



AI and Images Collections: Two Perspectives

Jessica Bushey, Ph.D
ITrust AI International Symposium,
Honolulu, HI
February 23, 2024

Jessica.Bushey@sjsu.edu

I Trust AI

- AI & Archives
 - Accessing archives online using AI tools
 - RA01
- AI-generated Images
 - Recordkeeping AI-gen Images
 - CU08




QUEEN LILIUOKALANI PHOTOGRAPH EXHIBITION: LILIUOKALANI, 1890'S

To navigate the exhibition:


Clicking on each thumbnail will pop-up a preview of the photograph. Click on the right or left side of the photograph to scroll through this folder's contents or you can click on the x on the upper right corner or anywhere outside the photograph to return to the thumbnails. The thumbnail and pop-up preview can be saved individually by right-clicking on them and selecting "Save as...". Both versions are set to 72dpi for easy web viewing. Click on the links below the thumbnails for higher resolution copies. "Cropped" is a lightly edited version ready for printing, "Original" is unedited with color target. Both "Cropped" and "Original" are 1200dpi. "Reverse" is a PDF file of the back of the photograph provided for informational purposes. For citation help, check our [Citing Sources from the Hawai'i State Archives](#) page.

Please click on following links to browse through the exhibition.


Folder Name/Description	Folder Number
Liliuokalani, pre-1880	PP-98-10
Liliuokalani, 1880's	PP-98-11
Liliuokalani, 1890's	PP-98-12
Liliuokalani, 1900-1917	PP-98-13
Liliuokalani, Jewelry, Music, Piano, Misc.	PP-98-14
Liliuokalani, Withdrawn Photos	PPWD-16-3
Liliuokalani, Withdrawn Photos	PPWD-16-4
Back to Queen Liliuokalani Photograph Exhibition Home	
Back to State Archives' Home	




Call No.: PP-98-12-001
Cropped: [TIF \(60 MB\)](#)
Original: [TIF \(270 MB\)](#)
Reverse: [PDF \(93 KB\)](#)




Call No.: PP-98-12-002
Cropped: [TIF \(74 MB\)](#)
Original: [TIF \(458 MB\)](#)
Reverse: [PDF \(132 KB\)](#)




Call No.: PP-98-12-003
Cropped: [TIF \(119 MB\)](#)
Original: [TIF \(272 MB\)](#)
Reverse: [PDF \(78 KB\)](#)





Call No.: PP-98-12-004
Cropped: [TIF \(124 MB\)](#)
Original: [TIF \(466 MB\)](#)
Reverse: [PDF \(143 KB\)](#)





Call No.: PP-98-12-005
Cropped: [TIF \(223 MB\)](#)
Original: [TIF \(335 MB\)](#)
Reverse: [PDF \(102 KB\)](#)











Hawaii State Archives. Queen Liliuokalani Collection 1890s:
<https://ags.hawaii.gov/archives/about-us/photograph-collection/queen-liliuokalani-photograph-exhibition/queen-liliuokalani-photograph-exhibition-liliuokalani-1890s/>




Contribute

Information - Brief


Hierarchical level: Item
Date: ca. 1943-1965.
Reference: R112-2467-5-E, Volume number: 30448
Item no. (creator): ZK-1092-10
Type of material: Photographs
Found in: Archives / Collections and Fonds
Item ID number: 4233836

Context of this record: [Show detail](#)

Library Archives Canada. Department of Defense. Arctic Exploration Muskox:
<http://central.bac-lac.gc.ca/.redirect?app=fonandcol&id=4233836&lang=eng>



Link to this record



Add to My Research



RA 01 – Increase Access to Images & AV records through AI-generated Metadata

- Jan 2022 – Dec 2025
- Case studies:
 - Hawaii State Archives [Dr. Adam Jansen]
 - International Federation of Red Cross & Red Crescent, Film Archives, Geneva. [Grant Mitchell & Sarah-Joy Maddeaux]
- Activities:
 - Annotated Bibliographies for AI & Film and AI & Photo Collections
 - Literature Review of AI & Photo Collections
 - Testing COTS AI Video Indexing for Digital Video & Digital Film



RA 01

- Early Findings:

- Variety of AI technology and tools being used by archivists, COTS and bespoke software
- Goal is for AI to provide results from queries of image and AV collections
- Verry few available to the public, mostly using computer vision to describe content of images
- Training data is local (archives own)
- Labelling is either AI-driven (non-specific) or expert-driven
- Human oversight needed

RA 01

- Questions:

- How do these activities meet existing archival theory & standards of arrangement & description?
- As AI technology and tools become more available, how can archivists prioritize context over content?
- How can crowdsourcing or data curation be integrated into AI tools to improve instances after release?
- What is the impact on user satisfaction?
- What is the impact on Reference Services?



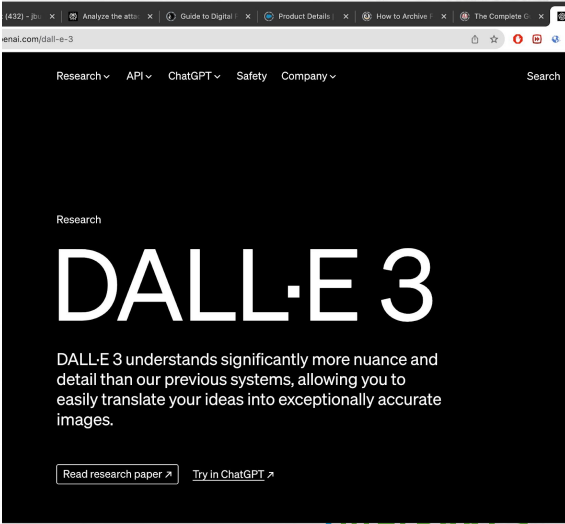
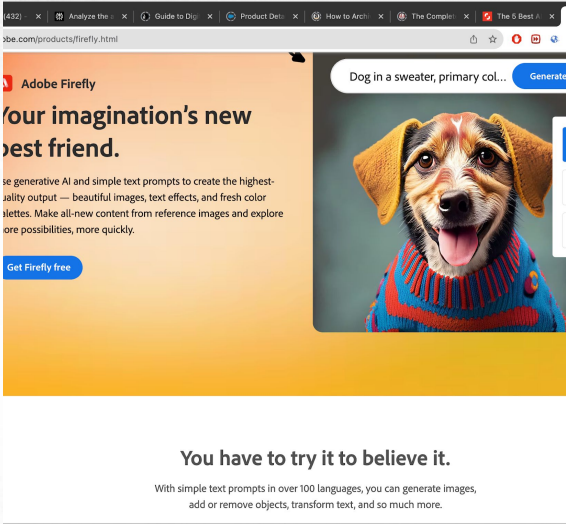
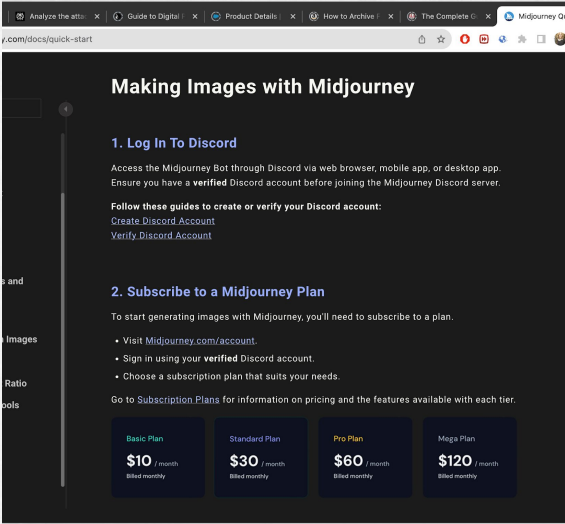
CU 08 - Recordkeeping Practices of Creators using AI to generate Images

- April 2023 - Dec 2025
- Activities:
 - Literature Review of Gen-AI Image creation and management
 - Journalism and Communications
 - Medical Diagnostics
 - Law Enforcement
 - Publication – “AI-generated Images as an Emergent Record Format” (Bushey, 2024).



INTRODUCTION

Generative AI, a subset of AI trained to produce new content, either randomly or based on prompts provided by users.



Can We No Longer Believe Anything We See?

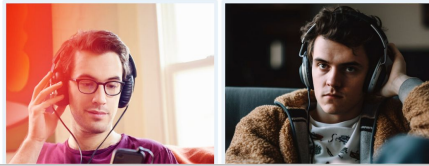
By [Tiffany Hsu](#) and [Steven Lee Myers](#)

April 8, 2023

Seeing has not been believing for a very long time. Photos have been [faked](#) and [manipulated](#) for nearly as long as photography has existed.

Now, not even reality is required for photographs to look authentic — just artificial intelligence responding to a prompt. Even experts sometimes struggle to tell if one is real or not. Can you?

Which image was created by artificial intelligence? Click on your guess



BUSINESS

An A.I.-Generated Spoof Rattles the Markets

A stock sell-off driven by a since-debunked picture underscored fears about how artificial intelligence could be used for nefarious purposes with big consequences.

By [Andrew Ross Sorkin](#), [Bernhard Warner](#), [Sarah Kessler](#), [Michael J. de la Merced](#), [Lauren Hirsch](#) and [Ephrat Livni](#)



An A.I.-generated image of Francis. The flood of papal

A.I. Muddies Israel-Hamas War in Unexpected Way

Fakes related to the conflict have been limited and largely unconvincing, but their presence has people doubting real evidence.

Share full article



A New York Times photograph shows thick smoke from Israeli bombs rising in Gaza on Thursday. Other authentic footage from the war with Hamas has fueled skepticism stemming from the rise of A.I. technology. Source: Abu Hana for The New York Times



By [Tiffany Hsu](#) and [Stuart A. Thompson](#)
Tiffany Hsu and Stuart Thompson monitor the spread of false and misleading information online.

Published Oct. 26, 2023 Updated Oct. 30, 2023



RATIONALE

The rationale for this research is to better understand the nature of AI-generated images as an emergent record format and to identify risks posed by their creation and use to establishing and protecting their trustworthiness as historical records.

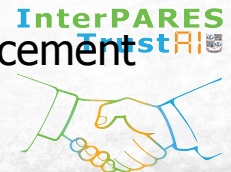
At this early stage we are gathering information about how AI-generative images are being created and managed in the domains of journalism & communications, law enforcement and medical diagnostics.

InterPARES
TrustAI



Literature Review

- KEYWORD SEARCHES
 - AI-generated images; generative AI; authenticity; trustworthy AI-generated images; synthetic images; and deepfakes.
- Library and Information Science Source (LISS) database [2014-2023]
- Factiva database [2022-2023]
- 61 sources [peer-reviewed journal articles, conference papers, project reports in arXiv, newspaper articles, and project websites]
- Analyzed the content
 - Journalism and Communications; Medical Diagnostics; Law Enforcement
- Thematic Categories



RESULTS

- Field of Medicine
 - Synthetic images produced by Generative Adversarial networks (GANs) to contribute to data sets to train ML tools for diagnosis
 - Use of medical images in datasets and risks to privacy
 - Role of Digital Imaging and Communication in Medicine (DICOM) metadata / ISO 12052 standard
 - Artificial Intelligence Industry Innovation Coalition (AI3C) launched 2022
- Field of Law Enforcement
 - Facial Recognition Technology (FRT) issues with bias in datasets to train ML
 - Focus on detecting deepfake images and videos [fake evidence to manipulate legal proceedings & disinformation]
- Field of Journalism and Media Communications
 - Loss of public trust & challenge to veracity of documentary news images
 - Detection of fakes, verification of authenticity, establishing persistent provenance
 - Role of Coalition for Content Provenance and Authenticity (C2PA) metadata manifest [industry led]
 - Copyright of AI-generated works



AI-gen Metadata

- **trainedAlgorithmicMedia**: The image was created using a model derived from sampled content.
- **compositeSynthetic**: The image is a composite that includes synthetic elements.
- **algorithmicMedia**: The image was created purely by an algorithm not based on any sampled training data (for example, an image created by software using a mathematical formula).



Beta: Currently, this property is in beta and only available for IPTC photo metadata. Adding this property makes your image eligible for display with an AI-generated label, but you may not see the label in Google Images right away, as we're still actively developing it.



DISCUSSION

- Authenticity and Verifiability
- Manipulation and Misinformation
- Bias and Representation
- Attribution and Intellectual Property
- Transparency and Explainability
- Ethical Considerations



Authenticity and Verifiability

- Approaches and activities that focus on contextual information that contributes to the creation and use of AI-generated images as trustworthy records.
- **C2PA metadata specification through open-source provenance embedding tools (e.g., content credentials in Adobe photoshop and Leica cameras)**



Manipulation and Misinformation

- Approaches and activities that contribute to identifying images that have been intentionally altered or created as fakes.
- **Establishing trust credentials and definitive statements about the reliability and accuracy of born-digital images, either through guidelines or established procedures.**



Bias and Representation

- Approaches and activities that focus on training datasets for generative AI.
- **Emerging policies aimed at the importance of the images and labels used in training tools and efforts to identify the provenance of training data to reduce racial– and gender-bias.**



Attribution and Intellectual Property

- Approaches and activities specific to the rights and responsibility of AI-generated images.
- **Current public consultations with Federal Copyright Offices to determine:**
 - Use of copyright-protected works in the development of AI systems;
 - Potential for AI-generated outputs to infringe existing copyright-protected works;
 - Role of AI systems in generating content and approaches to accurately attributing works and determining copyright protection; and
 - Lack of practical enforcement remedies for rights holders.



Transparency and Explainability

- Concerns specific to the proprietary nature of AI technologies and tools.
- **Complexity of AI processes and systems is a challenge to understanding the processes by which images are being created. COTS applications are being adopted by organizations and individuals with little knowledge of ongoing access, dependencies and future interoperability, technological obsolescence etc.**



Ethical Considerations

- Concerns and approaches specific to privacy requirements.
- **Issues raised regarding personal identifying information (PII) in DICOM and C2PA metadata linked to images that could be included in training data.**



RESEARCH QUESTIONS

1. How are individuals and organizations using AI-generated images?
2. What AI tools and technologies are being used to create, manage, and store AI-generated images?
3. What actions are being taken by individuals and organizations to identify AI-generated content?
4. What standards and/or policies are guiding procedures for creating, using, and preserving AI generated images?



NEXT STEPS

- Analysis of the C2PA specification using the IP Trust Creator and Preserver Guidelines
- Analysis of evolving Intellectual Property Rights for AI-generative images
- Analysis of emerging industry-specific policies and procedures for AI-generated images
- Collaboration and/or participation in industry-led guidelines and best practices for creation, use and preservation of AI-generative images
- Survey and Interviews with creators & users of AI-generative images



Mahalo

Jessica.bushey@sjsu.edu

InterPARES
TrustAI

