https://doi.org/10.2352/issn.2168-3204.2026.23.1.5
This work is licensed under the Creative Commons Attribution 4.0 International License.
To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Mobile Imaging and Metadata Capture for Cultural Heritage Preservation

Dean Serrentino, Historiq; Peter Burns, Burns Digital Imaging, LLC; Lilli Cooper, Historiq; Michael Sao Pedro, Historiq; Jay Bosanko, Historiq; Juilee Decker, Rochester Institute of Technology; Don Williams, Image Science Associates, LLC; Thomas Rieger, Rochester Institute of Technology

Abstract

Large-scale public participation, using everyday tools like smartphone cameras alongside metadata captured at the moment of documentation, may be one of the most practical ways to preserve vast amounts of undigitized cultural heritage. We describe how contemporary mobile devices can support selected cultural heritage digitization tasks within a practical quality framework consistent with ISO 19264 and FADGI. We then outline how digital collections, particularly for small-to-medium-sized institutions, can complement mobile-device capture with automated metadata generation and cataloging.

Introduction

The global capacity to digitally preserve our shared cultural heritage in centralized facilities is much less than the rapidly expanding amount of heritage materials being produced. As a result, only highly curated materials can be digitally included in the collections that will become our shared legacy. Enabling local capture of material valuable to future understanding of heritage can foster greater documentation and collective heritage preservation efforts.

Mobile devices (phones and tablets) now combine high-resolution sensors, computational imaging, and increasingly mature capture controls that are portable and already familiar to many users. Although mobile-phone cameras are being used for digital collections, there can be a general reluctance to general adoption.

For archives, museums, libraries, and smaller repositories, particularly those in underserved or resource-constrained communities, this raises a compelling question: can a mobile device support any part of a serious digitization program, and if so, under what conditions? The question matters not only for cost but also for access.

Smaller institutions, distributed projects, field documentation, and rapid-response capture often lack practical and cost-effective access to dedicated reprographic systems. As the need for phone-based capture grows, its imaging performance evaluation has been addressed. Imaging guidelines, such as FADGI focus on measurements rather than technology, so mobile-phone capture falls within this framework. Previous efforts to evaluate mobile imaging performance have been reported in Ref. 2 for image sampling and resolution, and Ref. 3 for color capture.

However, such devices introduce variables that complicate direct comparison: computational image processing, limited operator control, proprietary image pipelines, changing firmware behavior, and inconsistent metadata exposure across devices. In addition, optical limitations due to the reduced size limit imaging performance. In some projects, it may be necessary to adjust imaging standards when time and resource constraints limit ideal conditions, particularly when access, preservation, or documentation needs take precedence over technical perfection.

For example, supervised interns may undertake capture to support the work of trained staff. A second example may be when fieldwork in off-site or remote locations may require working without full studio equipment or controlled environments. Similarly, in contexts where there is urgency to document at-risk materials, due to environmental, social, or logistical challenges, the priority may shift toward capture and access, even if this results in lower-resolution or less standardized outputs.

However, such devices introduce variables that complicate direct comparison: computational image processing, limited operator control, proprietary image pipelines, changing firmware behavior, and inconsistent metadata exposure across devices and applications.

One arrangement for the use of mobile phones could be as part of a *hub-and-spoke* model, as shown in Fig. 1. The spokes represent the secure exchange paths for transferring images and metadata to the hub. The hub processes incoming content, performs image post-processing, and validates image quality and metadata before archiving.

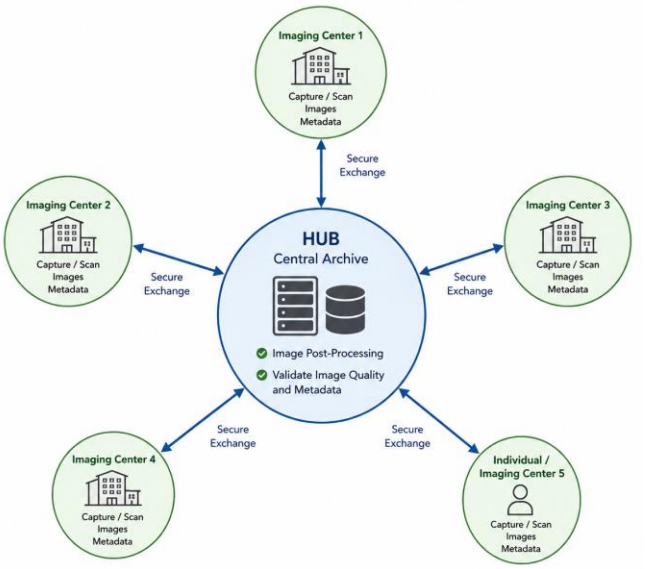


Figure 1: Hub-and-spoke model for building a digital archive

This general arrangement is not new to archiving. However, placing the image processing function in the hub, rather than at imaging workstations, could be important in situations with limited technology infrastructure. Alternatively, a *lending-library* model would have complete portable imaging workstations returned with their captured content before redeployment.

We address two aspects of their use for current cultural imaging applications. First, we address imaging performance evaluation of mobile-phone cameras in the context of the FADGI Guidelines. This was done as a limited pilot study of several modern mobile phones selected to represent a practical range of current devices.

Secondly, we report on automated cataloging and metadata generation that will be important to the scaling up of mobile-phone adoption.

Imaging Performance Guidelines

The FADGI Guidelines were among the first to quantify the performance of imaging processes for the purpose of building a digital library of our shared cultural heritage. The guidelines served as a benchmark for institutions to reference when creating digital images for inclusion in this emerging global digital library. The guidelines are intentionally technology-agnostic, anticipating future advances in imaging and archiving technology. [1]

The root of the FADGI guidance is now almost two decades old, and imaging products and services have significantly changed since. Initially, only the finest professional equipment could achieve the goals outlined in the guidance. Today, many more products can achieve the spirit, if not the letter, of guidance. The most significant new entry is the mobile phone camera.

FADGI Guidance now allows spatial and color-image processing to enhance image capture. Color-image processing with custom ICC profiling can dramatically improve color capture accuracy, meeting the FADGI guidance. However, some spatial processing in current cell phone cameras deviates from FADGI recommendations, thereby suggesting the need for new guidance specific to collections and uses.

Camera Evaluation

For mobile-device cameras, there are several common technologies used in current designs. For example, the CMOS sensors, color-filter arrays, and small optical lenses could lead to similar imaging performance. However, differences become evident when we include autofocus, object-oriented signal processing, and computational photographic methods for scene rendering.

Here our task is to evaluate several consumer-designed systems for use with copy stands, with controlled-illumination imaging of mostly flat objects. We will concentrate on the color (encoding) accuracy and imaging resolution results.

A limited evaluation of seven current mobile device cameras was conducted. The devices were: iPad v5 Air tablet, iPhone 14 Pro, iPhone 15 Pro Max, iPhone 16 Pro. Max, iPhone 17, Samsung Galaxy A17, and Xiaomi Leica Leitzphone. For comparison with a full-frame mirrorless camera was included, the Sony a7RV, with a Sony FE 55mm F1.8 ZA lens.

The copy stand arrangement was used, as shown in Fig.2, with an experienced operator. Several test targets were captured with the camera set to the maximum image sampling resolution, an uncompressed DNG or TIF image format, when available.

We will base our reporting primarily on the FADGI ISO 19264 test target of Fig. 3. Test images of this target can be used to evaluate several imaging performance parameters. Color (encoding) accuracy evaluation is based on comparing the image pixel values for the color patches to corresponding values for the reference test target. The capture of fine detail is measured using several edge features at the center and near the four corners of the imaging field. In addition, the small black circles can be used to evaluate the optical distortion due to lens aberrations.

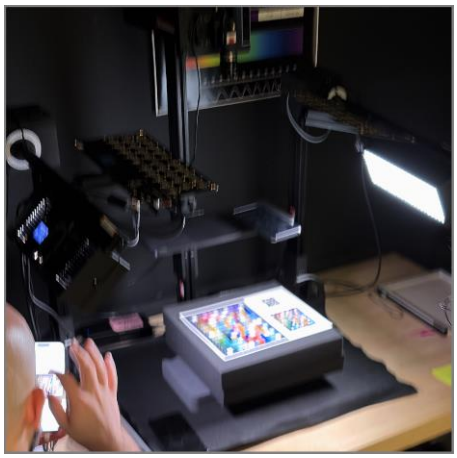


Figure 2: Copy stand used for imaging performance evaluation, showing color test chart

Figure 3: FADGI ISO 19264 test target

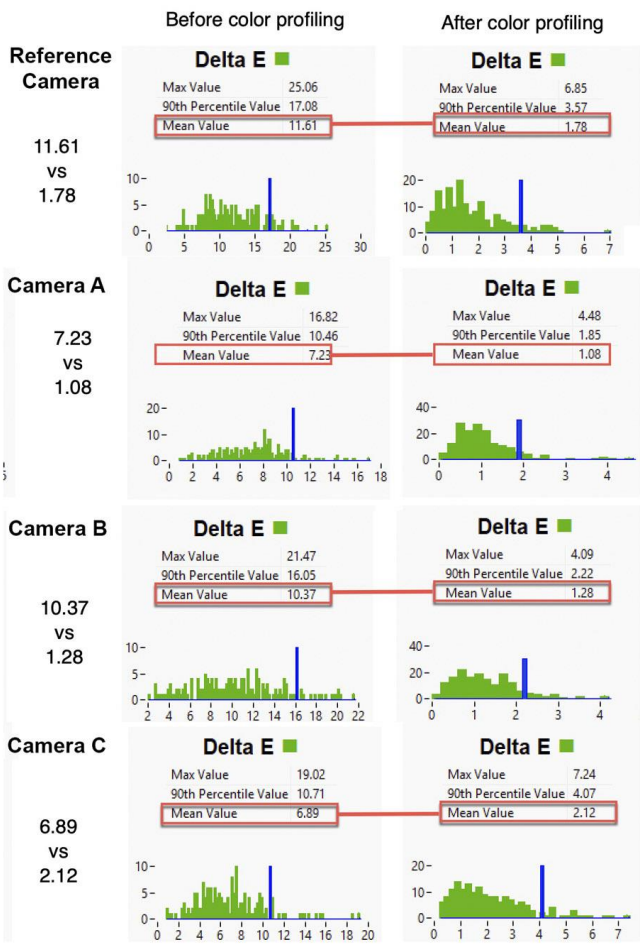
Illumination Nonuniformity Correction

One of the first observations was the apparent non-uniformity of the captured images. This was likely due to the combination of copy stand lighting and optical fall-off of the lens. Although not planned for, a shading-correction operation was performed, based on several areas of the grey background of the target of Fig, 3. A third-order, two-dimensional polynomial function was fitted to these values, and a pixel-by-pixel multiplication (shading) correction was applied.

Color Encoding Accuracy

Color coding accuracy as described here is with respect to the image color encoding of the object. There is no assumption on how the captured image will be rendered to either print or display, but rather how those colors are best encoded relative to the original target's color content defined by a given working color space. We adopted ΔE_{00} as the color-goodness measure, consistent with the FADGI Guidelines. The lower the ΔE_{00}, the more accurate the color encoding. We have chosen to focus on Adobe RGB 1998 color space for its popularity.

As mentioned above, the original test images suffered from illumination nonuniformity, so it is no surprise that the color errors for the four original candidate cameras are high. These ΔE_{00} color error results are shown in the center column of Fig. 4



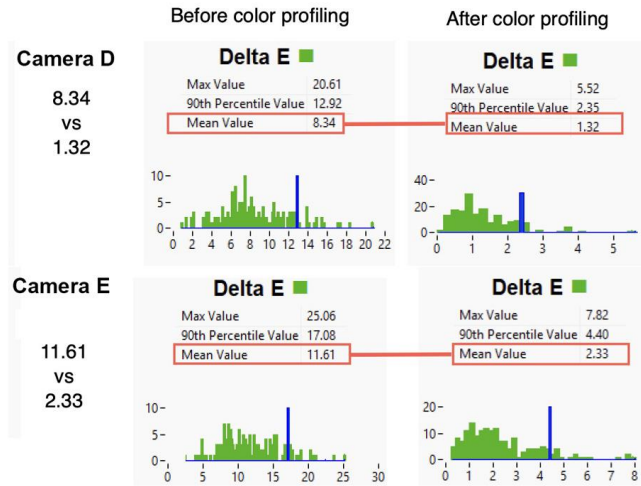


Figure 4: Color encoding error results for the reference Sony camera, and the mobile devices tested

Nevertheless, these errors can be minimized by color correction using ICC color profiles. For each target test image, an input device ICC color profile was generated, and then assigned and converted to Adobe RGB (1998) encoding. The results after this correction are shown on the right-hand side of Fig. 4. The differences between the original non-profiled and profiled results are notable and demonstrate the extent to which ICC color profiles can remediate poor color performance, within reason. That's the good news. Note the similar mean color error values for the mobile devices listed.

Spatial Frequency Response

The method used to measure the capture of object or scene details by a camera is the Spatial Frequency Response (SFR). This is based on the analysis of edge features in the test target and expressed as a function of spatial frequency. The test target of Fig. 1 is designed with five edge features, allowing the SFR to be evaluated near the center and edges of the optical field. This is important since optical falloff in performance from center-to-edge is common. Figure 5 shows cropped image edge regions in a typical test image from our study.

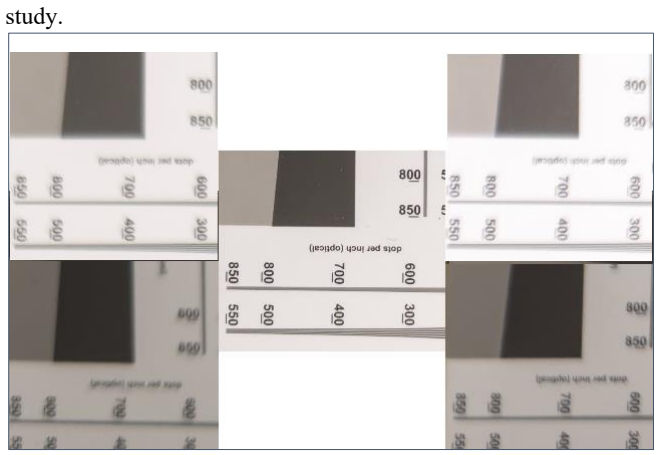


Figure 5: Cropped sections of the FADGI target edge areas showing intra-image variation in sharpness

Note the sharp features in the center compared to those near the edges of the image. The corresponding SFR results are shown in Fig. 6. The five locations are: Center, northwest (NW), northeast (NE), southeast (NE), and southwest (SW).

This behavior was observed to varying degrees for the mobile devices tested. For reference, the Sony camera showed much less SFR falloff, as shown in Fig. 7. We would expect this, based on the optical size and requirements. Here we have simplified the reporting by showing the average SFR for the edge features, rather than each edge.

We can further express this SFR falloff in terms of the ratio of edge to center results. In Fig. 8, the results are shown for the devices evaluated in the vertical direction. The results for one device were unreliable and omitted, possibly due to apparent high digital image sharpening. Note that a theoretically perfect result would be a ratio of 1.0, and the reference Sony camera approaches this for a wide range of spatial frequencies.

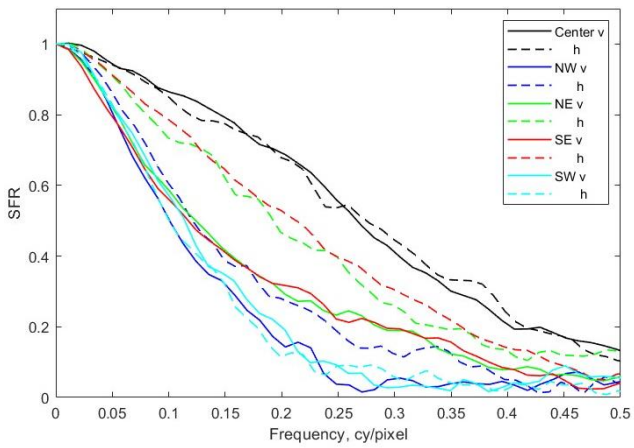


Figure 6: Example SFR results for the edges shown in Fig. 3 for the five locations (see text) in both vertical and horizontal directions.

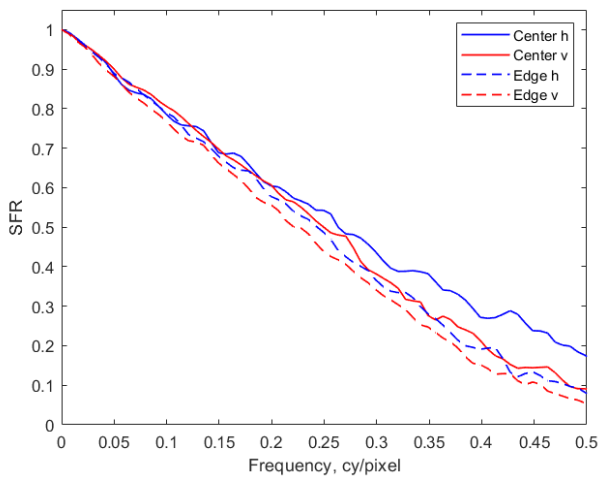


Figure 7: SFR results showing consistent center-to-edge performance for the Sony a7R camera

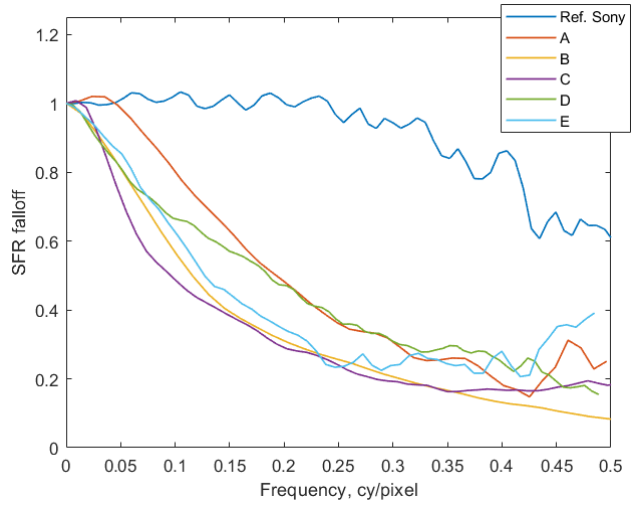


Figure 8: SFR fall-off expressed as an SFR ratio edge/center

Operational Mobile Workflows for Metadata, Cataloguing, and Archival Processing

The significance of mobile devices in our context lies not only in image capture, but in the combination of widespread availability, portability, voice input, local review, network connectivity, and structured data entry. These characteristics make mobile devices useful for archival processing, object cataloguing, and field documentation, especially where specialized workstations are impractical. Mobile workflows therefore function as a complementary tier of cultural heritage imaging and documentation, alongside higher-control imaging systems used for preservation-grade digitization. The following examples illustrate recent progress. We describe case studies that used mobile phone capture and AI-enhanced archival arrangement and archivist-created description using software developed by Historiq. Historiq helps archives move collections from boxes to discovery faster, pairing hands-on archival processing with technology-enabled workflows that keep archivists at the center.

General Federation of Women's Clubs

Historiq's work with the General Federation of Women's Clubs (GFWC, Washington, DC) [4] provides one example in archival processing. More than 50 cubic feet of records documenting GFWC's past leadership were processed at approximately 1 hour per cubic foot, including box- and folder-level inventory, rehousing, labeling, description, and finding aid creation. Nearly all stack-side work was completed on an iPhone using voice-based description, with no digitization or AI-generated metadata; only final finding aid finalization and quality control were completed on a laptop.

Compared with a 12.2-hour-per-cubic-foot composite benchmark derived from published processing-rate guidance from RIT [5], MIT [6], and Yale [7], this represents approximately 92% less time per cubic foot. Because those manuals use different units, processing levels, and assumptions, the composite should be understood as a practical comparison, not a universal processing standard.

Fort Ticonderoga: object-based cataloging

This workflow model also applies to object-based cataloging. In June 2026, our work with Fort Ticonderoga (Fort Ticonderoga, New York) used iPad tablet-based voice and camera capture for non-archival object cataloging. [8] Each object was described by voice and imaged with the tablet camera. The resulting images served as reference documentation but also supported real-time OCR/HTR and visual analysis. This improved the association between the object, the voice transcript, and the structured metadata for the entry. Here, digitization is not only a downstream access activity; it becomes part of live cataloging and metadata creation.

St. Catherine's School

A related case at St. Catherine's School (Richmond, Virginia) demonstrated a mixed mobile and desktop workflow. [9] In less than three weeks, two interns processed three archival collection areas and documented nearly 300 individual items, including publications, theater materials, books, trophies, founder-related objects, art, furniture, and other artifacts. Most cataloging was completed on laptops, while phones were used to photograph objects in place and attach images to records.

This case is not directly comparable to the above GFWC processing-rate benchmark, but it illustrates how mobile capture can support distributed documentation, preserve location context, and make structured collections work more repeatable.

These examples suggest that the value of mobile workflows is not limited to faster data entry. Mobile devices can reduce the distance between handling, imaging, description, labeling, and review. Furthermore, they enable greater technological accessibility, where the purchase of additional hardware was not necessary to achieve highly efficient cataloging and digitization.

Conclusions

Digitization of cultural heritage materials by established methods can only preserve a small fraction of the world's cultural heritage. Collectively, cultural heritage, especially in underserved communities, will be lost without embracing a model that includes distributed acquisition using the device that has revolutionized photography, the mobile phone camera.

In archives, this can support faster box- and folder-level processing. In object cataloging, it can connect voice description, reference imaging, OCR/HTR, and metadata generation within a single workflow. The same pattern may also apply in adjacent fields. For example, in libraries (particularly special collections), museums where cataloging of unique materials is paramount to discoverability, archaeology, and records management, where staff must document materials accurately while the physical context is still present. Mobile capture is therefore best understood as an additional tier of cultural heritage documentation: useful for access, triage, metadata enrichment, and distributed work, while complementing higher-control imaging systems when preservation-grade capture is required.

Our investigation of current mobile devices indicates several areas where they can meet the needs of image capture for collections, particularly in the above situations. However, we find limitations on their imaging performance. As for all copy stand work, attention should be paid to, and compensation made for, illumination uniformity/lens falloff. In addition, employing custom ICC color profiles for color correction may be needed to meet color-accuracy goals. With a nod to Jonathan Swift, the color profiles here seem to indicate that, indeed, a silk purse can be made out of a sow's ear.

However, without uniform (or corrected) illumination, these color improvements are constrained to the existing patch locations. They cannot be taken as applicable for all colors. For example, if the target were rotated at 90-degree increments with the same lighting set up and target patch locations, each capture could yield different color error results. This would not be the case with a nearly uniform lighting environment.

Optical falloff in image resolution/sharpness was observed for all mobile cameras, although the degree varied between devices. This effect can easily be evaluated using the currently adopted FADGI method. It is recommended that potential mobile cameras are evaluated for these characteristics, to select the maximum field of view to be used before image sharpness loss becomes severe.

Acknowledgements

Many thanks to Yosi Pozeilov and Alec Rouleau for helping with the image capture.

References

[1] FADGI Still Image Working Group, ed. T Rieger, US Library of Congress, http://www.digitizationguidelines.gov/

[2] P. D. Burns and D. Williams, Going Mobile: Evaluating Smartphone Capture for Collections, *Proc. IS&T Archiving Conf.,* 2016

[3] L. Humenuck and M. Assefa Abebe, Evaluating Color Transformation Quality and Accuracy of Prosumer and Mobile Phone Cameras for High Dynamic Range Cultural Heritage Documentation, *Proc. IS&T's Archiving Conf*. 273-1-6, 2024.

[4] Historiq, "Advancing Archival Processing with the General Federation of Women's Clubs."

[5] Rochester Institute of Tech. Archives, RIT Archives Processing Manual, 2024. https://www.rit.edu/archives/sites/rit.edu.archives/files/2024-09/RIT%20Archives_Processing%20Manual_092724.pdf

[6] MIT Distinctive Collections, Processing Rates, Archives Processing Manual, 2022. https://wikis.mit.edu/confluence/display/ARCHIVESPROCESSING/Processing+Rates

[7] Beinecke Rare Book and Manuscript Library, Yale University, 2.5.2 Rates of Processing, Beinecke Manuscript Processing Manual, 2021. https://msu-processing-manual.beinecke.library.yale.edu/

[8] Historiq, Advancing Archival Access with Fort Ticonderoga, https://historiq.com/case-studies/fort-ticonderoga

[9] Historiq, St. Catherine's School Turns Intern-Led Collections Work Into Archival Momentum, https://historiq.com/case-studies/st-catherines-school

Author Biographies

Dean Serrentino is Founder and CEO of Historiq, a company dedicated to expanding access to history through innovation in archival processing and digitization. He has nearly 15 years of

experience in artificial intelligence, data science, and engineering, building applied AI systems and workflow platforms.

Peter Burns, Ph.D., is a consultant working in digital image evaluation, system monitoring, and image processing. A frequent presenter, he has experience in several areas of imaging: digital photography, mobile imaging, and cultural heritage.

Jay Bosanko is Chief Archivist at Historiq and a former Deputy Archivist of the United States at the National Archives and Records Administration (NARA). A third-generation archives professional with more than 30 years of experience, he has held senior leadership roles, including Chief Operating Officer at NARA, and brings deep expertise in archival operations, policy, and large-scale collections management.

Lilli Cooper leads institutional partnerships at Historiq, working with museums, libraries, and archives to advance large-scale digitization and access initiatives. Former Vice President of Sales at a leading special collections digitization provider, she has supported projects involving millions of objects across diverse cultural heritage institutions.

Michael Sao Pedro is a technologist specializing in artificial intelligence and full-stack product development. He is the former CTO and co-founder of Apprendis, where he led engineering and product and served as principal investigator on multiple U.S. Department of Education SBIR grants.

Juilee Decker, Ph.D., is a professor of history at Rochester Institute of Technology, where she co-directs the Cultural Heritage Imaging Lab, a public-facing interdisciplinary research, education, and training partnership between the museum studies and imaging science programs at the university. Her work focuses on education and training, particularly in relation to the NEH-funded MISHA program (National Endowment for the Humanities Grant PR-268783-20 and Grant PE-303737-25).

Don Williams is the founder of Image Science Associates, a digital imaging consulting and software group. He has worked for several large cultural heritage institutes in the practical implementation of image quality controls and is the prime architect for the *GoldenThread* image quality evaluation tools. He has taught short courses for many years and contributes to several imaging standards activities.

Thomas Rieger is a Professor of Practice at Rochester Institute of Technology, concentrating on the development of next-generation imaging technology and workflows for the cultural heritage community and expanding the deployment of MISHA systems worldwide. Tom is a Fulbright specialist currently assisting with a digitization and preservation strategy for Indonesia. Previously, he served as the Chief of the Collections Digitization Division of the Library of Congress, and head of the Federal Agencies Digital Guidelines Initiative Still Image Working Group (FADGI).