

Jumping In – Testing Mobile Capture in Workflows Where Production and Preservation Collide

D. Michael Hansen, Global Vendor Manager; FamilySearch International; Salt Lake City, Utah/United States

Abstract

FamilySearch International, formerly the Genealogical Society of Utah, established in 1894, has been actively capturing images of records and books of genealogical value for nearly 90 years in microfilm or digital format worldwide. As new technologies have emerged in AI and image processing, several FamilySearch digitization projects have begun successfully implementing mobile phone solutions where practical camera work has become difficult to expand or initiate. This paper highlights a few mobile digitization projects in South Africa, Japan, and a home-grown solution.

Brief History of FamilySearch

Not long after our first microfilm project began in 1938, FamilySearch's collection in 1954 exceeded 100,000 microfilm rolls. Our first digital camera operation started capturing images in 1998. Then, in 1999, we launched our new FamilySearch.org website and began digitizing microfilm in 2000. Our last microfilming project finished in 2012. Later, in September 2021, we completed converting 2.5 million rolls of microfilm to digital images. Our current collection of over 20 billion images comprises records from over 170 countries all over the world. Recent years have driven 285 million visits to our website each year.

We do not operate for profit, nor are we a cost recovery organization. FamilySearch International is sponsored by The Church of Jesus Christ of Latter-day Saints which provides resources to help millions of people around the world discover their heritage and connect with family. FamilySearch's resources are free of charge to everyone, regardless of tradition, culture, or religious affiliation. We help people discover their family's history through our website, mobile apps, and in-person help at over 5,000 local family history centers, including the largest genealogical library in the world in Salt Lake City, Utah. We are dedicated to preserving important family records and making them freely accessible online, subject to local laws. Our work is worldwide and involves collaborations with more than 10,000 organizations in over 100 countries.

Our Standard Preservation Workflow

FamilySearch has deployed nearly 600 camera workstations throughout 46 countries for records digitization projects in archives worldwide. Many projects rely on a single camera operator, a computer, and a hard drive of images shipped back and forth either to Salt Lake City or a local ingest center. Conversely, some projects have many cameras, a fiber network connection, and upload images and metadata directly to a receiving bucket within our Amazon s3 cloud platform. In 1998, we started with the latest 16 MP Red Lake cameras, then improving to more sophisticated 29 & 50 MP Ilumis sensors. As color capture became more of a requirement driven by demand from record custodians, we introduced Nikon D610, D810, D850, and now the Z7ii camera models.

Primary specifications for digitization have typically remained fixed at 300 pixels per inch minimum resolution, color or grayscale,

in an uncompressed TIFF master image format, unless a different resolution or specification is required by the archives. We spend considerable time with each archive determining the feasibility of their required specifications, and what we can deliver within the desired timeframe for each project. Our camera workstations are run by a software platform created internally by FamilySearch called DCam and installed on location at the archives (Figure 1).



Figure 1. FamilySearch DCam camera workstation. (Image courtesy FamilySearch.)

DCam is used to automatically and consistently calibrate the camera, review image quality (Figure 2), and store each image in a specific folder structure called a "natural group," or the smallest grouping of collected items that someone could naturally hold in their hand, such as a box of loose sheets, a folder within a box of folders, or a book on a shelf. DCam checks for quality issues, such as clipping or focus issues, and provides tools for the operator to manually review, such as a histogram, and retake necessary images.

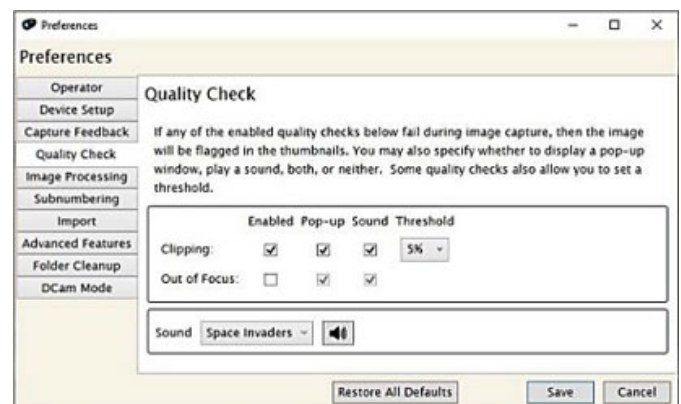


Figure 2. Adjusting quality settings in FamilySearch DCam digitization software. (Image courtesy FamilySearch.)

DCam requires certain fields of metadata for camera operators to correctly specify the record type being captured, the year of the event, and the location of the event (Figure 3). Metadata is saved as a custom METS XML schema verifying that names of record types and places have been validated against FamilySearch's standardized databases. All images and metadata for the captured batch are then transferred locally to a hard drive shuttle and physically shipped to FamilySearch in Salt Lake City, Utah, USA, or to a local ingest center equipped to connect the shuttle and upload images and metadata to FamilySearch's servers.

Use information from folder
 (None)

Cover Tag
 Select a Cover Tag

Record Title
 Inventories, Index T601

Volume Designation

Dates
 1699-1965 ☐ Month ☐ Day
 Year Year
 1699 To 1965 Add
☐ These dates are confirmed to be correct.

Archive Reference Number
 If there is no Archive Reference Number for this Folder, type - instead

Locality
 Search (Remote Database) Id: 395453
 Baltimore, Maryland, United States
 ✓ Standardized Place: Baltimore, Maryland, United States (County, 1776-Present)
 Add More Specific Place

Language
 Primary Language: English Other Languages:
 View in DCam Language

Record Type
 Search Add Record Type Id: 126786
 Legal Records, Court Records, Probate Records, Probate Proceeding Records
 ✓ Standardized Record Type: Legal Records, Court Records, Probate Records, Probate Proceeding Records

Figure 3. Metadata fields required for camera operators using FamilySearch DCam digitization software. (Image courtesy FamilySearch.)

Going Mobile

Preservation vs. Practical

Our standard approach is to digitize on-site at the location where records of genealogical value are housed rather than try to create regional digitization centers across the world. For one thing, it is common that records cannot leave the archives' premises.

Furthermore, in many countries it is difficult or expensive to procure camera equipment, service the equipment when issues arise, ship computer equipment across borders without large fees, or find a stable internet connection, or even a reliable power grid. For these reasons, a growing percentage of our digitization projects are now being done beyond our standard DCam workflow, now relying on vendors' own digitization personnel, skills, equipment, and expert processes.

In today's world where mobile phones are easily accessible, portable, and readily connected online, the appeal of using a cell phone for digitization can seem both inexpensive and convenient when compared to traditional camera capture methods. Our goal in cultural heritage digitization is to capture a preservation-quality master image, which then enables FamilySearch to achieve its innate goal of identifying a person's name, event, place, and date on that image with considerable quality. It is our experience that with careful consistency, mobile phones can indeed pass our minimum specifications for image quality, but with some operational rigor and implementing best practices.

Adopting a new digitization workflow where images and metadata can be submitted to FamilySearch from a variety of imaging devices, such as mobile phones, means that we need to be able to work with a variety of changing conditions. New questions are posed that make us carefully consider:

- Are we willing to accept a high-resolution JPEG instead of a TIFF? If yes, then under what conditions?
- Can we work with other high-definition raw formats for mobile devices, such as HEIF or HEIC?
- At what point in our downstream process is TIFF still necessary, especially in a workflow that relies on JP2 format for long-term preservation?
- Is there a way to use a mobile phone to capture consistent results, or do auto-settings make it impossible to control performance and quality?

These may all seem like simple questions with straightforward, scientific answers, but in the arena of real-world digitization where weekly production matters, where cameras are expected to produce between 7,000-10,000 images weekly, where AI analysis will now check image quality for 100% of incoming images without having to sample the image set, and as full-text searching within images becomes more thorough, accurate, and a robust tool to be used in so many other ways – indexing, redaction, advanced search – how much do we need to know before we can even begin a project? Do we have to change our entire downstream process to handle all the new change?

Thanks to vendors worldwide working on FamilySearch projects, our ability to stand on the sidelines and debate what's coming is like being on a train that has already left the station. We have already leaped into a preservation workflow change internally and created new paths for vendors to digitize images and submit images and metadata to us using newly engineered processes and tools that do not rely on our older DCam workflow. We can now accept images in many formats provided they pass a FamilySearch minimum standard threshold. We have a growing but more limited number of tools that will map metadata into the schemas we need to publish records online. Vendors themselves have been innovating new solutions on their own for submitting images, metadata, and indexes to FamilySearch beyond our standard publication path, and we ourselves have had to develop new tools to keep up. With so many devices being used in so many ways, it is also our experience

that using a target-based workflow can help ensure that multiple devices produce comparable results, especially when not everyone around the world relies on FADGI or Metamorfoze preservation guidelines and standards.

Project Highlights, New Solutions

FamilySearch has over 180 cameras in Africa digitizing records in South Africa, Nigeria, Ghana, and several other countries. One project in a South African archive houses three rooms dedicated to FamilySearch's camera capture project digitizing government records using a total of 16 DCam-Nikon Z7ii camera workstations side-by-side. The third room contains a vendor-created solution of 10 innovative mobile phone camera stations (Figure 4). Starting at the beginning of 2025, all 26 stations have been working on the same primary project digitizing over 16M images (approx. 6M records), and the project is forecasted to be completed later this summer.

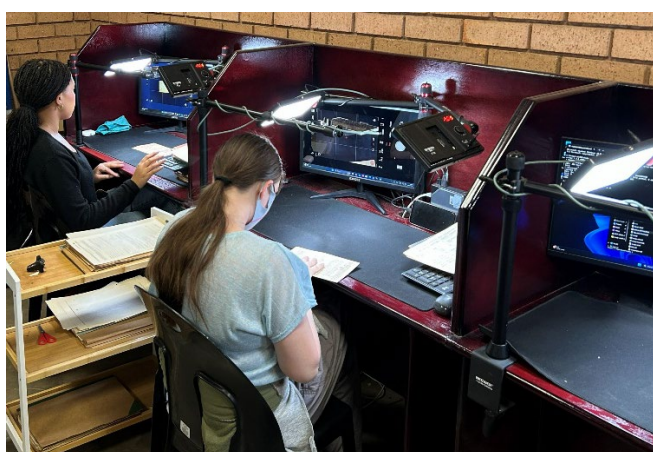


Figure 4. Mobile phone digitization workstations in South Africa. (Image courtesy FamilySearch.)

While many Apple iPhone® models are touted specifically for their image quality, some FamilySearch vendors are branching out and adopting the use of different mobile phone models from other manufacturers, such as the Xiaomi 14 Ultra®, which uses a 50MP sensor and a large, Leica Vario-Summilux® lens (Figure 5).



Figure 5. Mobile phone with 50MP camera used in records digitization project. (Image courtesy FamilySearch.)

This particular project location became ideal for testing the performance of not just one camera workstation in isolation, but for

the possibility of comparing multiple devices with one another. Each camera workstation was prepared to capture a variety of industry-standard imaging targets. Each workstation was comprised of a custom-built stand holding the camera parallel to the imaging surface, and two Sutefoto T18® LED panels illuminating the imaging area (Figure 6).



Figure 6. Top-down view of the mobile phone digitization workstation. (Image courtesy FamilySearch.)

At each camera workstation, several targets were imaged, including a ColorChecker 24, a line resolution target, and a GoldenThread object-level target. Image samples were also tested with and without a representative document in the image (Figure 7).

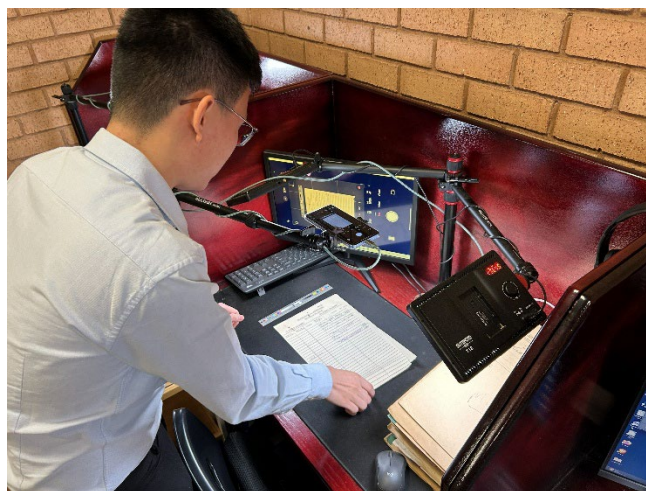


Figure 7. A vendor employee positions the document next to a GoldenThread object-level target. (Image courtesy FamilySearch.)

Once the imaging test was complete for each of the ten mobile phone workstations, sample images were then uploaded to a proprietary online portal where FamilySearch vendors doing FamilySearch digitization projects can test their own samples and perform analysis using a variety of industry-standard targets, including the well-known GoldenThread target. Ideally, sample images would then meet or exceed specifications important to FamilySearch's minimum image quality profile. Otherwise, recommendations could then be made to change settings or

parameters of the equipment based on the analysis of the target (Figure 8).

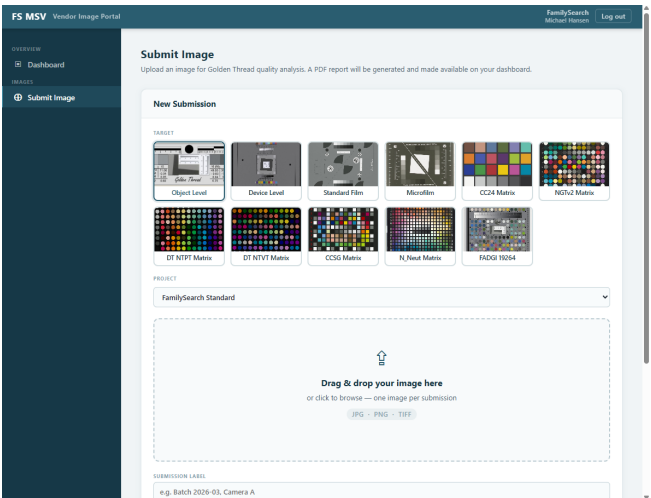


Figure 8. FamilySearch’s online portal for analyzing sample images from FamilySearch digitization projects. (Image courtesy FamilySearch.)

Interpreting the Results

Initial imaging analysis showed positive results in terms of white balance and DeltaE for color consistency, as well as overall color registration (Figure 9). This provided confidence that the mobile phone device was capable of capturing accurate color and with good RGB channel alignment. However, many grayscale patches came out light, which seemed to indicate a potential issue due to exposure clipping or LED illuminant intensity than perhaps just the mobile phone sensor itself. Additional analysis showed the sampling frequency, or optical resolution, to be in excellent position to capture at least 400 pixels per inch. Looking deeper into the reproduction scale accuracy, this metric relied simply on the EXIF metadata embedded within the data of the image headers, which, unless the user purposefully overrides the data manually with the desired setting for pixels per inch, displays an industry-standard value of 72 – useful only if downstream processes require it. FamilySearch metadata standards in fact do not use this EXIF metadata value, so the 72 value is meaningless for our internal processes. Whether we decide to remove this particular metric or not label it a contributing factor when analyzing images against a FamilySearch profile is still to be determined.

Other metrics showing positive results were the mobile phone camera’s ability to capture sharp images and clarity without the need for applying any level of enhanced sharpening, even minimal application. This was likely due to the performance of the Leica lens used on this particular device. The most concerning metric was inevitably sampling efficiency, or the ability to detect a specific level of detail before interpolating with digital blur. What this particular metric revealed is that the imaging environment was prone to motion blur – blur from the camera stand or desk itself. Creating a more sturdy environment that could withstand minor vibrations would show an improvement in this particular metric.

Overall, the image seemed to meet the FamilySearch minimum standard provided that the operator was able to address the motion blur – especially where production statistics approaches 10,000 images weekly. In this case, the vendor was able to quickly address the sturdiness of the tables and improve this particular metric.

Determining which metrics are absolutely essential for image quality vs. identifying which metrics are not as critical is also a subjective process. Making just a few recommended adjustments to the FamilySearch minimum standard profile to reduce unnecessary metrics would undoubtedly increase the percentage of compliance rating. In the end, with this project, FamilySearch was able to successfully jump into the world of mobile phone digitization thanks to vendor innovation and being able to take target-based measurements along the way.

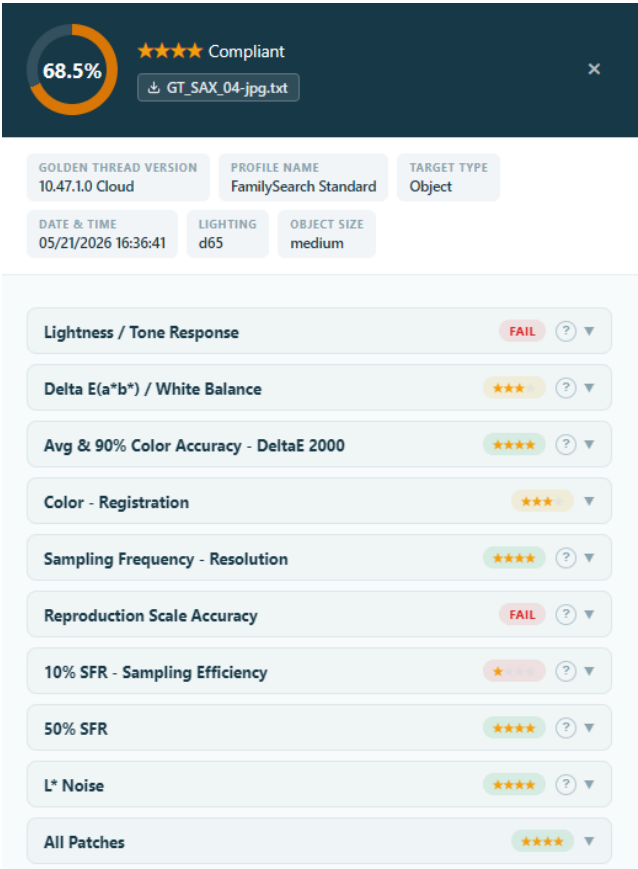


Figure 9. Initial analysis using GoldenThread object-level target for one workstation. (Image courtesy FamilySearch.)

Receiving Data for FamilySearch.org

Gathering metadata to describe the collection being digitized with specific, required fields can often be problematic for mobile phone capture. Vendors already familiar with the FamilySearch DCam workflow have created ways to transfer images from the mobile devices, then divide their digitized collection into the required natural groups, along with the minimum required metadata fields, including record type, year, locality, and any additional volume descriptors or custom fields required by the archive. FamilySearch also now requires metadata at the image level to describe the specific event, the event year, and the location of the particular event.

Images and metadata are then uploaded to a FamilySearch Amazon s3 storage location where artificial intelligence reviews each image and is currently being programmed to identify the following categories of image quality automatically:

- Motion blur
- Out-of-focus
- Improperly cropped
- Obscured information
- Wrong image orientation
- Incorrect number of pages captured in one click
- Reflection, glare
- Extra material present in the imaging area
- Exposure issues, lighting
- Missing pages
- Defective, corrupt data
- Images out of sequence
- Multiple natural groups present
- Skewing
- Image duplicates
- Compression, artifacting

Images and metadata that pass the final quality control step are then forwarded through to FamilySearch.org for publication online to be viewed and searched by patrons using the site. Data rejected are flagged as rework for the camera operator to recapture, recreate, and resubmit.

Japanese Tombstone Project

Without the use of a portable RTI imaging environment, FamilySearch is currently investigating the use of mobile phones for upcoming projects to digitize family tombstones in outside environments in Japanese cemeteries – numbering around 1M images – then systematically apply the use of artificial intelligence internally to extract the written data from the headstones, using the data to build structured family records. Such a unique project brings new challenges and new specifications for mobile imaging, including the ideal time of day when to capture, where to position the camera to avoid parallax issues while maintaining optimal resolution of the smallest engraved text, and the number of images to capture a complete surrounding representation of the tombstone, whether rectangular or rounded (Figure 10).



Figure 10. Japanese family tombstone listing deceased family members. (Image courtesy FamilySearch.)

FamilySearch has encouraged interested vendors to investigate the following camera models ideal for the project:

- **Nikon D850** (8256 x 5504 pixels = 45.4 MP) with a 24-70 mm zoom lens
- **Nikon Z7ii** (8256 x 5505 pixels = 45.4 MP) with a 24-70 mm zoom lens
- **Canon EOS R5 Mark II** (8192x5464 = 44.7 MP) with a 24-70 mm zoom lens
- Mobile: **iPhone 16** main or ultra-wide (48 MP)
- Mobile: **iPhone 17 Pro / Pro Max** main, ultra-wide, telephoto (48 MP)
- Mobile: **Xiaomi 14 Ultra**, Leica Summilux lens (50 MP)

Home-Grown Solution

When projects are less than 10,000 images, perhaps considered too small or located too remotely to send someone onsite, FamilySearch has created a portable, easy-to-use system relying on whichever mobile phone or portable camera the operator owns or prefers to use. Scan-n-Share (Figure 11), developed and refined over several years by Daren Hatch, FamilySearch Imaging Technician, gives FamilySearch vendors an option when resources are tight. Scan-n-Share allows users with their own mobile phones to have a reliable and consistent imaging environment for record digitization without having to be an expert in digital imaging.



Figure 11. Scan-n-Share system developed in-house at FamilySearch by Daren Hatch. (Image courtesy FamilySearch.)

This device is only used internally by FamilySearch or through FamilySearch-affiliated projects. However, the benefits include a wide platform size of 79 cm and 50 cm deep (31" x 19.5"), four LED light panels positioned for even, uniform lighting. The system structure with lights and carrying case weigh a total of 8 kg (17 lbs), and the time to setup or remove the system is about five minutes. A special camera mount installed on the top of the machine allows the user to lock a mobile phone's position in place, remove the phone, and perfectly position the device back in place to resume imaging should the session be interrupted.

Some vendors have even created an app that the user can install on iOS or Android operating systems, and images are then uploaded from the mobile phone to the vendor's central processing servers where they can be audited and enhanced with the correct metadata before being finally uploaded to FamilySearch.

Conclusion

FamilySearch digitization projects continue to grow at an ever-increasing rate, and digitization vendors are becoming more sophisticated with image capture technologies, new image formats, image processing algorithms and artificial intelligence, and software that can replace much of FamilySearch's home-grown DCam-Nikon camera workflow. Mobile phone digitization workflows are becoming more widely used, and several recent mobile phone imaging devices through FamilySearch vendors offer promising results after much rigor and testing. We continue to strive in building a preservation workflow that could receive images from a variety of devices that meet minimum specifications for publication on FamilySearch.org.

Author Biography

D. Michael Hansen lives in Utah and received his MS in Print Media, specializing in imaging science and color science, from the Rochester Institute of Technology in Rochester, New York, in 2006. Michael has been involved in cultural heritage imaging projects for over 20 years, including working with Roger L. Easton Jr. on images of the Archimedes Palimpsest. Michael is the Global Vendor Manager for FamilySearch, where he started in 2015.