Digital Asset Management Functional Requirements Report

Go For Broke National Education Center

Prepared by Savannah Lake
June 2020

A. Organization

Go For Broke National Education Center (GFBNEC) is a museum and archive that seeks"to educate and inspire character and equality through the virtue and valor of our World War II American veterans of Japanese ancestry" ("Our Mission"). GFBNEC's educational and archival work is directly tied to this mission, as the organization collects and highlights stories often excluded or minimized in the dominant narrative of World War II, creating a more accurate and inclusive historical record. The center also maintains the Go For Broke Monument, which features the names of over 16,000 Japanese Americans who served during World War II. Archival collections at GFBNEC are available for research, featured within the museum itself, and incorporated into education materials and outreach.

GFBNEC has digital assets by virtue of its archival holdings, which are partly born-digital and have been partly digitized to promote both access and preservation. GFBNEC also has digital assets by virtue of it being a modern organization, with the majority of its work being completed digitally. However, this report will cover the archival digital assets only, per the recommendation of the senior archivist at GFBNEC. Currently, each department—such as finance, human resources, and development—is fairly siloed. Trying to force everyone's diverse workflows and types of assets onto one system would require too much change management and training (Do). Further, there would be minimal added benefit of having reports, financials, and other documents on a digital asset management (DAM) system, as there are less concerns with these documents about long-term preservation, description, and reproduction rights. These documents could be well organized and managed within the shared drive and backed up in Dropbox, following a clear folder structure. Instead, this report will focus on finding an appropriate DAM for the digital assets generated by and/or maintained by the archives department, as these assets require long-term preservation, and require robust description in order to facilitate discovery, preservation, and legal reuse.

With regard to content, GFBNEC primarily generates digitized surrogates of their analog collections as well as audiovisual oral histories. The digital surrogates include scanned photographs, correspondence, government records, patches, scrapbooks, and illustrations, all coming from their collections of materials from World War II Japanese American veterans. The oral histories are a major part of GFBNEC's archival holdings as well, and were created in the late 1990s through approximately 2012. The oral histories feature veterans speaking to their life experiences, and include videos and at times transcripts. And finally, a significant part of their digital assets consist of items that should be incorporated into their institutional archives; namely photographs of events and community outreach.

B. Requirements
1. Users

The staffing and structure of GFBNEC is an important factor for determining both the needs and capacity of the organization with regard to digital assets. GFBNEC is a nonprofit organization with eleven full-time operational staff, as well as three museum front desk attendants ("Staff"). Accordingly, each department is fairly small, comprised of one to three people. The organization also has a strong volunteer base to help with projects and events.

Most relevant for the administration of a DAM are the archives and information technology departments, both of which are comprised of one staff member each. However, there are a number of departments that use the digital archival assets, including education, exhibits, and communications. Users of the DAM are as follows:

- *Senior Archivist* - the senior archivist has the most knowledge about and responsibility for the digital assets that would live in the DAM. As such, the senior archivist would be responsible for the implementation and ongoing administration of the DAM. Accordingly, the senior archivist would have the broadest rights within the DAM, including administrative rights (the right to create users and set rights), ingest rights (assets and metadata), search rights,

download/publish rights, and preservation rights (convert formats, fixity checks, delete assets). As the senior archivist is directly responsible for the accession paperwork and deeply knowledgeable about the archival assets, they would also have legal rights (intellectual property assignment); in fact, because the intellectual property status of an item can have legal and financial implications if mistakenly applied, the senior archivist (as well as the president and CEO of the organization) should be the only ones with legal rights. This will help ensure responsible and knowledgeable stewardship. Similarly, the senior archivist, along with the president & CEO, are the only users that should have permissions to delete files, as this right can have serious ramifications and should be limited as much as possible.

- *Communications Department: Public Relations Manager, Development & Events Manager, and the Vice President of Development and Strategic Initiatives* - the communications department will use the DAM to source content for GFBNEC's outreach, including their email newsletter and social media platforms. Accordingly, it is important for the communications department to be able to search and source relevant assets, confirm that they are cleared for reuse, and download and publish them. While the senior archivist would be primarily responsible for asset and metadata ingest, there are current GFBNEC events (such as galas) that are photographed, and considered to be institutional archives. Since members of the communications department may interface with event photographers, they would also be given asset and metadata ingest rights, with the senior archivist assigned to review their submissions and approve them before ingest (Do). The communications department would not have legal or preservation rights.
- *Education Department: Director of Education & Exhibits and Education Manager* - the education department uses digital archival assets within exhibits and within educational outreach (such as school field trips and online resources). Similar to the communications department, they need to be able to search assets, confirm they can be used, and download and publish them. Additionally, because there may be photographs taken of current exhibits or education events, the education department should be given asset and metadata ingest rights, with the senior archivist assigned to review their submissions and approve them before ingest (Do). The communications department would not have legal or preservation rights.
- *President & CEO* - the president & CEO of the organization sometimes utilizes assets for speaking engagements (Do). Accordingly, it is important that the president & CEO have search and download/publish rights. Additionally, given the president & CEO's high position in the organization and involvement should be an issue with legal reuse, they should also be given legal rights. While it is unlikely that the president & CEO would be using such rights on a day-to-day basis, assigning said rights provides a back-up, should the senior archivist ever be away from the office or unavailable. Similarly, the president & CEO is the only other user aside from the senior archivist with the right to delete files, given the serious implications of that right being mistakenly or purposely misused.
- *Website/IT Coordinator* - the website/IT coordinator would be the sole technical support for the DAM, and would serve as consult to the DAM administrator (i.e. the senior archivist) on all systems issues. Accordingly, the website/IT coordinator would have administrative and preservation rights, as well as the basic rights afforded to all users of the DAM (ingest, search, download/publish). The website/IT coordinator would not, however, have legal or delete rights, as the senior archivist and president & CEO are the ones ultimately responsible for the legal stewardship of the assets.

In a larger organization, some of the more sensitive rights, such as legal rights and delete rights, would be spread across several staff with the same responsibilities (i.e. 2-3 IT staff or 2-3 archivists) in

order to create checks and balances that would mitigate any intentional misuse. However, given the small size of GFBNEC, this is not an option. Just as the analog collections are entirely under the purview of one staff member, the rights concerning the digital assets will similarly all reside with the senior archivist.

2. Content Assumptions
There are 94,352 digital assets related to the archives on the shared drive. More specifically, the "Archives Workspace" folder on the shared drive contains digitized analog collections and materials related to the oral histories, totalling 24,633 files and approximately 398.453 GB. Approximately 90% of these files are photographs (359.832 GB), with about 20% of the photographs related to oral histories and the remaining files consisting of digital surrogates of the archival assets. Only 3% of the files (11.591 GB) are videos related to the oral history project and a 2014 Japan trip, meaning the video assets are mostly not on the shared drive. The access copies of these videos are on the organization's Omeka page, while access and preservation copies are hosted by the University of Southern California as part of community partnership (Do). The remaining 7% of the files (27.03 GB) in the "Archives Workspace" folder are various spreadsheets, documents, presentation slides, database files, and other miscellaneous files used operationally by the archivist.

In addition to the "Archives Workspace" folder on the shared drive, the "Photos" folder contains relevant digital assets—mostly more contemporary materials, such as photographs of GFBNEC galas and outreach events. This folder contains 69,719 files and is approximately 893.027 GB. The vast majority of this content is event photos, comprising 93%, or 831.203 GB. The remaining files consist of database files for the photos, short videos for events, presentation slides, and other miscellaneous files.

3. Formats Supported
Formats
There are 83 file formats on the shared drive. Below is a snapshot of the twenty most prevalent in terms of storage size.

Format	Number of files	Size (bytes)	Percentage of storage
tif	19316	803.62 GiB	66.85%
jpg	63705	300.32 GiB	24.98%
mov	130	21.86 GiB	1.82%
mp4	73	16.50 GiB	1.37%
avi	62	14.05 GiB	1.17%
bkf	1	9.61 GiB	0.80%
cr2	258	7.76 GiB	0.65%
psd	509	5.40 GiB	0.45%
jpeg	466	5.26 GiB	0.44%

mdb	16	3.98 GiB	0.33%
bak	11	2.83 GiB	0.24%
asc	4	2.35 GiB	0.20%
nef	204	2.18 GiB	0.18%
pdf	739	1.98 GiB	0.16%
flv	192	825.79 MiB	0.07%
db	982	504.72 MiB	0.04%
pub	2	471.57 MiB	0.04%
rm	234	351.64 MiB	0.03%
mts	2	345.14 MiB	0.03%
eps	71	255.60 MiB	0.02%

Similarly, below is a snapshot of the twenty most prevalent in terms of the number of files.

Format	Number of files	Size (bytes)	Percentage of storage
jpg	63705	300.32 GiB	24.98%
tif	19316	803.62 GiB	66.85%
doc	1770	210.57 MiB	0.02%
xls	1366	212.38 MiB	0.02%
txt	1125	38.99 MiB	0.00%
db	982	504.72 MiB	0.04%
pdf	739	1.98 GiB	0.16%
csv	542	3.80 MiB	0.00%
xml	514	2.91 MiB	0.00%
psd	509	5.40 GiB	0.45%
jpeg	466	5.26 GiB	0.44%
docx	394	64.02 MiB	0.01%

dat	273	1.91 MiB	0.00%
cr2	258	7.76 GiB	0.65%
rm	234	351.64 MiB	0.03%
nef	204	2.18 GiB	0.18%
flv	192	825.79 MiB	0.07%
ini	187	25.93 KiB	0.00%
mov	130	21.86 GiB	1.82%
mht	119	16.46 MiB	0.00%

I reviewed codecs for the two most common audiovisual files, .mov and .mp4 files. The .mov files are comprised of ProRes video codecs and PCM audio codecs. The .mp4 files are comprised of Advanced Video Codec (AVC) video codecs and Advanced Audio Codec Low Complexity (AAC LC) audio codecs.

Directory conventions

There are some opportunities for improving the folder directory within the "Photos" folder. At the first level, the folders are arranged by year, which is a helpful enough facet, so long as staff have a strong institutional knowledge of what year an event happened. However, the second level has some variation, with some being organized by photographer ("2015\Photos by Jason Kusagaya\Rose Parade GFB_1-1-2015"). It is less likely that the average staff member would know which photographer photographed what event; instead, listing the event first and the photographer after facilitates better browsing ("2015\EOA\shane sato"). Further, not all folders have the photographer's name. While having the photographer's name is helpful in that it allows you to know who to attribute, full reuse rights by photographer and by event should be listed in detail in a centralized place, as not enough context is truly given by inserting a photographer's name into the file directory structure. And finally, when including dates in folder and file names, it is best to format it by year, month, date in order to ensure effective item sorting (eg. YYYY-MM-DD or YYYYMMDD).

The "Archives Workspace" directory structure is helpful in that it allows for browsing by collection and separates access copies from preservation copies (Archives Workspace\Digital Collections\2016.006 - Gohata\Album 3\TIFs). However, best practice is to eliminate spaces, to enable any scripting and autogeneration of metadata and information. This practice should be applied across the drive.

File naming conventions

With regard to file naming conventions, there is something of a spectrum with how uniquely and descriptively items have been named. Items within the "Photos" folder exhibit the most issues, with many instances of names like "D6C_0007.JPG, D6C_0008.JPG, etc" and "IMG.0295.jpg, IMG.0296.jpg, etc." This is likely because although these assets are considered institutional archives, they have been largely handled by nonarchival staff who are not trained in the importance of unique identifiers and file naming conventions. These file names should be updated to include important information that would help sort and differentiate the files, such as the date and a unique identifier for the event type.

Assets within the "Archives Workspace" folder have a bit more structure, but there is still room for optimization. File names for the digital surrogates of analog collections incorporate unique identifiers that link the item to its parent collection. While a file name is not a true unique identifier as it can be changed and thus does not have fixity, it does incorporate a unique descriptor that is used across resources—including file names, spreadsheets, and finding aids—that connects all related items together (Tadic, Class 5 - DAM vendors and open source options; OAIS; Digital workflows). For the collection items, this unique descriptor is the accession number of the analog collection, which is included in each file name of the digital surrogates.

There is some variation in naming across collections that could be standardized; for example, newer file names include whether or not the file is an access, mezzanine, or preservation copy, as well as the folder within the collection that the item came from. For example, "2016.013_010_007_acc.jpg" has the collection's unique descriptor, the accession number (2016.013); the folder number (010); a number for the scan itself (007); and a designation as an access copy (acc). Older names, such as "2014.001_160.JPG" do not reveal folder number or what type of copy the item is. It is possible there was only one folder within the analog collection, but that is unlikely at 160 items, and in either case should be indicated, for both clarity, context, and easier retrieval if the analog version is ever requested.

However, not all digitized collections follow this structure. File names like "SKSB_0640.tif, SKSB_0641.tif, etc." can be found, likely in the state it was received from the imaging company and not yet changed due to a backlog of work. More robust file naming, akin to what was done with the 2016.013 collection, should be applied across all collections to ensure discoverability.

Generally speaking, names should not have any spaces or special characters, which can be found throughout both the "Archives Workspace" and "Photos" folders ("Dad's resume picture187.jpg" and "T.Sgt. James Mitsuda(Co.C) home in Hilo.jpg", for example). Separating aspects of a file name by an underscore allows you to create scripts to automatically extract metadata about files, which could be especially helpful for ingest into a DAM.

4. Metadata

Much like the file naming and directory conventions, practices surrounding metadata vary significantly between the archival assets within the "Archival Workspace" folder and the institutional archival assets within the "Photos" folder, with more standardization and attention paid to the former as these assets are managed by the senior archivist while the "Photos" folder is largely managed by nonarchival staff.

Metadata conventions

Assets within the "Photos" folder are primarily managed by nonarchival staff, resulting in less consistent and robust metadata. The file naming within the folder gestures at metadata, capturing the event date and at times the photographer (Do). However, this practice is inconsistent, and critical technical, administrative, and descriptive metadata is missing, including rights for reuse, description, and format. No metadata standards or controlled vocabularies are used, and the metadata lives only in the directory names.

Within the "Archival Workspace" folder, assets primarily comprise of digitized analog collections and oral histories. Metadata for the digitized collections includes:

- *A master accession log*, with descriptive and administrative metadata for each collection (such as provenance, donor, access rights, and where the collection is within the digitization process) on an Excel spreadsheet. Per the senior archivist, the accession log incorporates several elements of Dublin Core, such as creator, contributor, identifier, and date (Do). There are a few relevant Dublin Core elements missing, such as description, subject, and rights.

- *Container lists*, or brief surveys of the what is in each analog collection (which are later digitized), saved in Word documents. No metadata standard is followed.
- *Finding aids* developed for analog and digitized collections, in PDF, Word, and EAD formats. Finding aids are created with ArchivesSpace, hosted locally and online through the Online Archive of California (OAC), and follow the Describing Archives: A Content Standard (DACS) metadata standard.

Metadata for the oral histories includes:
- *Transcripts* of the oral history, saved as txt, Word, and PDF files.
- *Catalog sheets*, with information about the interviewee, saved as an Excel document.
- *A log of the oral history interviews*, saved as an Excel spreadsheet and including information such as duration of interview, digital file name, technical format, interviewee and interviewer name, date and location of the interview, if the interviewee was incarcerated and where, the army unit interviewee served with, and what campaigns the interviewee served in. The spreadsheet follows a homegrown schema, specialized to the experiences of these veterans and using attributes GFBNEC staff would like to search with.

With regard to controlled vocabularies, the Library of Congress Subject Headings (LCSH) are used within the finding aids. GFBNEC has also developed their own controlled vocabularies for cataloging relating to the Japanese American experience during World War II, as such terminology is absent from LCSH (Do). This includes names of incarceration facilities, campaigns, and military units, and the vocabularies are primarily used within the finding aids and the organization's Omeka website.

While there is robust metadata for archival materials that lives on the shared drive, communications and education staff primarily search for assets through the organization's Omeka website, where GFBNEC publishes their digital archival assets for the public to access. The Omeka website uses Dublin Core, but several fields that could facilitate search and reuse—including description, subject, and rights—are not consistently or thoroughly cataloged. The director of exhibits & education noted that "I know there's a lot more [material] in the Omeka but that gets by the wayside if someone doesn't know about it" (Keller). Essentially, because descriptive metadata is not systemized or robust enough within Omeka, some assets cannot be surfaced by search alone, but require someone to already know that the asset exists.

Staff from the communications team described similar issues with finding archival resources. For example, the public relations manager found that it can be difficult to find an asset "mainly because I'm not 100% familiar with all the different resources we have at our disposal" (Tsuchida). Insufficient metadata within Omeka means that users who are not intimately acquainted with the collections have trouble finding relevant assets. For the public relations manager, what happens then is that they either enlist the senior archivist to find resources, or they will use assets found "mainly in the interest of time, but not necessarily because it's the most powerful image or video" (Tsuchida).

Currently, the senior archivist is relied upon to bridge gaps created by insufficient metadata. For example, the public relations manager said that if an asset does not have enough description or context, they will usually run the asset by the senior archivist to confirm its meaning (Tsuchida). While the senior archivist is able and willing to help because they have this expertise, relying on this expertise instead of documenting this knowledge within a DAM creates duplicative work, unnecessary burdens on the senior archivist's time, and is not the most sustainable solution should the senior archivist ever be out of the office or otherwise unavailable to support staff in this way. A DAM with more consistent and robust metadata collection could mean that staff would better be able to source information independently.

Further, the distributed nature of metadata practices at GFBNEC, in which databases of metadata are created for specific types of resources like digitized collections or oral histories, means that some resources are overlooked and thus not cataloged or accounted for. The director of exhibits & education

recalled an exhibit they were preparing for about Nisei women in which at the very last minute, just a few days before the exhibit opened, the senior archivist found a DVD about Nisei women that they were able to incorporate into the exhibit (Keller). It took the education and archival staff so long to find the DVD because it wasn't incorporated in the Omeka or the shared drive. Having more extensive metadata coverage all centralized, in a DAM for example, would have saved a lot of time and effort in discovering the resource.

Recommendations

A more unified metadata strategy across archival assets would improve the discoverability, legal reuse, and preservation of resources. As a first priority, metadata should be created for assets within the "Photos" folder as much as is possible, including the event name, date, event description, and terms around reusing the photos. This is a significant amount of work, as it will have to be completed for years' worth of events, and information will likely need to be tracked down within email correspondence, event flyers, and photographer contracts. Dublin Core is an appropriate schema for these assets, as its minimal, lightweight requirements will facilitate this workflow, acknowledging both GFBNEC's smaller staff and the limiting nature of retroactively assigning metadata years after the fact, while also covering the bases of critical information needed to steward resources.

The digitized archival collections that live within the "Archives Workspace" folder are already cataloged with a modified Dublin Core schema within the accession log and in Omeka, with finding aids following DACS and incorporating both LCSH and homegrown controlled vocabularies. It is important that both the accession log and Omeka records better and more consistently comply with this schema; in many instances in Omeka, for example, descriptions, subjects, and rights status are absent. These absences account for many of the issues staff in education and communication departments were having when attempting to source records. The accession log also lacks explicit rights metadata that defines in what contexts the collection can be accessed (e.g. online, in person only).

More consistent application of Dublin Core and GFBNEC's modified Dublin Core—across all assets—will facilitate better discovery and preservation while protecting the organization from accidental misuse of assets. A DAM would help enable this, centralizing the diverse metadata resources into one system that could be searched. In addition to being able to support Dublin Core and GFBNEC's modified Dublin Core, the DAM system ideally should be able to support both LCSH and GFBNEC's homegrown controlled vocabulary. Priority should be given to GFBNEC's homegrown schema, as it is more comprehensive and reflective of the Japanese American experience captured within their assets.

5. <u>Ingest</u>

GFBNEC already has extensive metadata for many of its assets, which took time and labor to build out. Any new DAM system should be able to ingest this existing metadata, to prevent duplicative work. The Dublin Core records from the Omeka website seem the most relevant for this process, since they are item-level records. Collection-level metadata is available through the accession log, finding aids, container list, and oral history log. Such collection-level records should be available and searchable within the DAM, in addition to mapping collection-level metadata that is relevant to individual item-level assets, such as format, incarceration camp, military unit, and campaigns.

As for the mechanics of integrating and ingesting the records, Omeka allows for a CSV export of their data ("CSV Export Format"). The DAM system should thus allow for batch ingest of CSV files, as well as mappings from which elements of the collection-level metadata applicable to all items, such as rights information and incarceration facilities, are able to be mapped to the relevant "description" or "rights" field within the item-level record.

With regard to asset ingest, the DAM should provide for both batch and one-off imports, as both methods will likely be used throughout the lifetime of the DAM.

6. Editing/Input

In addition to batch ingest, the DAM system should also allow for editing and data entry within the system itself, to account for one-off submissions into the DAM as well as any necessary edits to incorrect metadata within the DAM. Authorization to input and edit specific fields is listed more extensively within the "Users" section of this report. Generally speaking, the senior archivist should have access to edit and authorize all fields, while communications, education, and IT staff should have access to all of the descriptive fields save for rights. Fields for legal rights, however, should be limited to the senior archivist and the president & CEO, given the financial and legal ramifications of if this field is entered incorrectly.

While the ability to edit assets within the system would be an added bonus, there is not enough of a benefit to make this a strict requirement of the DAM system or pay a substantial fee increase for this functionality. Any edits to the assets, such as minimal video editing or edits to documents like the master list of controlled vocabularies, would be made by the senior archivist alone. Accordingly, since such edits do not require collaboration with multiple staff, the ability to have a DAM as a centralized location for edits is less of a need. Instead, having some versioning functionality could be helpful, with edits made to identically named assets being tracked within the DAM for transparency and clarity.

7. Search

Despite the substantial metadata built out for resources, staff still have trouble finding relevant assets. This is in part because metadata is distributed and uncentralized. For example, there isn't a centralized way to search the accession log, container lists, and finding aids when wanting to learn more about the digitized archival collections. The director of exhibits & education recognized this, saying that "it would be wonderful to be able to search a term, and get all the various assets associated with that at once—that would be a dream" (Keller). Right now, some assets are on the Omeka while some are solely on the shared drive. Having a centralized repository with a DAM would retrieve all relevant results and make search more meaningful.

Another issue with search are the shortcomings of Omeka. As mentioned in the "Metadata" section of the report, item records within Omeka often lack fields critical to search, such as description, subject, and rights. Further, the Omeka search function is not especially robust or user-friendly. While the advanced search allows users to craft searches that combine specific keyword and boolean searches for each field, it is in a complex, composite form likely intimidating to a lot of users. Further, there is no way to filter results once retrieved; an entirely new search would have to be crafted. This prevents simple browsing of materials, an important discovery method for staff—so important, in fact, that the director of exhibits & education says that they browse for assets more than they conduct directed searches, while the public relations manager said that they browse approximately 40% of the time (Keller, Tsuchida). A DAM with browse functionality could enable staff's preferred discovery method; for example, the director of exhibits & education mentioned that they would appreciate browsing by incarceration site (Keller). Here, again, the tie between metadata and search is called to the forefront, as an important metadata field as part of the modified Dublin Core would be "incarceration site." Including such concepts unique to GFBNEC's mission within the metadata schema will improve search and asset discovery. Core metadata fields users would want to search by include creator (i.e. the veteran's name), description, subject, unit served with, incarceration site, campaigns, and rights.

The director of exhibits & education noted that the shortcomings of Omeka's search functionality means that they often have to involve the senior archivist when seeking assets. For example, while they can easily find assets tied to a name of a veteran, it is harder to find assets tied to a concept, such as unit affiliations. To find those assets, they have to talk with the senior archivist to brainstorm relevant collections (Keller). Workflows within the DAM will be discussed at length in the "Workflow" section

below; with regard to search, though, more robust search and browse functionality within the DAM would relieve some of the burdens on the senior archivist's time, as staff will be more empowered and able to source assets independently.

8. Display and Access

Currently, staff are primarily searching for assets through the organization's Omeka, which is publicly accessible as well. Access copies are available on the Omeka, which means that if staff want higher-quality preservation copies, they either have to know where they are stored within the shared drive or consult with the senior archivist. This means that staff are currently passing off hard drives to one another or sharing assets via Dropbox (Do). A centralized DAM with access and preservation copies for staff to access will standardize and streamline this process. As for external access, there ideally should be a public interface to the DAM as well so that the public can access it. Omeka serves as this public access point currently. This should be transitioned to and mirrored with the DAM.

With regard to languages, the DAM needs to support Latin and Japanese characters, as materials within the collection are in English, Japanese, and Hawaiian Pidgin English.

A "user space" in which internal users of the DAM could annotate assets is desirable, but not a deciding factor for the DAM. The senior archivist noted that this functionality could be especially helpful for videos; when staff would like a certain clip from the video, they could let the senior archivist know which clip directly within the DAM (Do). Doing this through a centralized user space as opposed to decentralized email correspondence would create a history staff could refer back to later, to learn how and where the video was previously used.

Assets within the DAM should be both streamable and downloadable, as it will help both internal and external users confirm if an asset is appropriate for use. However, the senior archivist noted that while external users should be able to download assets (if the rights of the asset allow it), external users should only be able to do so with access-quality copies. If it is a more professional broadcast or commercial endeavor needing higher resolution assets, GFBNEC should be involved; for research, personal, or academic work, lower-resolution copies should suffice (Do).

9. Workflow

The senior archivist described current workflows as being "very frazzled," saying that currently storing, sourcing, and delivering assets can feel "all over the place" as the senior archivist has to juggle Google Drive, Dropbox, and hard drives (Do). The senior archivist stated that "I'm kind of a slave to the moment," and that a DAM would centralize things and make project management easier (Do). While the diffusion of metadata complicates search, the diffusion of where assets are saved complicates the asset delivery process. A DAM could address both of these issues, both of which could help the senior archivist, who is heavily relied upon currently to source and deliver assets to other staff.

With regard to asset discovery, the senior archivist is relied upon extensively by both the communications and education departments, the two primary users of digital archival assets aside from the senior archivist. The director of exhibits & education says that finding assets often involves looping in the senior archivist, to either download/cut clips or to find assets since he has a deep understanding of the collection (Keller). The public relations manager also mentioned that if they are unable to find what they need on the website or in the collections, they will turn to the senior archivist, and that it usually takes a day or two to find the asset (Tsuchida). The senior archivist recognized the demand on his time, stating that even with a DAM he would likely still help with cutting clips, but ideally staff would rely less on him to source relevant assets due to better metadata and searching functionality (Do).

The "Users" section of the report lays out in depth which users and departments would be involved in asset creation and approvals. Generally speaking, archival, communications, and education departments should all be given asset creation rights. While the senior archivist is the DAM administrator

and ultimately responsible for the stewardship of the archival assets living within the DAM, given that there is only one archival staff member, having education and communications staff assist with asset creation could be a great help, especially as these two departments in select cases may be more directly acquainted with the creation of that asset (for example, communications staff and the photographs created of a gala).

While all three departments would have creation rights, approvals would all be handled by the senior archivist, so that assets could be checked for compliance with metadata standards, file naming conventions, and controlled vocabularies. This is an important step as consistent and complete metadata is necessary for reliable discoverability, preservation, and legal reuse of assets; an archivist is trained in these information management concepts, and can best ensure compliance.

Further, at this phase the senior archivist would be able to complete the rights information, a critical aspect of asset description as correct rights information facilitates legal reuse of assets. Currently, both the education and communications department are unaware of the rights status for assets, resorting to either consistently asking the senior archivist if it is okay to use or assuming the usage is fair use (Keller; Tsuchida). The senior archivist is well versed in provenance and the rights status of items, and is the best resource for ensuring accurate rights management. However, the current model involves legal risk, duplicative work, additional strain on the senior archivist's time, as the archivist is asked to repeatedly confirm the rights status of an item, or the rights status is not confirmed at all. A DAM would correct this, making the rights status of items transparent to all DAM users.

Similarly, the senior archivist will be in control of all preservation aspects of assets, including checksums and redundancy measures. This, again, is because the senior archivist is the DAM administrator, and is trained in understanding the importance, mechanics, and implications of archiving and preservation. More details on this can be found in the "Digital Preservation" section.

10. Publishing Content

Currently GFBNEC publishes content onto their website, Omeka, social media (Instagram, Facebook, and Youtube) and outward communications (principally their weekly email newsletters). Social media postings occur roughly three to four times a week (Tsuchida). The ability to publish content from the DAM to either the website or their social media platforms would be nice to have, but not essential (Do). The added step of downloading an asset from the DAM in order to post it on another platform, while not the most streamlined, is not costly enough in time and labor in order to necessitate this functionality. For the public-facing DAM interface, it could be helpful to have functionality to automatically share an asset on social media, so that public users can easily share content.

11. Digital Preservation

A DAM system could standardize and optimize digital preservation measures. Currently, redundancy for assets is not systematic, with some on the shared drive, Dropbox, Amazon S3, and external hard drives (Do). The Amazon S3 storage has several oral histories as well preservation copies of digitized analog collections (Do). Given the low costs to upload but the high costs to download, the Amazon S3 is a cold storage option intended to be something of a last resort, to use only when all other redundancies have failed. While assets from the "Archives Workspace" folder have this somewhat inconsistent redundancy treatment, assets in the "Photos" folder have even less redundancy, with the only backup being the work the website/IT coordinator does to back up the shared drive (Do).

Instead, GFBNEC should follow the 3-2-1 backup rule, which states there should be at least three copies of assets, on two different media, with one of them offsite (Tadic, Class 7 - Metadata Modelling). This applies to both access and preservation copies. In addition to the DAM, assets should be saved in two other locations to follow this best practice. One should be Amazon S3 cold storage, as they already practice this (albeit inconsistently). The third option could be either cloud storage with more immediate

online access than cold storage or hard drives. Copies should also be geographically dispersed and should not all be from the same vendor. This means confirming that the Amazon S3 server is not located in southern California. Further, the third copy of assets, whether it be in accessible cloud storage or hard drives, should again ideally be stored outside of southern California to mitigate risk against natural disasters happening in the region. If this is not possible with the hard drives, at the very least they should not be stored in the same office.

A preservation priority is also digitizing all tapes, as this physical media is deteriorating and will cause data loss. While the majority of the tapes have been digitized, approximately 10-15% need to be digitized, most of which are MiniDV. While MiniDV is a newer format dating from 1995 to the late 2000s, its small size and thinness of the tape presents a deterioration risk ("Videotape"). Further, as a newer format it is unclear how long it will be supported, and it requires proprietary tape ("Videotape"). All of these factors put its long-term support at risk and digitization should be prioritized, ideally to an open format that is not owned by anyone and supported by standards organization that will ensure its long-term management, such as uncompressed 8 or 10 bit or JPEG 2000 (Tadic, Class 3 - Technical and Preservation Metadata).

And finally, a DAM could facilitate digital preservation actions. Given the small size of the organization and the archives department specifically (one staff member), extensive preservation responsibilities are not sustainable (Do). However, the DAM could automate and facilitate the most pressing digital preservation actions, such as creating and verifying checksums and monitoring format obsolescence.

12. User Rights and Security

Since no personnel, human resources, or financial files will be on the DAM, restricting rights to specific assets will be a lower priority. There is the possibility that some archival collections, as part of their accession agreement, will have restricted digital access, but then it is unlikely that GFBNEC would have prioritized that collection for digitization. The priority for user rights and security, then, will be limiting the users' ability to mass delete or mass change assets (Do). A detailed delineation of user rights can be found within the "Users" section. Generally speaking, while nonarchival staff will have metadata and asset ingest rights, all ingest by these staff need to be checked and approved by the senior archivist. Deleting assets should be restricted to a select few who can use the right deliberately and carefully; it is recommended the senior archivist and president & CEO alone should have this right.

Similarly, legal rights should be restricted to these two staff. While the senior archivist mentioned that he would be comfortable with other staff assigning legal rights so long as he could override without intervention, the financial and legal ramifications of an error (either from misinformation in input or accidentally changing the field) are large enough that the risk should be minimized as much as is possible.

13. Interoperability

GFBNEC does not have plans to get a collections management system, so it is not necessary to select a DAM with integration capabilities with a collections management system (Do). However, the senior archivist noted that it would be helpful if the DAM could be integrated with a basic cloud system, such as Dropbox or Google Drive, to help deliver assets from the DAM to staff (Do). Having this functionality would support staff who extensively use Google Drive or Dropbox to continue to operate with these tools (Do).

14. Technical

Currently, GFBNEC uses both Windows and Mac (with the majority of the staff on Windows, and archival staff on Mac). The senior archivist specified that there are no underlying databases for the archival assets, but GFBNEC generally uses MySQL (Do). Workstations run on a Linux server. With

regard to open source software, the website/IT coordinator is often open to approving the integration of open source software if it does not pose any threats (Do). However, given the potential maintenance and patching involved, and that there is only one IT staff responsible for all of the technical platforms and hardware within the organization, a turnkey solution—that still allows for data exports in the event of any future change—may be more sustainable.

15. Example Systems

Choosing a DAM system will require carefully reviewing GFBNEC's specific needs and budget. This document, along with the accompanying Excel spreadsheet, are helpful tools for understanding GFBNEC's digital asset needs and the specific functionalities a DAM should include. With regard to the budget, further research will be necessary to evaluate the costs of the current asset discovery and preservation model against adopting a DAM system. Return-on-investment (ROI) calculators are freely available, where GFBNEC can calculate the costs of the current model (which can include inefficient searching, duplicative work, missing files, rights infringement, and lost revenue opportunities) against the costs of a DAM system (which can include software, training, and technical overhead) (Philson).

Adopting a DAM system will help GFBNEC realize the full benefits discussed in this report, including optimized search, centralized storage, and digital preservation support that helps ensure the authenticity of their materials. Potential options identified as being especially suited for smaller organizations include Daminion, Canto, and Libris ("Daminion"; "DAM"; "Simplify Your Storage With Canto's Organization Features"). Pricing is not available online, but these would be platforms to share the functional requirements spreadsheet with, and solicit quotes.

However, if after doing the ROI calculations and receiving quotes from DAM providers, it is determined that GFBNEC cannot afford to implement a fully integrated DAM system, there are still measures they can take to address some of the inefficiencies and preservation issues mentioned in this report. GFBNEC already has a Dropbox account that can hold up to two terabytes for only $120 per year. The digital archival assets are approximately 1.2 terabytes, projected to reach 1.8 terabytes in five years. Dropbox would thus be a feasible storage option for the digital archival assets. This solution, combined with strategic and consistent directory structure and file naming conventions as well as Excel spreadsheets of metadata that directly link metadata records to assets with unique file names, could create a lower cost asset storage and discovery system. The redundancy measures described in the "Digital Preservation" section could also take place with minimal cost. There are a few obvious drawbacks to such a set-up: first, separating metadata from the assets prevents visual searching and browsing; second, this solution does not provide for digital preservation measures like fixity checks and checksum verification; and third, this model still requires significant upfront work to build out and standardize metadata. It is, however, an option if the budget will not allow otherwise.

Works Cited

"CSV Export Format." *Omeka*, https://omeka.org/classic/plugins/CsvExport/. Accessed 6 June 2020.

"DAM." *Libris Systems*. *librisdigital.com*,
https://librisdigital.com/software-solutions/dam-digital-asset-management/. Accessed 8 June 2020.

"Daminion." *Multi-User, Server-Based Digital Photo Management Software*, https://www.daminion.net/. Accessed 8 June 2020.

Do, Gavin. *Interview with Gavin Do, Senior Archivist at Go For Broke National Education Center*. Interview by Savannah Lake, 27 Apr. 2020.

Keller, Megan. *Interview with Megan Keller, Director of Education & Exhibits at Go For Broke National Education Center*. Interview by Savannah Lake, 8 May 2020.

"Our Mission." *Go For Broke National Education Center*, http://www.goforbroke.org/about/mission.php. Accessed 2 May 2020.

Philson, Lauren. *Essential DAM Implementation: A New Project Guide and Toolkit*. 2015, https://www.iointegration.com/io/essential-dam.

"Simplify Your Storage With Canto's Organization Features." *Canto*. *www.canto.com*, https://www.canto.com/features/organize/. Accessed 8 June 2020.

"Staff." *Go For Broke National Education Center*, http://www.goforbroke.org/about/staff.php. Accessed 2 May 2020.

Tadic, Linda. *Class 3 - Technical and Preservation Metadata*. UCLA IS 289 Digital Asset Management.

---. *Class 5 - DAM vendors and open source options; OAIS; Digital workflows*. UCLA IS 289 Digital Asset Management.

---. *Class 7 - Metadata Modelling*. UCLA IS 289 Digital Asset Management.

"Videotape." *Preservation Self-Assessment Program*, 6 June 2020, https://psap.library.illinois.edu/collection-id-guide/videotape#minidv.

FUNCTIONAL REQUIREMENTS CHECKLIST

#	Function				
1.0	**USERS**	**Quantity: Year 1**	**Quantity: Year 5**		
1.1	internal (i.e. GFBNEC staff)	10	15		
1.2	external (i.e. photographers or videographers hired by GFBNEC)	0	5		
1.3	public	GFBNEC needs to track with Google Analytics/Omeka plug-in	Can estimate after reviewing Google Analytics/Omeka plug-in results		
2.0	**FILE FORMATS**				
	There are 83 file formats total; this is a snapshot of the top ten most prevalent in terms of storage size, enumerated by number of files and storage size.	**Quantity: Year 1**	**Quantity: Year 5**	**Storage: Year 1 (GB)**	**Storage: Year 5 (GB)**
2.1	tif	19316	28974	803.62	1205.43
2.2	jpg	63705	95558	300.32	450.48
2.3	mov	130	156	21.86	26.232
2.4	mp4	73	88	16.50	19.8
2.5	avi	62	74	14.05	16.86
2.6	bkf	1	1	9.61	9.61
2.7	cr2	258	387	7.76	11.64
2.8	psd	509	764	5.40	8.1
2.9	jpeg	466	699	5.26	7.89
2.10	mdb	16	19	3.98	4.776
	TOTALS	84536	126719	1188.36	1760.818
3.0	**CONTENT TYPES**				
	Enumerated by percentages and storage size.	**Quantity: Year 1**	**Quantity: Year 5**	**Storage: Year 1 (GB)**	**Storage: Year 5 (GB)**
3.1	Event photos	64.36%	65.78%	831.203	1246.8045
3.2	Digital surrogates of archival collections	22.29%	22.78%	287.866	431.799
3.3	Oral history photographs	5.57%	3.80%	71.966	71.966
3.4	Various documents and operational files for "Photos" folder	4.79%	4.89%	61.824	92.736
3.5	Various documents and operational files for "Archives Workspace" folder	2.09%	2.14%	27.03	40.545
3.6	Oral history videos	0.90%	0.61%	11.591	11.591
	TOTALS	100.00%	100.00%	1291.48	1895.4415
4.0	**METADATA**	**Supported?**	**Notes**		
4.1	Must support Dublin Core metadata structure and a modified Dublin Core				

#	Function				
4.2	Must support customizable metadata fields				
4.3	Must support homegrown controlled vocabularies				
4.4	Does the DAM support Library of Congress Subject Headings (LCSH)?				
4.5	Can metadata be exported/extracted out of the DAM? In what formats?				
4.6	Can items be linked to collection-level metadata records and resources?				
4.7	Can collection-level metadata be mapped to specific fields in item records?				
5.0	INGEST	Supported?	Notes		
5.1	Must support CSV and Excel metadata batch ingest				
5.2	How can new metadata be imported into the DAM?				
5.3	How can legacy metadata be imported into the DAM?				
5.4	Does the DAM support batch ingest of assets?				
5.5	Does the DAM support importing individual assets?				
5.6	Can the DAM extract technical metadata from files upon ingest?				
6.0	EDITING/INPUT	Supported?	Notes		
6.1	Must be able to create new records				
6.2	Must be able to edit records				
6.3	Must be able to restrict editing/input capabilities by user				
6.4	Does the DAM support editing assets within the DAM?				
6.5	Does the DAM support versioning? How?				
7.0	SEARCH	Supported?	Notes		
7.1	Must support keyword search				
7.2	What restrictions does the DAM have with regard to indexing metadata fields?				
7.3	Does the DAM support Boolean search?				
7.4	Does the DAM enable browse (faceted search) based on metadata fields?				
7.5	Does the DAM enable browse (faceted search) based on controlled vocabularies?				
8.0	DISPLAY AND ACCESS	Supported?	Notes		
8.1	What are the DAM's display resolution options (high-res, middle, low)?				
8.2	Does the DAM have an public access interface that can be set to provide access to low-resolution items only?				
8.3	Does the DAM support item-level access authorization?				
8.4	Does the DAM support Latin characters?				

#	Function				
8.5	Does the DAM support Japanese characters?				
8.6	Does the DAM support an internal user space for annotations?				
8.7	Assets must be downloadable from the DAM				
8.8	Assets should be streamable within the DAM				
9.0	**WORKFLOW**	**Supported?**	**Notes**		
9.1	Does the DAM support asset creation approvals?				
9.2	DAM must support restricting certain rights to specific users				
10.0	**PUBLISHING CONTENT**	**Supported?**	**Notes**		
10.1	Does the DAM support publishing onto social media platforms (Instagram, Facebook, Youtube)?				
10.2	Does the DAM support publishing onto websites?				
10.3	Does the DAM support publishing onto enewsletters?				
10.4	Does the DAM support the public sharing content from the DAM? In what way (e.g. email, direct link, social media)?				
11.0	**DIGITAL PRESERVATION**	**Supported?**	**Notes**		
11.1	Can the DAM generate checksums?				
11.2	Can the DAM verify checksums?				
11.3	Does the DAM monitor for format obsolescence?				
11.4	What is the long-term health/sustainability of the DAM software company?				
12.0	**USER RIGHTS AND SECURITY**	**Supported?**	**Notes**		
12.1	Does the DAM allow the DAM administrator authentication rights by function (e.g. ingest, edits)				
12.2	Does the DAM allow for specific fields to be edited by specified users?				
12.3	Does the DAM allow for specific functions to be limited to certain users (e.g. deleting assets)				
13.0	**INTEROPERABILITY**	**Supported?**	**Notes**		
13.1	Does the DAM interoperate with cloud storage systems (such as Dropbox, Box, and Google Drive)?				
14.0	**TECHNICAL**	**Supported?**	**Notes**		
	List the following technical specification for your system.				

#	Function				
14.1	Operating system: does the DAM work with Windows 10?				
14.2	Operating system: does the DAM work with Mac? (*retrieve specific operating system once GFBNEC office is open again*)				
14.3	Database: does the DAM work with MySQL?				
14.4	Server platform: does the DAM work with Linux?				
14.5	Which browsers does the public interface of the DAM work with?				
14.6	If cloud-based, which browsers does the internal interface of the DAM work with?				
15.0	**HIGH-LEVEL PRICING (provided by DAM vendor)**				
	The pricing estimated here is not binding.				
15.1	Initial license				
15.2	professional services for customization				
15.3	maintenance - 1st year				
15.4	3rd party apps initial license				
15.5	software configuration				
15.6	training/roadmap toolkit				
	TOTAL high-level start-up costs				
15.7	Annual maintenance				
15.8	Annual maintenance - 3rd party applications				
15.9	annual user support				
	TOTAL annual ongoing costs				