Syreeta Gates

Digital Preservation

Assignment 2

What is The Gates Preserve Archive Museum and Library?

The Gates Preserve: Archive, Museum, and Library (TGPAML) is currently a 501(c)3 non-profit organization that supports the documentation, digitization and exhibition of hip hop culture, with a primary commitment to provide access to the public. TGPAML focuses on hip hop culture from the 1990s and beyond. TGPAML's mission is to collect, exhibit, organize, preserve, and provide access to materials relating to hip hop that will be housed in physical and digital repositories. The library will also have collections of creative and informational sources for use in the study, research, or leisure around six areas of interest (business, travel, tech, history, fashion, and psychology) as they pertain to the broader hip hop context.

Project Overview

Over the years I have had a variety of different insights based both what a collection's assessment and digital preservation could shape into. I explore various formats in the collection below, and have done the research so I am thoroughly informed on the recommendation that I have provided. The overview of the project will entail figuring out the best systems and protocols for both the collection's assessment and digital preservation plan. With hopes to use the recommendations that I've found to directly implement the status of the collection.

Background of the Collection

The content of the collection consists of books, audiovisual materials such as: U-matic tapes, VHS, 12 inch and 45 inch vinyl, and cassette tapes, born-digital materials, ephemera, photographs, sculptures, magazines, and visual art. These items have been collected over the last ten years, and have been apart of my personal collection. Up until recently have I decided to create an institution, in which I will donate that collection to.

Status of the Collection

Currently, approximately 70% of the items have been indexed in a spreadsheet. Data will soon be uploaded to a Digital Asset Management System (DAMS). The print materials, in particular books and magazines, are housed in acid-free bags, mylar sleeves, and acid-free boxes. All materials are stored in a private residence, away from sunlight. Magazines are stored upright in mylar sleeves. The collection features a growing volume of born-digital materials. Currently, this includes about 500 born-digital photographs and over 6 TB of video footage. Future plans include digitizing a large magazine collection, audio and video content, and archiving living collections.

RECOMMENDATION

- Perform a physical assessment of collections materials, their storage environment, and handling procedures.
- See Preservation Self-Assessment Program (PSAP) for more information.[1]
- Assessment should examine:
  - Condition: Survey of collection condition (either item-level or a sampling)
  - Storage: Assessment of storage environment
  - Handling: Description of handling practices, amount of handling, types of use (scanning, active reading, etc.)

Collections Care: Environment

There are three primary environmental factors that can pose a threat to long-term preservation:

1. Light
2. Temperature/humidity
3. Dust/dirt

Light can cause fading and discoloration if exposure is prolonged. Light damage is cumulative over time and irreversible. A hot and humid storage climate can promote deterioration and mold growth. The ideal environment is about 60-70° F, medium relative humidity (30-50%). Dust is not only unpleasant to work with, but is actually acidic and can damage materials if it sits on the surface over time.

RECOMMENDATIONS

- Store items in boxes. This is a simple solution, but boxes goes a long way. Boxes protect against light, dust, temp/humidity fluctuations, and provide an extra layer of protection against water/fire damage. Specifically for the epherma content
- If possible, store oversized media (posters, prints, artwork) flat in drawers or boxes
- Store vinyl records with the spine vertical
- Store materials off the ground (at least a few inches) to mitigate flood risk
- Keep away from sources of water and heat, such as radiators, pipes, sprinklers
- Don't display valuable items permanently: rotate display or display a reproduction
    - Display away from windows
    - You can use UV-protected "museum glass" frames for long-term display
- Use a humidity monitor to check levels in your location
    - If fluctuations are extreme, try to seal windows better or use a dehumidifier to moderate levels

Collections Care: Handling

Handling inherently puts materials at some risk. The best preservation method for original materials is to reduce the amount of handling overall. Whenever possible, make a facsimile (if it can be done easily without damaging the original). This is particularly true for fragile materials such as artworks, photographs, film, and negatives. The PSAP Collection ID Guide is a great resource for identifying formats, preservation risks and recommendations for a wide variety of materials. See entries for U-matic tapes, VHS tapes, and compact cassette tapes, for example[2]

RECOMMENDATIONS

The following are recommendations for handling specific formats in the collection:

Photographic prints, film, negatives

- Use nitrile gloves
- Always handle media by the edges (do not touch the surface)
- To clean, use a dust blower or very gentle antistatic brush
- Photographs, especially color photos, are very sensitive to light
    - Store in acid-free boxes or PVC-free plastic sleeves
    - Store prints in flat boxes to prevent curling

Ephemera materials

- Store in acid-free containers only
- Remove rusty clips and staples, replace with plastiklips
- Remove ALL rubber bands
- Don't use gloves if handling brittle pages (they crumble more easily)
- Gently use a document cleaning pad to remove surface dirt

Audiovisual materials

- Risk and deterioration varies by format, but by its nature, most audiovisual material must be played back to use the content. This playback damages the media over time:
    - i.e., in the case of cassettes: *"Frequent playback wears on the media and degrades the sound quality over time. This medium is especially susceptible to damage from playback as it may jam in the playback deck and be "eaten," which can cause crimping and breaking during playback. Internal cassette elements like pads and rollers are susceptible to damage."*[3]

- In addition to handling, obsolescence is a risk. Audiovisual formats change frequently, and it can be very difficult to find equipment to playback older media

Collections Care: Supplies

Most of these can be found on Amazon or places like B&H Photo, which is where I have purchased materials in the past. Supplies for preserving specific materials can be found at industry suppliers like Gaylord Archival (which I currently use), Hollinger, and Talas (though at higher cost).

RECOMMENDATIONS

Basic supplies for handling, cleaning, storage:

- pH pen (for testing acidity of storage materials)
- Microspatula (removing paper clips or staples)
- Humidity monitor
- Gloves (for photographs)
- Cotton gloves (for artwork, artifacts)
- Dust blower
- Antistatic cleaning cloth (for flatbed scanner)
- Document cleaning pad
- Acid-free boxes / folders / backing board / envelopes of appropriate size and orientation
- Acid-free tissue (for interleaving pages, wrapping or packing)
- PVC-free plastic sleeves such as polyester or polyurethane (aka mylar)[4] [5]

Arrangement

Collection arrangement should make it easy to keep materials organized and facilitate easy browsing and retrieval. If when a collection is acquired its arrangement shows the creator's intent, it's usually preferable to maintain the original order. For example, if a journalist donates a box of research files organized in folders, maintain the original order of the files as is. For collections that grow over time as

new items are added (like the magazine collection), arrange materials in a way that promotes ease of use.

Physical Control

Once collections are arranged, housed, and stored, it's crucial to keep an accurate record of storage locations. This can be done in a spreadsheet or Content Management System. Ensure that containers and individual folders are labeled and identified.

Intellectual Control

To enable the public to find and use TGPAML materials, create collections guides. These can be as simple as inventory lists, collections summaries, or basic finding aids. To identify individual items, it would be useful to assign each item a unique ID. This will help with inventory, logging and tracking use over time, and intellectually linking digital copies to originals.

RECOMMENDATIONS

- Create public-facing collections guides to be made available in order to allow would-be users to find relevant materials
- Develop an ID numbering schema to give each item a unique identifier. Keep it simple and have it suit your needs. IDs should be alphanumeric (hyphens and underscores are okay, no periods or spaces). Here are a few different approaches:
    - Simple numbering: 00001, 00002, etc.
    - Format: photo-0001, photo-0002, etc.
    - Collection/date (magazines): TheSource-1996-June, Vibe-2003-November-cover
- The same ID should be used as the filename any time an item is digitized
    - Vibe-2003-November-cover.tif

Metadata

Useful description of the collections should facilitate control, storage, access, and digital use. The fields provided in the existing TGP Inventory List Google Sheets doc (under the categories Administrative, Descriptive, Technical, and Preservation) are suitable for collections management. In general, it helps to keep data clean for consistency. For example, Duration/Size should list units consistently (i.e. 24 x 24 in. or 48 min.). It should be catalogued consistently and as the most preferred name (i.e. “Warhol, Andy” not “Andrew Warhola”). There is an option to sort and filter data in Excel or OpenRefine to check for inconsistencies. When selecting a DAMS, consideration should be given to how customizable and extensible its data fields are. Can they easily accommodate select fields from your existing metadata?

RECOMMENDATIONS

For the following fields, see notes below. Note, it is recommended that some basic fields are made mandatory. If the cataloger needs to supply data, such as a descriptive title, use [brackets].

FIELD	RECOMMENDATIONS
Unique ID	Make required. Assign each item a unique ID once a schema is developed.
Title	Make required. If no title exists, supply a descriptive one in [brackets]. i.e. [Northeast corner of Gates Avenue and St. James Place] or [abstract illustration with geometric figures]
Creator (new field)	Make required. Separate from the current field People for clarity.
Subject (new field)	This could list themes, placenames, or people (other than the creator).

Date	Make required. If approximate date known, use circa dates or a range. i.e. “c. early 1980s” or “1971-1974”
Format	Make required. If an item is digitized, add the digital copy here too. For example, if you scan a photo, Format could be “Photographic print (original); TIFF (digital)”.

Acquisitions and Collection Development Policy

When formally accepting donations, a Deed of Gift needs to be signed by the archive and donor. The Deed of Gift should establish the transfer of ownership of the materials themselves, ownership of the rights (if the donor is the copyright holder), and other conditions for how the materials can be used and accessed.

See Deed of Gift examples at Society of American Archivists Museum Archives Section Standards and Best Practices Resources Guide.[6]

A Collection Development Policy can help define the scope of your collections and outline your collecting rationale. This will keep collection objectives focused, encourage relevant donations from the public, and allow you to tactfully refuse unwanted gifts.[7]

RECOMMENDATIONS

- Draft a Deed of Gift (would be worth reviewing with a legal advisor)
- Write basic collection development policy, considering factors like:
  - TGPAML users and audience
  - Rationale for collecting
  - Scope of collections - format and types of materials, time period covered, geography covered, subject areas of focus
  - What other archives have overlapping objectives?

Digital Preservation: Background

Archives maintain records of permanent value created by organizations, institutions, groups, or individuals. As records are increasingly produced in digital form, there is a clear need for archives to support their long-term digital preservation. Recommendations in the following pages come from sources noted. For further background and information about implementing best practices, see the following resources:

- Preserving Objects With Restricted Resources (POWRR)[8]
- Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC)[9]
- Digital Preservation Business Case Toolkit[10]
- Library of Congress Personal Archiving[11]

RECOMMENDATION

To start, gather information about what exactly your digital assets are. This can be accomplished with an inventory, which will account for digital assets in the archive. You can use this to account for the extent and makeup of collections (file sizes and formats), and to identify and prioritize materials for of value. Make note of any potential problems, like obsolete or hard-to-open proprietary file formats or errors in filenames.

An inventory could include:

- Category of materials (Papers, Photographs, Interviews, etc.)
- Associated projects/events/publications
- Creation date(s)
- Location (where are they currently stored?)
- Extent (file sizes, number of files)
- File formats
- Other file characteristics or notes (i.e., some files won't open)

Digital Preservation: Preferred File Formats

The following outlines common types of digital records the archive might want to collect. File types that have poor long-term viability (proprietary formats like Photoshop's .PSD files, for example), should be normalized into preservation formats for long-term storage.[12]

Media type	Original file format	Normalized version for preservation	Notes
MS Office (2007 and after)	DOCX, PPTX, XLSX	*Use original format*	These files are XML files, so they work well for preservation.
MS Office (pre-2007)	DOC, PPT, XLS	*Use original format*	Proprietary format, but they show no signs of being dropped from MS Office, so they continue to be acceptable.
Graphics - sustainable	JPG, TIFF, PNG	*Use original format*	
Graphics - low sustainability	PSD, BMP, GIF	TIFF	Flatten any layers

Adobe Illustrator	AI	PDF/A or High Quality PDF	Save as “high-quality” PDF. After saving as high quality PDF, use Adobe Acrobat to convert to PDF/A.
Audio	AC3, AIFF, MP3, WAV, WMA	*Use Original Format, see note.*	Monitor formats for obsolescence issues. Migrate to WAV if format becomes endangered.
Video	AVI, MOV, MP1, MP2, MP4, SWF, WMV, QT, M4V	*Use original format,* see note	Migrate to FFV1/LPCM in MKV if format becomes endangered. (further reading)[13]
Camera raw	CR2, RAW, DNG	TIFF	Use Photoshop or other tool that reads these files.
Websites	HTML, web assets	WARC, screenshots	Create WARC files using Webrecorder. Use Webrecorder Player to play back. Additionally, use a browser plugin to save screenshots as PNG files (i.e. FireShot)

Digital Preservation: Digitization

The goal behind digitizing collections should be “capture once, use many times”. Scanning is resource-intensive in terms of time and cost, and adds wear to the life of collections materials. For this reason, any digitization project should set out to capture “master” files that can be preserved and be used to create derivative copies for any uses going forward.

RECOMMENDATIONS

Recommended flatbed scanners:

- EPSON Perfection V600 Photo Scanner ($229)
- EPSON Perfection V800 Photo Scanner ($899)
- EPSON Perfection 12000XL Graphic Arts Scanner ($3299)

The following specs should be used for scanning most materials:[14]

- Resolution: 600 dpi for photographs, 300 for magazines (simply because print publication image quality is already low)
- Format: TIFF
- Mode: RGB
- Bits: 24-bit for color, 8-bit for grayscale

When creating derivative copies, save as JPEGs. You can make the resolution smaller for these versions (300 dpi is more than enough), depending on the use. If scanning very small items, scan at a higher dpi. If the item you’re scanning is about 2 x 2”, scan at 1200 dpi. If it is a 35mm slide, try scanning at 2000 dpi.

Several vendors with relatively affordable rates:

- Hudson Archival (professional, great service, higher cost)
- Dijifi (NYC-based, medium cost)
- FotoBridge (mail-in service, low-medium cost)
- METRO NY Library Council (it appears in-studio equipment can be rented at a good rate, especially for members; they also have audiovisual conversion equipment for digitizing tapes)

Digital Preservation: Digital Storage

Digital storage organization can mirror arrangement of physical collections. A possible arrangement of folders in the archive directory for a born-digital photo collection might look like:

>Archives

>John Smith Photograph Collection

>2009

>June Magazine Shoot

>MASTERS

JohnSmith-0001.tif

JohnSmith-0002.tif

JohnSmith-0003.tif

>COPIES

JohnSmith-0001.jpg

JohnSmith-0002.jpg

JohnSmith-0003.jpg

>123 Gallery Opening

>2010

>2011

Digital Preservation: Security

The ideal, effective storage solution for The Gates Preserve would provide a repository for storing "master" copies of digital files with long-term value. The infrastructure would allow for regular access, but its primary attributes would be data security, reliability, appropriate cost, and ease of use.[15]

Storage locations

One of the best ways to secure data is storing copies in more than one location. Archives data should be stored in at least two different locations, and ideally three (two physical devices and one in cloud storage). Each copy should be regularly backed up. Backups should be recorded in a log. Check backups regularly to ensure their integrity. Any important data on external media (like CDs, thumb drives, or floppy disks) should be copied and stored elsewhere as this media becomes obsolete and risks data corruption over time.

Migration

It's important to migrate all archives data to new storage media every few years. Even secure, external hard drives fail or become obsolete eventually, so it's an essential part of a long-term sustainability plan to migrate data storage systems at least every five years.

Web apps (Google Suite, etc.)

Web apps like Google Sheets, Drive, and Dropbox are fine for working files, but shouldn't be considered a long-term solution for storing data. Among other factors, there is a high risk of accidental editing or deletion.

Digital Preservation: Quality Control

Data quality

To ensure data quality, try to maintain controlled vocabularies. These can be authoritative lists of people's preferred names, place names, or subject terms. Use existing vocabularies (LC Subject Headings, Getty Art & Architecture Thesaurus or Union List of Artist Names) or maintain in-house lists. You can include variants (pseudonyms, married names, alternate spellings).

Use Excel or OpenRefine to check datasets for inconsistencies.

Scanning quality

Keep the scanning area clean (no food/drinks near the materials, use clean hands, don't touch the scanner glass). After talking with Kelly Haydon she recommended I reimagine how I've been thinking about quality control and quality assurance.

When digitizing magazines or other print publications, the image may appear distorted by a dot pattern called moiré. Use the scanner's DESCREEN setting to mitigate it.[16]

General media quality control

Never work directly from master digital files when editing.

Regularly check quality of media (during a large-scale scanning project, for example), either by sampling or reviewing individual digital files.

Digital Preservation: DAMS and CMS

DAMS and Content Management Systems (CMS) range from open-source programs to expensive enterprise systems. Selecting one depends on the scope and size of digital collections, user needs, and available resources.[17] Aside from the cost of licensing or purchasing software, consider the resources needed in terms of time and technical expertise to install, maintain, backup, and troubleshoot a system. For a small organization, a low-cost, secure, low-maintenance system that allows for efficient description and retrieval is recommended.

Collective Access (example)

- Pros
  - Free, open-source (but need to manage your own server or pay for hosting)
  - Good for showcasing digital collections
  - Customizable fields
- Cons
  - Cost of hosting service

Omeka.net (example)

- Pros
  - Easy to use
  - Good for publishing digital collections and exhibits online
  - Affordable hosting included
- Cons
  - Appearance customization is limited
  - Storage cap might limit audio/video

Access to Memory (AtoM) (example)

- Pros
  - Web-based and open-source tool for archival description
  - Affordable hosting available through LibraryHost
- Cons
  - Somewhat inflexible design

PastPerfect Museum Software (example)

- Pros
  - Supports core functions of archival administration
  - Good for description, collections management, and inventory control
  - Allows for publishing online using PastPerfect Online (cost extra)
- Cons
  - Support costs extra
  - Dated interface design

Extensis Portfolio (example)

- Pros
  - Flexible DAMS that allows for adding metadata in batches
  - Modular, configurable service
- Cons
  - Much higher cost

Bibliography

[1] https://psap.library.illinois.edu/manual#assessments

[2] https://psap.library.illinois.edu/collection-id-guide

[3] https://psap.library.illinois.edu/collection-id-guide/audiotape#compactcassette

[4] Personal Archiving Workshop, led by Mike Satalof

[5] https://hsp.org/historical-heritage-organizations-0/hidden-collections-initiative-for-pennsylvania-small-archival-repositories/resources-for-small-archives

[6] https://www2.archivists.org/groups/museum-archives-section/2-appraisal-and-acquisitionaccession

[7] https://hsp.org/sites/default/files/images/HCIPSAR/nextstepshcipsar.pdf

[8] https://digitalpowrr.niu.edu/

[9] https://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/trac

[10] http://wiki.dpconline.org/index.php?title=Digital_Preservation_Business_Case_Toolkit

[11] http://digitalpreservation.gov/personalarchiving/

[12]

http://files.archivists.org/groups/museum/standards/8.%20Workflow%20Documentation/Preferred%20file%20formats_BardGraduateCenter.pdf

[13]

https://www.fiafnet.org/images/tinyUpload/Publications/Journal-Of-Film-Preservation/Matroska-and-FFV1_Kromer_JFP96.pdf

[14] https://www.in.gov/iara/files/Digitization_Standards.pdf

[15] https://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/trac

[16] https://www.scantips.com/basics6b.html

[17] https://www.nationalarchives.gov.uk/documents/archives/cms-dams-options-for-archives.xls