Preservation Plan

African Library Project's Digital Collection

"To really be preserving something you need to be able to make it discoverable and accessible."

- Trevor Owens, The theory and craft of digital preservation

Rational for digital preservation

Digital preservation involves the management and maintenance of digital objects to ensure the authenticity, accuracy, and functionality of content over time in the face of technological and administrative changes. (Rieger, 2018)

Simply put, digital preservation is about making the best use of your organization's resources, making those resources accessible not just for the current designated community—African Library Project (ALP) staff and core volunteers such as container managers, country liaisons, and board members—but for the future members of ALP staff and volunteer network. In the following preservation plan, I will discuss the organization's current preservation model, how it can be improved, and the action items that should be taken to establish a more sustainable preservation practice.

There are risks in adopting a new digital preservation plan. Data can be lost when migrating to another storage system. There also may be problems when onboarding staff and volunteers to new guidelines that may be established in this digital preservation plans. These risks can be mitigated, and I will discuss how below. It's important to keep in mind the benefits of implementing a digital preservation plan that is sustainable and efficient far outweigh the risks.

How does African Library Project preserve its digital information?

Currently, born-digital assets are submitted to ALP in three ways. ALP African partner's use WhatsApp to document the arrival of ALP libraries, including but not limited to unloading books, schools receiving books, and pre and post-library spaces. When that content is submitted through WhatsApp, it's stored in the internal application. Later, it is stored in Google Drive by ALP staff. There are no naming conventions for these files. The other way ALP receives born-digital assets is through their volunteer book drive organizers (BDOs). BDOs document their book drive processes to be shared with ALP staff and volunteers via email or social media. The final way ALP receives born-digital assets is through their staff and core volunteers. Assets submitted by the staff and core volunteers are stored in Dropbox and on Flickr.

Locations of data (Unstructured, structured, and semi-structured)[1]

1. WhatsApp
2. User or volunteer-generated (social media, email)

[1] Structured data is data whose elements are addressable for effective analysis. It has been organized into a formatted repository that is typically a database. Semi-structured data is information that does not reside in a relational database but that have some organizational properties that make it easier to analyze. Unstructured data is not organized in a predefined manner or does not have a pre-defined data model.

3. Dropbox (direct uploads)
4. Flickr

What will a digital preservation plan accomplish?

Digital content that is easier to find and understand. No more vagueness about what a piece of content is or where it is from. Consistency in storage of digital content across the organization indefinitely. Reduce the amount of time users search for data for the purpose of storytelling. Reduced time searching for data often leads to a higher return, so we can spend less time searching through tons and tons of content for the right image or the right video and more time telling people our story.

Organizational statement of commitment

The core volunteers and board members of ALP will make up the committee for the digital preservation action plan and will review updates to the plan each month as they are submitted into the board status update by the led project manager, who in this case is the marketing manager of ALP. The marketing manager and marketing intern, both of who have a background in library work and library and information science, will have to collaborate with core ALP volunteers and the board of directors to ensure a long-term preservation plan.

After the preservation plan is reviewed, the marketing manager will begin implementing digital preservation actions (see *Preservation Actions and Quality Control*). The first step is to back up all data on an external solid-state hard drive that will be stored with the marketing manager for the duration of the project before being transferred to the ALP board chair. Following this step, the marketing manager focus on reorganizing current data within ALP's digital collection using the agreed-upon standards.

The committee members, on a volunteer basis, will be responsible for testing workflows and metadata standardization for ALP's collection and approving project hours and any additional project expenses. They will also be responsible for restricting access to the collection to only the project manager while the collection is undergoing essential transformation prior to the first step. The committee members must commit to maintaining their role to ensure a sustainable digital preservation plan for the duration of their directorship with ALP.

Statement of financial commitment

In order to complete and maintain digital preservation actions on behalf of ALP, the organization will have to make certain financial commitments. ALP is a small organization, so it's important to optimize the way information is being disseminated and learned as well as manage information in the most efficient way possible. Investment, financial and otherwise, will be required to keep staff and volunteers up to date with the latest in digital preservation. Staff and volunteers will also need to undergo training on how to manage resources efficiently within the new digital preservation framework.

Though now this plan requires a lot of staff resources it's our hope that over time less resources will be needed because the way we store digital-born content submissions will evolve, and ideally, resources will be managed through a single platform with automated backups to maintain data security and ensure file fixity.

National Archives and Records Administration (NARA) Levels of Digital Preservation

This model is used as a guide for institutions and organizations to start digital preservation. It can also help institutions and organizations "evaluate how they are doing in terms of mitigating risk of loss [of their resources] and identify concrete technical next steps they can take to move all or part of their operation to the next level." (Phillips, 4)

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
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 storage media and what 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 copies in 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 fixity info if it wasn't provided with the content	- Check 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 to 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

Figure 1. National Archives and Records Administration Levels of Digital Preservation

Preservation Actions and Quality Control

Some of the needs of the organization include fundraising, marketing and public relations, ensuring book drive registrations and incentivizing volunteers to become repeat book drive organizers. How do we do that? Storytelling. Storytelling happens through content. How do we optimize our content? We start here:

NARA Levels of Digital Preservation - Level 1 (Protect Your Data)

- Storage and Geographic Location
- Metadata
- Information Security
- Fixity and Data Integrity
- File formats

1. Surveying current digital assets within one storage type (Dropbox)
    a. ALP Digital Content
2. Content survey assessment and suggestion. This assessment will determine the state of current digital content on a microlevel, the tools (Internet Archive) needed to bridge gaps in digital preservation (bulk rename software) and information architecture or how content should be stored based.
3. Create naming conventions (and ensuring data standardization) based on researched and proven standards within the digital preservation[2]
4. Create ingest guidelines[3] on how assets should be stored and what storage includes (Submission Information Package (SIP))
5. Secure files by restricting access to who can move and delete files and ensuring backup and non-collocation of inventory.
6. Organize and rename files for easier data management based off of naming conventions decided and formalized by ALP board and staff
7. Archiving previous digital assets used for communications, fundraising, or marketing purpose for future finding and use (blogs, newsletters, press releases for or about African Library Project, previous summit or Harambee material(s)) using established naming conventions and ingest guidelines

---

[2] Case study: File naming, Stanford University - https://library.stanford.edu/research/data-management-services/case-studies/case-study-file-naming

[3] Ingest guidelines must include how to create descriptive metadata (location, school, person(s)) that will be used to in file naming. Administrative metadata (who uploaded file, date, file format) created automatically by Dropbox.

These steps will allow us to practically start carrying out our digital preservation plan. To ensure authenticity of the resources that ALP is preserving, we will use a fixity tool 'Fixity' to monitor the bitstream encodings that makeup each asset. 'Fixity' will allow us to evaluate any changes to our data. However, it is not the only criteria we will use. See *monitoring and review* for more ways ALP will ensure the digital preservation plan's success.

Metadata creation

ALP will include use, administrative, descriptive, and administrative metadata in order to preserve its digital assets. This metadata will be stored in an excel spreadsheet with limited access. The following is a chart that delineates the types of metadata needed based how where born-digital assets are being submitted to ALP.

Where data is being submitted	**Who is submitting data**	**Metadata created**	**Metadata needed**
WhatsApp	African partners	Creator, location, time, technical metadata	Use metadata (rights, contributors), Language
Email / Social	Book drive organizers	Creator, location, time, technical metadata	Use metadata (rights, contributors)
Dropbox	Core volunteers, staff	Location, technical metadata	Use metadata (rights), administrative metadata (source)
Flickr	Core volunteers, staff	Location, technical metadata	Use metadata (rights), administrative metadata (source)

Roles and responsibilities

As agreed upon by ALP board, primarily this effort will be led by the marketing manager. Twenty percent of marketing manager time has been allocated to establishing digital preservation plan as a part of knowledge management duties.

There are many ways to marketing manager can invest time in realizing the best practices of digital preservation, such as subscribing to trade publications and attending industry conferences

or workshops. In order to make ALP's digital preservation efforts sustainable long-term, core volunteers and staff alike must invest time and collaborate in order to learn the value of digital preservation and its practices within the nonprofit space. All ongoing developments will be disseminated at each monthly board meeting by the primary project leader, the marketing manager.

Monitoring and review

Upon completion of action items, the marketing manager will send out google form to designated communities (ALP staff and volunteers) to gather feedback on ease of use and ease of accessibility of digital content. The ease of using ALP assets and submitting assets for ALP's designated community will be evaluated and will ultimately determine whether this project was successful or not. As stated in the early section, *Organizational statement of commitment,* the status of this project will be reviewed each month during ALP's monthly board meetings. Updates will be listed on board status update, and reviews will encompass evaluating the digital preservation plan's progress based on project timeline.

Sources

Owens, Trevor. The Theory and Craft of Digital Preservation. Chapter Five: The Craft of Digital Preservation. Johns Hopkins University. Baltimore, Maryland.

Phillips, M., Bailey, J., Goethals, A., & Owens, T. (2013) The NDSA Levels of Digital Preservation: An Explanation and Uses. National Archives and Records Administration, Washington DC.

Rieger Y. Oya. The State of Digital Preservation in 2018: A Snapshot of Challenges and Gaps. ITHAKA S+R. Online. https://sr.ithaka.org/publications/the-state-of-digital-preservation-in-2018/.

Stanford Libraries, Data Management Services. “Best Practices for File Naming.” Stanford, California. library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-naming.

Stefano De. Paula (al et.), NDSA Publication. Staffing for Effective Digital Preservation: Checking Your Digital Content. What is Fixity, and When Should I be Checking It? Online. URL: http://hdl.loc.gov/loc.gdc/lcpub.2013655117.1