ACM Multimedia 2026
Rio de Janeiro, Brazil

Foundations and Frontiers of Watermarking

A unified end-to-end treatment of watermarking



Dr. Vishal Asnani

Adobe Research



Dr. Shruti Agarwal

Adobe Research



Prof. Benedetta Tondi

University of Siena



Dr. Pierre Fernandez

Meta AI (FAIR)



Prof. Furong Huang

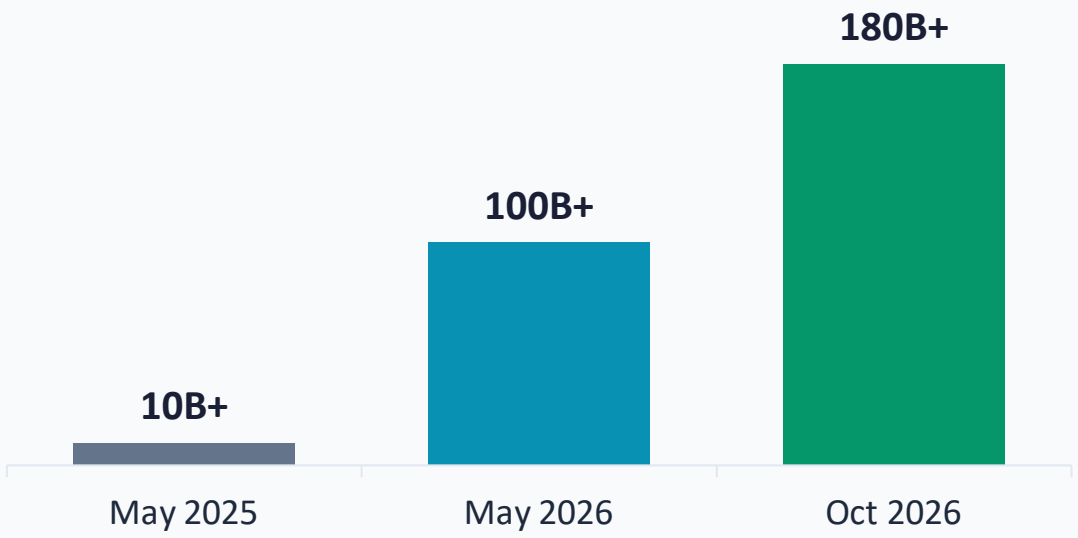
University of Maryland

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Watermarking and provenance at internet scale

Where the signals stand today (October 2026)

Images and videos watermarked with SynthID (billions)



180B+
images and videos carry a SynthID watermark, up from 10B+ in May 2025

240,000
years of audio watermarked with SynthID

1M+
SynthID checks every day in Search, the Gemini app and Chrome

Sources: Google DeepMind (Oct 2025 and Oct 2026 posts; May 2026 figure as reported at Google I/O); C2PA Deployment Guidance 1.0 (Jul 2026); CAI (Jan 2026).

Adobe Meta Università di Siena 1240 University of Maryland

Who is on board with provenance

Labeling, watermarking and Content Credentials programs in production



image source: chatgpt.com

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Who is on board with provenance

Labeling, watermarking and Content Credentials programs in production



Adobe
Content Credentials, TrustMark

Google
SynthID, Content Credentials

Meta
Meta Seal

OpenAI
C2PA, SynthID, textGrain

TikTok
Watermarking program

Apple
SynthID partner, coming soon

image source: chatgpt.com

Watermarking for GenAI Labeling

What is GenAI Labeling?

The challenge of identifying and marking content that has been generated by AI systems. As generative models become ubiquitous, distinguishing AI-generated images from real ones is critical for trust and transparency.



Why Do We Need It?

Regulatory compliance is driving adoption:

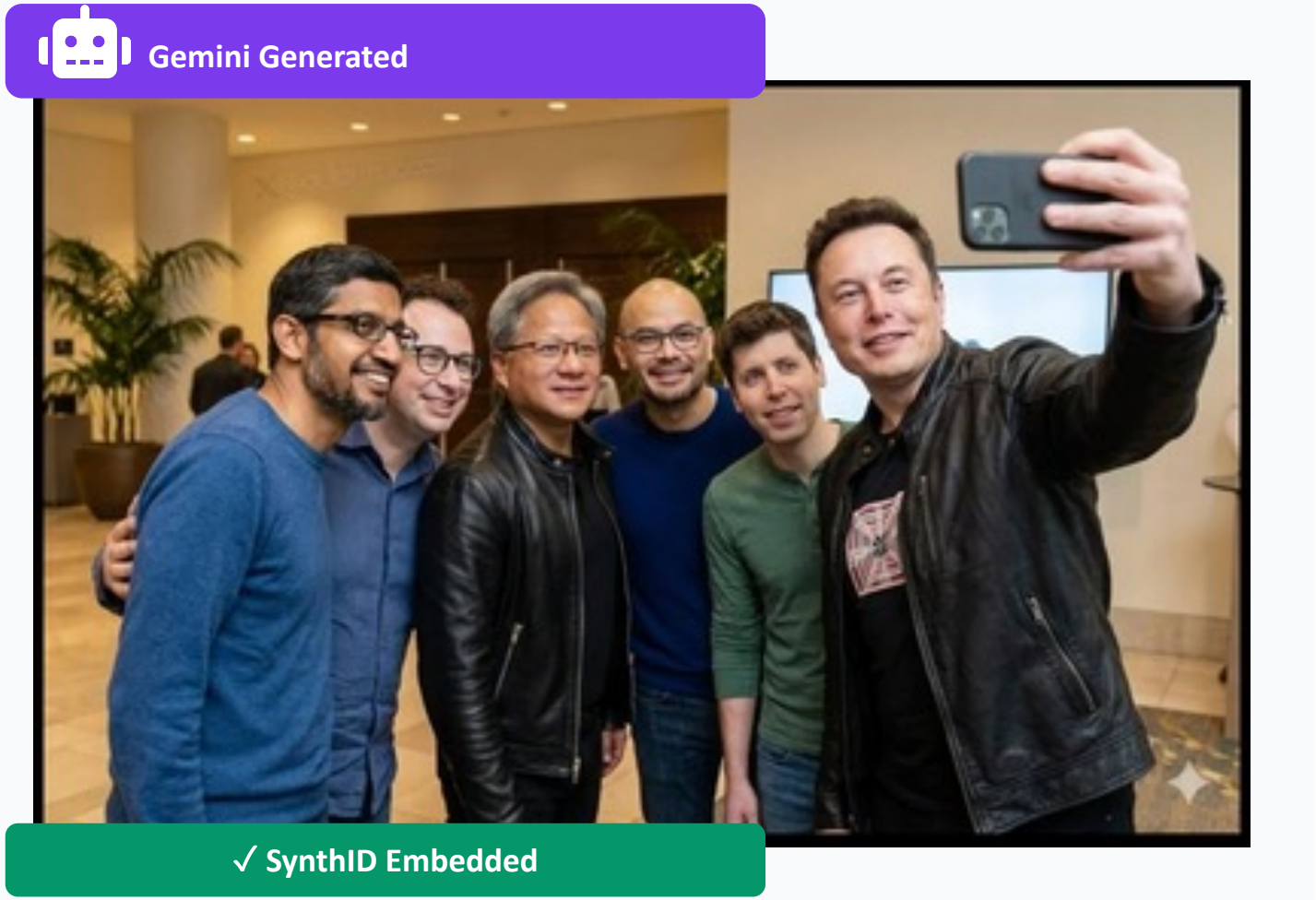
- EU AI Act
- US Executive Order on AI
- China’s AI labeling requirements

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Gemini

1. Generate
Gemini generates an image with embedded SynthID watermark

https://gemini.google.com



Gemini

2. Detect

Watermark decoded to verify AI-generated origin

AI Image Origin: Google AI/Gemini



is this image created by gemini?

Analysis & 1 more

Based on an analysis of the image, it has been determined that most or all of this content was generated or edited using Google AI.

This particular image, which features tech leaders Elon Musk, Sam Altman, Jensen Huang, Demis Hassabis, and Sundar Pichai, went viral as part of a series of AI-generated "tech CEO" photos. Online reports and social media posts from late 2025 identify it as a creation of Google's image generation models (specifically a version of Gemini).

Gemini

3. Edit/Share

Media shared on YouTube or social media platforms

Cropped



Center crop – 60% of original

JPEG Compressed



Heavy compression – quality 15%

Screenshot



Browser screengrab capture

Social Upload



Re-encoded for social platform

SynthID watermark detected in all transformations — robust to cropping, compression, screenshots, and social media re-encoding

Adobe Meta Università di Siena 1240 University of Maryland

SynthID is becoming a shared watermark

Who embeds it and who can check it (October 2026)

Google

Embeds and detects

Gemini, Imagen, Veo and Lyria: images, video, audio and text.

OpenAI

Embeds

Media outputs carry SynthID and can be checked at synthid.com. Text uses its own textGrain.

NVIDIA

Embeds

Cosmos world-model video on build.nvidia.com, a partner since 2025.

Kakao

Embeds

Generative media from Kakao's AI products.

Apple

Announced, coming soon

Named by Google as a partner. Rollout across Apple's products is still pending.

GetReal Security

Detection partner

Content-verification platform that can detect SynthID, partnered since May 2025.

Source: Google DeepMind, "We're making it easier to identify AI-generated content globally" (Oct 7, 2026, updated Oct 29) and SynthID Detector announcement (May 2025).

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Google SynthID Detector: a public verify site

Open to everyone, globally in English, since October 7, 2026

1 **Upload**
Drop in a JPG, PNG, MP4 or MP3 file.

2 **Scan**
Checks for SynthID from Google, OpenAI, NVIDIA and Kakao (Apple coming).

3 **Result**
Watermark detected, or "no confident signal".

Not a general AI detector. No signal does not mean human-made, and heavy edits can weaken the watermark. Meta and Microsoft content is not covered.

synthid.com

Scan to try it on your own files

Source: Google DeepMind (Oct 7, 2026); InfoQ (Oct 21, 2026); The Decoder (Oct 7, 2026).

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Watermarking synthetic content marks AI-generated media at the point of creation. But how far does this approach go as a transparency mechanism?

Advantages

- Standalone & Platform-Independent
- Regulatory Alignment
- Invisible & Non-Disruptive
- Scalable Deployment

Limitations

- Only Marks AI — Not Everything Else
- Vulnerable to Determined Adversaries
- No Authenticity for Real Content
- Incomplete Transparency

Watermarking is necessary but not sufficient — a broader provenance framework is needed for complete transparency.

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Beyond GenAI Labeling - Provenance

A digital nutrition label for everything you see.

Content Credentials

Content Credentials securely attach verifiable provenance information to content (photos, videos, audio, generative AI), proving the origins of media, how it is created, composited, and edited. This information travels with the content and can be displayed on apps and websites to be verified by anyone.

Answers: How? + Who? = Trust

- 54 products from 31 companies passed C2PA conformance (May 2026)
- 6,000+ members in the Content Authenticity Initiative (Jan 2026)

Full Transparency

Tracks not just AI-generated content but all content — camera captures, edits, composites, and re-distributions.

Verifiable Origin

Cryptographically signed provenance that proves who created the content and how, resistant to tampering.

Travels with Content

Provenance metadata is durably bound via watermarking so credentials persist even when metadata is stripped.

Adobe Meta Università di Siena 1240 University of Maryland

ChatGPT

create the same image but for a white background ppt. This image is looking very sharp contrast on a hite background image



Adobe FireFly

Google Gemini

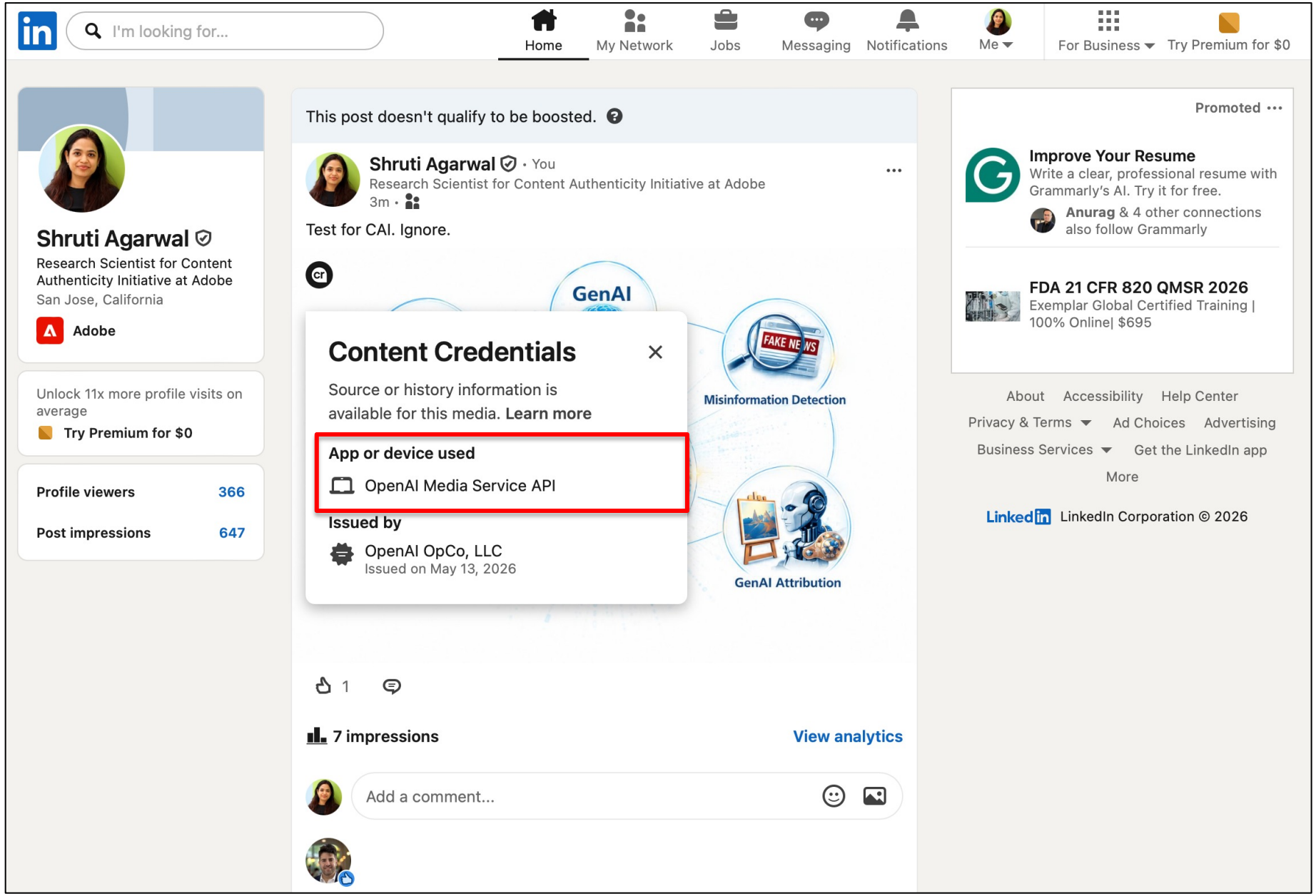
Adobe

Meta

UNIVERSITÀ DI SIENA 1240

UNIVERSITY OF MARYLAND

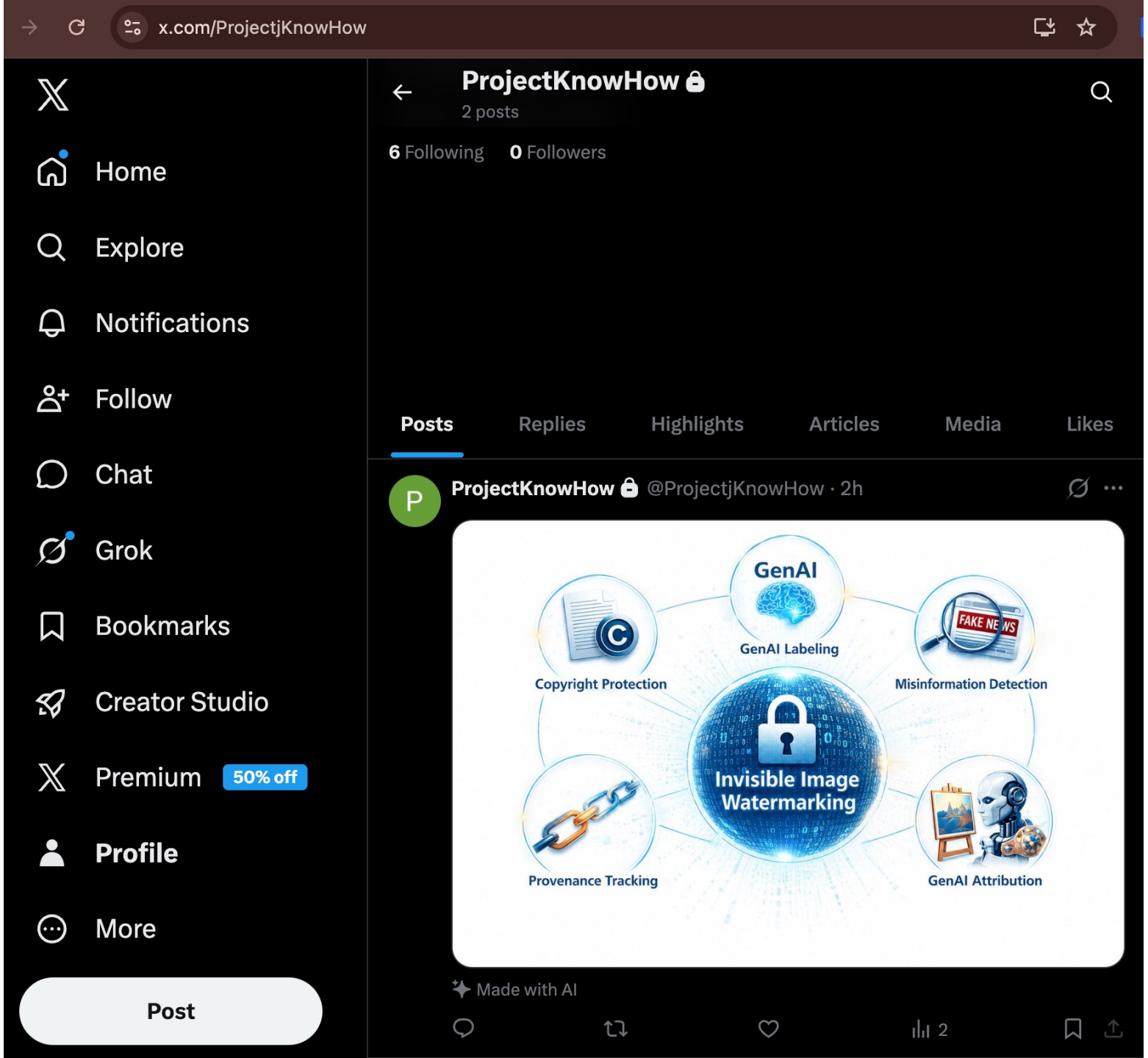
LinkedIn


I'm looking for...
Home
My Network
Jobs
Messaging
Notifications
Me
For Business
Try Premium for $0
Shruti Agarwal
Research Scientist for Content Authenticity Initiative at Adobe
San Jose, California
Adobe
Unlock 11x more profile visits on average
Try Premium for $0
Profile viewers 366
Post impressions 647
This post doesn't qualify to be boosted.
Shruti Agarwal • You
Research Scientist for Content Authenticity Initiative at Adobe
3m •
Test for CAI. Ignore.
GenAI
Content Credentials
Source or history information is available for this media. Learn more
App or device used
OpenAI Media Service API
Issued by
OpenAI OpCo, LLC
Issued on May 13, 2026
FAKE NEWS
Misinformation Detection
GenAI Attribution
1
7 impressions
View analytics
Add a comment...
Promoted
Improve Your Resume
Write a clear, professional resume with Grammarly's AI. Try it for free.
Anurag & 4 other connections also follow Grammarly
FDA 21 CFR 820 QMSR 2026
Exemplar Global Certified Training | 100% Online| $695
About Accessibility Help Center
Privacy & Terms Ad Choices Advertising
Business Services Get the LinkedIn app
More
LinkedIn Corporation © 2026

X.com

Credential Lost

Meta-data lost on X.com



OpenAI

Log in

Try ChatGPT ↗



Generated with OpenAI tools

This content was generated using OpenAI tools. The signals below provide supporting evidence for this result.

✓ **SynthID detected**

We found a SynthID watermark that originated from OpenAI.

✕ **Content Credentials not detected**

We did not find a trusted C2PA manifest originating from OpenAI.

Hide details

Verify another image

https://openai.com/research/verify/

Three Pillars of Provenance

Metadata — Secure provenance data accompanying image (C2PA manifest)

Fingerprint — Robust, tamper-sensitive matching of pixels to recover manifest

Watermark — Robust identifier in pixels matching to recover manifest

	Legacy stripping	Attacker spoofing	Deterministic lookup	Hard Binding	Client-side detection	Standalone (no network)
Metadata	✗	✗	✓	✓	✓	✓
Fingerprint	✓	✓	✗	✗	✗	✗
Watermark	✓	✗	✓	✗	✓	✗

"To Authenticity and Beyond: Building Safe and Fair Generative AI on the Three Pillars of Provenance". Collomosse & Parsons. IEEE CG&A. May 2024.

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

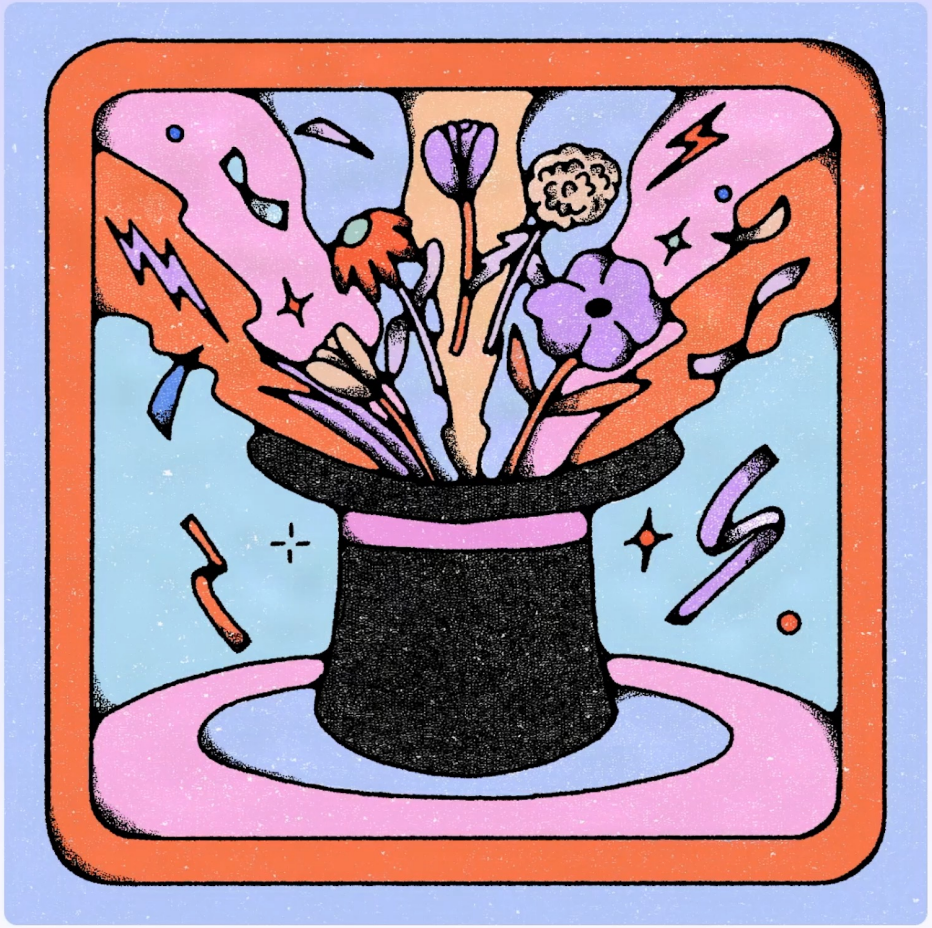
contentauthenticity.adobe.com

Adobe Content Authenticity Beta Preferences Apply Manage Inspect Help Join Discord Add Chrome extension

Sign your work with Content Credentials

Content Credentials help you protect your content before sharing it online. Connect your verified name, social media accounts, and more to your work, anywhere it's published.

Get started

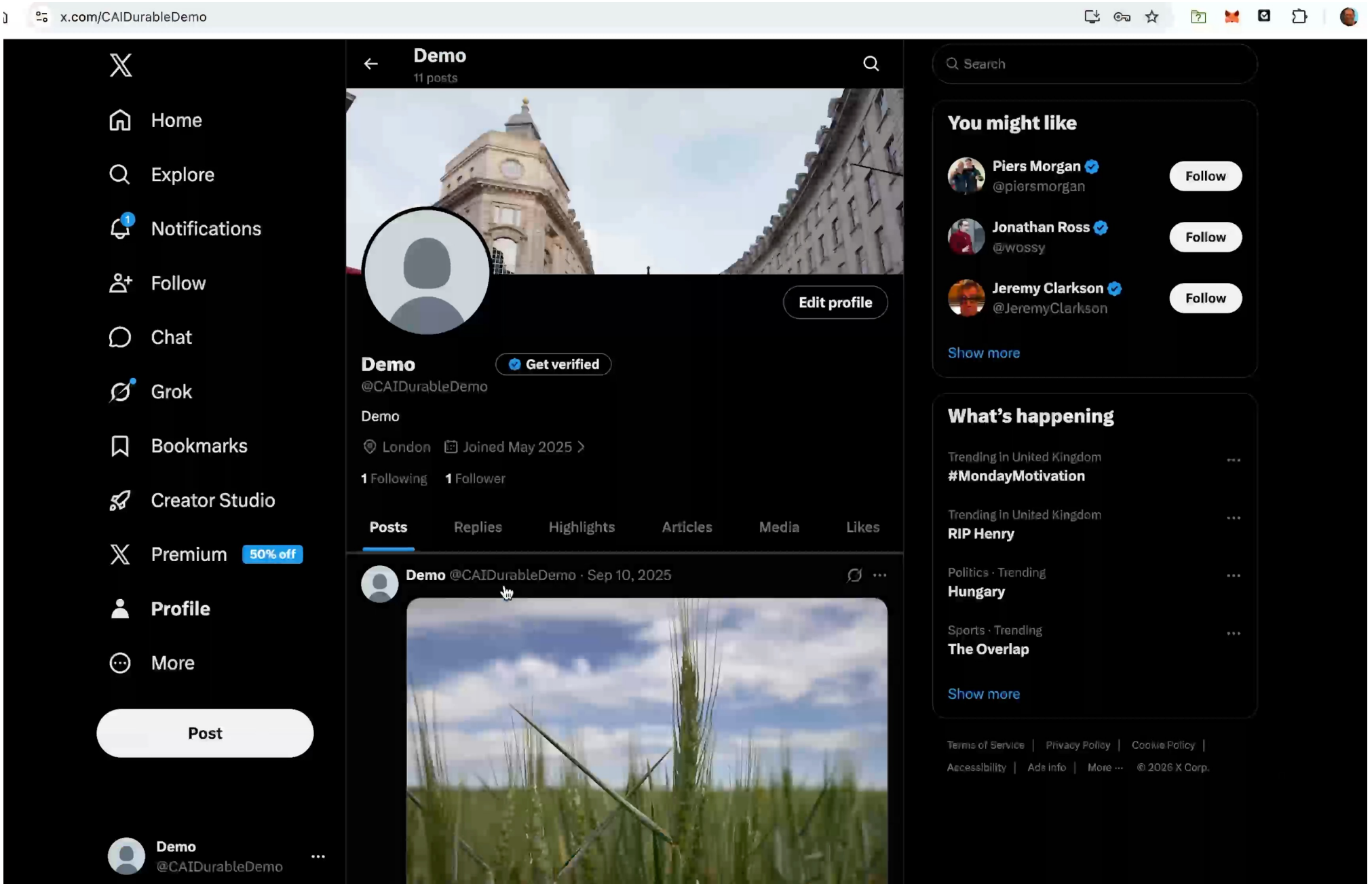


This is how your Content Credentials may appear to others.

https://contentauthenticity.adobe.com

Adobe Meta Università di Siena 1240 University of Maryland

https://contentauthenticity.adobe.com/inspect



Under the hood: signing with all three pillars

Public SDKs only: c2pa-python for the manifest, TrustMark for the watermark

SIGN

Watermark

Embed an ID

TrustMark writes a random ID into the pixels.

Fingerprint

Hash the pixels

A 64-bit perceptual hash of the saved image.

Metadata

Sign the manifest

c2pa-python signs it with the ID as a soft binding.

Lookup service: watermark ID → signed manifest + fingerprint

VERIFY A COPY THAT LOST ITS METADATA

Watermark

Decode the ID

TrustMark reads the ID from the pixels.

Metadata

Recover the manifest

Look the ID up in the registry.

Fingerprint

Check the match

Close hash: accept. Far apart: reject a copied watermark.

Embed the watermark before signing: the manifest hard-binds the exact file bytes.

Adobe Meta Università di Siena 1240 University of Maryland

Demo: sign and verify with c2pa-python

Runs in Google Colab. Scan the QR code to open the notebook.

```
# 1. Watermark: embed a random ID
wm_id = random_bits(tm.schemaCapacity())
tm.encode(img, wm_id, MODE="binary").save(p)
# 2. Fingerprint: 64-bit perceptual hash
fp = phash(Image.open(p))
# 3. Metadata: sign manifest holding the ID
with Builder(manifest_for(wm_id), ctx) as b:
    b.sign_file(p, "signed.jpg", signer)
# Lookup service: ID -> manifest + fingerprint
registry[wm_id] = {"manifest": m, "fp": fp}
```

Case	Manifest	Watermark	Fingerprint	Result
Signed original	Yes	Yes	Match	**Verified**
Metadata stripped, JPEG q60	No	Yes	Match	**Verified**
Resized to 75%, JPEG q70	No	Yes	Match	**Verified**
Watermark copied to another image	No	Yes	Mismatch	**Rejected**

Expected results of the four steps in the notebook.

Run it

pip install c2pa-python trustmark

Open in Google Colab

github.com/agarwalShruti15/three-pillar-colab

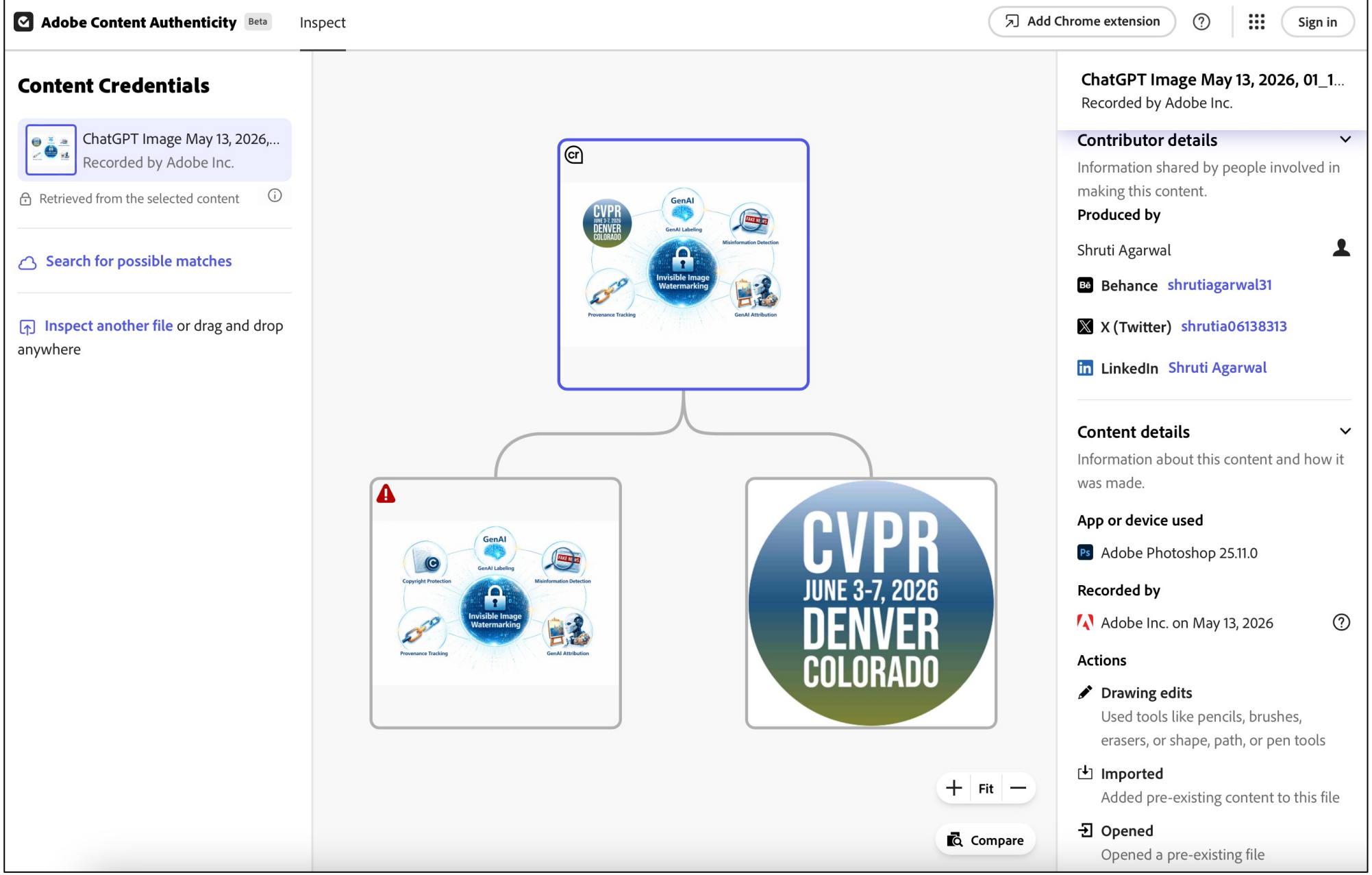
Adobe Meta Università di Siena 1240 University of Maryland

Durability - after edit

Edit

Edited the image using Photoshop



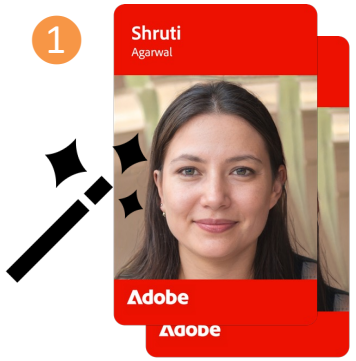


https://contentauthenticity.adobe.com/inspect

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Durability - after print

1



Digital ID creation

2



Apply Content Credentials to images (Digital IDs)

3



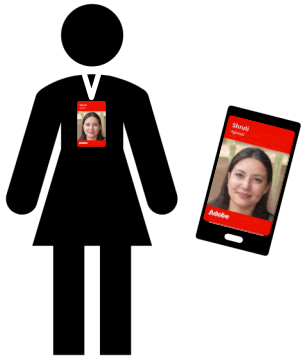
Images (Digital IDs) with Content Credentials and watermark applied

4



Print physical ID cards

5



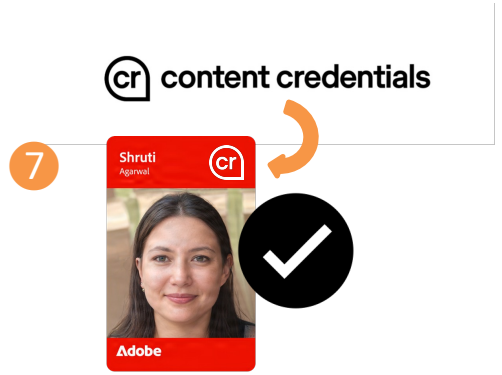
Resident captures a physical ID card using mobile app or camera

6



Captured image is verified for authenticity

7



Content Credentials of the captured image are recovered using the watermark for authenticity verification

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Try it yourself

Scan with your phone camera. Everything here is free to use.

SynthID Detector

Check an image, video or audio file for SynthID.

synthid.com

Adobe Inspect

See Content Credentials, including ones recovered from the watermark.

contentauthenticity.adobe.com/inspect

TrustMark

Embed and decode an image watermark. pip install trustmark

github.com/adobe/trustmark

TextSeal Colab

Watermark and detect text in your browser. No API key.

github.com/facebookresearch/textseal

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Resources

Google SynthID

- https://gemini.google.com
- https://ai.google.dev/responsible/docs/safeguards/synthid
- Dathathri, Sumanth, Abigail See, Sumedh Ghaisas, Po-Sen Huang, Rob McAdam, Johannes Welbl, Vandana Bachani et al. "Scalable watermarking for identifying large language model outputs." Nature 634, no. 8035 (2024): 818-823.

OpenAI Verify: https://openai.com/research/verify/

Adobe Content Authenticity

- https://contentauthenticity.adobe.com

Durable Credentials

- J. Collomosse & A. Parsons, "To Authenticity and Beyond: Building Safe and Fair Generative AI on the Three Pillars of Provenance". IEEE CG&A. May 2024.
- T. Bui, S. Agarwal, and J. Collomosse. "Trustmark: Robust watermarking and watermark removal for arbitrary resolution images." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2025.
- pip install trustmark

Interoperable Watermarking

- A. Petrov, S. Agarwal, P. Torr,, A. Bibi, & J. Collomosse. “On the coexistence and ensembling of watermarks.” *Advances in Neural Information Processing Systems*, *38*, 107058-107094, 2026.

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND

Resources (continued)

New since June 2026

Google SynthID Detector

https://synthid.com

Google DeepMind, "We're making it easier to identify AI-generated content globally", Oct 7, 2026:

https://blog.google/innovation-and-ai/models-and-research/google-deepmind/synth-id-ai-content/

Meta TextSeal and TextSeal2

https://github.com/facebookresearch/textseal

T. Sander et al. "TextSeal: A Localized LLM Watermark for Provenance & Distillation Protection." arXiv:2605.12456, 2026.

"TextSeal2: Distortion-Free LLM Watermarking with Copula Mixing and Memoryless Races." arXiv:2610.40613, 2026.

OpenAI textGrain

OpenAI Developer Community, "OpenAI's approach to EU text provenance rules", Oct 2026:

https://community.openai.com/t/openais-approach-to-eu-text-provenance-rules/1403521

Anthropic text watermark verification

https://www.anthropic.com/news/claude-text-watermark

Content Credentials and TrustMark

https://c2pa.org

https://contentauthenticity.adobe.com/inspect

https://github.com/adobe/trustmark

Three-pillars demo

https://github.com/contentauth/c2pa-python

https://github.com/adobe/trustmark/tree/main/c2pa

Adobe Meta UNIVERSITÀ DI SIENA 1240 UNIVERSITY OF MARYLAND