IS&T Archiving Conference (ARCHIVING 2024)

Washington, DC, USA 9-12 April 2024

ISBN: 979-8-3313-0107-1

Printed from e-media with permission by:

Curran Associates, Inc. 57 Morehouse Lane Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright[©] (2024) by Society for Imaging Science & Technology All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

For permission requests, please contact Society for Imaging Science & Technology at the address below.

Society for Imaging Science & Technology 7003 Kilworth Lane Springfield, Virginia 22151 USA

Phone: 703-642-9090 Fax: 703-642-9094

info@imaging.org

Additional copies of this publication are available from:

Curran Associates, Inc. 57 Morehouse Lane Red Hook, NY 12571 USA Phone: 845-758-0400 Fax: 845-758-2633 Email: curran@proceedings.com Web: www.proceedings.com Washington, DC APRIL 9-12

ARCHIVING202

PRESERVATION / ARCHIVING

Session Chair: Hana Beckerle, Library of Congress (US) 10:10 – 12:20

10:10 Heritage Science - Perspective, Provenance, Preservation,

Fenella G. France, Library of Congress (US) Often the perspective of heritage science can seem opaque, with the perception of science being only for "scientists". The fairly recent move to using the term "heritage science" is better known in Europe, but there is some confusion as to what this term encompasses. Heritage science refers to any of the multidisciplinary fields that contribute to the discovery, security, and preservation of a diverse range of cultural heritage materials. Many heritage collection items are complex multi-composites with convoluted preservation needs and degradation pathways. One of the areas that seems to be less well understood and appreciated is the capacity for heritage science to add new layers of knowledge to collection items, as well as the capacity for re-interpretation through this new information. Further areas of heritage science that support humanities include confirming provenance, the ability to link and reconnect separated collections, and the utilization of new technology to provide levels of security, an area of support greatly needed in the current environment for extensive trafficking of heritage.

10:30 - 11:10 MORNING BREAK / EXHIBITS OPEN

11:10 Making Sense of Bureaucratic Documents – Named Entity Recognition for State Authority Archives, Venla Poso¹, Mikko

Lipsanen², Ida Toivanen¹, and Tanja Välisalo¹; ¹University of Jyväskylä and ²The National Archives of Finland (Finland) **6**

The usability and accessibility of digitised archival data can be improved using deep learning solutions. In this paper, the authors present their work in developing a named entity recognition (NER) model for digitised archival data, specifically state authority documents. The entities for the model were chosen based on surveying different user groups. In addition to common entities, two new entities were created to identify businesses (FIBC) and archival documents (JON). The NER model was trained by fine-tuning an existing Finnish BERT model. The training data also included modern digitally born texts to achieve good performance with various types of inputs. The finished model performs fairly well with OCR-processed data, achieving an overall F1 score of 0.868, and particularly well with the new entities (F1 scores of 0.89 and 0.97 for JON and FIBC, respectively).

11:30 Hyperspectral Database of Synthetic Historical Inks,

The aim of this work is to provide the cultural heritage community with a comprehensive hyperspectral image database of handwritten laboratory samples, including various writing inks commonly found in historical documents. The database contains 195 samples registered in the VNIR (400-1000 nm) and SWIR (900-1700 nm) spectral ranges, along with complete information about the ink recipes (components and concentrations used for each ink and mixture), and their corresponding Ground Truth images. The database is now publicly available as part of a bigger database related to the Hyperdoc project and can be used to perform different tasks. We present here one example: the classification of iron gall vs non-iron gall inks.

12:00 FILM2PAINT: Transforming Photographic Documentation on Reversal Film into Paintings' Accurate Colors, Irina-Mihaela

Displaying the past appearance of artworks by reversing degradation phenomena holds significant value for art historians, conservators, museum curators, educators, and the wider public, as it seeks to estimate the original artist intention. In this work, we aim to restore the past colors of a painting from documentary records done on reversal film photographs. The challenge with these photographs is that due to film-specific chromogenic processes, their colors are inaccurate with respect to the captured object. For this reason, we test the performance of four color correction methods in compensating for the color distortions inherent to each film type by using a dataset of reversal films of two color targets, X-Rite ColorChecker Digital SG and Coloraid IT-8. Furthermore, we apply the same method to detect changes due to aging and/or conservation treatments in the painting Junger Proletarier (1919) by Paul Klee, by comparing a color corrected film record from 1995 with a more recent digital capture of the painting from 2005. Our results indicate that the method which best accounts for the film chromogenic processes to reveal the actual colors of the photographed object is based on non-linear optimization using a neural network.

EXHIBITOR PROFILES

Session Chair: Robert Kastler, MoMA (US)

12:20 - 12:40

Archiving 2024 exhibitors Crowley, Digital Transitions, Image Science Associates, nextScan, and Picturae share information in 3-minute profiles.

12:40 – 14:00 LUNCH BREAK ON OWN

SHORT PAPERS

Session Chair: Meghan Hill, Library of Congress (US) 14:00 – 15:00

14:00 "Less is More": How Understanding the Process of Motion Picture Film Scanning Can Make Your Life Easier, Alice Plutino, University

The processing operated by scanner systems is not always openly available. The various digitization hardware and software can lead to different approaches and workflows in motion picture film preservation and restoration, resulting in inconsistency among archives and laboratories.

This work presents an overview of the main approaches and systems used to digitize and encode motion picture film frames to explain these systems' potentials and limits.

14:15 Wide-format High-speed Film Scanner, Matt Anderson and

Rich Chaney, nextScan (US) A-1 USA government agency project to design and implement a scanner to digitize wide-format film over 10" in width at a high-rate of speed.

14:30 Multispectral Imaging as a Preservation and Valuation Tool at the Minas Gerais Public Archive, Brazil: A Case Study on an 18th-century Illuminated Manuscript, Alexandre Cruz Leão, Alexandre Oliveira Costa, Marcia Almada, Rodolpho Zanibone,

Multispectral Imaging has become an indispensable tool for Cultural Heritage materials and objects analyzing, documenting, and visualizing. This study delves into applying this technique to an 18th-century illuminated manuscript at the Minas Gerais Public Archive, Brazil. Currently undergoing restoration intervention at The Federal University of Minas Gerais, the manuscript exhibits faded writing due to moisture, and the application of multispectral imaging (UV, Visible, and IR) with seven different wavelengths proves highly effective in recovering lost information. The light source was provided from LED and Halogen Lamp, and the result shows a very clear text on the manuscript after the digital processing by ImageJ and PCA. The Minas Gerais Public Archive, established in 1895, plays a vital role in safeguarding the state's documentary and historical heritage. It was the first-time multispectral imaging was applied to cultural objects in Minas Gerais.

14:45 Examining and Exulting: Multimodal Imaging and Collaborating

Through a New Lens, Kayla Kee, J. Paul Getty Trust, and Olivia

This contribution describes the inspiration for, and beginning stages of, important conversations around revitalizing the way routine imaging operations and distribution procedures are carried out in support of ongoing internal object-based research projects at Getty. As part of somewhat routine technical imaging of a late-18th-century work on paper, a variety of multimodal images were captured to support the technical study of this object. The simple act of distributing this image set as a registered, layered Photoshop file enabled visual examination and comparisons of the data that would not have been possible had they been delivered as separate, unrelated files, as is the normal mode of operation for multimodal capture campaigns like this. This easy change in the delivery and utility of the delivered file, and the enthusiasm it was met with, has had a positive impact on our interdepartmental communication and has reinforced the need to carefully consider the best ways to support and meet the needs of all project stakeholders.

With this contribution, we seek to generate conversation and advice around continuing to improve our own intra-institutional collaborations with respect to multimodal data interrogation, and more broadly, building robust frameworks of communication between imaging, conservation, and science colleagues working toward common goals.

DIGITIZATION PROJECTS

Session Chair: Charles Walbridge, Minneapolis Institute of Art (US) 15:00 – 17:40

15:00 Digitizing the Library of Congress Hebrew Manuscripts

15:20 United States National Archives and Records Administration

The United States National Archives and Records Administration (NARA) has issued new regulations that establish standards for the digitization of US government records. The regulations are part of an effort to transition to a fully electronic government, and allow US federal agencies the authority to digitize and destroy source records and the electronic version become the recordkeeping copy. The specifications draw upon established international digitization standards such as ISO 19264, Metamorfoze, and FADGI guidelines.

By adopting the image quality specifications found in ISO 19264 and the image analysis method described by FADGI, NARA has effectively defined the minimum requirements for a digital surrogate to serve as legal and evidentiary purpose as the source record. This paper presents the records management context of digitization, as well as discussing the quality management, documentation, image and metadata specifications, and validation requirements.

15:40 - 16:20

AFTERNOON BREAK / EXHIBITS OPEN

16:20 Reframing the Conservation: A Case Study in using 3D Imaging Technology to Restore Carved Wooden Elements on a Painting Frame, Chris Heins and Deepa Paulus, The Metropolitan Museum

This paper will present the story of a collaborative project between the Imaging Department and the Paintings Conservation Department of the Metropolitan Museum of Art to use 3D imaging technology to restore missing and broken elements of an intricately carved giltwood frame from the late 18th century.

16:50 Building a Program-from 0 to 26,000,000, Kristin A. Phelps and

17:10 How to (Mass-)Digitize Newspapers in Switzerland — The Swiss National Library Approach — Renewed, Martina Hoffmann,

ARCHIVING20

Washington, DC APRIL 9-12

17:30 Closing remarks

CONFERENCE RECEPTION

18:00 - 20:00

Join colleagues at DC's iconic Woolly Mammoth Theatre Company, 641 D St. NW, located two blocks from NARA for food, drink, and lively conversation. Please wear your badge to enter.

THURSDAY 11 APRIL 2024

PANEL: EMPOWERING CULTURAL HERITAGE ORGANIZATIONS TO PRIORITIZE SUSTAINABILITY 9:00 - 10:20

This panel explores innovative strategies for cultural heritage organizations to embed sustainability at the core of their operations, emphasizing these institutions' crucial role in leading environmental stewardship. It aims to empower attendees with actionable insights and collaborative frameworks to integrate sustainable practices that safeguard our cultural legacies for future generations.

Moderator

• Carolina Gustafsson, Stiftelsen Föremålsvård i Kiruna (Sweden)

Panelists

- Jerry Foust, cultural resource consultant, and adjunct professor, Georgetown University
- Eliana Glicklich-Cohn, senior manager, real estate & sustainability, MoMA
- Stephanie Shapiro, co-founder and managing director, Environment & Culture Partners

10:20 - 11:00

COFFEE BREAK / EXHIBITS OPEN

MANAGEMENT / ASSESSMENT

Session Chair: Bethany Scott, University of Houston Libraries (US) 11:00 – 12:10

11:00 Micrio as an Ultra-Resolution Story Format Smorgasbord,

and workflows allowing ever greater throughput, the user experience for accessing high and ultra-high-resolution imagery does not advance at the same rate. This paper describes how Micrio is used by Galleries, Libraries, Archives and Museums to deliver their high-resolution imagery to a plethora of devices. It then focuses on the evolution of storytelling formats that the project-based development of the platform has originated starting from multimodal 2D images, to immersive environments and, finally, 360-degree views on heritage objects.

11:30 How Long Is Long-Term? — An Update, Barry M. Lunt, Daniel Kemp, Matthew R. Linford, and Wood Chiang, Brigham Young

At the Archiving 2011 Conference a paper was presented titled, "How Long Is Long-Term Data Storage?", in which the author summarized the state of archival digital data storage. This paper covered causes of failure, the life expectancy of the data with the storage methods available at the time, format obsolescence, and current research into archival storage. With that paper being about 13 years old now, it is time for an update. This paper will discuss the same topics covered in the previous work, including failure, life expectancy, format obsolescence, as well as current research in all 4 of the primary data storage methods used today (harddisk drives, solid-state drives or flash memory, optical discs, and $\frac{1}{2}$ " magnetic tape. It ends with a discussion of our current research on a humanreadable, high-density storage medium.

11:50 Revolutionizing Archival Internships: Unlocking Global

Collaboration, *Christie Jovanovic, Holocaust Museum LA (US)* . **60** Amidst the challenges of a global pandemic, Holocaust Museum LA redefined their traditional model of an internship in the Archive, creating a virtual program with lessons learned in remote collaboration. Beyond the hurdles of diverse time zones and required skill sets, this initiative significantly increased the team's capacity. With standardized processes and meticulous quality checks, interns worldwide contribute to processing and cataloging museum collections. Looking ahead, their work lays the foundation for a public-facing database for the collection, elevating accessibility and fulfilling our core archival mission.

12:10 - 13:30

LUNCH ON OWN

DIGITAL QUALITY

Session Chair: Justyna Badach, National Gallery of Art (US) 13:30 – 14:40

13:30 A New Transmission Target for Cultural Heritage Imaging,

Roy S. Berns and David R. Wyble, Gray Sky Imaging (US) ... **64** The accurate digitization of film using high-resolution digital cameras, especially historic positive and negative film, presents a difficult challenge for cultural-heritage imaging. Approaches used for reflecting materials—e.g., profiling using color targets—are difficult to apply to transparent materials due to a paucity of film-specific targets, measurement challenges of small patch sizes, and the inadequacy of these targets for historical films and negatives. Research was carried out to design, construct, and verify a new transmission target. Simulation was used to select 80 filters, optimized from a 476-filter set of absorption filters with criteria including colorimetric performance for the 80 filters and four validation spectral datasets, color gamut, and spectral diversity. A prototype target was constructed, measured, and imaged. All criteria were met. Future research will refine the target and validate its performance using independent targets and color-challenging photographs.

14:00 How Understanding the Primary Function of the Human Eye in Low Light Can Benefit Digital Imaging Operators, Nora Ibrahim,

Osher Map Library & Smith Center for Cartographic Education, University of Southern Maine (US)

Color accuracy is essential in digital imaging for Cultural Heritage when reproducing an object into digital format. As a practitioner in the field, curiosity rose as to what the long-term effects are of working in a low-light environment after chronic fatigue and strained eyes became the daily norm. Multiple literatures were reviewed to extract the groundwork of color vision in low light and what achieves proper and consistent information processing; however, there were no known specific studies for the impetus for this topic. This research aims to explore color vision in low light to recognize potential short- and long-term physiological vision changes. Studies uncovered multiple variables impact human vision when processing an image or scene. This paper investigates the structure of the eye by comparing human vision to the structure of a camera system, the processing of color in the retina, the recommended viewing environment for cultural heritage capture, different conditions that impact perception, and solutions for regular eye maintenance. As is often said, awareness is the first step to prevention.

14:20 Beyond RGB 2.0: Further Improvements to a Free, Opensource, Spectral Image Processing Software Application for Workflow, Analysis, and Repeatability, Leah Humenuck and Susan P.

Farmand, Rochester Institute of Technology (US)71 Beyond RGB is a recently developed software application for colorimetric and spectral processing. Two sets of RAW RGB images (a dark current image, a flatfield, target, and artwork), taken under two different lighting conditions are used as the inputs. After processing, the software provides a color calibrated RGB image, along with the calibration accuracy, a spectral stack, and if applicable estimated spectral reflectance of regions of interest may be selected. The updates for this version of the software includes: dual screens for simultaneous file comparisons, updated a*b* and L*C* plots showing the actual and expected values, batch processing, meta data allowing for precision reproducibility for color calibration,

BEHIND-THE-SCENES TOURS

and user interface upgrades.

Tour times vary. Tours are included in the registration fee. Visit the registration desk to confirm or change your tour.

Archiving 2024 thanks the following institutions for providing Behind-the-Scenes tours to this years attendees:

- Library of Congress, Preservation Research and Testing Division (PRTD)
- Library of Congress, Digitization Operations Center
- National Gallery of Art, Department of Imaging and Conservation
- National Museum of African American History and Culture, Family History Center
- Smithsonian Libraries and Archives, Digitization and Access, Joseph F. Cullman 3rd Library of Natural History, National Museum of Natural History Library

EVENING ON OWN

Enjoy an evening on your own or meeting up with other attendees for some informal fun.

ARCHIVING2024 Washington, DC APRIL 9-12

FRIDAY 12 APRIL 2024

CLOSING KEYNOTE

9:00 - 10:10

Session Chair: Carolina Gustafsson, Stiftelsen Föremålsvård i Kiruna (Sweden)

The Content Authenticity Initiative - Fighting Misinformation Through

Digital Content Provenance, Santiago Lyon, head of advocacy and education for the Content Authenticity Initiative, Adobe (US)

In an increasingly confusing media landscape, the Content Authenticity Initiative (CAI) is working to implement the industry wide C2PA standard securely establishing the origins of digital files, any changes made to them in editing and sharing that information with the viewer/consumer through Content Credentials, the digital equivalent of nutrition labels. This technology has direct applications in the archiving and DAM environments. This plenary provides an overview of the CAI to familiarize attendees with the philosophy and methodology of the program.

DIGITAL ADVANCED TECHNIQUES / 3D

Session Chair: Kurt Heumiller, National Gallery of Art (US) 10:10 – 12:40

10:10 **Kintsugi 3D: An Empirically-based Photogrammetry Production Pipeline,** Lou Brown and Charles Walbridge, Minneapolis Institute of Art; and Michael Tetzlaff, University of Wisconsin-Stout

10:40 – 11:20 COFFEE BREAK

Selecting the optimal resolution and post-processing techniques of 3D objects for cultural heritage documentation is one of the most distinguishable challenges within 3D imaging. Many techniques exist to document a tangible object at very high objective accuracy, but there also exist techniques that can visualize a similar perceptual accuracy without documenting the objective values. The application difference between storage of complex geometric data and the visualization of it could be fundamentally

different, and if the two methods are not disassociated it could lead to either false or inaccurate digital documentation of a cultural heritage object. In this investigation we compare several different metrics for evaluating the quality of a 3D object, both objectively and perceptually, and look at how the different approaches might report greatly different outputs based on the post-processing of a 3D object. We also provide some insight in how to interpret the output of various metrics, and how to compare them.

11:40 Standards Based Authentication System and Method Using

Physical Characteristics of an Object, Larry Kleiman, Spectral

The uniqueness of an object, such as an original work of art, can be demonstrated by accurately and precisely measuring "visible electromagnetic energy," VIS, found between 360 and 830 nm, reflected off the "original" specimen, the Reference, R, a specific area of the original, following a standardized measurement and transformation procedure defined by the Commission Internationale de 'Eclairage (CIE). The measured R spectra are transformed into CIE X, Y, Z values (see: CIE31) and then into CIEL*a*b* (see: CIE76) values and stored in a database protected by a block chain mechanism. To authenticate, the stored R Lab values are compared to a comparable set of Lab values, S, generated by measuring the same specific area of the specimen to be authenticated as was originally measured for R. S is compared to R via the CIE DeltaE 2000 algorithm (see CIEDeltaE00), generating a set of values, DeltaE which measure the difference between S and R for each pixel in the measured area. The mean and standard deviation of the DELTAE values over the whole measured area is found and used to create a unique Authenticity Factor, AF, the probability that the difference between any pair of corresponding cells in R and S is less than 1, the Just Noticeable Difference (JND), the CIE threshold that determines if the two specimens match. The AF is found from the cumulative probability function of the normal distribution of the DELTAE values of R and S. If AF=1, the sample perfectly matches the reference, denoting the sample is 100% authentic. AF<1 gives the confidence in the match between the sample and reference, ie, AF=.99 indicates there is a 1% chance the sample is not the same as the reference, or there is a 99% chance the sample is "authentic".

The purpose of this work is to present a new dataset of hyperspectral images of historical documents consisting of 66 historical family tree samples from the 16th and 17th centuries in two spectral ranges: VNIR (400-1000 nm) and SWIR (900-1700 nm). In addition, we performed an evaluation of different binarization algorithms, both using a single spectral band and generating false RGB images from the hyperspectral cube.

12:20 Digital Rescue of the Nordic Viking Era, Bjarte Aarseth, Museum

This project focuses on creating high-quality digital backups of valuable Norwegian cultural heritage through the surveying of the most irreplaceable and iconic objects, including Viking ships, historical artifacts, church portals, and stave churches. Utilizing advanced technology, particularly powerful industrial 3D scanners, attention is directed towards objects that form the core of Norwegian cultural history. With my background as an experienced master woodcarver actively integrated into the process, emphasis is placed on sensitivity and precision. The goal is to extract maximum information, preserve the unique craftsmanship, and create digital representations of superlative quality. This documentation is shared with conservators, souvenir producers, and the public, both online and through 2D drawings.

The project identifies the most valuable and irreplaceable aspects of Norwegian cultural heritage, while the revenue generated from the surveying proves to be as significant as the costs associated with scanner equipment and labor. This underscores the economic value of preserving and documenting our cultural heritage, forming the basis for a comprehensive strategy for the conservation and dissemination of Norway's cultural treasures to future generations.

12:40 - 14:00 LUNCH BREAK ON OWN

"TO INFINITY AND BEYOND!": FUTURE PREDICTIONS FOR CHI

Session Chairs: Archving 2024 Committee 14:00 - 15:00

What does the future look like for CHI? Here's your chance to participate in an open discussion. Join the Archiving 2024 Conference Committee as we explore the advancements and directions in this field.

15:00 - 15:20 COFFEE BREAK

INTEGRATING AI: SUCCESSES, ETHICS, AND THE FUTURE FOR CULTURAL HERITAGE 15:20 - 16:40

This panel provides insights into Artificial Intelligence's dynamic integration and use cases within cultural heritage institutions, highlighting successful implementations, navigating ethical considerations, and forecasting future advancements. Experts discuss how AI projects are conceived, tested, and rolled out while navigating known and unknown issues.

Moderator

• Robert Kastler, MoMA

Panelists

- Abbey Potter, senior innovation specialist, Library of Congress
- Robert Stein, chief information officer, National Gallery of Art
- Mike Trizna, data scientist, Smithsonian Institution

16:40 Closing remarks