

Developing Digital Preservation Strategies for the Baltimore Community Museum

Emily Flint
MLIS Candidate/Digital Preservation Consultant
University of Maryland
December 6, 2018

Table of Contents

Executive Summary.....	3
Initial Survey.....	4
Next Steps Plan.....	7
Digital Preservation Policy.....	11
Works Cited.....	14

Executive Summary

The documents in this report were created as part of the Digital Preservation Consultant Project assignment for INST 784: Digital Preservation at the University of Maryland's iSchool. Jessica Kunkler Shaw, the director of the Baltimore Community Museum in Baltimore, Ohio, expressed interest in partnering with a MLIS student to improve the museum's digital preservation practices. This project appealed to the museum because the organization is currently in the process of inventorying its collections and digitizing significant documents and photographs. The Baltimore Community Museum also previously digitized a large collection of local newspapers. However, there were no policies in place to manage digital content or to keep track of what items had been scanned thus far. The director of the museum and her interns were having trouble coming to a consensus on how to handle their content. Staff did not feel prepared to make decisions about digital preservation and were concerned about losing access to their files.

Throughout the fall 2018 semester, I worked with the Baltimore Community Museum to identify digital preservation strategies that would suit this small organization with limited staff. The Initial Survey of the museum's collections and practices reveals some areas that needed improvement. Staff were not consistently creating multiple copies of their files and did not know how to conduct fixity checks. In general, the museum was struggling to gain control over its collections due to their size. Files were not organized in any particular fashion after they were created; staff relied on memory to track their work and locate files. The Next Steps Plan provides three levels of recommendations based on the National Digital Stewardship Alliance Levels of Digital Preservation. The museum intends to follow the "beginner" plan at this time due to limited resources. However, it is possible that the museum will be able to explore the "intermediate" and "advanced" options after staff begin to feel more comfortable with digital preservation. The Digital Preservation Policy includes the most pertinent recommendations from the Next Steps Plan. In particular, the policy asks staff to make multiple copies of files, store these files in different geographic locations, conduct basic fixity checks, and create a basic inventory of their files. The museum will continue to use common formats like JPEGs for their files and will ensure that only staff have access to these files.

The guidelines in the Digital Preservation Policy will help the Baltimore Community Museum to begin actively managing its digital content without overwhelming staff. The museum's director will be able to use these recommendations to train interns and volunteers to complete digital preservation tasks. Eventually, museum staff hope to be able to share more of its digital collections online to benefit local genealogists and other users. By establishing procedures for making copies and checking file fixity, the museum will help to ensure the sustainability of its digital content for future projects. The Baltimore Community Museum will review the Digital Preservation Policy annually and update the document as needed.

Initial Report

This semester, I will be serving as a digital preservation consultant for the Baltimore Community Museum, which documents the history of the small town of Baltimore, Ohio and the surrounding areas. According to the director of the museum, the organization is “dedicated to the preservation of local historical artifacts and the exploration of the history and culture of Baltimore for the betterment of the community.” To support this mission, the museum collects documents from the township, papers of prominent citizens, and photographs. The museum’s collections also include a broad variety of historical artifacts like dresses, dental equipment, and even an airplane. Currently, museum staff are working to scan documents and photos from the collections. Issues of a local newspaper, the *Twin City News*, have also been digitized. Thus, much of the museum’s digital content is fairly recent and is still being created. The museum is more focused on digitization at this time, but they do also have born-digital objects like cassette tapes that contain oral histories. However, they are not sure what is on some of the tapes and do not have a cassette player. In the future, the Baltimore Community Museum would like to continue to collect similar resources related to the history of the area. The organization also hopes to strengthen its family history collections in order to assist members of the community working on genealogy projects.

When I interviewed the director of the museum, Jess Kunkler Shaw, and one of her interns, they emphasized that the museum’s collections are quite extensive for a town with a small population because the community is passionate about documenting local history. However, the people who originally founded the Baltimore Community Museum did not have a background in museum work and did not know the best way to organize and preserve the museum’s holdings. The end result is that present-day museum staff members are still trying to figure out what they actually have in their collections. Staff are also encountering items that are in poor condition and have had to prioritize some emergency situations. For example, Max, the museum’s intern, shared that some items from the collection had to be removed from a humid basement. The Baltimore Community Museum is in the process of inventorying its holdings and uses PastPerfect to keep track of its analog collections. These ongoing projects are critical for the organization as it continues to gain control over its physical collections, but this also means that staff may not have a great deal of time to manage and preserve digital objects.

Jess and her interns are scanning items as they come across them during their inventory process. Some photographs were also scanned by the museum’s previous director. After attending workshops and conducting research, Jess determined that scanning needed to be a priority for the museum because of the large amount of township records, documents, and photographs that exist in the collections. The Baltimore Community Museum is prioritizing documents and photos that are damaged and items that are of great importance to the history of the community.

Jess also stated that organization was a concern for the Baltimore Community Museum. While she generally knows where things are located, efforts to inventory the collections are still ongoing and formal documentation does not always exist for every item. There is also no real system in place to keep track of what has been scanned thus far. After items are scanned, staff save the files to a folder on the organization's laptop. The museum is also using Google Cloud to store some items. However, the organization does not have multiple copies of their files and does not have a process to back up the content. Staff are concerned that they would lose access to the files if the museum's laptop were to break. Staff members typically scan and save items as JPEG files. They have also attempted to use PDFs in the past, but have encountered some difficulties with maintaining a quality image of the documents being scanned. The museum does not conduct fixity checks on the files and staff do not know how to begin doing this. Jess and her interns are the only people who can read, modify, and delete the files on the laptop. Currently, museum staff members are the only people who have direct access to the Baltimore Community Museum's digital content. Researchers would have to visit the museum to use the staff laptop to view the majority of the items that have been scanned thus far. CDs of the digitized *Twin City News* content have been created and are being sold to the public in order to generate revenue.

The organization is pursuing grant funding to develop a genealogy research center based in the museum. This genealogy hub would be a place where people could access physical resources like family genealogies that have been donated to the museum, cemetery records, township records, and other documents that can help researchers track their ancestors to a given place and time. Visitors would also be able to access Ancestry.com and other similar sites. Staff believe that focusing on genealogy would serve the needs of the community and would bring more people into the museum. Eventually, the organization plans to have physical resources digitized and made available on their website so that people who live further away from Baltimore can still access the valuable family records in the Baltimore Community Museum's collections.

The director of Baltimore Community Museum is the organization's only permanent staff member. Jess started in this role about a year ago. She works part-time and reports to a board of nine people who are supportive of her work. In addition to running the museum, Jess manages rentals of the museum's facilities for local events. Rentals generates valuable revenue for the museum, but because Jess has to coordinate these rentals herself, she does not always have enough time to work on collections management tasks. Over the summer, Jess began working with a group of about five interns and continues to work with two of the interns this fall. The interns receive a stipend. The museum does not currently collaborate with any volunteers from the community. In the past, people have expressed interest in working with the museum but have not followed through. The museum has created an application for interested parties to fill out and would like to bring in volunteers in the future.

As a result of this semester-long consultant project, Jess hopes to be able to gain knowledge about digital preservation strategies so she can train interns and volunteers to assist

her with these sorts of tasks in the future. Staff realize that the Baltimore Community Museum's current system for managing digital content is not as strong as it could be. Jess is concerned that if the museum starts to take action now, they might do something incorrectly and end up having to redo the work. As the semester progresses, I hope that we will be able to identify meaningful first steps for the Baltimore Community Museum.

Next Steps Plan

Introduction

The Baltimore Community Museum documents the history of the small town of Baltimore, Ohio and the surrounding areas. The museum's collections include documents from the township, papers of prominent citizens, photographs, and historical artifacts. The museum's collections are extensive for a town with a small population and staff are currently working to gain better control over the museum's holdings through an inventory project. The director of the museum and her interns are scanning noteworthy items as they come across them while working on this inventory. The Baltimore Community Museum is primarily scanning documents and photos that are damaged and items that are of great importance to the history of the community. Staff estimate that they have created about 600 files thus far. Issues of a local newspaper, the *Twin City News*, have also been previously digitized. Staff would like to ensure that they don't lose access to this valuable content and hope to be able to make the scanned items available on the museum's website in the future. The museum is particularly interested in developing resources for genealogists.

Current Practices

The National Digital Stewardship Alliance Levels of Digital Preservation provide tiered recommendations for organizations interested in preserving digital content. The NDSA divides its recommendations into five categories and suggests four levels of action for each category (Philips, Bailey, Goethals, & Owens, 2013). By analyzing the Baltimore Community Museum's current practices and comparing them to the suggested practices in the NDSA levels, we can identify opportunities for future growth. The Baltimore Community Museum is approaching Level 1 for the Storage and Geographic Location category. The museum uses Google Cloud for photographs and Dropbox for documents, but the materials do not overlap. Staff are not sure how to conduct fixity checks on the files they are creating, so the museum does not currently meet the requirements for Level 1 of the File Fixity and Data Integrity category. The museum director and her interns are the only people who can read, modify, and delete files. While members of the community are able to come in and view the files on the staff laptop, this does not happen frequently. Staff supervise visitors who are using the laptop, so it is unlikely that visitors would accidentally delete or change files. Therefore, the Baltimore Community Museum meets the requirements for Level 1 of the Information Security category. While staff generally know what has been scanned thus far, there is no formal inventory of the digital content, so the museum is not quite at Level 1 for the Metadata category. The museum typically creates JPEG files and has also used PDFs in the past. Because this is a limited set of formats, the museum has reached Level 1 in the File Formats category. Fortunately, there are some simple steps that the museum can follow to achieve higher NSDA levels. In the following sections of the plan, I will describe a range of actions that staff could take to better manage the Baltimore Community Museum's digital content.

Beginner Plan

The National Digital Stewardship Alliance Levels of Preservation stress that organizations should keep multiple copies of their content to prevent loss due to bit rot or storage system failure. Files should also be stored in multiple geographic locations to protect against a natural or man-made disaster in a particular region. Google Cloud and Dropbox are both solid options for the Baltimore Community Museum for the time being. These cloud storage providers can sync files to the museum's staff laptop. It is also likely that these systems are storing files in a location outside of the Ohio region, though this process is not entirely transparent. Currently, the museum saves some files in Google Cloud and others in Dropbox because the free version of Dropbox has limited storage. One option might be to combine all of the files in Google Cloud in order to simplify tasks like fixity checks and completing inventories.

It will be essential for the Baltimore Community Museum to begin checking the fixity of its files. Fixity refers to “the property of a digital file or object being fixed or unchanged” (National Digital Stewardship Alliance, 2014, p. 1). Fixity can also be thought of as a “digital fingerprint” that serves as evidence that the museum's files are the same as they were before (Owens, 2018, p. 60). The simplest way to monitor fixity is to keep track of the number of files being created and their expected file size. If the file size or file count unexpectedly changes, this can be a sign that there is a problem. In general, it is a good idea to check fixity information when the content is first created and before and after it is transferred to a different storage system. (National Digital Stewardship Alliance, 2014, p. 3-5). If museum staff do not have time to produce a full inventory of the digital content with detailed metadata, keeping track of the file size and file count numbers will still be better than nothing. One way to simplify this process would be to schedule a regular time each month to update the inventory and check fixity information. This would be a quick way to meet the Level 1 requirements in the File Fixity and Data Integrity category.

Currently, access to the Baltimore Community Museum's digital content is restricted to the museum's director and interns. The museum is also using a limited set of common file formats. Because these are stronger areas for the organization, staff may not need to make as many changes to improve their practices. The museum should consider creating formal documentation that describes the current access restrictions to reach Level 2 in the Information Security category of the NDSA levels. When interns leave the organization, any shared passwords should be changed. It would also be useful to maintain a log that employees update whenever they delete or move files. This would help the museum to meet the Level 3 requirements for Information Security. To reach Level 4 in this area, the museum's director can perform audits of the security logs.

The museum's director should also keep track of the file formats that the museum uses and encourage new staff members to continue creating JPEGs when scanning items from the collections. Normally, an organization would need to create an inventory of all the file formats

that are in use to reach Level 2 in the File Formats area. Right now, the museum mainly uses JPEGs and PDFs, but they could create an inventory in the future if they start using a larger set of formats. To arrive at Levels 3 and 4, organizations need to monitor file formats for obsolescence and should be prepared to migrate or emulate files. However, since JPEGs and PDFs are so commonly used, these additional steps might not be necessary for the Baltimore Community Museum. The staff should use their limited time to strengthen practices in other areas.

Intermediate Plan

Some National Digital Stewardship Alliance members have expressed concern that cloud storage systems do not allow organizations to maintain full control over their content (Altman et al., 2013). In addition to storing files in a service like Google Cloud, the Baltimore Community Museum may want to create an additional copy of its files that is not housed in a third-party system. The museum's files will likely fit on a USB drive or external hard drive. This additional copy could be kept in a place that is only accessible to the museum's director for further security. The museum could also consider finding a "backup buddy" in another part of the country. This would involve trading external drives with another institution to minimize the risk of losing all of the museum's copies in a regional disaster. No matter how the organization chooses to proceed, it will be important to document the storage system and to ensure that staff members know how to access all of the various copies. If the museum is able to maintain access to three complete copies in different geographic locations, they will satisfy the Level 2 requirement for the Storage and Geographic Location category.

In addition to keeping track of the expected file size and file count, the Baltimore Community Museum can also use cryptographic hashes to monitor file fixity. Cryptographic hash functions like MD-5, SHA-1, and SHA-256 are algorithms that "[take] a given set of data (like a file) and computes a sequence of characters that then serves as a fingerprint for that data" (Owens, 2018, p. 109). This may sound daunting, but it is possible to automate this process. AVP's [Fixity](#) tool is a free service that can scan folders or directories and check for fixity issues. The museum can ask Fixity to monitor files on a monthly basis and send email reports when it detects changes to the files (AVP, 2018).

Advanced Plan

In addition to maintaining multiple copies of its files, the Baltimore Community Museum could also consider uploading items to the Internet Archive, which offers free storage and some support for digital preservation. This option would be appropriate for content that is in the public domain and would require museum staff to add some metadata to the items it uploads (Schumacher et al., 2014). The Internet Archive could facilitate public access until the museum is able to add more collections to its website. While more advanced software solutions like Preservica would provide greater functionality, they would also be more expensive. Because the museum is not creating a massive number of files, free options like Google Cloud and Dropbox

should serve the organization's needs for the time being. An option like Preservica may be worth considering if the museum greatly expands its digitization program.

One of the Baltimore Community Museum's biggest challenges is figuring out how to organize and keep track of its digital content. Establishing regular file naming conventions could help to solve this problem. In order to improve practices in the Metadata category of the NDSA levels, organizations are supposed to store administrative, transformative, technical, descriptive, and preservation metadata. However, many of these types of metadata do not need to be created manually. For the time being, the museum could focus on creating an inventory of its files with the file location, fixity information, and descriptive metadata. Maintaining a log of fixity information is also one of the requirements for Level 3 of the File Fixity and Data Integrity category. If staff have time, they could begin adding individual scanned items to PastPerfect to gain better intellectual control over the museum's digital content.

Conclusion

Some of these options are more labor-intensive than others. Even if staff only have enough time to pursue the simplest recommendations, this would still be an important step towards actively managing the Baltimore Community Museum's digital content. Once the museum establishes basic procedures for making copies and checking file fixity, it will likely become easier to implement some of the other suggestions.

Digital Preservation Policy

Introduction

The Baltimore Community Museum is responsible for documenting the history of the town of Baltimore, Ohio and the surrounding areas. Because the museum is currently undertaking digitization projects, digital preservation is becoming increasingly important for the organization. Staff need to be able to maintain the valuable digital files they are creating and hope to be able to make the scanned items available on the museum's website in the future. The Baltimore Community Museum recognizes that digital preservation can be challenging because the policies and methods developed for analog preservation may not always be applicable. Because of the rate of technological change, digital objects require more active management. This digital preservation policy will help the museum to gain better control over its digital content today so the organization is better equipped to tackle new challenges (such as new media formats) in the future. This policy outlines procedures and responsibilities for the Baltimore Community Museum in order to ensure sustainable access to the museum's digital content. In the long run, these guidelines should help to clarify workflows for staff and prevent unnecessary stress. The actions listed in this policy are intended to help the museum reach higher levels of the National Digital Stewardship Alliance Levels of Digital Preservation.

Scope

This policy applies to the digital content created by museum employees during digitization projects. Currently, staff are working to scan documents and photos from the Baltimore Community Museum's collections. Issues of a local newspaper, the *Twin City News*, have also been digitized. Going forward, the museum will continue to select items to be scanned. The Baltimore Community Museum will prioritize documents and photos that are damaged and items that are of great importance to the history of the community. The organization also hopes to select and preserve items that would potentially assist members of the community working on genealogy projects. This includes family genealogies that have been donated to the museum, cemetery records, and township records.

Storage

The Baltimore Community Museum currently uses Google Cloud for photographs and Dropbox for documents. Going forward, the museum will maintain at least three copies of its files to prevent loss due to bit rot or storage system failure. At least one copy of the files will be stored in a different geographic location to protect against a regional disaster. The museum will continue to use cloud storage providers but will also store one copy of the scans on an external hard drive that will be kept in a secure location.

File Fixity

The museum will monitor the fixity of its files to ensure that digital files have not changed or degraded over time. The simplest way to monitor fixity is to keep track of the number of files being created and their expected file size. Each month, staff will update the total file count and file size figures. As part of this monthly check-in, staff will verify that fixity information has not changed unexpectedly. Staff should also double check fixity information after transferring any files to a new storage system. For additional peace of mind, staff can use AVP's free [Fixity](#) tool to scan folders or directories and check for fixity issues, but this does not need to occur on a monthly basis.

Metadata

Museum staff will create an inventory of its scans that is updated at fixed intervals. The inventory will at least include the file names, file locations, a description of each object, and any available fixity information. If time permits, staff will also begin upload information about scanned items into PastPerfect like they would for physical artifacts.

Information Security

Access to the Baltimore Community Museum's digital content should be restricted to the museum's director, interns, and other staff. Members of the community may view the files on a staff laptop with supervision, but passwords should only be entrusted to museum staff members. The museum will also maintain a log that employees update whenever they delete or move files. This log should be audited on a quarterly basis.

File Formats

The museum typically uses common file formats like JPEGs and PDFs for its scans. New staff members should be instructed to continue using these formats. If the museum decides to expand its digital collections in the future, staff will create an inventory of all of the file formats that are in use and monitor file formats for obsolescence.

Roles and Responsibilities

The museum's director will be responsible for leading digitization projects, securing funding for resources like external hard drives, and training other staff members on digital preservation tasks. Other staff members will assist with tasks like making backup copies of files, checking fixity information, and updating inventories.

Policy Review

Best practices in digital preservation continue to evolve, so this digital preservation policy may also need to be revised in the future. The policy will be reviewed annually to ensure that the document still meets the organization's needs. The policy should also be updated whenever there is a substantial change to the scope of the Baltimore Community Museum's digital collections (for example, if the museum begins digitizing audiovisual materials). As part

of the review process, the museum's director should consult with interns and volunteers to identify any workflows that need be modified.

Works Cited

- Altman, M., Bailey, J., Cariani, K., Gallinger, M., Mandelbaum, J., & Owens, T. (2013). NDSA Storage Report: Reflections on National Digital Stewardship Alliance Member Approaches to Preservation Storage Technologies. *D-Lib Magazine*, 19(5/6). <http://doi.org/10.1045/may2013-altman>
- AVP. (2018). Fixity User Guide Version 1.2. Retrieved from https://www.weareavp.com/wp-content/uploads/2018/07/Fixity_v1.2_UserGuide.pdf
- National Digital Stewardship Alliance. (2014). What is Fixity, and When Should I be Checking It? Washington, D.C. Retrieved from <http://www.digitalpreservation.gov/documents/NDSA-Fixity-GuidanceReport-final100214.pdf>
- Owens, T. (2018). *The Theory and Craft of Digital Preservation*. Baltimore: Johns Hopkins University Press.
- Phillips, M., Bailey, J., Goethals, A., & Owens, T. (2013). The NDSA Levels of Digital Preservation: An Explanation and Uses. IS&T Archiving, Washington, USA. Retrieved from http://www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf
- Schumacher, J., Thomas, L.M., VandeCreek, D., Erdman, S., Hancks, J., Haykal, A., ... Spalenka, D. (2014). From Theory to Action: Good Enough Digital Preservation for Under-Resourced Cultural Heritage Institutions (Working Paper). Retrieved from <http://commons.lib.niu.edu/handle/10843/13610>