StratComm Digital Preservation Project

Collected Works

Andy Cleavenger

University of Maryland

Executive Summary

The assessment of StratComm's digital holdings and preservation state detailed in this digital preservation project reveals an organization that has many beneficial practices already in place. Among these is a basic scheduled backup system; a logical folder structure and file naming conventions; regular application of technical, administrative, and descriptive metadata; and a well-informed approach to file format choices. The largest areas for potential improvement are data integrity and information security. StratComm is best advised to perform fixity checks at a minimum before and after transfers are performed or copies made. Ideally, these fixity checks would also be performed annually for all holdings no matter their usage status. In the area of information security, network permissions for staff access to the multimedia server deserves more careful consideration of the group that is allowed access as well as a tightening of the activities that group is permitted to perform in that area. StratComm is additionally advised to take advantage of easy steps to improve their preservation status in the areas that they are already doing well. Examples of such measures include trading backup copies between site offices to establish a third copy of files in a different location; performing an inventory of content and documenting the network topology; and performing an inventory of file formats in use and monitoring their obsolescence issues.

As a high volume production house it may not make sense for StratComm to strive for the highest levels of the NDSA Levels of Digital Preservation rubric because it is not within the scope of their corporate function. StratComm is best advised to pursue closer partnerships with the Corporate Archives and Enterprise Computing divisions to provide for the curatorial and preservation needs of the image collection. This could potentially be addressed by forming a Digital Preservation Committee comprised of R504 and R301 leadership, Corporate Archives staff, StratComm photography staff, and representatives of Enterprise Computing. This group should establish an annual schedule for meeting, and a set agenda of topics to cover at each meeting.

The following documents begin with a Survey and Report that analyzes the scope of StratComm's image collection; its current content management practices; its staff's perceptions of self-efficacy; and an appraisal of the resources that are potentially available for improving tools and processes. Following this, the Next Steps Digital Preservation Plan provides low, medium, and high resource remedies for each focus area of the NDSA Levels of Digital Preservation rubric. And the final document, the Digital Preservation Policy, provides a notional set of policies that could be adopted by StratComm to address the majority of their digital preservation challenges. This policy document includes a statement of purpose and definition of the intended user population, as well as outlining the asset lifecycle, the applicable activities required at each stage, and the responsible parties in each division.

TABLE OF CONTENTS

Digital Preservation Survey and Report

Introduction

I have chosen to focus my project on the digital image collection of the Strategic Communications (StratComm) department. StratComm is an in-house communications department for a not-for-profit company with offices in the Washington D.C. and Boston areas.

While the primary focus of StratComm is in communicating externally with industry and the public, the context under which they collect and preserve their image collection is more internally focused. Their primary impetus for both the curation and preservation of their digital assets is to empower StratComm's staff, as well as the company's employee population, to use the image collection to effectively communicate about its work program. It is especially important to StratComm that this is done in a manner consistent with the new brand guidelines that are currently being developed.

While this mission is focused on short-term communication needs, it is highly dependent upon prior curatorial and preservation groundwork being laid. For example, the ability to locate existing imagery featuring a particular research or technology requires applying descriptive metadata, a Level 3 requirement of the National Digital Stewardship Alliance's (NDSA) Levels of Digital Preservation rubric (Phillips, Bailey, Goethals, and Owens, 2013, p.3). Another example would be maintaining access to files saved in obsolete or proprietary file formats. This requires ongoing monitoring of industry trends and a program of format migration to convert obsolete file types to formats that are accessible to modern tools. These are Level 3 and Level 4 activities on the NDSA Levels of Digital Preservation rubric (Phillips et. al., 2013, p.3). Without these activities StratComm will at best continue to experience difficulty in searching the early digital image collection, and at worst may lose access to that collection permanently.

The following assessment is based on a series of interviews performed in October of 2018 with members of StratComm leadership and Corporate Archives.

Scope of Digital Holdings

StratComm's image collection technically starts in the 1950's with the founding of the company. From this time up until 1995 all images were shot on film and printed on paper. These traditional film-based images are all currently housed and preserved by the organization's Corporate Archives department.

In 1995 StratComm (called Communications and Publications at that time) transitioned to digital imaging systems with the purchase of a Kodak DCS 465, which was a digital back designed to be mounted on a Hasselblad 500 series film camera. This purchase was followed in short order by the additional purchase of a Kodak DCS 460, which was a digital back permanently screwed onto the back of a Nikon N90s. These two cameras began StratComm's digital image collection.

In the early days CDs were the primary storage medium for all image files. The Washington office currently holds about 800 CDs and DVDs. Boston also had a CD collection, however it was significantly smaller, and reportedly much less organized. Boston transitioned to using external hard drives in the early 2000's and continued this practice until the introduction of the present server-based infrastructure. These hard drives were transitioned to the Corporate Archives department upon the retirement of the Boston photographer.

The Washington office's choice of CD as a storage medium remained prevalent until about 2009 when StratComm began to take advantage of the lower prices for external hard drives. Over time the Washington photographer gradually morphed from using an external RAID array to a centralized server system available to the entire multimedia department. Since the servers are used as a platform for video editing high network bandwidth is important, making it necessary for each branch office to maintain their own local server. However, each office is capable of accessing the other's server if necessary. Today, the combined size of the online portions of the image collection in both Washington and Boston is roughly 8TB.

The current state of StratComm's storage system described above does not yet meet requirements for any of NDSA's levels for Storage and Geographic Location category, however the general trend away from optical media and towards a server-based infrastructure is positive progress towards Level 1 (Phillips, 2013, p.3).

Current Content Management Practices

On an OS level, the organizational practices between locations are best characterized as partially divergent. There is little consistency apparent in the upper level folder structures on each location's multimedia server. StratComm leadership and design staff frequently have trouble locating the correct directory to find recent photography files without the assistance of photographers. However, once the correct directory is found the folder structures are mostly similar. Projects are stored in chronologically ordered folders, each with the date and project or customer described in the folder name.

Metadata

The photographers in both locations primarily interact with the image collection through Lightroom. Lightroom is both a RAW processor as well as a file management tool. The Boston photographer tends to add metadata to image files solely through Lightroom. This includes the conveyance of the camera's automatically captured EXIF metadata, as well as authorship, copyright, and rudimentary descriptive metadata in the form of keyword tags. The Washington photographer uses a combination of Lightroom and Adobe Bridge for applying metadata to images. Bridge functions more like a high-powered browser window, allowing the user to see previews of certain types of graphics files, as well as allowing the user to edit the embedded metadata of those files. The Washington photographer uses Lightroom to apply authorship data, camera EXIF data, and copyright data. Descriptive metadata is then applied in Bridge, including a title, description in paragraph form, as well as the application of keyword tags, which have been constructed to mirror the organization's taxonomy of search terms. Bridge is used for this because it allows a more granular capability of applying sub keywords without necessarily applying the parent node as well.

StratComm's current metadata practices described above qualify as the storage of standard technical and descriptive metadata; a Level 3 requirement of NDSA's Metadata category. However, requirements for Levels 1, 2, and 4 are currently not met. These levels require an inventory of content and storage locations; storage of administrative and transformative metadata; logging of events; and storage of standard preservation metadata (Phillips, 2013, p.3).

EnterMedia DAM system

StratComm uses a digital asset management (DAM) system called EnterMedia as a means of establishing access for all company employees to search the image collection. It is important to note that the bar for inclusion in EnterMedia is that the images have at least some reuse value. This means that some images that are not related to the company's work program, such as retirement and birthday parties, may not be included. Management did not wish to choke the system with a large amount of material of limited or no interest. As such, the images uploaded to EnterMedia do not encompass everything shot, but are a hand-selected subset that is mostly chosen by the photographers.

The retirement of MediaBeacon – the previous DAM system – occurred in conjunction with the purchase of EnterMedia in 2015, however the implementation of EnterMedia took longer than anticipated. Only in the last year has EnterMedia begun to function as initially planned. The images from Washington created during this three-year gap formed a sizable backlog that has just recently been augmented with metadata and uploaded to EnterMedia. Boston also has a backlog of images, but is still early in the process of synchronization.

Also missing from the EnterMedia database are any of the images that currently reside on the CD and DVD collection. Since StratComm's communication needs tend to skew towards featuring the most current content possible the absence of these images from 1995 through 2009 is not as critical as the three-year gap discussed above. However, the value of these early digital images cannot be completely discounted.

Communications showing the history of a particular program, or retrospectives featuring the history of the organization often use this type of material.

File Migration Efforts

The files housed in the Washington CD and DVD collection have two major vulnerabilities:

1) The discs are rapidly decaying.
2) The RAW files produced from the early Kodak cameras required a Photoshop plug-in whose most recent supported version was Photoshop 5.5 on Mac OS 9.

The Washington photographer has begun the process of migrating files off of the CD collection and transferring them to the Washington multimedia server. At the time I write this most material between 1995 and 2006 has been successfully migrated. Only a few more years still remain to be migrated, however the much larger tasks of identifying a solution for format migration, as well as applying metadata to the images remains a significant need. Without these steps the images will remain unavailable for either access or searchability.

The activities described above do not yet meet any requirements for NDSA's Storage category, however the migration of files off of optical media represents progress towards achieving Level 1 (Phillips, 2013, p.3).

File Format Choices

The choices of file formats that are saved by each photographer and provided to the end-user are another point of divergence between the branch offices.

The Washington photographer retains all RAW files in their native proprietary .NEF format. Once he has edited and processed the images he wants he exports high quality JPEGs, which are then used as the master copy for any further manipulation in Photoshop or providing to the end-user. The retention of proprietary RAW formats is admittedly a vulnerability, however this choice was originally made with the expectation that RAW files would eventually be discarded after a pre-determined sunset date. In practice, this discarding of RAW files has not happened because the time necessary to locate and delete these files costs more than the additional storage needed to simply retain them. For delivery formats, JPEG saved at maximum quality in AdobeRGB was chosen because it was high enough quality for print and fast to export. JPEG also satisfied end-user requests for a format that was usable in a wider number of display platforms than TIFF. Washington used to export to TIFF many years ago when camera megapixel counts were much lower, however end-users had continual problems with not having the necessary software to open or use TIFFs. Additionally, the extreme megapixel counts of today’s cameras make TIFF significantly slower to output and manipulate, and they require more than 10 times the amount of storage space.

The Boston photographer converts all RAW files from Canon’s proprietary .CR2 format and saves them as .DNG. The Digital Negative (.DNG) format is a RAW format created by Adobe in an attempt to create a standard format for RAW camera data that will enjoy a comparatively longer period of compatibility with RAW converters. Technically it is still a proprietary format, however it is worth noting that Adobe has submitted .DNG to the International Standards Organization in an effort to establish it as an accepted standard ("Adobe seeks International recognition for DNG", 2018). How widely the .DNG format has been adopted is unclear.

The Boston photographer provides different formats to the end-user depending on their level of graphics expertise. If the requester is a StratComm designer he provides the raw .DNG file and lets them convert to their desired format. For all other users the photographer creates automated web galleries through Lightroom that feature thumbnails with reduced but sufficient resolution for utilizing in PowerPoint, which he feels is the most common need.

The file format practices described above come close to achieving a Level 1 status on NDSA’s File Formats category (Phillips, 2013, p.3). A major point that should be examined further is the use of .DNG as a RAW format. There are no non-proprietary RAW formats in existence, so nothing is a perfect choice. Adobe’s efforts towards making .DNG open and well documented are positive traits, however further research is warranted on current adoption levels and interoperability with present RAW processing tools.

Current Backup Strategy

My initial inquiries indicate that both branch offices back up their servers on a scheduled basis. Washington’s server is backed up to a server that is collocated with the primary server using software provided by the company’s IT department. The Boston server is backed up by software obtained by the multimedia staff and, like Washington, it is also backed up to a server that is collocated with the primary server. A third copy of files for either location does not exist, nor does there appear to be any capability for performing fixity checks to guard against file corruption.

The backup strategy described above does establish a second copy of files; a requirement of NDSA’s Level 1 of the Storage and Geographic Location category. However, because this copy is collocated with the original it is vulnerable to the same threat profile as the original. As such, current practice does not yet meet Level 1 requirements. In terms of fixity, StratComm’s current absence of any ability to check data integrity at any stage of use or storage does not satisfy the requirements for any of the NDSA levels in the File Fixity and Data Integrity category (Phillips, 2013, p.3).

Information Security

StratComm has some minimal information security protections in the form of password requirements for access to the multimedia servers, however some access granted to designers has caused issues in the past. One notable incident caused the unintended deletion of an entire folder of images. This event illustrates the need for tighter control of read/write permissions; a Level 1 requirement of NDSA's Information Security category (Phillips, 2013, p.3).

Staff Perceptions

Due to organizational restructuring efforts undertaken in the last year most members of StratComm leadership are recent new hires. As such, their level of self-efficacy in accessing and utilizing the digital image collection is presently limited. StratComm design staff also exhibit inconsistent levels of knowledge concerning the use of EnterMedia as an image search resource. This is partly due to lack of prior knowledge in the case of recently added contract staff, but permanent staff also exhibit this lack of knowledge due to the aforementioned delayed implementation of EnterMedia. The long gap between the retirement of MediaBeacon and the implementation of EnterMedia created a workflow dynamic that required designers to seek the assistance of photography staff to personally perform image searches. It has now become reflexive to continue this practice.

Except for the photographers themselves, Corporate Archives staff currently seem to have the most accurate insight into the state of the image collection, however this knowledge of status does not translate to access. They are aware of and familiar with StratComm's server workflow, the existence of the CD collection, and the existence of new EnterMedia DAM system, however they only enjoy access to EnterMedia. Of all the stakeholders, Corporate Archives staff know best where to go to look for images, but they also know how big the gaps are and how much of the collection is not physically under their control. This limits their ability to assist their customers with historical image searches, as evidenced by the occasional need to refer customers to StratComm photography staff for image search assistance.

It is important to note that despite StratComm leadership's relative unfamiliarity with the full scope of the image collection, they have expressed a high degree of interest in its value as a resource and the important role it will play in communicating about the work program of the organization to an external audience. Even more critically, the image collection will play a vital role in establishing the new visual brand both internally and externally.

Ancillary Content

StratComm leadership mentioned two types of digital content that they would be interested in collecting which they currently are not. The first is stock imagery that is purchased through corporate accounts with large stock image collections. These are downloaded from various websites on an as-needed basis. The challenge with this material is that the licensing prevents them from retaining that same file for a later use. Provided their account is paid and active, they are free to download the same file a second time for a single use, however this licensing rule precludes adding stock image content to EnterMedia.

The second type of content they are interested in collecting is their burgeoning set of iconography they are developing as part of the company-wide rebranding effort. They have not collected vector art before in a repository that made such content available to employees. They expressed a strong desire to create curated packages of these icons based on themes that mirror the company's research areas. They wish to empower employees to effectively leverage the company's new brand identity without necessarily requiring the assistance of StratComm. Such a self-service repository would be critical to that end.

In a similar vein, StratComm leadership expressed a great deal of interest in the creation of themed image galleries of photography specifically curated to align with these same research areas. Leadership also emphasized a desire to gain insight into usage. Version control, usage stats, and even usage permission were of particular interest.

Potential Resources for Effecting Change

As previously mentioned, StratComm's leadership is all newly hired in the last year. As such much of the status that is outlined in this document was new information to them at the time of the interview. Since budget planning is a process that is measured in quarters and years it is not surprising then that their response on the question of potential resources to devote to preservation sounded doubtful for the provision of new budget allocations or staff time in the short-term. However, it was mentioned that if they can find ways of displaying the increased value of curatorial and preservation activities they may be able to fold some of that work into the larger corporate modernization effort that is still under way. There is also a unique opportunity for the multimedia group to explore new options for their server infrastructure as they consider the replacement of their aging Mac servers. Both of these responses leave considerable room for hope.

As a short-term strategy, StratComm leadership feel that their most effective approach for addressing preservation needs will be to look for strategic partnerships with existing departments in the company, such as Corporate Archives and IT. These organizations may be employing solutions for their own needs that could also work for StratComm.

Conclusion

I concede that many of the concerns that were outlined during the interview were notably curatorial in nature. At first glance these issues might seem out of scope in an examination of strictly preservation concerns, however I have included them above out of recognition that curatorial and preservation activities enjoy a symbiotic relationship that is not easily separable. After all, it is difficult to establish access without having preserved the material.

In the next phase of this project – the Digital Preservation Plan – I hope to address both the curatorial concerns raised here as well as the more straightforward preservation needs.

References

Adobe seeks International recognition for DNG. (2018). Retrieved from https://www.dpreview.com/articles/8083544151/adobedng

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

Next Steps Digital Preservation Plan

Introduction

To provide a brief recap, my project is focusing on the management and preservation of the StratComm digital image collection. StratComm is an in-house communications department for a not-for-profit company with offices in the Washington D.C. and Boston areas.

In most cases the current preservation state of StratComm's digital image collection displayed aspects of Level 1 requirements, but did not meet them all. For example, under the category of Storage and Geographic Location, StratComm did meet the requirement for having at least two copies of its files, however those two copies are collocated with each other, making them both vulnerable to the same threat profile. In the category of Information Security, StratComm does require a password to log on to the multimedia server, however once logged on all users possess read, write, and delete access. This makes the original files vulnerable to accidental deletions or changes that overwrite the original.

The two areas in which StratComm earns drastically different rankings is in File Fixity and Metadata. StratComm currently does not have any tools for performing fixity checks to ensure data integrity during any stage of use or storage. As such they do not meet any requirement for any level in this category. However, in the category of Metadata StratComm has been applying standard technical and descriptive metadata embedded in the images for a number of years, albeit with differing approaches that have evolved with the introduction of new tools and new guidance from changing staff. The application of this type of metadata satisfies the requirements for Level 3 of the Metadata category.

The following digital preservation plan will be broken down according to the categories detailed in the National Digital Stewardship Alliance's Levels of Digital Preservation rubric. While these requirements may seem prescriptive it should be noted that effective digital preservation is performed in the context of the organization's preservation intent, and that it is an incremental process (Schumacher et. al., 2014, p.15). That is to say that choices that might make sense for the Library of Congress may not make sense for a multimedia production house like StratComm, and vice versa. It also means that organizations are best advised to focus on the next logical and achievable steps to improve preservation rather than stress about trying to meet the highest standard all at once. Forward progress is best maintained by concentrating on the next achievable steps. Towards that end each section below will include recommendations that require low, medium, and high resource expenditures. I recommend concentrating on the low and medium recommendations as more immediate actions to take. The recommendations requiring higher resource expenditures are included for purposes of long-term organizational planning and to highlight opportunities to partner with other groups whose functional mandate may be more appropriate for addressing the preservation need in question.

Storage and Geographic Location

At present, StratComm does not quite meet the criteria for Level 1. In most cases, two distinct copies of all files do exist, but they are collocated on a backup server that resides in the same location as the original server. While this does protect StratComm from the possibility of hard drive failure it leaves them vulnerable to disasters that could potentially affect both systems, such as floods or fires. Washington's photographer recently completed a migration of all data residing on optical media for the 1995–2009 segment of the image collection. This is a positive step, however that data too is now collocated on the same servers. Technically the copy on optical media does mean that three distinct copies exist for that segment of the collection, but this media should not be relied on due to its age, and that optical media is physically located on the same floor of the same building as the multimedia server that houses the other two copies, giving it the same threat profile.

Table 1: the Levels of Digital Preservation (Storage and Geographic Location)

	Level 1 (Protect your data)	**Level 2 (Know your data)**	**Level 3 (Monitor your data)**	**Level 4 (Repair your data)**
Storage and Geographic Location	**- Two complete copies that are not collocated - For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system**	**- At least three complete copies - At least one copy in a different geographic location - Document your storage system(s) and storage media and what you need to use them**	**- At least one copy in a geographic location with a different disaster threat - Obsolescence monitoring process for your storage system(s) and media**	**- At least three copies in geographic locations with different disaster threats - Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems**

Note. Adapted from 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

Next Steps

Low resource measures. Probably the shortest line to achieving Level 1 status in the Storage and Geographic Location category would be for StratComm to take advantage of their bifurcated office configuration and begin to back up each location's multimedia server to the other location's multimedia server. This would establish a third copy of every file, but more importantly, at least two of those copies would exist in different locations with different threat profiles. In a single stroke this would satisfy the highest Level 4 requirements for number of copies and location of those copies.

The remaining requirements for Level 2 call for an audit of the current storage system and what is required to maintain access to it (Phillips, Bailey, Goethals, and Owens, 2013, p.3). This is a simple matter of a small amount of staff time to perform the inquiry and requires a fairly low expenditure of resources.

Medium resource measures. The remaining requirements for Levels 3 and 4 require ongoing, but still minimal amounts of staff time for monitoring and planning activities. For example, to meet Level 3 requires the establishment of a process for monitoring obsolescence issues associated the organization's storage systems (Phillips et. al., 2013, p.3). This could come in the form of a simple annual audit of StratComm's current storage status and an analysis of current trends in storage options.

To meet Level 4 requirements requires the creation of a comprehensive plan to maintain both files and metadata on an accessible media or system (Phillips et. al., 2013, p.3). While this is not a terribly time intensive effort it probably represents the highest degree of time spent in active monitoring of current technological and industry trends and careful analysis of how those trends fit into the context of StratComm's preservation intent and workflow. The inter- and intra-departmental collaboration required to form this plan would likewise require a moderate expenditure of staff time.

High resource measures. StratComm could further diminish the threat profile of its three separate copies by establishing a third full copy of all files in a commercial cloud server such as Amazon Web Services. This would place a third copy in a third location with a different threat profile from the previous two copies. This diversified threat profile also establishes a new controlling organization as an additional level of potential protection. For example, if some catastrophic network event were to cripple all of StratComm's computing power the copy in the cloud would likely remain accessible.

I have this option listed as a high resource expenditure due to the complications surrounding StratComm's public release process and the effects that process has on the company's approval of cloud storage. To make this option happen would likely require a fee for the commercial service as well as the engagement of multiple employees across many departments including StratComm, IT, and Information Security to gain corporate approval and establish non-disclosure agreements with the vendor.

File Fixity and Data Integrity

At present StratComm does not employ any measures for checking file fixity or data integrity. As such they do not meet the requirements for any level of the National Digital Stewardship Alliance's Levels of Digital Preservation rubric.

Table 2: the Levels of Digital Preservation (File Fixity and Data Integrity)

	Level 1 (Protect your data)	**Level 2 (Know your data)**	**Level 3 (Monitor your data)**	**Level 4 (Repair your data)**
File Fixity and Data Integrity	**- Check file fixity on ingest if it has been provided with the content - Create fixity info if it wasn't provided with the content**	**- Check fixity on all ingests - Use write-blockers when working with original media - Virus-check high risk content**	**- Check fixity of content at fixed intervals - Maintain logs of fixity info; supply audit on demand - Ability to detect corrupt data - Virus-check all content**	**- Check fixity of all content in response to specific events or activities - Ability to replace/repair corrupted data - Ensure no one person has write access to all copies**

Note. Adapted from 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

Next Steps

Low resource measures. StratComm could avail itself of free cryptographic hash generators such as onlinemd5.com to both create checksums for their data as well as verify the existing checksums to ensure data integrity has not been compromised. This hash could be stored in a simple text file inside the folder of final images. One consideration is that the high volume of images produced by StratComm may preclude performing fixity checks on an item level. It may be more time efficient to package all final files into a zip file and perform the fixity check on that. But aside from the staff time involved with manually performing the tasks, all of these efforts employ freely available or already existing tools.

Medium resource measures. StratComm might be able to use corporate initiatives such as the "Meaningful Work" program to engage with programmers in other areas of the company that are still waiting for sponsor tasking. These programmers could automate certain tasks, such as the regular intervals of fixity checks required of Level 3, or perhaps the checking of fixity before and after transfers. Transfers of data from one system to another are one of the biggest points of vulnerability during which degradation can occur (National Digital Stewardship Alliance, 2014, p.3). Checking the fixity before and after such events is a Level 4 requirement (Phillips et. al., 2013).

Another option would be to approach EnterMedia's software engineers about the possibility of building fixity checks into StratComm's present DAM system. StratComm already owns this tool, and the vendor is particularly engaged and open to producing custom solutions to such problems.

High resource measures. There are tools available, such as AVP's Exactly, which will help with the automation of many of the tasks associated with checking fixity. These are open source options, and many are even free of charge, however as Owens notes in *Theory and Practice of Digital Curation*, these should be regarded as "free puppies", not "free beer" (Owens, 2018, p.116). Open source solutions may not cost anything to download, but they require a great deal of setup, care, and maintenance, often from multiple departments across the organization. Additionally, since these solutions are often the product of non-profit, cooperative industry partnerships it is important to factor in the active role that StratComm may need to assume in the community stewardship of that tool in order to help ensure the tool's continued existence and make sure that it continues to serve StratComm's goals (Owens, 2018, p.116).

A tool such as AVP's Exactly may be overkill for StratComm to procure and stand up on its own, however it might make sense for StratComm to partner with Corporate Archives to lobby for the resources necessary to implement such a system. This shared capability would benefit the collections held by both departments. The significant need for a budgetary increase has been hidden for many years by the organizational gap between the two departments. On one side, StratComm recognizes the vast scale and importance of their collection yet tends not to view preservation as a role within their purview, so no formal process for transitioning content to Corporate Archives exists. On the other side, Corporate Archives does view preservation as their function, however they have no current foothold in any stage of StratComm's digital asset lifecycle, so they are in no position to properly assess or appropriately plan for the care of such a large collection. A StratComm partnership with Corporate Archives would present a unified voice to articulate the true scope of the problem and the resources needed to address it.

Information Security

StratComm does presently require a password to access the multimedia servers, however no additional security provisions are in place to restrict access to the photography portions of each server, or to restrict changes and deletions to the files that reside there. For the most part the only individuals that access these servers with any regularity are the photographers and other multimedia staff, but login information has occasionally been provided to design staff in the past with deleterious effects. For example, it was discovered a few years ago that someone had moved an entire folder of images to a new location on the same server, but could not remember the new location. This broke the link that Lightroom had with the original files, making them unavailable to photographers for further use, effectively deleting them.

While the incident described above was the only example of data loss that I could find it does highlight a significant vulnerability as well as a gap in understanding among the multimedia and design staff about how to safely interact with the files on the server.

Table 3: the Levels of Digital Preservation (Information Security)

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Information Security	**- Identify who has read, write, move and delete authorization to individual files** **- Restrict who has those authorizations to individual files**	**- Document access restrictions for content**	**- Maintain logs of who performed what actions on files, including deletions and preservation actions**	**- Perform audit of logs**

Note. Adapted from 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

Next Steps

Low resource measures. The requirements for meeting Levels 1 and 2 of the Information Security category involve simple OS-level changes to read/write settings and perhaps the creation of a department policy that stipulates safe avenues of access for design staff (Phillips et. al., 2013). Such efforts are free of charge and would satisfy all requirements for Levels 1 and 2.

Medium resource measures. To meet the requirements for Levels 3 and 4 would require the acquisition of new software, or the augmentation of current software, to monitor changes and produce activity logs (Phillips et. al., 2013). One example of such an option would be to engage with EnterMedia's software engineers to customize a solution that would generate audit logs.

While this is an option of medium-level cost it is worth mentioning that it may provide very limited return on investment. EnterMedia already does not provide access to originals, it simply allows the user to download a copy. Thus the original is never in any particular danger. Plus StratComm's desire to shift the primary access point for all users to the EnterMedia DAM system provides its own level of protection for the original files that reside on the multimedia server. For the assets on the multimedia server it might be worth exploring the option of write blocking software under the condition that the implementation of such a tool must not present any undue complications to StratComm workflow.

High resource measures. Following on the discussion from the medium resource measures above, it bears mentioning that a significant restructuring of interdepartmental processes and staff roles should be considered. The functions outlined in Levels 3 and 4 would most logically reside with Corporate Archives. This is another example of where it would be beneficial for StratComm to partner with

Corporate Archives to lobby upper management for an increase in resources to build a shared, corporate-level ability to manage and preserve digital assets.

Metadata

Photographers in both Boston and Washington currently apply standard technical and descriptive metadata to all images, albeit with some variations in method. This satisfies the requirements for Level 3 of the Metadata category, but leaves the requirements for Levels 1,2, and 4 unmet. It is noteworthy to point out that several aspects of the requirements for Levels 1 and 2 are points of crosspollination with the Storage and Information Security categories discussed above. As such, measures taken to satisfy those requirements would perform double duty in satisfying requirements in this category as well.

Table 4: the Levels of Digital Preservation (Metadata)

	Level 1 (Protect your data)	**Level 2 (Know your data)**	**Level 3 (Monitor your data)**	**Level 4 (Repair your data)**
Metadata	**- Inventory of content and its storage location - Ensure backup and non-collocation of inventory**	**- Store administrative metadata - Store transformative metadata and log events**	**- Store standard technical and descriptive metadata**	**- Store standard preservation metadata**

Note. Adapted from 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

Next Steps

Low resource measures. Performing a formal inventory of content and its storage location would be a relatively low expenditure of effort for a staff member to perform. It could even be argued that the Survey and Report that preceded this document may actually satisfy the requirement. However, since these locations do not change drastically very often, this inventory would at most need to be done annually. This action, combined with another previously mentioned low resource measure – namely, the trading of backup copies between the branch offices – would quickly achieve a Level 1 status for StratComm.

Medium resource measures. StratComm could approach EnterMedia's software engineers to inquire about the platform's ability to track transformative and preservation metadata. If this customization to an existing StratComm platform is possible, it would perform the necessary functions for Levels 2 and 4. However, these changes might be outside the realm of possibility for EnterMedia, and it is also worth considering if EnterMedia occupies the most optimal point in department workflow to act as a gatekeeper for tracking change history and version control. It may be more logical to explore options that integrate into StratComm's workflow management tool, Wrike.

High resource measures. There is a dearth of good options that meet both the strict archival standards outlined by the NDSA while also meeting the needs of a high-volume production house like StratComm whose high rates of transformative activities create version control issues that would overwhelm archival tools meant to catalog static files. Open source, archival platforms such as DKAN could easily meet all of the standards outlined in each level of the Metadata category, however such tools are not very interoperable with the graphics platforms used by StratComm, making it likely that design staff would simply sidestep the DKAN interface and change histories would be lost. There are proprietary commercial tools such as Adobe Experience Manager (AEM) designed to perform all of these functions, however AEM's cost is extreme (in the millions) and the proprietary nature of the platform makes the formation of an exit strategy to a subsequent tracking system exceptionally difficult. Considering these factors, StratComm's acquisition of either DKAN or AEM, while technically an option, is honestly not recommended.

A more logical and cost-effective option to explore would again be to partner with Corporate Archives to advocate for the acquisition of a shared archival tool to enhance the corporation's capability to preserve digital object metadata. Part of this would necessarily involve a focus on the creation of preservation policy to formalize a pipeline for content to flow from StratComm to Corporate Archives.

File Formats

StratComm currently satisfies portions of each level of the File Formats category, but does not completely satisfy any one of them. For example, it could be argued that StratComm has chosen their currently used file formats based on an informed assessment of user community needs and long-term access issues. However, all of the formats they use are technically proprietary. It could also be argued that the Survey and Report document that preceded this paper could be considered an inventory of formats in use. However, the scope of that document was more of a broad overview than an in-depth analysis. Achieving Level 2 will likely require a more dedicated audit of file formats, as well as a deeper examination of the specific issues associated with each one, particularly the use of DNG and the larger role that RAW files play in StratComm's preservation intent and service model.

StratComm comes closest to meeting Level 3. Both photographers remain abreast of current industry trends in digital imaging, and both make long-term access a key driver for their format choices. The largest point of diversion is in RAW formats: a class of files for which no non-proprietary choice exists and longevity of access is an open question for all. The photographers' diverging approaches in this area are each supported by thoughtful reasoning, but are equally debatable and based on significantly different preservation intents. This lack of consistency is another argument for the need to perform an in-depth inventory of file formats, a re-examination of StratComm's preservation intent, and a subsequent synchronization of policy.

Table 5: the Levels of Digital Preservation (File Formats)

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
File Formats	**- When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs**	**- Inventory of file formats in use**	**- Monitor file format obsolescence issues**	**- Perform format migrations, emulation and similar activities as needed**

Note. Adapted from 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

Next Steps

Low resource measures. The actions required for meeting levels 1, 2, and 3, are all fairly modest efforts. In fact, the in-depth inventory of file formats recommended above would satisfy Level 2, and go a long way towards informing the best choices available for satisfying Level 1.

High resource measures. StratComm's legacy holdings are a complex panoply of image file formats, made all the more challenging by the occasional lack of file extensions and a high prevalence of an obsolete Kodak RAW format. Tools such as DROID (Digital Record Object Identification) will help identify files without extensions, and tools such as Preservica or Rosetta would help facilitate format migrations to maintain functional access to files. However, such commercially available tools require ongoing fees and a high degree of effort to install, configure, and maintain. Additionally, since the last known RAW converter for the Kodak DCS465 required an obscure plug-in for Photoshop 5.5 on Mac OS 9 the effort and expense of emulating that environment for the purpose of migrating those files may be considerable.

Due to the high level of effort involved with these measures, it is again recommended that a partnership with Corporate Archives be considered to build out a shared capability for digital preservation that meets the needs of both departments.

Conclusion

As stated in the introduction, digital preservation is a process that is necessarily informed by the context of the organization's preservation intent, and it is incremental in approach (Schumacher et. al., 2014, p.15). As a high volume production house it may not make sense for StratComm to strive for the highest levels of the NDSA Levels of Digital Preservation rubric because it is not within the scope of their corporate function. At the same time the interests served by striving for those upper levels of the rubric do represent a critical value to the company StratComm serves. As such, it is recommended that StratComm focus the bulk of their efforts on the curatorial and preservation activities that directly enable their business goals of improved image search. Activities outlined in this report that most directly serve that goal are improvement of metadata and the generation of metrics on image usage, including change histories and version control. A reappraisal of the collections featured on EnterMedia is also warranted to bring them in alignment with current corporate structure in a way that new employees would find logical. One possible means of doing this would be to switch to a system that tags images according to sponsor rather than by center. The centers change every time the company goes through a reorganization, whereas the sponsor remains much more persistent.

Addressing the more complex preservation needs really deserves a more concerted approach that answers the needs of the company as a whole, not just the needs of any one department. StratComm has an opportunity in this area to help advance the Corporate Operations center by lending its considerable voice to the call for increased resources dedicated to the long-term management and preservation of digital objects. And the most logical place for this capability to reside is Corporate Archives.

References

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-Guidance-Report-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

Digital Preservation Policy

I. Purpose

This document describes the activities and responsibilities required for the preservation of Strategic Communications' (StratComm) digital image collection. It is intended to outline the asset's needs at each stage of the asset lifecycle and to diagram the organizational relationships necessary to execute each stage.

II. Mandate

StratComm creates and manages visual assets for its own use in communicating about the company's work program, and also to enable the company's employee population to produce their own communications products via self-service tools. This insight and self-efficacy in the use of the image collection at every level of the company directly strengthens StratComm's brand objectives.

III. Definition of User Group

StratComm maintains the collection for its own use, and on a broader level, to enable company employees to communicate about their work. In that context, it is the goal of StratComm to preserve functional access to visual assets. As such, migrations of storage media and file format may be undertaken to ensure persistent access to these items is maintained in future computing environments.

Documenting the history of the company is also a priority, and StratComm partners with Corporate Archives in facilitating the retirement and long-term preservation of visual assets for that purpose.

IV. Scope

This policy applies to the following materials:

- All photography produced by StratComm staff.
- Some photography submitted by company employees, provided it has some reuse value in communicating about the company work program.
- Some vector graphics produced by employees, provided they have reuse value and align with the company's brand guidelines.

V. Collaboration

A Corporate Digital Preservation Committee will be formed in order to bring together the various stakeholders in the Strategic Communications, Corporate Archives, and Enterprise Computing departments. The group will meet at least annually to perform any scheduled audits and to evaluate opportunities to advance the search capability of the image collection.

The committee will include R504 and R301 leadership, Corporate Archives staff, StratComm photography staff, and representatives of Enterprise Computing capable of facilitating storage and backup services.

VI. Asset Lifecycle

StratComm's visual assets require different stewardship activities at each stage of their lifecycle. These stages and needs are outlined below.

A. Creation

1. Preferred Formats

For final, edited images uncompressed TIFFs are optimal, however provided no exposure adjustments are needed, they are indiscernible from JPEGs saved at the highest possible quality setting. As such, JPEGs are acceptable.

Despite the more open documentation of Adobe's DNG RAW format, its low adoption levels and scant software support make it somewhat less attractive as an archival format. In fairness, the ubiquity of Nikon and Canon native RAW formats offers no greater guarantee. In the interest of preserving maximum interoperability with today's available RAW processing tools, native Nikon and Canon raw formats (NEF and CR2) are preferred, however DNGs are acceptable.

For vector graphics, SVGs are preferred, however AI and EPS files are acceptable.

2. File Naming Conventions

Imported RAW files should retain the camera-assigned sequential number, and ideally, also apply the creation date to the beginning of the file name in order to minimize the likelihood of duplicate filenames. They should look similar to the following standard:
YYYY-MM-DD_D4S####.NEF

Exported final files should add the project or customer name to the beginning of the standard name outlined above for RAW files. This standard should like similar to this:
Project_Name_YYYY-MM-DD_D4S####.NEF

B. Management

1. Storage

a) Number and Location of Copies.

Three copies of final edited files should be established in different geographical locations in order to diversify StratComm's disaster threat profile. To accomplish this, each location will save all originals to their site's multimedia server. This copy will get duplicated to a corporate-level backup server at the same site, and then each location will send a third copy to a server at the other site to protect against any widespread disaster effecting the original site.

b) Folder Naming Conventions.

The photography section of the multimedia server should be organized by year. Each year folder will contain job folders organized chronologically by date, after which each will contain a few words explaining the project or customer. They will adhere to the following standard: YYYY-MM-DD_Project_Name

Folders for each month are acceptable practice for individual use in Lightroom, but should be avoided on the server where browsability by the group is of greater importance than brevity.

2. File Fixity

Checksums should be generated on a collection level for each job after all manipulation and processing of those files is complete. All files from that job will be zipped into a single package and a checksum will be generated for that package.

File fixity should be checked prior to, and after, any transfers or copies are made of the final edited files. The checksum should be stored in a text file accompanying the zip file that houses the collection of images, and it should be supplied to Corporate Archives for verification of fixity upon retirement of the assets.

3. Information Security and Access

Read, write, and delete access to the photography portion of each site's multimedia server should be restricted to the photography staff and their immediate supervisors. The wider multimedia and design teams may have read-only access to this area, but should be encouraged to use the EnterMedia DAM system as the primary search tool.

4. Metadata

StratComm will record metadata at various points during the asset lifecycle to optimize the asset's useful potential.

a) Technical

Cameras automatically record technical metadata associated with each capture, and record it as EXIF metadata. Photographers will ensure that EXIF metadata is conveyed from the RAW capture through to the exported final image.

b) Descriptive

Photographers will apply descriptive metadata on a collection level for each job using Lightroom and Bridge prior to uploading into the EnterMedia DAM system. This will consist of a title, description, and keyword tags that are aligned with the company's subject taxonomy terms.

c) Access and rights

Photographers will program their cameras and metadata templates in Lightroom and Bridge to apply authorship, copyright, and usage guidelines automatically during import and export functions. This will include photographer name, contact info, copyright, and sensitivity level.

d) Preservation

Preservation metadata will be created in collaboration with Corporate Archives staff at the time of retirement.

C. Distribution

The EnterMedia DAM system will act as the primary access point for both StratComm staff and the wider population of company employees. This tool will feature custom, shareable image galleries for delivery of images to customers in a manner that is browsable and also enables download of the high resolution file.

D. Retrieval

The EnterMedia DAM system will also feature both keyword search functionality as well as curated collections of images to facilitate easy browsing of common subject matter.

StratComm will also partner with KICS to optimize the contents of EnterMedia to make them interoperable with the company's EnterpriseSearch capability.

E. Archiving

1. Sunset Date

StratComm's desire to use the most recent material possible in most communications products results in a diminished prevalence of usage with age. As such, StratComm will transfer all files to Corporate Archives once they surpass 5 years of age in order to preserve these objects for historical purposes and to make room for new assets. This transfer to Archives will be performed annually.

2. Selection and Accessioning

At time of retirement Corporate Archives and StratComm photography staff will meet to determine which files are appropriate for long term preservation and which can be safely deaccessioned. This may or may not include the deletion of RAW files. Corporate Archives will have ultimate say in the selection of items for long-term preservation.

If StratComm believes an item – or collection of items – should be retained for active usage longer than the 5-year sunset date, or in a non-archival format not supported by Corporate Archives, StratComm will assume the responsibility of maintaining that library of assets for nearline access.

3. Transformation and Migration

At time of retirement Corporate Archives and StratComm photography staff will collaborate on any required format migrations necessary to normalize the collection to an archival standard.

Corporate Archives reserves the right to perform transformations of the original files in order to maintain functional access in future computing environments.

VII. Audits

Annual audits should be performed to check fixity of currently held assets, to facilitate the retirement of assets to Corporate Archives, and to review this policy document for any needed changes. The parties responsible for each of these efforts are outlined below.

A. Fixity Audits

StratComm photography staff will assume full responsibility for performing all audits of file fixity, both before and after transfers, as well as annually to protect against bit rot.

B. Retirement

Corporate Archives and StratComm photography staff will work in tandem to perform annual reviews of materials eligible for retirement to Corporate Archives

C. Policy Review

The full Corporate Digital Preservation Committee, including R504 and R301 leadership, Corporate Archives staff, StratComm photography staff, and representatives of Enterprise Computing, will annually meet to evaluate the need for any changes to this document.