FINAL PROGRAM AND PROCEEDINGS

Washington, DC **APRIL 9-12**

ARCHIVING2024

General Chair: Robert Kastler, MoMA

IS&T
imaging.org

www.imaging.org/archiving2024

Sponsored by the Society for Imaging Science and Technology

The papers in this volume represent the program of Archiving 2024,
held April 9–12, 2024 in Washington, DC.

Society for Imaging Science and Technology
7003 Kilworth Lane • Springfield, VA 22151 USA
703/642-9090; 703/642-9094 fax
info@imaging.org; www.imaging.org

ISBN Abstract Book: 978-0-89208-366-2
ISSN Print: 2161-8798
ISSN Online: 2168-3204
https://doi.org/10.2352/issn.2168-3204.2024.21.1.0

CONFERENCE SPONSOR

creeksideDIGITAL

CONFERENCE DONOR

HIP

CONFERENCE EXHIBITORS

CROWLEY
DIGITAL PRESERVATION SOLUTIONS

20 YEARS

ISA
Image Science Associates

nextScan
Scanning Technology Working For You

PICTURAE

WITH SUPPORT FROM

NATIONAL ARCHIVES

WELCOME TO ARCHIVING 2024

On behalf of the organizing committee, I extend a warm welcome to all attendees of Archiving 2024. Hosted at the prestigious National Archives in Washington, DC, this year's gathering promises to be exceptional, bringing together a diverse group dedicated to the vital work of digitizing, preserving, and sharing our collective cultural heritage materials. This common purpose unites this group into a community where we freely and openly share our knowledge.

The theme of this year's conference, "Science, Sustainability, and Security," reflects our commitment to advancing Archiving through innovative scientific approaches. As we look deeper into the ever-important and multifaceted roles that sustainability and security play in our field, institutions, and society, I'm proud that those in our community are the thought leaders focused on preservation. Sustainability and security must go hand in hand with preservation. Sustainability and security are already a part of our discussions, but I challenge us all to widen our lens on these fundamental pillars to expand our thinking, discussions, and collaborations.

With its rich history and symbolic significance, the National Archives and Records Administration serves as the perfect backdrop for our gathering. It reminds us of the immense responsibility the Archiving community bears. Through our collective efforts, we not only preserve the past, but also inspire future generations to explore, understand, reflect, and appreciate their history.

As we come together to share insights, I encourage you to engage fully in the opportunities that Archiving 2024 presents. You will find as much value in a coffee break conversation as any other experience. Introduce yourself and join the conversations, you'll be glad you did.

Thank you for being part of Archiving 2024, and welcome to our conference.

Sincerely,

—Robert Kastler
General Chair, Archiving 2024

CONFERENCE COMMITTEE

General Chair
Robert Kastler, MoMA (US)

Program Chair
Carolina Gustafsson, Stiftelsen Föremålsvård i Kiruna (Sweden)

Short Course Chairs
Kristin Phelps, Library of Congress Copyright Office (US)
Laura Margaret Ramsey, The Metropolitan Museum of Art (US)

Past Chair
Sony George, Norwegian University of Science and Technology (Norway)

AV Chair
Alexandre Leão, Universidade Federal de Minas Gerais (Brazil)

COOPERATING SOCIETIES

american institute for conservation
Preserving Cultural Heritage

Council on Library and Information Resources

cni
Coalition for Networked Information

DigitalPreservationCoalition

THE RPS
ROYAL PHOTOGRAPHIC SOCIETY

REVIEW COMMITTEE

Tarek Abu Haila, TU Darmstadt (Germany)
Ottar A.B Anderson, Intermunicipal Archive of Møre og Romsdal (Norway)
Judith Andrews, Smithsonian Institution, National Museum of African American History and Culture (US)
Hana Beckerle, Library of Congress (US)
Michael J. Bennett, University of Connecticut (US)
Roy S. Berns, Gray Sky Imaging Inc. (US)
Peter Burns, Burns Digital Imaging (US)
Antenie Carstens, National Library of South Africa (retired) (South Africa)
Elizabeth Chiang, George Eastman Museum (US)
Terrance D'Ambrosio, Northeast Document Conservation Center (US)
Lavinia de Ferri, University of Oslo (Norway)
Hilda Deborah, Norwegian University of Science and Technology (Norway)
Daniel Dennehy, Minneapolis Institute of Art (US)
Denis Doorly, Museum of Modern Art (US)
Roger Easton, Rochester Institute of Technology (US)
Rebecca Frank, University of Tennessee, Knoxville (US)
Luiza Gonçalves, Archivo La Esquina (Argentina)
Federico Grillini, Agrosense AS (Norway)
Jon Yngve Hardeberg, Norwegian University of Science and Technology (Norway)
Lei He, Library of Congress (US)
Javier Hernández-Andrés, Universidad de Granada (Spain)
Kurt Heumiller, National Gallery of Art (US)
Meghan Hill, Library of Congress (US)
Martina Hoffmann, Schweizerische Nationalbibliothek (Switzerland)
Chrissy Huhn, University of California Berkeley Library (US)
Anssi Jääskeläinen, South-Eastern Finland University of Applied Sciences (Finland)
Muhammad Arsalan Khawaja, University of Burgundy (France)
Olivia Kuzio, Getty Conservation Institute (US)
Alexandre Leão, Universidade Federal de Minas Gerais (Brazil)
Mikko Lipsanen, National Archives of Finland (Finland)
Ana B. López-Baldomero, Universidad de Granada (Spain)
Ana López-Montes, Universidad de Granada (Spain)
Vincent Wai-Yip Lum, The Chinese University of Hong Kong (Hong Kong)
Tabita Lumban Tobing, Norwegian University of Science and Technology (Norway)
Ramamoorthy Luxman, Universite de Bourgogne Franche Comte (France)
Dipendra J. Mandal, Norwegian University of Science and Technology (Norway)
Giacomo Marchioro, University of Verona (Italy)
Miguel Ángel Martínez Domingo, Universidad de Granada (Spain)
Kit Messick, Smith College (US)
Phil Michel, Library of Congress (US)
Jeanine Nault, Smithsonian Institution (US)
Adam Neese, Museum of Modern Art (US)
Taren Ouellette, Library of Congress (US)
Thierry Paquet, University of Rouen Normandie, LITIS (France)
Marius Pedersen, Norwegian University of Science and Technology (Norway)
Laura Margaret Ramsey, The Metropolitan Museum of Art (US)
Jamie Rogers, Florida International University (US)
Lynda Schmitz Fuhrig, Smithsonian Libraries and Archives (US)
Carla Schroer, Cultural Heritage Imaging (US)
David Schuster, Binghamton University (US)
Bethany Scott, University of Houston Libraries (US)
Steven Simske, Colorado State University (US)
Robert Sitnik, Warsaw University of Technology (Poland)
Dina Sokolova, Columbia University (US)
Markus Sebastian Storeide, Norwegian University of Science and Technology (Norway)
Jim Studnicki, Creekside Digital (US)
Todd Swanson, Getty Research Institute (US)
Giorgio Trumpy, Norwegian University of Science and Technology (Norway)
Mahsa Vafaie, FIZ Karlsruhe (Germany)
Eva Maria Valero, Universidad de Granada (Spain)
Corine van Dongen, Koninklijke Bibliotheek (the Netherlands)
Marie Vans, Colorado State University (US)
Christoph Voges, consultant (Germany)
David R. Wyble, Avian Rochester, LLC (US)

CONFERENCE SCHEDULE
SHORT COURSE AND TECHNICAL PAPERS PROGRAM

TUESDAY 9 APRIL 2024

SHORT COURSE PROGRAM

8:00 – 10:00

SC04: Contracting and Managing an Outsourced Digital Imaging Project
Instructors: Martina Hoffmann, Swiss National Library, and David Walls, US Government Publishing Office

SC05: Advanced Concepts in Color Measurement
Instructor: David R. Wyble, Gray Sky Imaging, Inc.

8:00 – 12:15

SC01: Photographic Lighting Principles for Glass Objects
Instructors: Andrew Fortune and Bryan H. Buchanan, The Corning Museum of Glass

SC02: Imaging and Other Optical Tools for Discovery and Preservation of Special Collections
Instructors: Fenella France and Meghan Hill, Library of Congress

10:15 – 12:15

SC06: Introduction to 3D: Planning and Consequences
Instructor: Kurt Heumiller, National Gallery of Art

SC07: Can/Should AI Do That? Heritage Digitization Use Cases for AI
Instructors: Julie McVey, National Geographic Society, and Doug Peterson, Digital Transitions

SC08: Long-Term Planning: Visioning for a Speculative Digital Future
Instructors: Leah Weinryb Grohsgal and Lydia Heaton, Library of Congress

SC09: Improving Color Accuracy of Challenging Materials
Instructor: Roy S. Berns, Gray Sky Imaging

13:30 – 15:30

SC14: 2D from 3D: Photogrammetric Challenges and Solutions
Instructors: Carla Schroer and Mark Mudge, Cultural Heritage Imaging

13:30 – 17:45

SC10: Multispectral Imaging: Low-Cost and Low Barrier-to-Entry Methods
Instructors: Juilee Decker and Roger L. Easton, Jr., Rochester Institute of Technology

SC11: Beyond the GUI: Using the Command Line, Open Source Tools, and Spreadsheets on Imaging Projects
Instructors: Jim Studnicki, Creekside Digital, and Noah Durham, National Archives Records Administration

SC12: Colors Revealed: Mastering Cultural Heritage Photography Through Effective Color Control
Instructor: Ken Fleisher, National Gallery of Art

SC13: Precious Plates: Identifying, Handling, and Photographing Daguerreotypes and Cased Photographs
Instructors: Tricia Zigmund and Margaret Wessling, National Gallery of Art

15:45 – 17:45

SC16: Capturing Specularity with Kintsugi 3D and Camera-mounted Flash
Instructors: Michael Tetzlaff and Darcy Hannen, University of Wisconsin – Stout; Carla Schroer, Cultural Heritage Imaging; and Charles Walbridge, Minneapolis Institute of Art

SC17: OpenDICE for Imaging Quality Assessment
Instructor: Lei He, Library of Congress

WELCOME RECEPTION

17:45 – 19:30

Join colleagues at Penn Quarter Sports Tavern, 639 Indiana Ave NW, located catty-corner across Pennsylvania Avenue from NARA for happy hour. Weather permitting, we'll be outdoors, so dress accordingly.

WEDNESDAY 10 APRIL 2024

WELCOME AND OPENING KEYNOTE

Session Chair: Robert Kastler, MoMA (US)
09:00 – 10:10

Seeing Lost Enclaves: Archivally Based 3D Immersive Reconstructions of Historic Asian American Spaces, *Jeffrey Yoo Warren, artist educator and 2023 LOC Innovator in Residence (US)*

Seeing Lost Enclaves is part of Jeffrey Yoo Warren's work as the Library of Congress's Innovator in Residence, using archival photographs and maps to virtually reconstruct erased communities of color. The process of reconstructing each neighborhood's buildings and streets weaves together and interlinks the few images remaining of these once-vibrant enclaves. Beyond spatial reconstruction, the project serves to honor and support a deeper understanding of these communities, and what their story means for us today, especially among Asian Americans. Building on speculative historical work and the critical fabulation of Saidiya Hartman, as well as the practice of re-existencia coined by Adolfo Albán Achinte, this creative approach to reconstruction is called relational reconstruction, as presented in a toolkit, video series, and workshops over a two year period. This talk presents some of the motivations for this form, discusses issues of invisibility and archival exclusion, as well as commemoration, audience, and personal connection in relation to the work. Relational reconstruction addresses the limitations of the archive as a means of accessing erased moments, experiences, and spaces and seeks to correct erasure, but through an imaginative, immersive, social, personal, and creative approach, in which visitors are enclosed in a gentle and nourishing space, rather than enclosing an artifact of history in a diorama at a museum or on a plaque.

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)* **1**

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 Poso[1], Mikko Lipsanen[2], Ida Toivanen[1], and Tanja Välisalo[1]; [1]University of Jyväskylä and [2]The 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,** *Ana B. López-Baldomero, Eva M. Valero, Anna S. Reichert, Francisco Moronta-Montero, Miguel A. Martínez-Domingo, and Ana López-Montes, University of Granada (Spain)* **11**

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 Ciortan and Giorgio Trumpy, Norwegian University of Science and Technology (Norway)* **17**

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 of Amsterdam (the Netherlands)* **23**

Since the advent of the Digital Intermediate (DI) and the Cineon system, motion picture film preservation and restoration practices overcame an enormous change derived from the possibility of digitizing and digitally restoring film materials. Today, film materials are scanned using mostly commercial film scanners, which process the frames into the Academy Color Encoding Specification (ACES) and present proprietary LUTs of negative-to-positive conversions, image enhancement, and color correction.

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, and Luiz Antonio Cruz Souza, Federal University of Minas Gerais (Brazil)* **28**
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 Kuzio, Getty Conservation Institute (US)* **A-2**
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 Collection,** *Ariel Segal and Hana Beckerle, Library of Congress (US)* **33**
This paper will present an overview of a project to digitize the Library of Congress Hebrew Manuscripts collection, which spanned from 2021 through spring 2023. It will describe the historical/cultural importance and breadth of the collection, as well as the workflow and processes used to digitize and display the manuscripts.

15:20 **United States National Archives and Records Administration Digitization Regulations,** *Michael Horsley, US National Archives and Records Administration (US)* **39**
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 of Art (US)* **43**
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 Xander Harcourt, US Copyright Office (US)* **A-5**
Since 2021, the United States Copyright Office, a sub-agency of the Library of Congress, has been digitizing its historical record books dating from 1870-1977. This collection, when complete, will be over 26 million images in size and is currently the Library of Congress's largest digitization project. In our presentation, we will give a brief overview of the history and strategy of the digitization of the collections held by the U.S. Copyright Office. We will also discuss the development and deployment of our in-house quality assurance process, which screens through approximately 100,000 images a week.

17:10 **How to (Mass-)Digitize Newspapers in Switzerland — The Swiss National Library Approach — Renewed,** *Martina Hoffmann, Swiss National Library (Switzerland)* **51**
The Swiss National Library (SNL) operates a variety of different digitization projects for different kinds of materials. The biggest part of the pie contains newspapers. Newspapers are of high interest to the public and to researchers in the Digital Humanities field. The effort must be made to put as much as possible newspapers online. The SNL as per her strategy is to take a leading role in this digitization effort. This paper will describe how the newspapers are digitized, what the SNL's role is and how the pipeline is structured from original to online presentation.

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,** *Erwin Verbruggen, Q42 (the Netherlands), and Marcel Duin, Q42 (Japan)* **A-7**
While digitization processes' technical requirements multiply, with setups 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 University (US)* **56**
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 (hard-disk drives, solid-state drives or flash memory, optical discs, and ½" magnetic tape. It ends with a discussion of our current research on a human-readable, 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)* **68**
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. Farnand, 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, and user interface upgrades.

BEHIND-THE-SCENES TOURS

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.

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 (US)* **76**

Advancements in accurate digitization of 3D objects through photogrammetry are ongoing in the cultural heritage space, for the purposes of digital archival and worldwide access. This paper outlines and documents several user-driven enhancements to the photogrammetry pipeline to improve the fidelity of digitizations. In particular, we introduce a new platform for capturing empirically-based specularity of 3D models called Kintsugi 3D, and visually compare traditional photogrammetry results with this new technique. Kintsugi 3D is a free and open-source package that features, among other things, the ability to generate a set of textures for a 3D model, including normal and specularity maps, based empirically on ground-truth observations from a flash-on-camera image set. It is hoped that the ongoing development of Kintsugi 3D will improve public access for institutions with an interest in sharing high-fidelity photogrammetry.

10:40 – 11:20

COFFEE BREAK

11:20 **3D Object Quality Metrics and their Differences: How Can We Evaluate Quality of Digitization?,** *Markus Sebastian Bakken Storeide, Sony George, Aditya Suneel Sole, and Jon Yngve Hardeberg, Norwegian University of Science and Technology (Norway)* **81**

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 Masters Digital Imaging, Inc. (US)* **88**

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".

12:00 **Evaluation of Binarization Methods for Hyperspectral Samples of 16th and 17th Century Family Trees,** *Francisco Moronta-Montero[1], Ramón Fernández-Gualda[1], Ana B. López-Baldomero[1], Marco Buzzelli[2], Miguel A. Martínez-Domingo[1], and Eva M. Valero[1]; [1]University of Granada (Spain) and [2]University of Milano-Bicocca (Italy)* **94**

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 of Cultural History, University of Oslo (Norway)* **101**

This project at the Museum of Cultural History in Oslo, utilizing advanced 3D scanning technology including Tritop, focuses on preserving Norway's iconic Viking artifacts, such as ships and historical objects. Our goal is to preserve and share these treasures, ensuring their legacy for future generations.

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**

IS&T CORPORATE MEMBERS

STRATEGIC PARTNERS

Adobe

Google

Meta

SUSTAINING

Qualcomm

xerox

SUPPORTING

Dolby

hp

RICOH

DONOR

APPLIED IMAGE®

COLORADO STATE UNIVERSITY

DXOMARK

Felix Schoeller Group
Best Performing Papers. Worldwide.

FUJIFILM

Image Engineering
MEMBER OF THE NYNOMIC GROUP

NTNU
Norwegian University of Science and Technology

OMNIVISION™

TELELUMEN
The Recording and Playback of Light

APPENDIX A

Wide-Format High-Speed Film Scanner

Matt Anderson; Lake Forest, IL, USA
Rich Chaney; Meridian, ID, USA

Abstract

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.

Scope of the Project

For nearly a century, the United States has been building a collection of large-format film from aerial and satellite missions. The sheer volume of physical film being archived and the tools for digitizing prior to Apex were slow, expensive and not capable of digitizing the entire collection of different wide format assets.

For nearly a century, the United States has been building a collection of large-format film from aerial and satellite missions. The sheer volume of physical film being archived and the tools for digitizing prior to Apex were slow, expensive and not capable of digitizing the entire collection of different wide format assets.

The Need

In 2018 the National Geospatial-Intelligence Agency sought to find a commercial solution to design, configure and build a high-speed film scanner with the following requirements, scan film sized from 70mm up to 10" wide, have a scan resolution of 4um spot size and to scan the film at 10 feet per minute, about 2 inches every second.

There were two main challenges that nextScan engineers identified, first, how to capture the images and second, processing and storing the images.

The Solution

To meet the specifications, four line-scan cameras to span the width of the film and stitch the images into a pixel-perfect composite in real time.

To accomplish the stitching accurately, precise mechanical alignment of cameras, mirrors and lenses. The camera mounting was designed with 3-axis alignment capability. Alternate alignment strategies were considered such as software image stitching but were rejected as they required altering the raw capture image.

Illumination of the film is also a critical design element. The Apex uses nextScan's patented LuminTec® lighting technology. This high-tech process synchronizes the strobe of the light line with the camera capture to freeze the image, providing a clear image.

Data Storage

As you can imagine, the wide-format film creates a tremendous amount of data. When scanning 10-inch film the Apex captures 819 million pixels per second which, capturing 12-bits of data, generates 1.53 gigabytes of data per second. This raw image data is stored on an internal 120 terabyte Ribbon Storage Device (RSD) which is networkable with via 10 GbE or fiber connectivity.

Full size images may be output in Big TIFF format. Other file types such as JPEG, BMP or PDF would be unable to save the full-size image due to their limitations.

Conclusion

NGA now enjoys their fleet of nine Apex scanners, working hard to complete their mission of getting more aerial film digitized and made available to researchers around the world.

Author Biography

Matt Anderson received his BS in communications from Bradley University (2002). He has worked as the VP of marketing for nextScan since 2019.

Rich Chaney received his BSEE from Texas A&M University (1994), an MBA from San Jose State University (1999), and multiple patents for data storage and semiconductor technology. Rich has held positions at American Semiconductor, Micron Technology, Quantum Corporation, Cypress Semiconductors, and Novellus Systems.

Science, Sustainability, Security ARCHIVING2024 Washington, DC APRIL 9-12

Examining and *Exulting*: multimodal imaging and collaborating through a new lens

Kayla Kee[1] and Olivia Kuzio[2]
[1] J. Paul Getty Trust; Los Angeles, CA/US
[2] Getty Conservation Institute; Los Angeles, CA/US

Abstract

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.

Motivation

Near the turn of the 19th century, the English printmaker William Blake (1757–1827) produced a group of thirty-three paper-based works of art that have been called his "large color-printed drawings" [1]. These objects are unique among Blake's work, because of the idiosyncratic manner in which he combined both printed and hand-applied media to create the overall compositions.

A new examination of the J. Paul Getty Museum's large color print, *Satan Exulting over Eve* (1795) (Figure 1), has recently begun, with the aim of providing scholars with a better understanding of the idiosyncratic technique by which Blake created these objects. The first step of this technical analysis involved scientific imaging of the object, including high resolution UV, visible, and IR photography, which informed subsequent macro X-ray fluorescence (ma-XRF) scanning.

These analyses resulted in a collection of images that individually contained valuable information for better understanding the materials and methods used. For example, the reflected UV photograph initially revealed reserves in some areas of the background that indicate compositional changes with respect to the posture of Satan (Figure 2. Additionally, the XRF maps differentiate the various colored inks elementally, and illustrate differences in their application and distribution across the surface of the print (Figures 3-4). When examined simultaneously and with respect to one another, details in each image aid in distinguishing different materials in the composition, which in turn inform inferences about Blake's process in creating the work.

Figure 1. *Color image of William Blake's Satan Exulting over Eve (84.GC.49). J. Paul Getty Museum, Los Angeles, CA.*

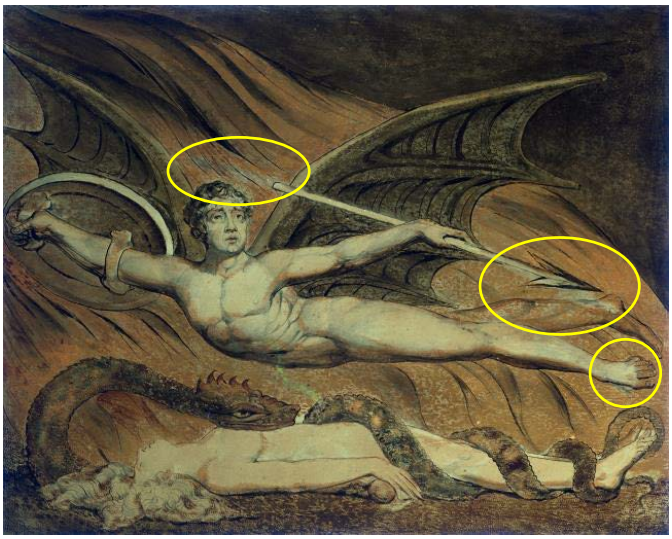

Figure 2. *UV image of Satan Exulting over Eve. Yellow circles denote areas of reserve in the red background that differ from the drawn placement of the spear and foot.*

Problem & Approach

While examining the images each on their own was useful to a point, it quickly became apparent that overlaying and toggling between them would be the only way to make the comparisons necessary to draw inferences about Blake's working methods. Based on the typical process of collecting and delivering these photography-based and XRF scan-based image sets—using two different imaging setups in two physically different locations and collected by two different departments—this is not how the images would be delivered to the conservator requesting these analyses. It followed that the somewhat trivial exercise—for a photographer—of creating a multilayered Photoshop file containing both the technical photography (UV, visible, and IR photographs) and XRF element maps became the crux of the first stage of the technical examination of this work, because it allowed for the visual correlation of evidence of material distribution provided by each technique.

The collection of images that resulted from carrying out each of these techniques were resized to a common scale and collated as layers in a single, registered Photoshop file. The ability to overlay the images in the stack, and to "flip through" them, like flipping through cards in a deck, provided a quick and convenient visual means of comparing the information conveyed by each imaging technique. This ability alone enabled an organized and informed way to infer more about Blake's working methods and the order in which he may have applied the different materials onto the substrate to achieve the overall texture and coloration of the work.

Results

Upon first look, *Satan Exulting over Eve* looks deceptively like a watercolor with outlines drawn in black ink. However, upon close visual inspection of the work on paper, it becomes clear that there are distinctly different textures present in the colored areas that indicate the colored media was applied by different methods. Some areas, like the green foreground, have a dappled appearance characteristic of application by printing, while other areas, like the fleshy highlights on Satan's body, are smoother look and appear to have been applied/manipulated by hand. A main goal of the technical imaging was to characterize and compare the inks present in the printed versus hand-brushed areas toward elucidating distinctions between the materials Blake printed versus applied by hand. Main takeaways about materials include elemental evidence suggesting that iron-containing materials, such iron-rich earth pigments (e.g. yellow ochre) were printed (Figure 3) while the mercury-containing pigment vermillion was hand-applied by brush (Figure 4).

While in many ways a simple solution, providing this easy-to-navigate, digital, co-registered multimodal image stack to the conservation stakeholders was a novel way to deliver these datasets internally at Getty. Because it was met with such enthusiasm, and because it so effectively increased our understanding of how these objects were made, we are now in the process of reconsidering our routine methods of capturing, combining, and delivering such multimodal datasets. Making standard such a deliverable will further, positively integrate the work of several departments at Getty, and serves as motivation to continue making small improvements to the ways in which we collaborate in order to see large, meaningful impact on the connections we are able to make with collections.

Future Action

At Getty our current internal DAM does not support viewing image files beyond JPEG or TIFFs. We are optimistic that this project will foster momentum toward making multimodal image stacks more accessible and viewable. In particular, this includes working with our software team to adopt an IIIF viewer(s) for files of this type. The image viewer ideally would be hosted on Getty.edu and be available for internal use, scholars, and public alike. The ability to view selected images captured across the spectrum would encourage not only more interest in critical looking and thinking around the works of art, but also more interest and visibility for the breadth of imaging work that we routinely utilize in our field.

Figure 3. *XRF map of iron illustrating its presence mainly in the interior of the wings, the serpent's scales, the lower part of the flames, and the ground—all of which are areas of the composition that appear to have been printed.*

Figure 4. *XRF map of mercury illustrating its presence mainly in the shading of the figure of Satan, some of the details of the flames, and in highlights in the serpent's facial features.*

Conclusions

While the creation of this kind of deliverable is perhaps unremarkable, it was the enthusiasm that this deviation from the norm was met with, and the conversations that ensued, that were extremely positive and noteworthy. This opened the door for new dialogue about the power of examining and storing multimodal image data in updated, relevant, useful ways for all project stakeholders, and has deepened interest in further developing interdepartmental communication and collaboration around this topic and beyond. We reiterate that 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.

References

[1] M. Butlin. *William Blake*. Vol. 1, 156-17.

Author Biographies

Kayla Kee is a photographer at the J. Paul Getty Museum. She photographs objects from each collection at Getty Museum, focusing on paintings, illuminated manuscripts, drawings, and photographic works. Additionally, she captures objects with scientific photography techniques such as UV, IR, RTI, etc. She has been with Getty since 2015, after graduating with a BFA in Fine Art from ArtCenter College of Design.

Olivia Kuzio is an assistant scientist in the Getty Conservation Institute's Technical Studies Research laboratory. She conducts technical studies on works of art to address questions of composition, artistic practice, and material degradation, with a focus on expanding the Institute's capabilities in spectral imaging techniques and developing standard protocols for scientific imaging of museum objects. She received a PhD in color science from Rochester Institute of Technology, as well as MS and BS degrees in chemistry from RIT and Pennsylvania State University, respectively.

Building a Program—from 0 to 26,000,000

Phelps, Kristin A., and Harcourt, Xander; United States Copyright Office; Washington D.C.

Abstract

Since 2021, the United States Copyright Office, a sub-agency of the Library of Congress, has been digitizing its historical record books dating from 1870-1977. This collection, when complete, will be over 26 million images in size and is currently the Library of Congress's largest digitization project. In our presentation, we will give a brief overview of the history and strategy of the digitization of the collections held by the U.S. Copyright Office. We will also discuss the development and deployment of our in-house quality assurance process, which screens through approximately 100,000 images a week.

Proposal

The United States Copyright Office (USCO) is unlike a traditional institution in the GLAM sector. While it holds several collections of historically important material, its collections are finite. However, since its collections are important to the public, a push has been made over the last decade to begin digitizing material for preservation and online access. The first foray into digitization for the Copyright Office was its collection of catalog cards, which numbered 41.5 million cards dating from 1870-1977. It was by no means a small effort and certainly not for an institution just beginning in digitization. The cards were digitized with the help of vendors working onsite from 2011-2014. The digitization project also included OCR capture. These cards were first available to the public as a proof of concept in 2018. The Virtual Card Catalog is available to the public here: https://vcc.copyright.gov/browse. The card project was not finished, however. It still needed to be a searchable collection on the Copyright Public Records System, or CPRS, which is currently in development. As a result, the cards are currently going through metadata capture and cleanup to allow for keyword searches of the digital images. These cards can be found here: https://publicrecords.copyright.gov

However, the catalog cards only represent a portion of Copyright's collection, which might be required to research specific items. Often, the cards need to be used in combination with the Historical Record Books, microfilm, and the Catalog of Copyright Entries (CCE). Fortunately, the CCE has already been digitized and is available thanks to the Internet Archive, which was supported by the Copyright Office. But the Historical Record Books, one of Copyright's most used assets, was not. Without the collection being digitized, researchers either needed to travel to the Copyright's reading room in Washington, D.C., or pay for the Copyright Office to perform the research on behalf of the researcher. This service normally begins at $200. With these two options being the only options available, the Copyright Office prioritized digitizing the Historical Record Books to make them freely available to the public.

How to approach a collection of 26,278 volumes containing an estimated 26,278,000 pages? Certainly, digitizing such a large collection is daunting to any organization. That was true not only for the Copyright Office, but also for our digitization colleagues at the Library of Congress. Fortunately, the two groups were able to come together to form a feasible path forward. During the planning stages, colleagues from the Library's Digital Scan Center provided guidance and valuable advice to help evolve an idea into a plan. This plan became the Historical Record Books Project, which is a multi-year project to create searchable digital images of pre-1978 copyright records. These records date from 1870-1977 and include registrations, renewals, assignments, notices of use of musical compositions, and other related records, including patent records and patent examiner indices from 1870-1940.Additionally, many of the works are still under Copyright protection.

The digitization phase of the project began early in 2021 and is anticipated to end in 2026. In order to make these important records available as soon as possible, nearly 10,000 record books are available with volume-level searching on the Library of Congress's online digital collections platform. During a second future phase of the project, the Copyright Office will gather record-level metadata to present the individual records and associated metadata in the CPRS. The Office is already doing some early experiments with machine learning to produce the record-level metadata.

Since the Copyright Office did not have a formal digitization program, it hired its first digitization manager and Quality Assurance (QA) manager. These two positions became responsible for making on-the-ground decisions and perfecting a workflow via trial and error. The Copyright Office is an Agile organization, and this was incorporated into the project workflow and project management, utilizing adaptive workflows, mini-pilot projects, and frequent team meetings. The Office contracted with trusted vendors who could meet the volume demand for image capture. And the Office created an in-house QA process that gave the ability to check and verify the vendor's images. The Office also worked with our vendor, the QA team, and OCIO to develop a streamlined workflow in which images were delivered bi-weekly, the QA team completes approximately one batch every week, the accepted batch is ingested into both presentation and long-term storage each week, and new volumes are available to the public approximately once a month.

Although the project was affected initially by the pandemic and suffered a few of the expected project start-up issues, the Copyright Office accepted its first batch of images from the vendor in March of 2021, and first public release of images occurred in February of 2022. All images are scanned to FADGI 3* levels, which are verified both by the vendor and the Copyright Office. As of October 31, 2023, 10.5 million images have been scanned, and nearly 10 million of those images are available to the public and can be viewed here: https://www.loc.gov/collections/copyright-historical-record-books-1870-to-1977/. The vendor generates 5 million images a year, which means that their team is producing somewhere between 45-52 images every minute during an average work week. In addition to their scanning duties, the vendor is also responsible for their own QA of images. Once the vendor has finalized the images, they are sent to the Copyright Office's QA team for another round of QA. Our QA team is staffed in cooperation with the Federal Research Division (FRD), a fee-based research and analysis unit within the Library of Congress. Because the images have gone through multiple rounds of QA, the

Office's QA team generally see an error rate with the images of less than 1%. The team have built-in time to ensure that if there is a delay at any part of the process, the project does not fall behind on its production rate.

The Copyright Office is approximately halfway through the first phase of the project imaging and still has a record-level metadata capture project to complete in the future. Additionally, the Office has already begun to assess several other collections for future digitization. The Office has already completed a condition assessment and inventory of 26,000 microfilm reels and is planning for partial digitization of the collection. The Office has also completed a 10% survey of our Patent Prints and Labels collection with the hopes of a conservation and digitization project in the coming years. So, the Copyright Office, in the period of a little over a decade, has truly gone from 0 to (almost) 26,000,000! USCO is committed to continuing to build our digitization program and already has plans to include other records and collections.

Such a project cannot be a true success without considering the lessons learned. Especially as the Copyright Office looks to digitize more in the future. First, no project with multiple teams performing multiple processes simultaneously can hope to be successful without a centralized production management system. The Office was fortunate that fellow colleagues in the Law Library of the Library of Congress had contended with this issue and created an Access database in response. The team has found that this database, the Copyright Office Production Management System or COPMS, has provided a way to "communicate" via data, generate reports, and keep track of our numbers. It took some modifications to make the database 'right' for the project, but once those changes were made, the database kept the team running on the same page. Second, understanding a project's entire resource needs is critical; this includes additional staff, IT-related needs and costs (including, but not limited to, hardware and software), and specialized skills. For much of the project's first two years, the project team spent time working through the realities of the project. The team began to understand how drastically understaffing could affect the workflow; if the project didn't have the right tools or fast enough processing power, it could end up stranded, and the team learned the importance of cross-training within the team for redundancy as well as a library of project written how-to manuals. Projects like this do not work without the right tools and people. Third, communication is key! Large-scale projects require continued communication regarding the overall strategy to keep all stakeholders in sync. This includes communication at the project team level and up. It also includes, as a publicly funded organization, communication with the general public. An added bonus of frequent communication is building a solid team that communicates and works well together. Fourth, Agile project management is beneficial for new programs and projects. Working in phases is especially helpful in conquering the learning curve. The team also continue to employ mini-pilot projects to test how changes may affect the timeline and output. Also, the emphasis on communication, as mentioned previously, is a cohesive element. Fifth, the team has found that simplifying processes by means of reducing requirements and streamlining the process made for a smoother workflow. One example of this is working with FRD for the QA team. The initial plan had called for filling positions on the team with full-time employees from Copyright who had their own tasks to complete. FRD works on a multiple project basis, so they were set up to come in and perform the QA. Another example comes from the workflow for pick up and return of the image drives from the vendor. Initially, this was split into a two-person job from Copyright, with one person picking up and another returning. It was determined that the pick-up and delivery could be a multitasked job and now a single person both picks up and returns the drives. It is a simple thing, but it makes the process more efficient and saves time. Finally, the team learned that building in time for research, analysis, content review, and assessment was critical. Figuring out the appropriate timelines for the QA and acceptance process changed over time, related to changes in the delivery. The team now receives multiple batches of images at a time but allows more time from delivery to acceptance. Also, a yearly lessons learned document helps the Office in making future decisions and to continue refining the team's process. Using time to research different workflows and looking at different technologies has been a positive use of the team's time over the first years of the project.

In conclusion, developing and running a project of this size is no small task. It takes time, expertise, passion, and a willingness to see it through. The Copyright Office has been supportive of the team's efforts and continues to be interested in not only this project but the future growth of digitization across the Office as well.

Author Biography

Kristin A. Phelps is the Digitization Manager at the U.S. Copyright Office. Previously, she has worked in cultural heritage imaging at Duke University, the British Library, and the British Museum. She has also in the past worked to develop a cultural heritage imaging and preservation program in Puerto Rico. She holds a degree in Photography and a Masters in Classical Archaeology.

Xander Harcourt has been with the Library of Congress – U.S. Copyright Office for over 20 years. He is a program analyst with the Office of Copyright Records and the Office's first Quality Assurance Manager for the Historical Public Records Program. Xander holds a bachelor's degree in business management, Agile Scrum Certification, and a host of other educational and professional certifications.

Micrio as an Ultra-Resolution Story Format Smorgasbord

Erwin Verbruggen; Q42; Amsterdam, Netherlands & Marcel Duin; Q42; Tokyo, Japan

Abstract

While digitization processes' technical requirements multiply, with setups 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.

Origins

Micrio is, in essence, a 2D image storytelling platform. It provides layers of storytelling that can be added by means of *markers*, marker tours, video tours, spatial audio and embedded media. Projects for museums, publishers, journalists, designers, NGO's, broadcasters and others have pushed the possibilities of the platform to include immersive environments, live-streamed interactive sessions and 360°-degree views of heritage objects.

The origins of the Micrio platform go back to 2015, when Dutch public broadcaster NTR & Pieter van Huistee Film commissioned digital product studio Q42 with an online experience around Jheronymus Bosch's painting *The Garden of Earthly Delights*.[1] In a web interface the visitor is taken on an audio-visual journey, including sound, music, video and images to enrich the storytelling. The visitor of the interactive documentary gets a better understanding of what it was like to live in the Late Middle Ages, and for example of what importance religion had in daily life. The interactive documentary can be read like a book, one can come back after a visit and pick up the book again from the shelf to further explore. The interactive documentary was part of a transmedia tryptich consisting of a documentary, the interactive documentary and a Virtual Reality documentary.

In the following years, Micrio would be used as a storytelling tool and a collection presentation platform.

***Figure 1**. The interface of the* Garden of Earthly Delights.

Development Trajectory

There are quite a few web libraries available that allow people to view and zoom through ultra-resolution images. Some of those also support markers, tours, etc. like Micrio does. Micrio was created entirely from scratch, having its own performance optimized algorithms and unique features such as positional audio and music.

What sets Micrio apart the most is that it's a full service and that it allows non-developers to easily create and share or embed their own interactive experiences. Micrio is a full-server environment that processes original ultra-resolution images (up to 10GP out of the box), provides globally optimized CDN-fronted hosting, has an online editor to enrich the images with markers, tours and audio, and much more.

In 2019, we added support for the International Image Interoperability Framework (IIIF). The Micrio Javascript client can be transparently used to view existing IIIF-hosted zoomable images. Micrio can be used as a versatile image server, the images therein hosted made accessible through the IIIF Image API 3.0. Micrio also offers support for hosting an organisation's data in their own managed Cloud hosting provider. When an image is uploaded in your organisation, it will be stored there, along with all processed IIIF data, uploaded assets, and published JSON data. This is particularly interesting when organisations have a large number of images, and don't want to rely on Micrio's own (Google-based) cloud hosting, or if they work with sensitive or private data and need to retain full data ownership and restrict access.
In 2020, we started experimenting with using edge-based computing to further optimize the speed with which images could be delivered, with results that our much to our satisfaction.[2][3]

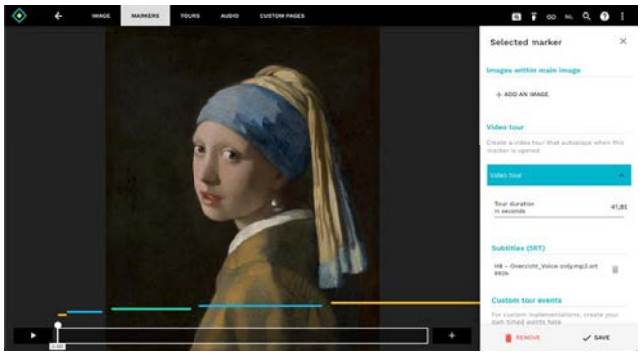

***Figure 1**. Micrio Dashboard view of a Videotour in the* Closer to Vermeer *Rijksmuseum special*

Storytelling Format Expansion

From the get-go, Micrio was used as, primarily, a storytelling platform and, secondarily, a collections presentation platform. In this section, we aim to highlight a number of practices and collaborator's requests that have pushed forward the boundaries of the tool.

From High- to Ultra-High Resolutions

In 2022, the The Operation Night Watch research team made available an extremely detailed photograph of *The Night Watch*. This 717 gigapixel image was made available for viewing on the Rijksmuseum website (and made it the museum's single most viewed web page up until that point).[4] It was "the largest and most detailed photograph of any artwork, and is four times sharper than its predecessor, which the Rijksmuseum published around 18 months ago".[5] The photograph was made as part of the research conducted by Operation Night Watch and pushed the limits of what Micrio could up until that point support.

From 2D spaces to Immersive Environments

With a flat image, a world can be represented. With a 360°-room image, that environment can be explored. With a series of 360°-images, those rooms can receive their own rhythm and purpose. Micrio developed its first 360°-image stories in 2018, after which the global pandemic brought forth further requests for immersive environments to experience and educate.

From Environments to Object Views

On request from a grassroot community project on fashion history, we started developing a 360-degree object viewer. In this case, ModeMuze's curators had gathered expertise to photograph fragile dresses on mannequins for digitization purposes. Shot in flat lighting, the mannquins were shot from different turning angles. Represented in the Micrio viewer, the dresses retain the ability to be seen from various angles, while allowing viewers to zoom in up close. It also offers the ModeMuze editorial team with storytelling capability on the twirling mannequins.

Conclusion

Throughout a series of collaborative projects, Micrio has grown to be a flexible platform for telling stories on ultra-high resolution imagery both on the web, on user's personal devices and in installation views in physical locations. The dashboard provides a host of storytelling formats, the availability of which is being driven forward by the projects we jointly undertake.

References

[1] https://archief.ntr.nl/tuinderlusten/en.html#.
[2] Duin, Marcel. 'From JavaScript to WebAssembly in Three Steps'. Q42 Engineering, 29 September 2020. https://engineering.q42.nl/webassembly/.
[3] Duin, Marcel. 'WebAssembly: Beyond the Browser'. Q42 Engineering, 24 March 2022. https://engineering.q42.nl/webassembly-beyond-the-browser/.
[4] https://www.rijksmuseum.nl/en/stories/operation-night-watch/story/ultra-high-resolution-photo
[5] https://www.rijksmuseum.nl/en/press/press-releases/rijksmuseum-publishes-717-gigapixel-photograph-of-the-night-watch

Author Biography

Erwin Verbruggen is a former audiovisual archivist and current technical project lead at digital product studio Q42. He is a graduate from the Preservation & Presentation of the Moving Image MA at the University of Amsterdam, interned at human rights organization WITNESS in Brooklyn (NY) and was a film programmer for the Open Air Film Festival Pluk de Nacht in Amsterdam.

With over 23 years of experience in web development, Marcel Duin has been a developer at Q42 since 2011. He founded the ultra -resolution storytelling solution Micrio in 2015, which he continues to develop and maintain.

ARCHIVING 2024 AUTHOR INDEX

*abstract only

IS&T BOARD OF DIRECTORS
JULY 2023 – JUNE 2024

imaging.org

Society for Imaging Science and Technology
7003 Kilworth Lane • Springfield, VA 22151 USA