capped at 2,000 minutes of video, or approximately 2 TB. Focusing on tapes from the Atari Coin-Op Division corporate records and the Ken Fedesna papers, the Archivist selected 120 U-matic tapes based on their labels. With this grant, The Strong was able to acquire necessary equipment (including a 9000 series U-matic player, CRT monitor, audio monitor, an a/v rack, software) and pay a part-time Digital Conversion Technician. Within a three-month window, more than 120 U-matic tapes were digitized, with associated metadata captured and ingested into Preservica.

Accessible Archival Media Formats

As of April 2019, the archivist had identified 6,464 components in processed archival collections as endangered media formats; as a result of the 2017-2019 grant-funded hardware/software acquisitions, fewer than 300 components remain inaccessible. We may continue to build this capacity with additional equipment, though certain low-quantity formats in our collections (e.g., Video8, Mini-DV) may be sent to outside vendors.

Born-digital and digital surrogate archival materials are being ingested into Preservica as of Fall 2019. This will allow the creation of virtual collections to be accessed by the public via Preservica's Universal Access module.

Making Physical into Digital

When digitizing physical materials for preservation, our goal is to capture the best quality and preferred file format the first time, which avoids excess wear and tear on the object and/or duplication of effort later. Our preferred file formats for digital preservation are as follows:

- Text files: PDF or TIFF
- Photographs, drawings, and still images: TIFF or JPEG2000 (for preservation); JPEG (for access)
- Audio: AIFF or WAV (for preservation); MP3 (for access)
- Video: AVI or MOV (for preservation); MPEG or MP4 (for access)

In-house equipment for physical-to-digital conversion includes scanners and audio and video conversion equipment. If conversion is needed for a media format (such as film) not covered by on-site equipment, there are local vendors available.

Providing reference materials to long-distance researchers usually involves creating and sending grayscale PDFs. If a folder or item is particularly fragile, ideally staff would create a preservation copy of the materials and then manipulate a copy of the resulting file for researcher use (e.g., convert to grayscale, optimize PDF, etc.)

Born-Digital Preservation

When preserving born-digital materials, it may be necessary to stabilize file formats and migrate files to the most up-to-date version that is computer-readable. As a rule, platform-independent, non-proprietary, well-supported open formats are preferred. Encrypted or password-protected files are difficult to fully preserve as are files with unembedded proprietary fonts. Our preferred file formats for common born-digital materials are:

- Text files: .PDF or .TXT
- Raster images: .TIFF or .JP2 (for preservation); .JPEG (for access)
- Audio: .AIFF or .WAV (for preservation); .MP3 (for access)
- Video: .AVI or .MOV (for preservation); .MPEG or .MP4 (for access)
- Spreadsheets or databases: .CSV or .TXT
- Computer programs: source code (for preservation); executable files (for access)

Preservica, the museum's Digital Asset Management System, is equipped to digitally preserve more than 1,200 different original file types. Stabilizing the formats for migration and access is still recommended for on-site use.

Processing digital archival collections also requires arrangement and description needed for a physical collection. It may be necessary to drill down to the item-level in order to fully comprehend the contents and context of a digital collection.

Limitations of Archival Digital Preservation

- Existing equipment:
  - May require access to both Windows and Mac workstations for present-day digital materials
  - May not be able to handle certain media formats on site; could require outside vendor collaboration
- Staff time:
  - Process of scanning physical materials requires additional staff time and effort
  - Process of digitizing audiovisual materials also involves a significant amount of staff time and effort, as digitization of media takes place in real time
  - Storage and migration of born-digital files requires significant tech-savviness and staff involvement
  - Downloading and copying digital files, as well as ingesting files into data asset management systems, requires significant time
  - Providing controlled researcher access to digital collections may be time-consuming
- Cost:
  - Digital storage remains costly
  - Proper digital asset management systems may be costly
  - Purchasing appropriate equipment, especially for obsolete formats, can be costly due to increasing scarcity

General Museum Collection

The Strong photographs museum artifacts in our in-house studio, using gray laminate seamless and tungsten lamps for optimum clarity and color balance. The resulting images are intended not only for internal reference documenting the physical appearance and condition of each artifact, but also for inclusion in online collections. Images from online collections provide a visual understanding of artifacts to users worldwide and are often requested for inclusion in publications and exhibitions in the museum.

Physical to Digital

Museum Objects— Images are captured as RAW files, imported to a local workstation and edited from CR2 files to 300 dpi TIFF images, renamed in accordance with the object ID number.

Editing of images is kept to a minimum, limited to cropping, color balance/correction, and clarity, to maintain the integrity of the object and accurately portray its condition. The final, edited, image files are linked to their corresponded Argus.net object record and will also appear in online collections.

Photos used for press/PR purposes, such as those for National Toy Hall of Fame, World Video Game Hall of Fame, or newsletter photos, are also shot in studio as RAW image files.

Online Collections

Once a web category and image files have been linked in Argus.net, artifact records qualify for upload to online collections.

Photographic Scanning

Two dimensional artifacts such as photographs, postcards, and game instructions are scanned in-house at a resolution of 300dpi or greater with the appropriate Epson scanners. Images are renamed and linked to the Argus.net artifact record following the same process for three dimensional artifacts outlined above.

Documentation

In 2018, The Strong began recording non-image based digital documentation for artifact records, loans, and conservation reports to link all pertinent information directly to each artifact record. As a result, all members of the collections team can easily access detailed information as opposed to the laborious sorting through hundreds of storage files to locate a singular document. This practice facilitates more thorough tracking of the provenance and condition of an object and provides additional copies of important records and details.

- **Object Records** include a variety of linked documents including Conservation Treatment Forms which provide a detailed account of what physical treatment(s) an artifact has undergone. Other possible linked documents may include instructions, manuals, or other contextual items that relate to the care, history, or provenance of the artifact. Under the documents tab, an image may be linked to provide greater context for the artifact, as opposed to visual documentation for online collections. One such example is a photograph of a donor with a sculpture he commissioned.

- **Party Records** include any information relevant to the donor/vendor. For instance, the obituary for a donor and prominent game designer is linked in the party record because it provides an overview of the game designer's life and career.

- **Activity Records** include any information related to a specific group of items acquired

- **Deed of Gift** documents provide the legal transfer of ownership for object(s) to The Strong once signed by the donor. This paperwork will also outline donor usage restrictions for items with intellectual property.

- **Incoming / Outgoing Loans** all receive individual activity numbers with pertinent information linked to the record, including (but not limited to) a copy of the loan agreement, proof of insurance coverage, shipping details, or any other relevant information.

Chronology of Technology Used for General Collections

Dates	Technology Capability	Purpose
1973 – 1989	Paper records	Curators wrote object record information on paper then sent these documents to University of Rochester for transcription (completed by work-study students in a 12AM-8AM shift). The transcription allowed data to be stored on servers at the UR, but much of the detail was lost or entered incorrectly during this process, resulting in hard to decipher skeletal records.
1980 – 1999	35mm Slides and Contact Prints	Contact prints from negatives were used for visual reference for artifacts. Slides were used for the most significant artifacts.
1980 – 1992	CATCLL	The Strong utilized an internal database CATCLL (short for Catalog of the Collection) to catalog artifacts and create object records.
1992	Collection Management Database Software	Major data conversion to the museum's first iteration of the Argus collections database. Although this software provided the opportunity to include more detailed information about each artifact, most records remained skeletal during the migration process.
2000 - Present	Digital Photography	Digital photography was implemented for collection artifacts. Photos were JPG quality, ranging from 72-180dpi over the years immediately following 2000.
2010 – Present	Digital Single-Lens Reflex (DSLR) camera and studio equipment	Digital SLR camera and updated studio equipment (lights, seamless, scanner) were acquired as part of an IMLS grant project focused on increasing the online presence for collection artifacts that pre-date 1930.
2015 – Present	Argus.net	Argus.net migration was completed in 2015. A browser-based, updated version of the museum's long-term database, the system can accommodate larger image files linked directly to artifact records, along with PDF's, JPGS, and other file types.
2018 – Present	Digital Single-Lens Reflex (DSLR) camera	Camera with greater capabilities was acquired to ensure high-quality images that can be shared through online collections, media/PR, and image requests for inclusion in publications, exhibits, etc.

Arcade Game Collection

Preserving Arcade Games

For general information on how arcade games are maintained and made accessible, see the ICHEG Conservation Policy. The summary included here describes digital components of arcade games that the museum staff preserves with backup copies.

Preserving a Hard Drive

When beginning the process to back up a hard drive image, care is taken to prevent any modification of the original drive. Any modifications could prevent the drive from functioning in the original machine. The use of a software or hardware write-blocker prevents the operating system from making any changes to the original drive during the copying process.

A sector-by-sector copy of the hard drive is made because it ignores the filesystem of the drive. This allows for a more accurate copy of the drive. The copying process generates a backup image that is the same size as the source drive, rather than just the storage space currently being used. Depending on the size of the drive, this process may take a long time. Do not initialize the drive when using Windows. A program such as HDD Raw Copy Tool on Windows will work but may modify the image with a header. Using DD in a Linux terminal will also create an accurate dump.

Preserving a ROM

An EPROM reader is used to backup ROMs on an arcade board. ROM dumps are labeled by location or label on the chip itself. The location may be listed on the motherboard itself. Selecting the proper chip type in the ROM reader is important, as improper selection could permanently damage the chip. Anti-static devices are used to limit potential physical damage to the chip.

Limitations of Arcade Preservation

Many arcade boards contain custom chips and devices that are not supported by a generic ROM reader. Security chips and other chips may contain important information needed to accurately preserve the game. Special tools may be available to read some of these chips, but it may require that the chip is de-soldered from the motherboard, which could lead to damage. All arcade games are treated on a case-by-case basis as it is impossible to make note of all hardware configurations.

Modern arcade games may also have network capabilities. No games in the current collection require a network connection to function, but additional features may be missing if the services cease.

Electronic Game Collection

Physical to Digital: Electronic Games

Digitization of electronic games on physical media ensures that a data backup will exist regardless of degradation of the original physical media, which occurs regardless of use. In some cases, as with floppy disks, the act of playing can modify or edit the original media, altering its condition. The migrated copy can become the means of access, using floppy drive emulators, flash carts, or emulation. The original media will no longer be used, preserving its original state.

When new electronic games are acquired, they are assessed using RAVE criteria. Games designated as at-risk due to their physical media should be prioritized for digitization. Prior to digitization, the media should be assessed for physical damage that may prevent proper digitization, including corrosion, mold, dust, and dirt.

New hardware devices can migrate data from its original console format to a digital format. When selecting new preservation hardware, we refer to outside video game preservation groups. The objective is to adapt to the best methods available and avoid using any destructive methods to achieve data migration.

Once a game is migrated, the object record is edited in the collections database to indicate a storage location and data asset archive. Many games have multiple revisions, which results in different data, so it is important to only link data for the specific object that is dumped.

Video Capture for Preservation

Captured video is incredibly important to preservation. It provides a record for how a game functioned using original hardware, which may cease to function in the future. This footage can be used to study the gameplay, to improve other preservation methods such as emulation, or to capture elements of games that may not exist in the future, due to their limited-time availability or the lack of networking servers.

It also is important to migrate other forms of audio-visual formats such as cassette tapes, VHS, U-matic video, and laserdiscs. Analog formats are endangered formats as defined by the RAVE standards. As time passes, playback equipment obsolescence will make it more difficult to use these formats. Digitization is needed to capture video game content and audiovisual media.

For preservation, video capture should be uncompressed at the native video output of the device being captured. The best available video output should be used from the device.

For older consoles, the hardware often outputs at 240p. Most capture devices will see this content as 480i, leading to interlacing artifacts that, when de-interlaced, reduce the framerate of the capture from 60 frames per second to 30 frames per second. Devices like the retrotink2x will scale the 240p output to 480p preserving the progressive 60fps content with minimal quality loss.

For RF-only consoles, a VHS player can be used to interface the RF video and convert it to composite. RF video should only be used if no other options are available, as the quality is extremely low.

For audio-video media, similar steps are necessary. Use the best quality output and ensure that the audio is not peaking. A time-based corrector may be used to eliminate errors from mechanical instability in analog media formats.

Capturing software such as VirtualDub or Open Broadcaster Software can be used. For uncompressed video, a hard drive RAID setup or solid-state drive may be necessary due to the amount of data being written to the computer. Most single hard drives are too slow to capture uncompressed video.

Preserving an Electronic Game Cartridge

A custom dumper, such as the Sanni Cart Reader, can help migrate cartridges from their original format to a binary format stored on a computer. The dumper devices typically interface with a PC through an onboard USB port or SD card. While the dumping procedure varies per format and device, the resulting binary is playable using emulation software.

Outside organizations involved in video game preservation have created databases in the form of DAT files that contain the checksums of the data dumped for games. Comparing the checksum contained in a DAT file against the newly migrated data ensures that the dumps are accurate.

In situations where no previous dump has been made, the DAT files will contain no checksum to compare against. To verify that the data is accurate, multiple dumps may need to be created with the cartridge connectors cleaned between attempts. Applications like HxD can be used to compare the resulting files and verify that the data is the same between dumps. The dumped backup should be played on original hardware or emulation and compared directly to the original artifact to help verify the results.

Preserving a Game on Optical Disc

Migrating a disc typically involves using a modern PC to create an accurate disc image (often a .ISO or .BIN/.CUE file) that contains the filesystem and data of the original optical disc.

While many optical disc formats may look the same and use similar physical media formats, developers can prevent certain formats from being read on devices that are not the original platform. Some formats like Microsoft Xbox and Nintendo Wii, despite using DVD media, are unreadable on a PC drive and require special hardware. Copyright protections can render a dump unusable, even in cases where the data on the disc is copied successfully.

DICUI software application handles a variety of different disc formats and attempts to keep copyright protections intact. As a result, these disc images can be used to play games with few modifications.

For CD-based games, Plextor drives can read data from discs that other drives cannot. These older, recommended drives are currently the only way to capture certain types of data.

Preserving a PC floppy diskette

The Kryoflux Floppy Drive Controller creates a flux preservation image from a floppy diskette. It samples the magnetic flux transition timings to create an accurate representation of the floppy diskette. This flux-level representation is unusable by many software emulators until it is converted to a different format.

Limitations: For floppy diskettes, the Kryoflux does not handle all formats well particularly “flippy-floppy” disks. For instance, Commodore 64 and Apple II disks should be converted with alternative specialized tools.

Limitations of Electronic Game Preservation

Hobby programmers create the software and hardware that preservation experts depend on. Support for such solutions can be non-existent and may disappear at any time. With any device used, software and supporting documentation should be saved upon acquisition, as both may become unavailable in the future.

Due to the reliance on hobby programmers, solutions for unpopular platforms may be limited. This can impact the ability to migrate data to the highest standards. The focus on non-destructive methods that preserve the original artifact can also create dumps that may not be 100% accurate, but the data is still usable and representative of the original software. This also allows for further backups if new methods are developed.

Copyright protections can impact the functionality of a digitized migration. In the case of PC games, the game may fail to without the original CD. In most cases, a home console will not boot a copied optical disc. Modifications to either the software or hardware may be necessary to play the digital backup. These modifications circumvent the copyright protection and allow these backups to be played.

For cartridge-based games, many devices exist to allow for ROM images to be run on the original hardware such as the EverDrive line of flash carts. These cartridges vary in quality, but they provide an authentic experience on original hardware while preserving the original artifact.

Preservation of Digital-Born Games

Digital-born electronic games are games that are released through the Internet, usually through a digital storefront, which do not exist on physical media until being downloaded. Access to these storefronts cannot be assured long-term, so local copies must be made to preserve the game. When backing up any digital-born game, the version should be noted. If it is unavailable, the date of the backup should be included with the downloaded files.

Preserving a mobile game

Backups of digital mobile games start with backing up the installation file. This can be downloaded directly from the digital storefront. This only captures a small portion of local data for some games, though. Upon booting the game for the first time, the game may download additional data. To backup this data, a rooted or jailbroken device is usually needed, as access to the files without is usually blocked. For a full backup, both the original installation file and user data should be backed up.

Preserving a PC game

Whenever possible, digital PC games should be purchased from a DRM-free storefront. This will allow for the download of installation files that can still be used regardless whether the digital storefront is still available. All files should be downloaded and stored together, including multiple versions of a game, instruction manuals, and additional files such as artwork.

If a DRM-free version is not available, the game should be installed using the digital storefront. A backup of the install directory should be made, which could help provide future play if the storefront disappears.

Preserving a console game

Unlike PC games, DRM-free storefronts do not exist for console games. It may also be impossible to copy the data from the console in any usable format. The game is downloaded and installed to a collection console.

Preserving a web-browser game

Not all web-browser based games can be easily download. If possible, the game and related files are downloaded to a folder. The website URL is noted, as some games may be locked to the specific website. If the game requires additional software to be download, such as Flash or Unity, a version of that software is downloaded and saved with the game. A note in the object record indicates the browser that is used. Confirmation that the game works in downloaded form is done by testing without an Internet connection.

Limitations of digital-born game preservation

The above highlights how to backup user-downloadable data for a game. It does not mean that the game will be accessible. Many digital-born games require online services and servers in the form of developer-run servers or cloud-based services (e.g., Google services, Facebook). These services are not preserved when backing up user-content, which means the game will no longer function. Having user-data provides access to artwork and data that can be used to help understand the game and can be used to help potentially for access in the future.

Digital-born games receive numerous updates, each of which can change the game. While one version of the game may be backed up, it may only represent a game that was available for hours or days due to the rapid changes.

Chronology of Technology used for ICHEG Preservation

Dates	Technology Capability	Purpose
2011 – 2015	Blackmagic Intensity Shuttle, Startech USB3HDCap, AverMedia CD311, Dazzle	Video Capture project funded by IMLS to document play of pre-2005 games.
2018 – Present	Pocket Programmer III, EPROM dumper	Allows for dumping and writing of compatible EPROM
2018 – Present	Sanni Cart Reader	Captures data from cartridges including Super Nintendo/Famicom, Sega Genesis/Mega Drive, Game Boy, Game Boy Color, Game Boy Advance, and Nintendo 64.
2019	Applesauce Floppy Drive Controller	Captures data from Apple II floppy diskettes
2019	Kazzo Cartridge Dumper	Allows for dumping of Nintendo Entertainment System and Famicom cartridges.
2019	Apple FloppyEmu	Allows for use of floppy disk images on original hardware using SD cards instead of original media
2019	Commodore 64 Ultimate II+	Allows for use of floppy disk images on original hardware using SD cards instead of original media, while also allowing for backing up of original disks and tape

Conclusion

Future Goals for Digital Preservation

The Digital Preservation Working Group has identified limitations in each aspect of digitization and preservation work in the library, archives, general collection, and electronic game collection. In a broad sense, these limitations can be categorized as limited staff time for digitization and metadata creation, lack of equipment to access and digitize the range of formats collected, and limited data storage to preserve digitized materials. Additional staff, equipment, and data storage would improve the preservation of the digital assets in the collection.

The digital preservation handbook describes what the Strong has the capability to digitize and preserve. The varied media formats, video game formats, and types of electronic and digital materials that the museum, library, and archives have collected require constant effort by the Digital Preservation Working Group to provide sustained preventive preservation care for the objects entrusted to us.

References

Beagrie, Neil and Daniel Greenstein. "A Strategic Policy Framework for Creating and Preserving Digital Collections." July 14, 1998. http://www.ukoln.ac.uk/services/papers/bl/framework/framework.html

Bogus, Ian, George Blood, Robin L. Dale, Robin Leech, and David Mathews. "Minimum Digitization Capture Recommendations." The Association for Library Collections and Technical Services Preservation and Reformatting Section, June 2013. http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations

Colloton, Edddy and Moomaw, Kate. "Rewind, Pause, Playback: Addressing a Media Conservation Backlog at the Denver Art Museum." The Electronic Media Review, Vol. 5, 2017-2018.

Erway, Ricky and Julianna Barrera-Gomez. "Walk This Way: Detailed Steps for Transferring Born-Digital content from Media You Can Read In-house." Online Computer Library Center Research, 2013. https://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf

Gough, Lui. "Sigh-Quests: More forgotten cartridges arrive, but fail to turn up data." February 1, 2016. https://goughlui.com/2016/02/01/sigh-quests-more-forgotten-cartridges-arrive-but-fail-to-turn-up-data/

Hodge, Gail. "Best Practices for Digital Archiving: An Information Life Cycle Approach." D-Lib Magazine, Vol. 6, No. 1, January 2000. http://www.dlib.org/dlib/january00/01hodge.html

Kuny, Terry. "The Digital Dark Ages? Challenges in the Preservation of Electronic Information." International Preservation News, No. 17, May 1998. http://www.ifla.org/V I/4/news/17-98.htm#2

Kussman, Carol, National Digital Stewardship Alliance (NDSA), Wayne Graham, Winston Atkins, Aliya Reich. "2019 Levels of Preservation Implementation Guide and Definitions." October 15, 2019. https://osf.io/qgz98/

McDonough, Jerome, Robert Olendorf, Matthew Kirschenbaum, Kari Kraus, Doug Reside, Rachel Donahue, Andrew Phelps, Christopher Egert, Henry Lowood, Susan Rojo. "Preserving Virtual Worlds Final Report." August 8, 2010.

Prael, Alice Sara. "Centralized Accessioning Support for Born Digital Archives." Code4Lib Journal, Issue 40, May 4, 2018. https://journal.code4lib.org/articles/13494

Rosenthal, David S.H. "Emulation and Virtualization as Preservation Strategies." October 30, 2015. https://mellon.org/media/filer_public/0c/3e/0c3eee7d-4166-4ba6-a767-6b42e6a1c2a7/rosenthal-emulation-2015.pdf

Schumacher, Jaime, Lynne M. Thomas, Drew VandeCreek, Stacey Erdman, Jeff Hancks, Aaisha Haykal, Meg Miner, Patrice-Andre Prud'homme, Danielle Spalenka. "From Theory to Action: Good Enough Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions. A Digital POWRR White Paper for the Institute of Museum and Library Services." August 2014. https://powrr-wiki.lib.niu.edu/images/a/a5/FromTheoryToAction_POWRR_WhitePaper.pdf

Weidner, Andrew, Sean Watkins, Bethany Scott, Drew Krewer, Anne Washington, and Matthew Richardson. "Outside the Box: Building a Digital Asset Management Ecosystem for Preservation and Access." Code4Lib Journal, Issue 36, April 20, 2017. http://journal.code4lib.org/articles/12342