CONTENTS

Executive Summary ................................................................................................................................. 2
Introduction to the Museum .................................................................................................................. 3
Survey of Digital Collections ............................................................................................................... 4
“Next Steps” Plan ............................................................................................................................... 8
Digital Preservation Policy .................................................................................................................. 12
Additional Resources ......................................................................................................................... 12
The purpose of this report is to give an overview of the digital holdings of the College Park Aviation Museum, to assess current preservation efforts, and to offer suggestions on how to expand and improve upon existing policies and practices. It begins with a summary of an interview I conducted with Curator of Collections Laura Baker, which was used as a baseline for understanding the museum’s digital collections. It is followed by a “next steps” plan based on the National Digital Stewardship Alliance’s Levels of Digital Preservation, which includes recommendations on storage, file fixity and data integrity, information security, metadata, and file formats. It concludes with a digital preservation policy that is intended to serve as a statement of basic principles to be used by museum staff and volunteers.

This report was created in 2018 as part of project for a graduate-level digital preservation seminar at the University of Maryland iSchool. It was written with input from Laura Baker and with feedback from iSchool Professor Dr. Trevor Owens, who is also Head of Digital Content Management at the Library of Congress.

Whenever possible, scalable or incremental solutions have been suggested so that the museum can start taking action with the staff and resources it has available. It is also important to emphasize that this report should be considered a living document, as digital preservation is an on-going, iterative process. Museum staff are encouraged to adapt the document as the museum’s digital collections grow and evolve, or in response to new technology or standards.
The College Park Aviation Museum (CPAM), located less than ten minutes from campus, is an institution dedicated to “promoting aviation innovations at College Park Airport and in Prince George’s County while fostering research, inventiveness, and lifelong curiosity about the history and science of flight.” The museum is on the grounds of the College Park Airport, which has the distinction of being the site where Wilbur Wright taught the first military aviators to fly in 1909, and because it has been in operation ever since, it is known as the world’s oldest continuously operating airfield. Inside the museum’s exhibition space, a visitor can encounter historic aircraft and artifacts related to local aviation history while also getting an unobstructed view of the airport’s runway through a large wall of windows. The museum is a great destination for families, with numerous hands-on activities for children and a balcony overlooking the runway where visitors can bring picnic lunches.

In keeping with its mission to preserve local aviation history, the museum has a small research library which contains materials focusing on the people, aircraft, and events associated with the College Park Airport and aviation in nearby Prince George’s and Montgomery Counties. Additionally, the library holds general materials on the development of civil and military aviation, significant aviators, and aviation technology, as well as a collection of children’s books dating from as early as 1910. Complimenting the books, magazines, and audiovisual material in the library, their archival holdings include photographs, negatives, newspaper articles, letters, logbooks, pamphlets, posters, scrapbooks, and other records. Significant topics represented include Airmail, the Benjamin Foulois Collection, The Cloud Club and Columbia Air Center, the Elaine Harmon Collection (Women’s Air Service Pilots), Engineering and Research Corporation (ERCO), Preservation of College Park Airport Collection, and the U.S. Army Signal Corps Aviation School.
To learn more about the digital holdings of the organization, I interviewed Laura Baker, who began her career at the CPAM as a Museum Educator two years ago and was recently promoted to Curator of Collections. Laura was very enthusiastic to have CPAM be one of our class’s digital preservation projects, because she wants CPAM to be a forward-looking, modern institution with as much of their holdings digitized and made accessible as possible. Baker is new to her position and has recently been focusing on staffing changes and important projects like the creation of a new air mail exhibit, so she has not yet had the time to do a full survey or audit of the museum’s digital holdings. Our interview was fortuitously timed because she said she is looking forward to focusing more on the collection and its priorities. The interview questions were a useful starting point for thinking about what digital material the museum has, how it is currently being managed, and what steps need to be taken to preserve it.

There is no complete inventory of library or archival holdings that can be consulted to identify the museum’s digital holdings, although individual inventories exist for specific categories or digitization projects. Laura reported that the bulk of the museum’s digital holdings are in photos. There are, for example, a large number of airmail photos that have been scanned at 600 dpi in conjunction with the recent exhibit. There is also a collection of historic photos that are lower quality jpegs and have been problematic for the museum because no one knows where these photos are originally from. They seem to be copies of photos from other institutions, and researchers will sometimes contact CPAM for copyright clearance for these photos, but the museum does not know the status of these photos or
who to contact for the proper permission. An educator is currently working to try to solve this problem. In addition to historic photos, there are also a number of contemporary photos of museum events, like the recent Airmail Centennial. These photos are stored on the common drive, along with a small amount of scanned documents like airmail pamphlets and letters. A limited number of these photos are also on the CPAM website, but the website is hosted through the Maryland-National Park and Planning Commission and CPAM does not have full control of their website. This puts some frustrating limits on how many photos they can share and how they can be displayed.

In the past, the museum partnered with Digital Maryland (DM) to digitize and host their ERCO photo collection. When I asked Laura about how this relationship worked and if it was something they would like to continue in the future, she said she did not know but was eager to investigate. The relationship with DM was lost after staffing changes, but the Digitization Supervisor, Linda Tompkins-Baldwin, recently visited the museum to explain how DM works and renew CPAM’s ability to post directly to the DM site. She also offered to help the museum in whatever way they needed throughout the digitization process. From the museum’s perspective, there are significant benefits to partnering with DM again. It allows CPAM to put more content online without hitting roadblocks from the Commission, it serves as an additional backup of digitized files, and it comes at no cost to the museum. DM will pick up the materials from the museum and do all the digitization work so that CPAM can focus on other projects. The downside to partnering with DM is that it is only a collections database and cannot serve as a platform for digital exhibits or other types of interpretive content that the museum would like to produce. Also, if the museum wanted to charge fees for use of the photos, DM works on an “honor system.” Laura does not see these issues as a reason not to partner with DM again, but if the museum upgrades from PastPerfect into something with more online capability, she believes they could have a more meaningful digital collections presence than what DM offers. She also cautioned that if the partnership is restarted, it needs to be
integrated into the daily operations for collection management so that it will survive any future staffing changes.

In addition to digital photos, the museum has a number of materials in other formats. There is a wall of VHS tapes, DVDs, and cassettes, which mostly chronicle events that have taken place on the airfield (like air shows) or movies shown at the museum for a program. The museum plans to migrate AV materials in obsolete formats to digital files in the future. Moreover, there are CD-ROMs with contemporary photographs of CPAM events and both floppy disks and USB-drives that contain unknown content. After our interview, Laura was able to locate an inventory of these various formats. The museum does not have any old hard drives or computers—Laura stated that it is their policy to get rid of old equipment when it is no longer needed, transferring the files stored there to newer systems. CPAM has been collecting oral histories in conjunction with the Library of Congress (LOC) Veterans History Project. When asked if the museum keeps a copy of those oral histories after submitting them to LOC, Laura consulted with the person in charge of the project. She discovered that the most recent oral histories are stored on the common drive and the older ones are stored on DVDs. There is a hard copy list of the oral histories with the names and dates of each.

When asked how the digital content is being managed, Laura said there currently is no plan in place for preserving digital content. After reviewing the National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation together, we concluded that the museum is not yet at level one in the five categories. Employees are generally conscientious about labeling files in ways that make them easy to find, creating file structures that make sense, and refraining from “cluttering” the drive with unnecessary files. Laura reported that moving forward, new policies are in place for naming conventions, and educators are sorting through the common drive to identify photos. There are no file fixity and data integrity checks on the museum’s digital content, and the files on the common drive are not set for restricted access because the museum has had problems in the past with employees restricting access to administrative files and not telling others how to access those files when they leave the museum’s employment. In regards to metadata, inventories of digital content tend to be created
on a project-by-project basis and it seems some content produced in the past may not be inventoried at all. They also use the collection management system PastPerfect, and Laura would like to see more complete and accurate records kept of all objects in the collection, admitting that this has not always been done well in the past.

On the museum’s perceptions of the state of their digital content, Laura said that right now the museum’s digital content is limited and not much is being done with it, but the institution is eager to do more. She assessed the creation and preservation of digital content as not quite “mission critical,” but “definitely a priority,” and actions are already being taken in that direction. The museum’s director believes that digital content is an important part of how museums operate today and wants to see CPAM catch up to what other institutions are doing. Meeting this goal, however, sparks important questions—how does the museum do it, how do they pay for it, and who is going to be the person(s) doing the work? These practical questions can sometimes constrain CPAM’s ambitions, as they (like nearly all museums) have only limited time, equipment, and funds.

When asked what digital content the museum is not currently collecting but would like to, Laura said the museum has made it a goal of diversifying their collections. They operate in a county where the majority of residents are African-American or Latino, and they would like to see more people of color represented in their collection. They also want to collect more material related to women and aviation. The goal of diversification is not one explicitly linked to digital content, but if there are digital materials they could be collecting that would make their collections more inclusive, that is something they would like to pursue. They would also be interested in archiving digital content related to exhibits and events that are happening at the museum and the airport. The museum does not have a blog, but they have written blog posts for the Smithsonian National Postal Museum that they would like to see preserved.

The museum does not have an archivist on staff, so the digital materials fall under Laura’s purview, but there are plans to hire a Collection Assistant and an Assistant Director to assist with projects that might include digital preservation in the future. Laura would also like to initiate a relationship with the University of Maryland to get one to three interns working in the museum on a semester-by-semester basis, although to do so she feels the museum needs more support and equipment. Right now, for instance, the museum only has one computer that would be available to interns. Laura would also like to get better digitization tools and train volunteers or interns on how to use them. There is only so much that the museum can put on their common drive, however, so they would also need more storage space if they were to significantly increase their volume of digital content. To both sustain existing content and to produce more, the museum would have to consult with Prince George’s County and various other funding streams, and perhaps look into applying for new grants. Given the current limits of their staff, equipment, and budget, Laura believes its necessary to get a grant to outsource the creation and maintenance of their digital content or to use a free service like DM.
Based on the information gathered from my interview with the Curator of Collections, Laura Baker, this section of the report will offer recommendations for the next steps the museum might take to improve their digital preservation strategies. The recommendations are based on the National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation, which provides succinct, clearly stated guidelines across five areas of concern for digital collections—storage and geographic location, file fixity and data integrity, information security, metadata, and file formats.

For each category, the NDSA Levels provide four levels of progressively more advanced suggestions for digital preservation. This is helpful for an institution like CPAM, which is in the early stages of developing a digital preservation plan, because it allows for incremental change and takes into account the fact that many institutions have only limited staff and resources with which to begin addressing the needs of their digital collections. Dr. Owens, our professor and one of the authors of the NDSA Levels, emphasized that just getting an institution to level one is a significant accomplishment, and so we shouldn’t think of an institution as only succeeding in digital preservation if they are at level three or four. With this in mind, I am recommending that the CPAM try to achieve level one in all five categories, while also offering suggestions for moving farther along the levels should they decide now or at some point in the future to adopt a more ambitious plan.

**Storage**

One of the basic principles of digital curation is to keep multiple copies of all digital files to protect against bit rot, system failures, or regional threats (like natural or man-made disasters). Anyone who has ever lost files after a computer or hard drive has failed will understand why this step is so important. CPAM relies on their common drive to store a large portion of their digital files. As a first step, CPAM should consult with their parent organization, the Maryland-National Capital Parks and Planning Commission’s (MNCPPC) Office of the Chief Information Officer (OCIO) to determine what measures are already in place to back up the contents of their common drive and how often these backups are performed. To satisfy level one of the NDSA levels, CPAM should ensure that there are at least two complete copies of the contents of the common drive and that these copies are not located in the same place.
CPAM should also aim to get the images, videos, or other data that now exists in various formats into their common drive. This is a crucial step because formats like CDs, DVDs, USB drives, and VHS tapes are not made to last forever and can become damaged or lost. Migrating the media on these formats to a storage system will enable them to be stored and backed up in a way that will improve their longevity. It will also help to unify the collection in one place to simplify its care and management. Laura expressed concerns during our interview about running out of space on the common drive, so it may be necessary to ask MNCPPC for more space. If necessary, the museum could continue to use the common drive for administrative files but use another storage system like Dropbox (which starts at $12.50 a month) for media files like photographs, videos, or scans.

The two steps listed above would meet the minimum threshold for NDSA level one, but by putting more of its collections online, which is already a stated goal of CPAM, the institution could reach level two or three. The most economical way to do this is through a partnership with Digital Maryland, which will collect materials from the museum, digitize them if they are not already digital, and host the scanned images on their site. CPAM should ensure that they get copies of all digital files so that Digital Maryland and CPAM can serve as safeguards for one another in case a back-up of a file is needed. The partnership with Digital Maryland can thus serve as a digital preservation strategy while also producing more online content.

If the museum wants to explore other options, it might also look into uploading some of its digital content to the Internet Archive or Wikimedia Commons, both of which would host the material for free. There would be more labor involved with uploading the files (and digitizing them if they are not already digitized), but these sites have the advantage of reaching a wider audience. CPAM might look to the Internet Archive’s American libraries collections or Wikimedia Commons collections to see options of how their material might be presented on these sites. The caveat to uploading more content online is that these files need to be in the public domain or the museum needs to clear any potential copyright issues.

Additional steps needed to complete level two and three would be to document storage system(s) and media formats and what is needed to use them (level two) and establishing a process for monitoring the obsolescence of the storage system(s) and media formats (level three).

**File Fixity and Data Integrity**

The concepts of “file fixity” and “data integrity” are not particularly well-known outside of the IT sector, but what they mean in layman’s terms is making sure that a digital file has not been altered or corrupted. In other words, is the institution preserving the file they intended to preserve? For a small institution like CPAM, which does not have a digital preservation specialist, it may initially seem challenging to know how to get started, especially when jargony terms like “check sums,”
“cryptographic hash function values,” or “digital signatures,” are bandied about. Fortunately, there are free tools that exist for automated monitoring and reporting on data integrity.

One service that Dr. Owens recommended is AVP’s Fixity. According to the website, “Fixity scans a folder or directory and creates a manifest of the files, including their paths and their checksums, against which a regular comparative analysis can be run. Fixity monitors file integrity through the generation and validation of check sums, and file attendance through monitoring and reporting on new, missing moved, and renamed files.” The user can use the tool as needed or chose to schedule these tasks daily, weekly, or monthly, setting the specific day and time that they automatically occur.

Since CPAM is new to this type of preservation work, I would recommend downloading Fixity, exploring the settings and experimenting with how it works, and then adopting a plan that seems realistic given their goals and priorities. To reach level one, CPAM would need to check file fixity when new content is added to their storage system if that information is available, and it should create the file fixity information if it doesn’t already exist. To reach level two, CPAM would need to check fixity on all ingests, use write-blockers on original media, and virus-check high risk content. Level three requires fixity checks at regular intervals, maintaining logs of fixity information, detecting corrupt data, and virus-checking all content. If these steps sound too demanding, Dr. Owens has said that doing a fixity check once a year is better than doing nothing at all.
Information Security

To reach level one on information security, CPAM should identify who has the ability to read, write, move, and delete individual files, and restrict those authorizations when appropriate. To illustrate why these steps are necessary, consider what might happen if an employee resized an image file to make it fit on the website or in an email newsletter but accidentally overrides the original file. Or maybe a volunteer saves a file with a file with a generic name like “Oral History,” and accidentally erases an older file that was on the common drive with the same name. Limiting the number of people who can perform these types of tasks with the digital collections will mitigate the possibility of accidents like these occurring. From my interview with Laura, I know that there is some hesitation to adopt file restrictions because there were past incidences in which files were restricted and were no longer accessible when an employee left the museum. These potential problems could be largely avoided, however, by documenting access restrictions (level two) and speaking with the MNCPPC’s OCIO to discuss how to override access restrictions associated with former employees.

A more rigorous approach would involve maintaining logs of who performed what actions on files (level three) and performing audits of these logs (level four).

Metadata

One of the challenges encountered in my partnership with CPAM is that there is no master inventory of CPAM’s digital collections. The museum has inventories of material digitized on a project-by-project basis, but because they are not unified in one place, finding the inventories may involve asking the person(s) in charge of the project to locate them, and it seems like some materials may have “fallen through the cracks” and may not be inventoried at all. My recommended first step is to track down as many inventories as can be located and put them into one folder on the common drive. Then create a list of material that still needs to be inventoried and put that task on a to-do list. The next step would be to see if these multiple inventories could become one master list, either through copying and pasting data into an excel workbook or logging material into past perfect. If this is not practicable, then maybe a “Guide to Digital Collections” could be created that would explain where to find each inventory and a summary of what each inventory contains.

Going to NDSA level two would require creating administrative metadata for digital files (such as when and how it was created, and who can access it) and transformative metadata (logging any changes to the file). Level three requires storing standard technical and descriptive metadata about the digital files.
Digital Preservation Policy

File Formats

When creating new digital files, it seems that the museum is inclined to use standard, popular formats which increases the likelihood that the museum will still have access to these files even if the formats becomes obsolete. If a popular format like PDFs become obsolete in the future, for example, IT specialists will have to invent tools for accessing and migrating these files because they are so common.

NDSA level one requires that the museum give input whenever possible into the creation of digital files to encourage the use of preferred formats. The Smithsonian has a policy on digital formats that includes a useful table that the museum might want to use as a model. To achieve level two, CPAM should inventory the file formats currently in the museum’s collections, and to reach level three, it should monitor file format obsolescence issues.

Conclusion

To conclude, I’d like to acknowledge that these suggestions may initially seem overwhelming for a museum with limited staff and resources, but that the museum should keep in mind that they are not expected to implement all these steps at once. While it would be ideal for the museum to reach level one in all five categories, even moving to level one in a few of the five categories would be better than doing nothing at all, and steps corresponding to levels two or three can be viewed as “stretch goals” for the future. Overall, these suggestions should be seen as flexible guidelines, and perhaps also as ideas for future internships or grant applications so that the museum’s current resources do not get overtaxed.
Purpose

As the College Park Aviation Museum (CPAM) seeks to expand its collections and facilitate greater access to its historic materials, a growing portion of its holdings will be either born-digital or digitized material. This document was created to establish guidelines on how to manage and organize these materials and to protect them against the risks of loss or technological obsolescence. In following these guidelines, the museum will be actively working to ensure that its digital collections continue to be available for future generations.

Scope

This policy covers both born-digital and digitized material. Born-digital refers to material originally created in digital form, such as photos uploaded from a digital camera or text files created using Microsoft Word. Digitized materials refer to any material that has been converted to a digital format, such as scanned photographs or documents.

While many of the museum’s digital files are on the common drive, this policy also encompass material currently stored in other formats such as USB-drives, floppy disks, CDs, and DVD+Rs. Although certain media items like VHS and cassette tapes are not typically put into the same category as digital files, these are “endangered” formats that are obsolete and have limited lifespans, so the best way to ensure ease of access and continued preservation is to convert them to digital files. With the expectation that their conversion will be a priority they should be considered to be within the scope of this policy whenever applicable.

Standards

This policy draws from the National Digital Stewardship Alliance (NDSA)’s Levels of Digital Preservation. The NDSA Levels were chosen as a basis for recommendations because it provides succinct, clearly-articulated standards and encourages an incremental, scalable approach to digital preservation. It recognizes the need to provide realistic options to institutions with limited time and resources, which is the goal of this policy as well.

Storage

CPAM will consult with their parent organization, the Maryland-National Capital Parks and Planning Commission (MNCPPC)’s Office of the Chief Information Officer (OCIO) to determine what measures are already in place to back up the contents of their common drive and how often these back-ups are
performed. CPAM will ensure that there are at least two complete copies of the contents of the common drive and that these copies are not located in the same place.

CPAM will make it a priority to get the images, videos, textual files, and other data that now exists in various media formats onto its common drive, which will simplify its care and management. It will phase out the use of formats like CDs, floppy disks, DVD+Rs, and USB-drives as storage, because those formats have a limited lifespan and are prone to damage or loss.

If CPAM runs out of storage space on its common drive and MNCPPC is unable to provide additional resources, CPAM will continue to use the common drive for administrative files and will use a secondary cloud storage system like Dropbox for media files like photographs, videos, and scans.

CPAM will continue to pursue a partnership with Digital Maryland (DM), which will enable them to continue to expand their digital holdings while also making them more accessible to the public. The files hosted by DM can serve as additional copies for preservation purpose, but CPAM will also keep two copies of all digitized files in its own storage systems.

**File Fixity and Data Integrity**

CPAM will use a tool like AVP’s Fixity to perform data integrity checks on all its digital files at least once annually, and will also aim to perform checks after large ingest or transfers of files. If the checks indicate that files are missing, unintentionally altered, or corrupted, the files will be restored using one of the museum’s back-up copies.
Information Security

CPAM will identify who has the ability to read, write, move, and delete files and restrict those authorizations when appropriate. Restrictions will be recommended for all historic collections, although there may be a desire for more flexibility with administrative files. These restrictions will be documented on a spreadsheet and periodically reviewed.

Metadata

CPAM will create a mastery inventory of CPAM’s digital files that are deemed worthy of preservation, including those on the common drive and on various other formats such as DVD+Rs, CDs, floppy disks, or USB-drives. CPAM will also inventory all media formats such as VHS tapes or cassettes which are either obsolete or in danger of obsolescence, so that these materials can be prioritized for digitization. This can take the form of a single spreadsheet or multiple spreadsheets that are organized within one folder.

CPAM will chose the level of description most appropriate to the material being described and the amount of resources it is able to devote to the task, which realistically may not always be a file-level description. For the initial inventory, it is sufficient to say “1 CD containing 150 images of CPAM events, circa 2005,” which can be expanded in more detail at a later date if desired. At this stage, it is more important to get a general understanding of the scope and contents of the collections (including file formats) than it is to create detailed file-level metadata.

CPAM will agree on naming conventions for all born-digital and digitized files that will be added to its collection in the future. If time and resources allow, it will also standardize its legacy holdings.

File Formats

CPAM will encourage the use of preferred file formats for the creation of new digital materials. Preferred formats are ones that are commonly used, widely accessible, and open-sourced. A guideline for determining what digital preservationists consider to be a preferred formats is the Smithsonian Institution Archives (SIA)’s Recommended Preservation Formats for Electronic Records.

For text documents, spreadsheets, and presentations that are in their final form (and will not be edited), the preferred formats are PDF or PDF/A. For images, the preferred format is TIFF (uncompressed), for audio it is BMF-Broadcast WAV (.wav extension), and for video it is Motion JPEG 2000, MOV, or AVI. For a more complete listing of both preferred formats and acceptable formats, CPAM employees are encouraged to consult the table in the SIA guidelines.
CPAM will create a list of file formats in its collection and monitor each for obsolescence issues. Converting files to the preferred formats is only recommended for files in formats that are not considered acceptable by SIA.

**Review**

Digital preservation is not something that is done once and then forgotten about, but rather it is an ongoing, iterative process. In recognition of this fact, the digital preservation policy should be reviewed annually by staff members of the CPAM and revised when necessary. The museum should anticipate that evolving technology and standards may create the need to amend the policy. It may also need to be adapted in response to changes in the museum’s goals or priorities.
**Articles**


https://www.windowscentral.com/how-rename-multiple-files-bulk-windows-10


“Recommended Preservation Formats for Electronic Records.” *Smithsonian Institution Archives*.  
https://siarchives.si.edu/what-we-do/digital-curation/recommended-preservation-formats-electronic-records

http://commons.lib.niu.edu/handle/10843/13610

**Tools**

Digital Preservation Step by Step (Orbis Cascade Alliance)  
https://orbiscascadecccd.github.io/digprezsteps/index.html

Dropbox  
https://www.dropbox.com/?landing=dbv2

Fixity (AVP)  
https://www.weareavp.com/products/fixity/