CAI, C2PA, and Identity

Technical overview for ICANN 79

Eric Scouten · 06 March 2024

CAI and C2PA overview

Overview of C2PA data model

Identity “2.0” assertion

Questions: At end, please

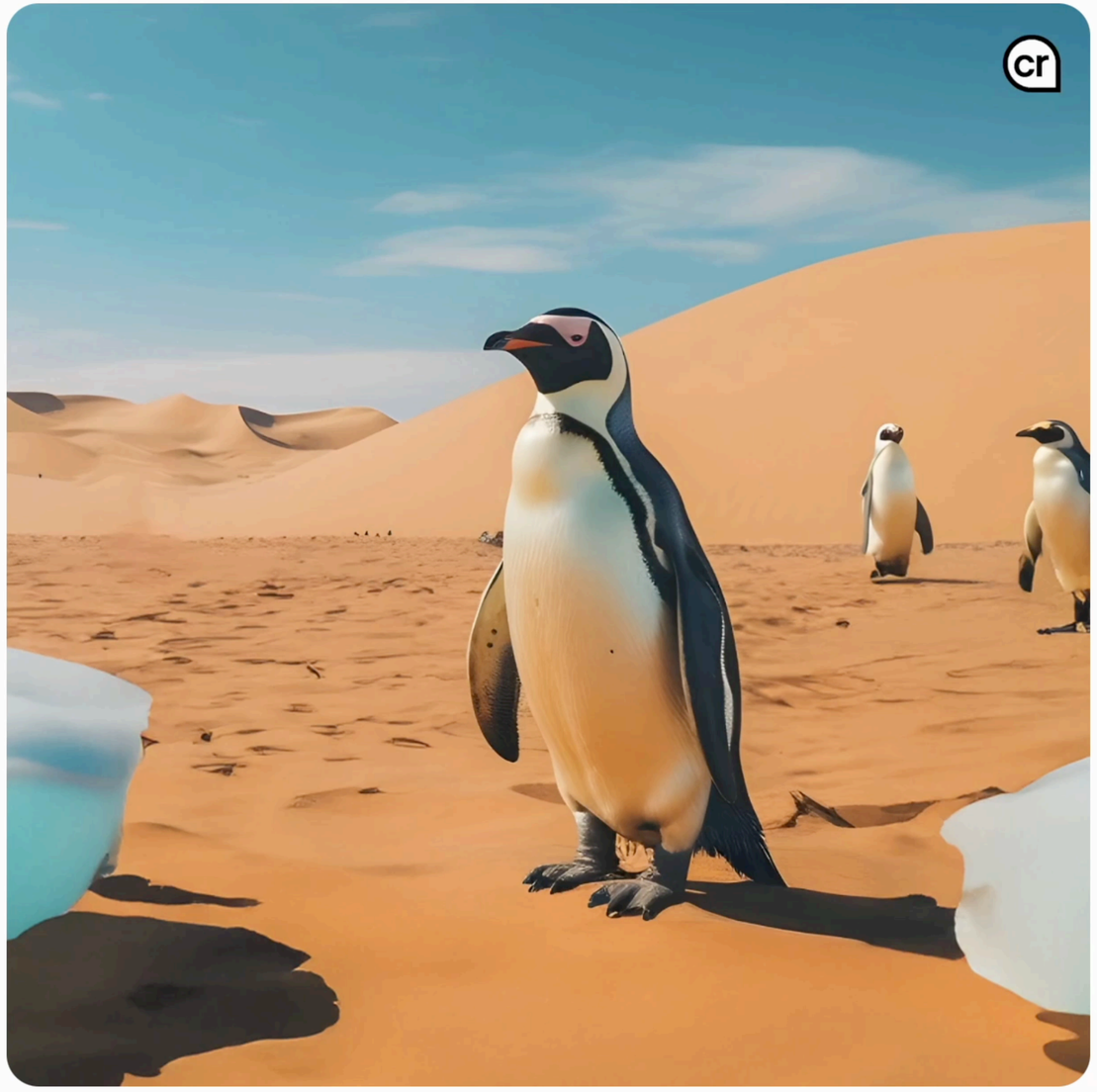
Content
Authenticity
Initiative

“We are a **community** of media and tech companies, NGOs, academics, and others working to promote adoption of an open industry standard for content authenticity and provenance.”

contentauthenticity.org

Wait, where did this image come from?

Deepfakes. Voice cloning. Synthetic media. It's hard to tell what's accurate and authentic these days.



The New York Times

Making Deepfakes Gets Cheaper and Easier Thanks to A.I.

Forbes

The Battle Against Fake News Enters The Age Of Deepfakes

CNN BUSINESS

Puffer coat Pope. Musk on a date with GM CEO. Fake AI 'news' images are fooling social media users

The Guardian

Elections in UK and US at risk from AI-driven disinformation, say experts

TIME

From Scams to Music, AI Voice Cloning Is on the Rise

CBC NEWS

With rise of AI-generated images, distinguishing real from fake is about to get a lot harder

Membership ~2500

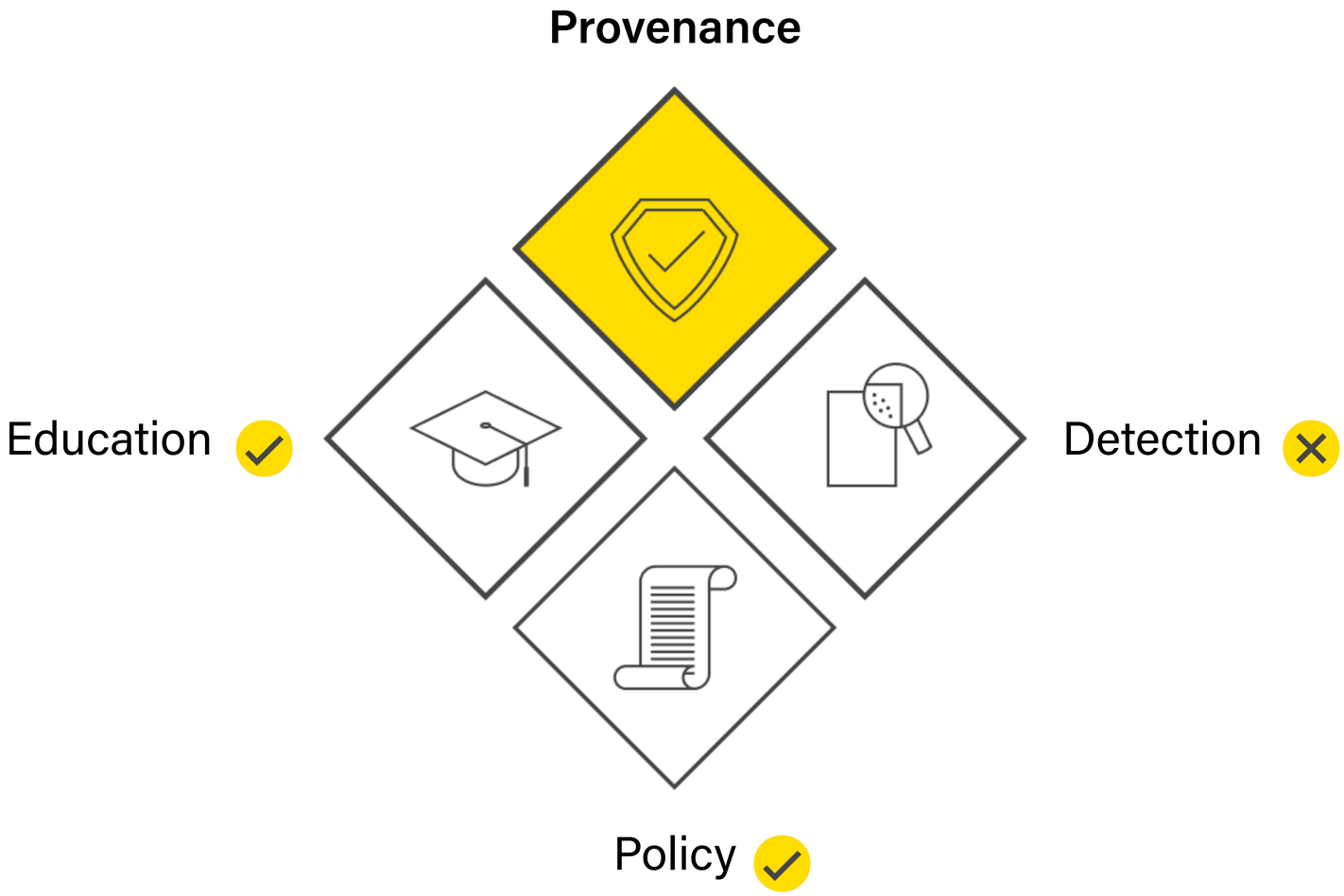
Nikon

Leica

SONY

Content
Authenticity
Initiative

Pillars of Transparency and Authenticity



Provenance

The basic trustworthy facts about the origins of a piece of digital content (image, video, audio recording, document)

Demo: How it works in Photoshop

2 Invalid video Content Credentials or assets

This sample shows how the Content Credentials player can surface important information to viewers when there are issues with a video's Content Credentials or segments within a video.

“C2PA addresses the prevalence of misleading information online through the development of **technical standards** for certifying the source and history (or provenance) of media content.”

c2pa.org

The Coalition for Content Provenance and Authenticity (C2PA)

- Standards organization
- Within Linux Foundation
- Adobe on Steering Committee & Chairs TWG
- V1.0 C2PA Specification Released on Jan 2022
- V2.0 C2PA Specification Released on Jan 2024

Meta

OpenAI

All have announced implementation

C2PA

Secure Metadata

Assertions

What was done to an image, by whom, when, etc.

Asset Hashes

Cryptographically bind assertions to image

Content Credential

Bundle of assertions, asset hashes, and digital signatures

Identity “2.0” assertion

Current work in progress

Identity “2.0” assertion

Why?

- **Ensure that identity is securely bound to content.**
  - The CreativeWork assertion in the C2PA 1.x specification did not provide this assurance and was deprecated for that reason.

- Provide new home for identity given new C2PA 2.0 core spec focus on machine-attestable content.

Identity “2.0” assertion

What?

> “This specification aims to allow an actor to document their relationship to a C2PA asset produced by them or on their behalf independently from the C2PA claim generator, and to allow consumers of a C2PA asset to independently verify that the received asset was in fact produced by the named actor and has not been tampered with.”

– Identity assertion specification “Design goals” section

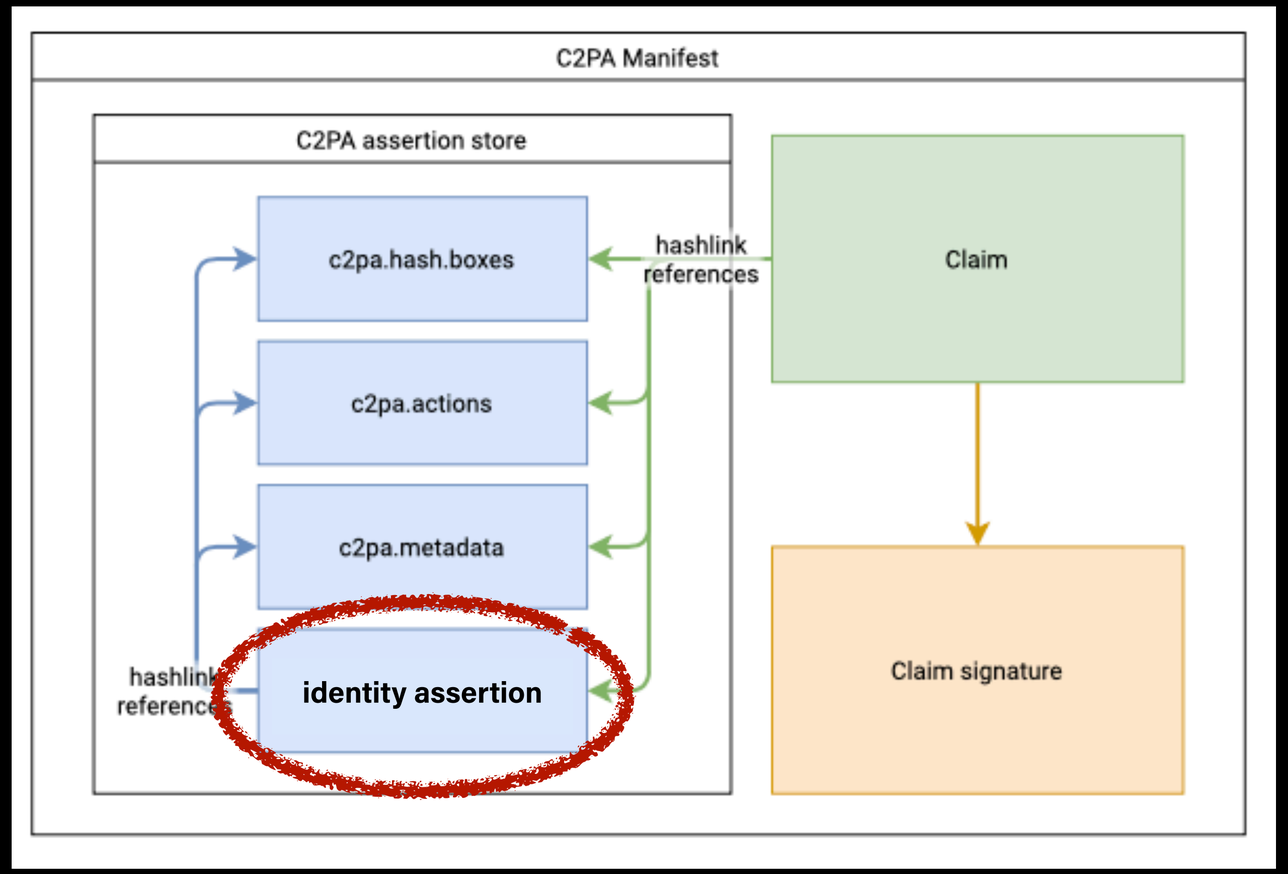
Identity “2.0” assertion

Technical overview

Identity assertion allows one credential subject to sign a **list of assertions.**

Meant to imply **subject’s authorization or active participation** in production of the asset.

Additional trust signal over and above the C2PA claim generator signal.



Identity “2.0” assertion

Introducing the Creator Assertions Working Group

- **https://creator-assertions.github.io**
- Governed under **Community Specification License** process
- Will also take up the following assertions deleted from C2PA 2.0
  - Training and data mining
  - Endorsements
  - Metadata

Identity “2.0” assertion

Please join and guide this implementation!

- **https://creator-assertions.github.io**

- Weekly meetings
  - Typically on Mondays at 0830 Pacific / 1130 Eastern / 1530 UTC
  - Contact me (scouten@adobe.com) for invitation

Questions and discussion ...