Image Archives in the Age of AI

From Data to Discovery

Presentation by Dr. Jessica Bushey
San José State University

Keeping Evidence: Archives, AI & Trust in Public Records
Library and Archives Canada, Ottawa, Canada
October 29, 2026

Photographic Records: Analogue to AI



Research Design: Four Interconnected Phases

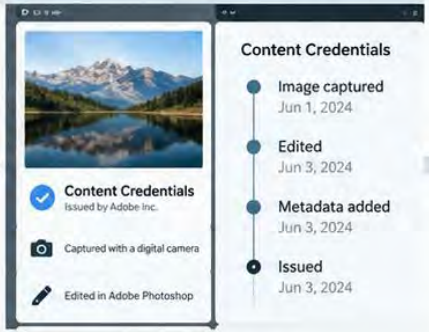
Recordkeeping Practices of Creators Using AI to Generate Images (CU08)

1 2023 Interdisciplinary Literature Review



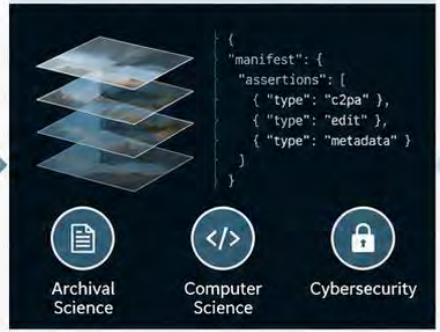
- Established current knowledge on AI-generated images
- Identified emerging concerns
- Formulated research questions

2 2024 Literature Review and Archival Diplomatic Analysis



- Expanded literature review
- Conducted archival diplomatic analysis of Content Credentials
- Examined provenance as an emerging technical approach

3 2024 Transdisciplinary Computational Study



- Investigated cryptographic provenance
- Explored operationalising archival identity and integrity
- Integrated archival, computer science and cybersecurity perspectives

4 2025 Survey of Professional Image Creators and Preservers



- Surveyed professional image creators and preservers
- Investigated engagement with generative AI
- Examined expectations for provenance and trust
- Analysed adoption of standards and verification tools

Each phase informed the next, providing a comprehensive understanding of the technical possibilities, professional practices and infrastructures necessary to support trustworthy photographic records.

C2PA

Coalition for Content Provenance & Authenticity

Content Credentials



Mountain Lake.jpg
JPEG Image
May 28, 2024 at 3:02 PM
1.2 MB • 4032 x 3024

Content summary

This image is AI-assisted
The majority of this content was generated with an AI tool.

About the creator

Name
Alex Morgan
Digital identifier
alexmorgan@example.com

About the creation

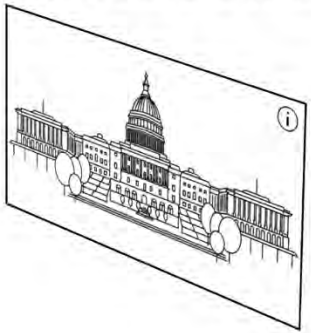
AI tool used
Adobe Firefly
AI tool version
Firefly Image 2 Model (2024)
Edited with
Adobe Photoshop 2024

Content history

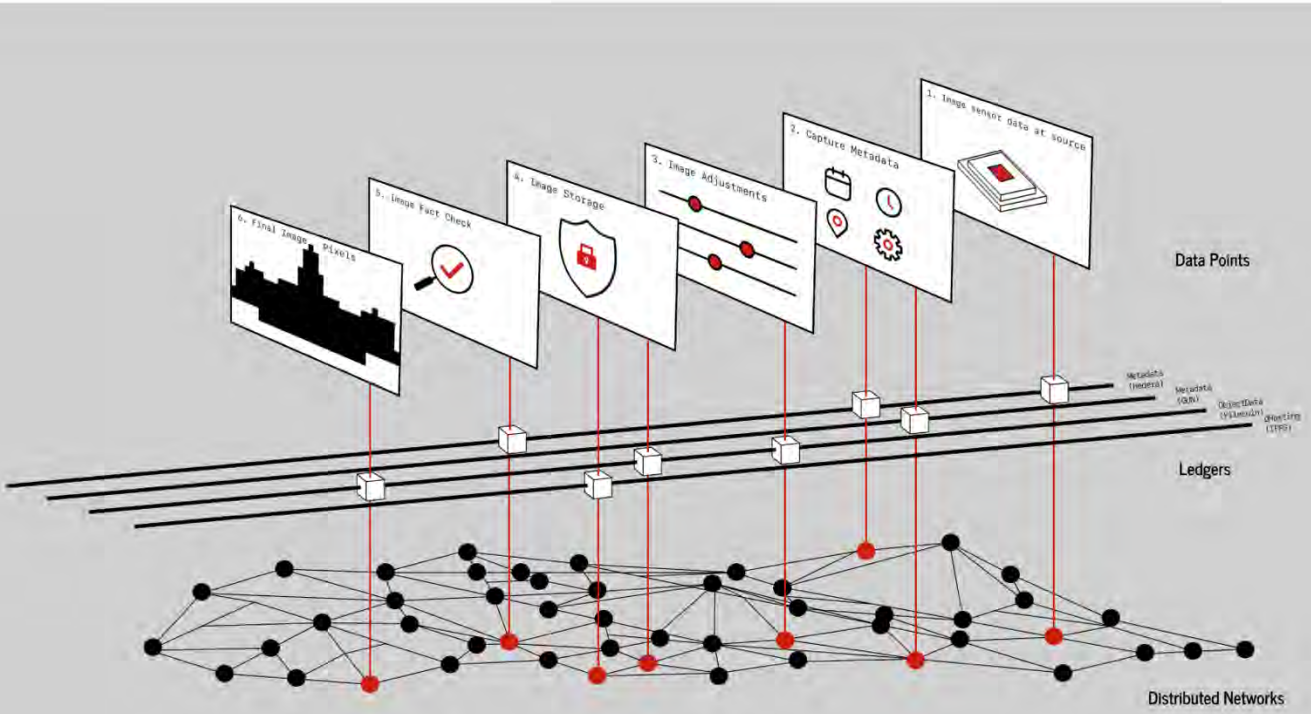
Created
May 28, 2024 at 2:45 PM
Edited
May 28, 2024 at 3:02 PM

Content Credentials are powered by Adobe and adhere to the Content Authenticity Initiative (CAI) standard. Learn more

THE PHOTO AS A DIGITAL REPOSITORY OF TRUST



- Using the new Content Authenticity Initiative's attribution standard, each photo becomes a container of image pixels as well as additional data and links.
- All these claims and assertions about a photo's veracity are securely embedded directly within the photo itself.
- Each travel with the photo wherever it is published on the web, and soon on social media.



78 days The Archive – Technology, Starling Lab 2021

https://www.starlinglab.org/image-authentication/

INTERPARES TRUST AI · CU08

Key Finding: Significant Correspondence

Archival requirements

InterPARES 2 Guidelines

Identity

- Persons and technologies involved in creation
- Dates
- Documentary form
- Intellectual property
- Source materials
- Versions

Integrity

- Subsequent actions
- Technical changes
- Responsible actors
- Mechanisms for identification and verification
- Preservation over time

Alignment

Correspondence in C2PA

C2PA specification (assertions)

Assertions supporting identity metadata

- Creator and organization
- Software and hardware used
- Creation date and time
- Content type and format
- Source materials (e.g., references)
- Versions and relationships

Assertions supporting integrity metadata

- Recorded actions (e.g., edits, AI generation)
- Technical changes and parameters
- Actors (human and computational)
- Cryptographic signatures
- Mechanisms for verification
- Linked manifests and asset history

Example: Content Credentials for an AI-generated image



AI-generated image
(example from CU08 analysis)



Identity information

- Creator and organization
- AI tool and version
- Dates and times
- Content type
- Source materials
- Versions

Integrity information

- Recorded actions
- Technical changes
- Actors (human and AI)
- Cryptographic signatures
- Verification mechanisms
- Linked history

Key finding

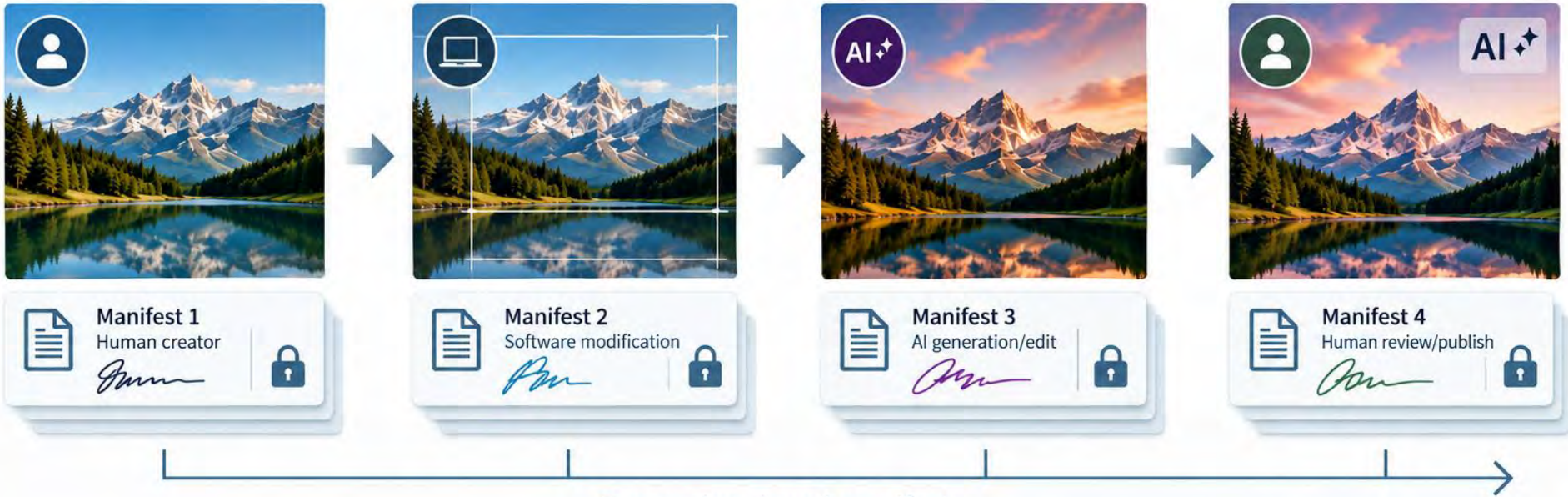
The analysis revealed significant correspondence between archival metadata requirements and C2PA assertions, including the potential to document computational actors and actions involved in AI-generated images.

Important limitation

Correspondence does not demonstrate that the information will necessarily be accurate, persist across systems or remain available for preservation.

INTERPARES TRUST AI · CU08

Key Finding: Documenting Change Over Time



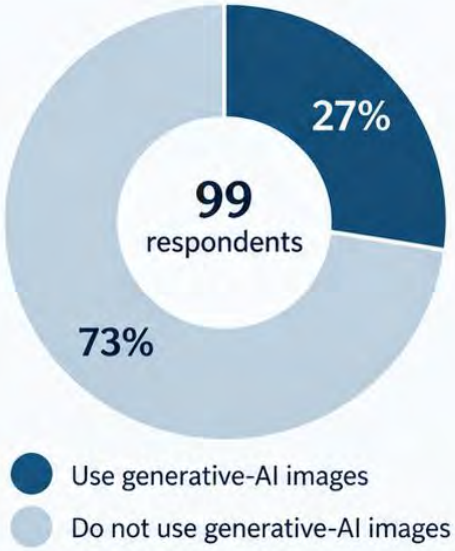
Key finding

A cryptographic provenance framework with multiple signed manifests can document change over time, supporting complex creation and modification workflows.

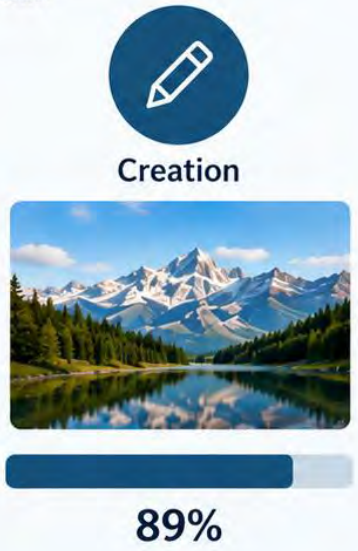
INTERPARES TRUST AI · CU08

Key Finding: Limited Adoption and Multiple Barriers

Use of generative-AI images



Generative-AI image use significant at creation



Reasons for non-use

51%
Attribution and copyright concerns

35%
Unfamiliarity with tools

33%
Professional field does not permit use

Key finding

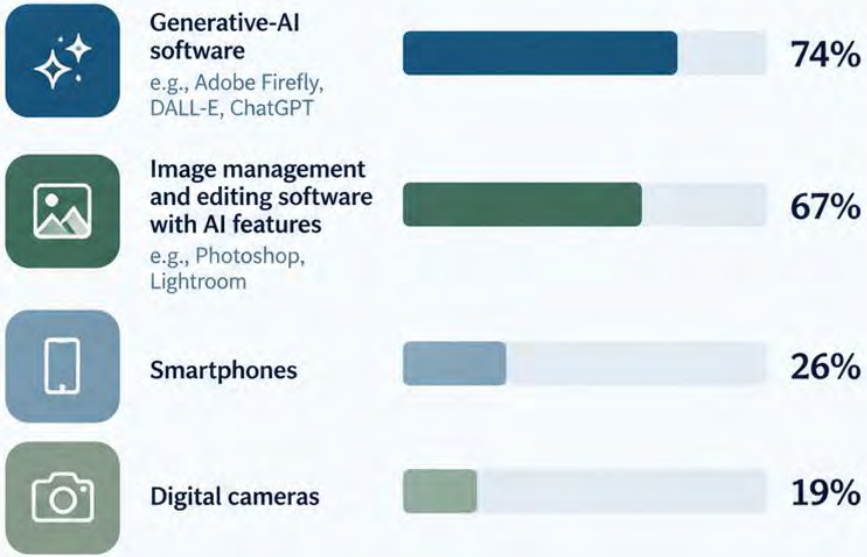
Generative-AI image practices remain limited among professional image creators and preservers, shaped by legal, ethical and institutional conditions.

INTERPARES TRUST AI · CU08

Key Finding: AI Use Concentrated in Creation and Post-Production

Tools and technologies used by AI-active respondents

(n = 27)



Generative-AI tool use across the image lifecycle

Mean reported frequency (n = 27)



Key finding | AI use among image professionals is concentrated in image creation and post-production, with lower use in access, dissemination and storage.

INTERPARES TRUST AI • CU08

Key Finding: High Expectations for Identification of AI-gen Images

High expectations for identification and provenance

% of respondents (n = 99)



Important to identify images created or modified using generative AI



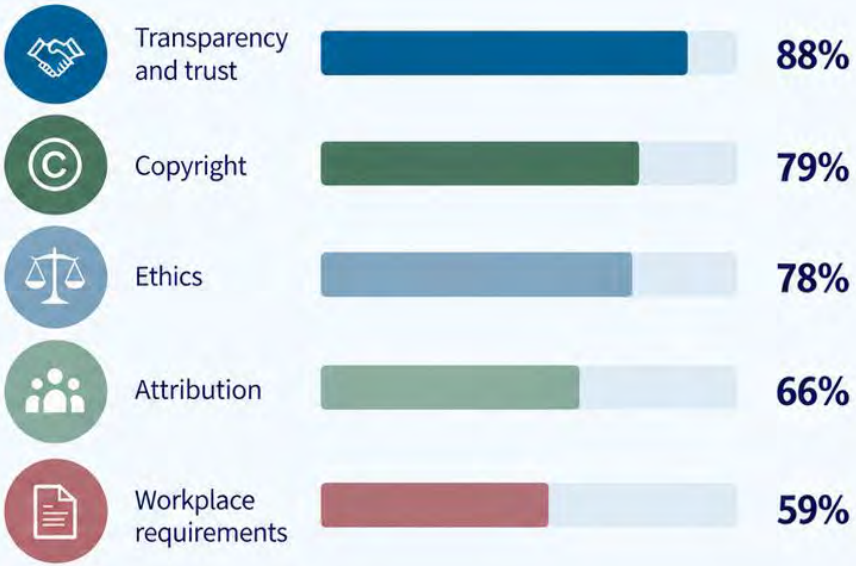
Very important that online images involving AI algorithms be identified and their provenance verified



Photo archives, stock image agencies and news media services should disclose whether generative AI was involved

Motivations for identification

% of respondents (multiple responses; n = 97)



Key finding

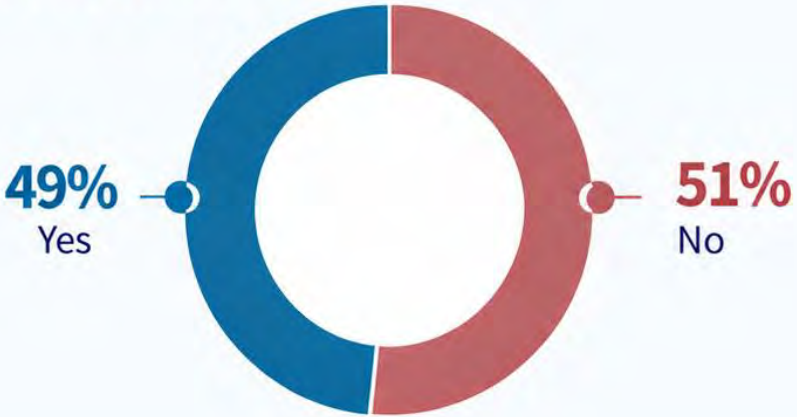
Professionals have strong expectations of Archives, news media and stock image agencies to identify and disclose AI-generated images.

INTERPARES TRUST AI • CU08

Key Finding: Standards and Policies for AI-generated Images

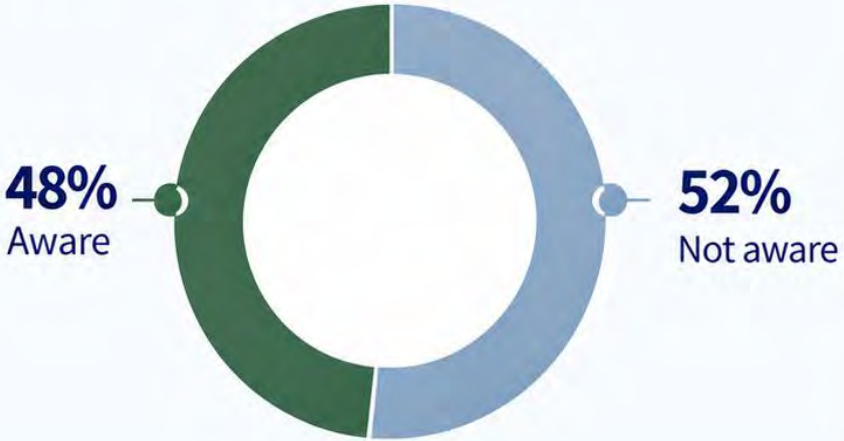
Many are not guided by standards or policies

% of respondents (n = 75)



Awareness of Content Credentials

% of respondents (n = 99)



Key finding

Adoption of standards on policies for AI-generated images and awareness of C2PA is an area for growth.

INTERPARES TRUST AI • RA01

Key Findings: Literature and Initiative Review

RA01 examined how AI technologies can support the management, description and discovery of existing archival image collections through computer vision and machine learning.

Literature Review

- Initiated in 2023 and expanded in 2024
- 29 sources selected for detailed analysis
- Focused on computer vision and machine learning applications

Predominant Approach

Computer vision, including:

- Object detection
- Image classification
- Image-similarity analysis



Example of computer vision annotation in a historical photograph (illustrative).

Key Potential

These technologies offer possibilities for analysing large photographic collections, particularly where item-level description is limited or inconsistent.

Key Concerns

- Training-data bias
- Facial recognition
- Domain expertise
- Need for interdisciplinary collaboration

Research Stage

- Much of the work remained dependent on grant-funded research projects and prototypes, leaving questions about feasibility for routine archival practice.

Overall Implication

Although RA01 did not progress to an applied case study, the literature-based findings contribute another dimension to our investigation: AI is not only involved in creating images but is increasingly involved in producing information about archival images.

INTERPARES TRUST AI

Emerging Initiatives and Next Steps

Library of Congress

Content authenticity and provenance for libraries, archives and museums.

- Initiating work on content authenticity and provenance.
- Community of practice exploring the implications of C2PA for cultural heritage institutions.



Murray, J., & Sternfeld, L. (2026). *Content Authenticity and Provenance in the Age of Artificial Intelligence: A Call-to-Action for the LAMs Community.*

Trust in Archives

Addressing the relationship between AI technologies and archival collections.

- A joint initiative of the Association of Moving Image Archivists and the Archival Producers Alliance.
- Explores AI-related issues, including the use and licensing of archival materials.

Trust in Archives | ASSOCIATION OF MOVING IMAGE ARCHIVISTS + ARCHIVAL PRODUCERS ALLIANCE

https://www.trustarchives.org/

Copyright and future research

Copyright emerged as an important concern among image professionals.

- Beginning an annotated bibliography on the legal and recordkeeping implications of AI-generated images.
- Further research will require a separate project and potentially additional funding.



Key point

These initiatives indicate growing engagement with the challenges identified in this research and provide opportunities for archivists to contribute professional knowledge and requirements.