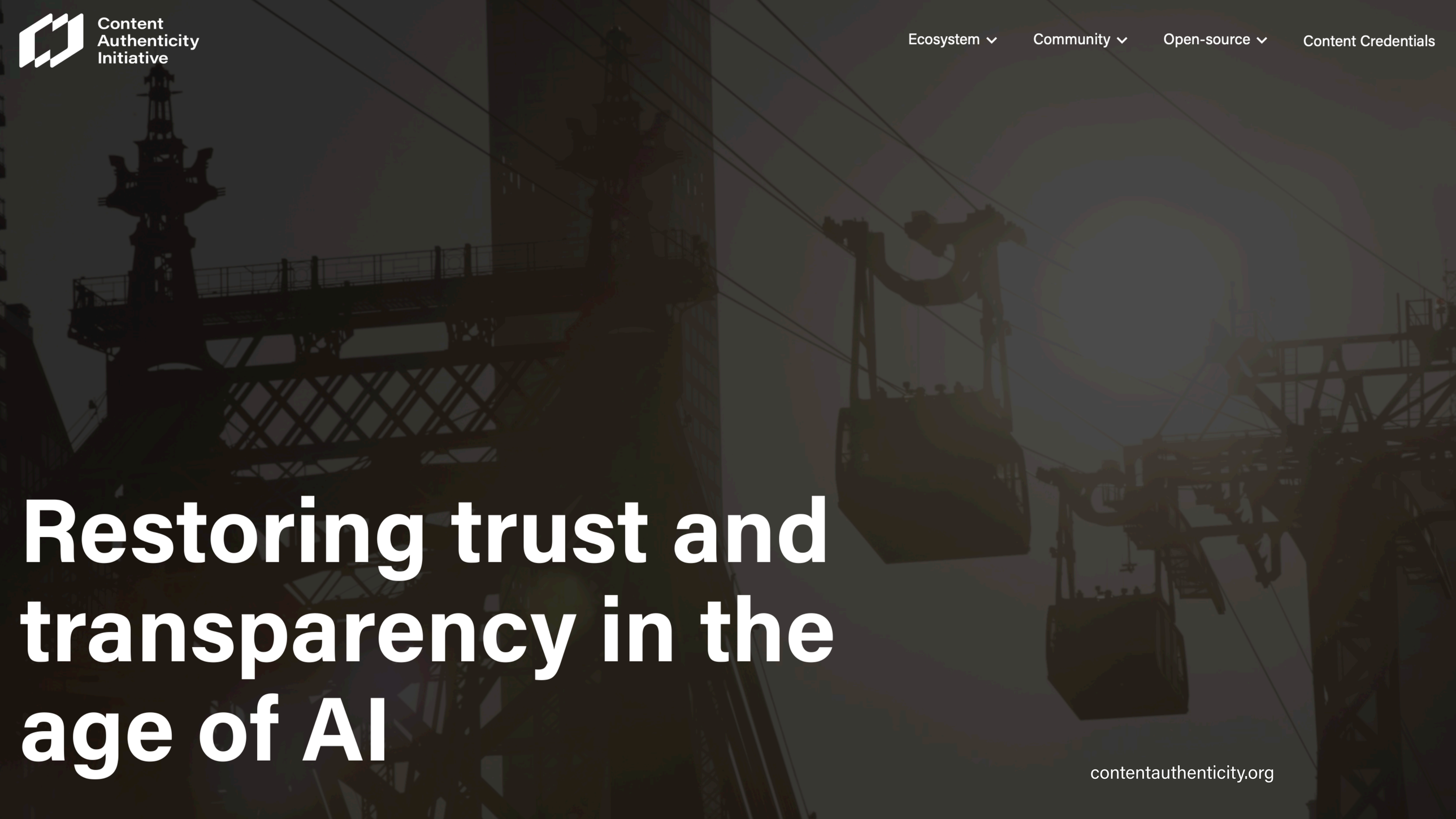
Creator Assertions Working Group

Content Authenticity 101 for Internet Identity Workshop 41

Eric Scouten · Identity Standards Architect · Adobe

21 October 2025


Content
Authenticity
Initiative
Ecosystem
Community
Open-source
Content Credentials
Restoring trust and transparency in the age of AI
contentauthenticity.org

The problem, in a nutshell

On the Internet …

digital media content

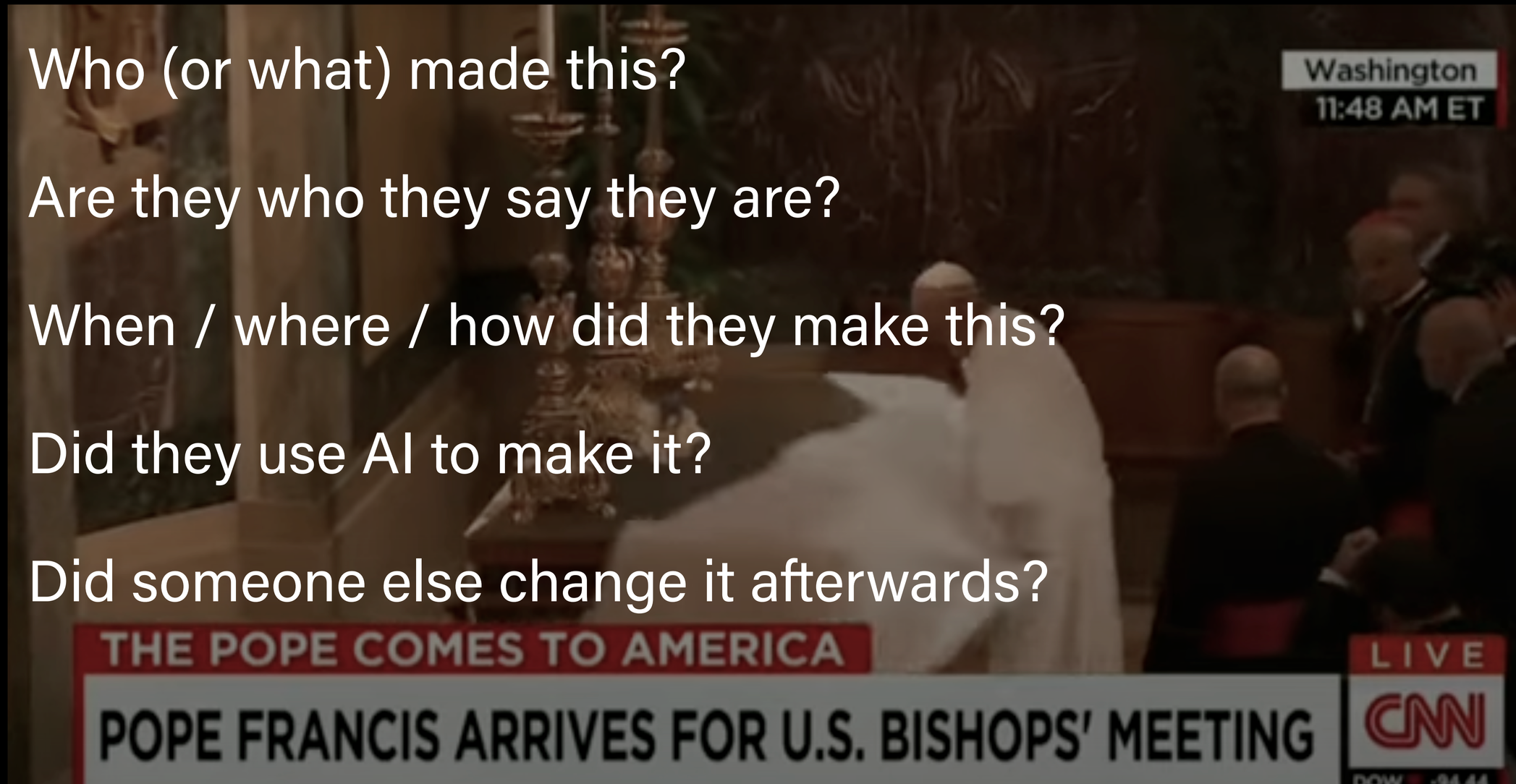
can travel from a **content creator**

to **unforeseen recipients**

via **unknown channels**.

You might ask yourself …

- Who (or what) made this?
- Are they who they say they are?
- When / where / how did they make this?
- Did they use AI to make it?
- Did someone else change it afterwards?



A tamper-evident digital “nutrition label”

We provide tools for digital **content creators** …

- Hardware and software **tool vendors** – and
- Individual and organizational **content creators**

… to describe and sign their work.

The three Cs … who does what here?

What and how
C2PA

Coalition for Content Provenance and Authenticity

c2pa.org

Who
CAWG

Creator Assertions Working Group *(part of DIF)*

cawg.io

Advocacy and education
contentauthenticity.org

Who is taking accountability?

C2PA claim generator

Hardware or software tool involved in creating the content.

CAWG named actor

Individual or organization involved in creating the content.

What are they taking responsibility for?

C2PA claim generator
can describe ...

- GPS data / time of capture (if known to hardware)
- Edit actions taken / AI used
- Ingredients incorporated into content

CAWG named actor
can describe ...

- Individuals or organizations involved in *creating* content
- Individuals or organizations *depicted* in content
- Metadata / context for content

Two responsible parties, two signatures

C2PA claim generator

- X.509 certificate / COSE signature
- **NEW** (July 2025): Certificates have C2PA-specific key usage, not interoperable with other purposes
- Issued to hardware or software that demonstrates compliance with C2PA rules

CAWG named actor

- **Flexible** *framework* for using multiple kinds of digital credentials
- Intended to bind credential to content
- Optional – for those that wish to identify themselves as content creator

C2PA data model

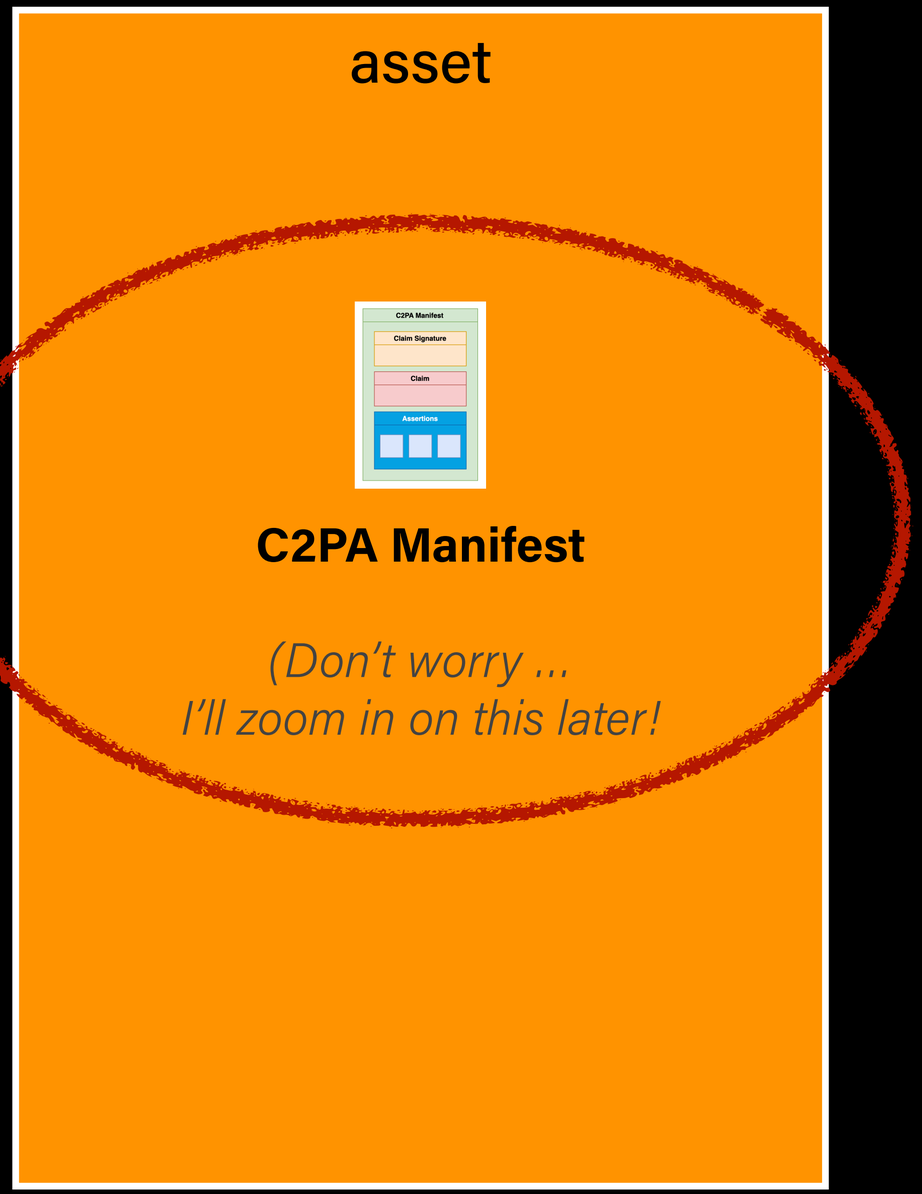
C2PA data model

Overview

An **asset** is any piece of digital media that we wish to describe.

Currently, we support still images, motion pictures, recorded, audio, documents (PDF), fonts, and more.

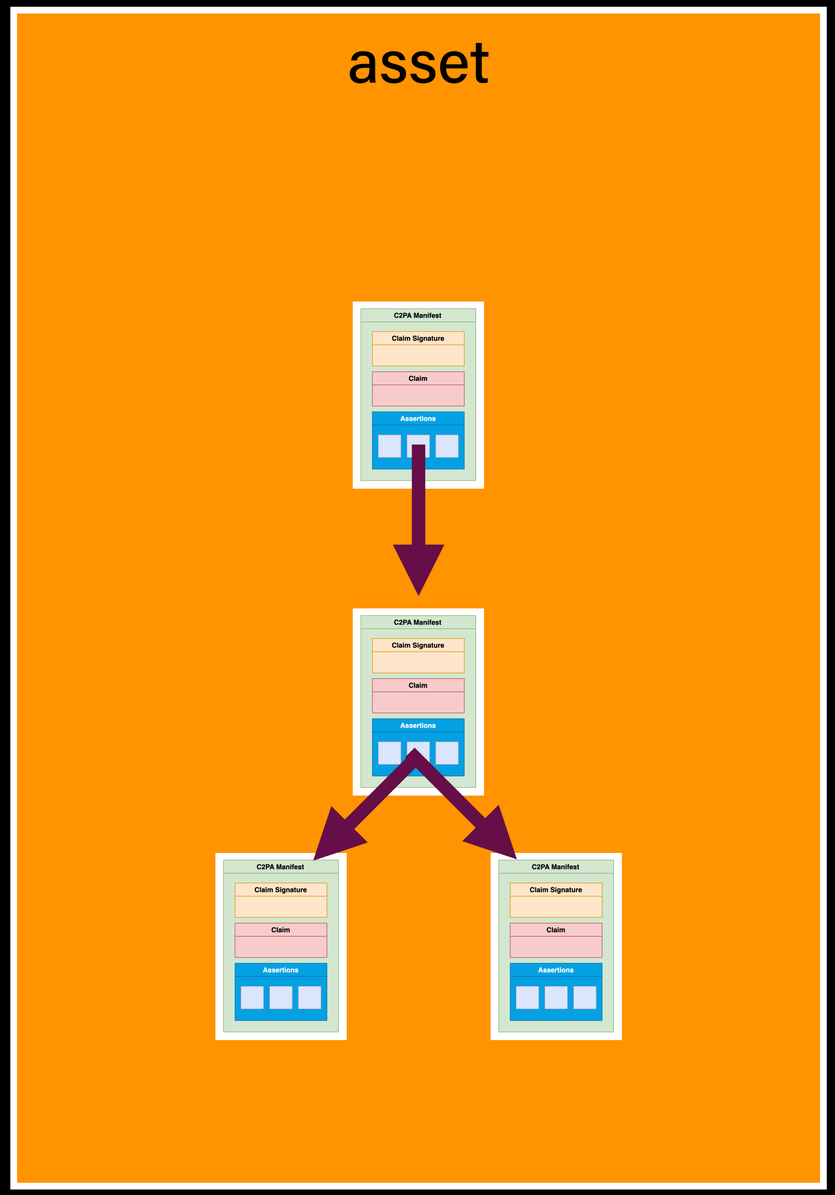
An asset is described by a **C2PA Manifest**.



C2PA data model

Overview

A C2PA Manifest can refer to any number of *ingredient manifests* when earlier content is incorporated and composed into a new asset.



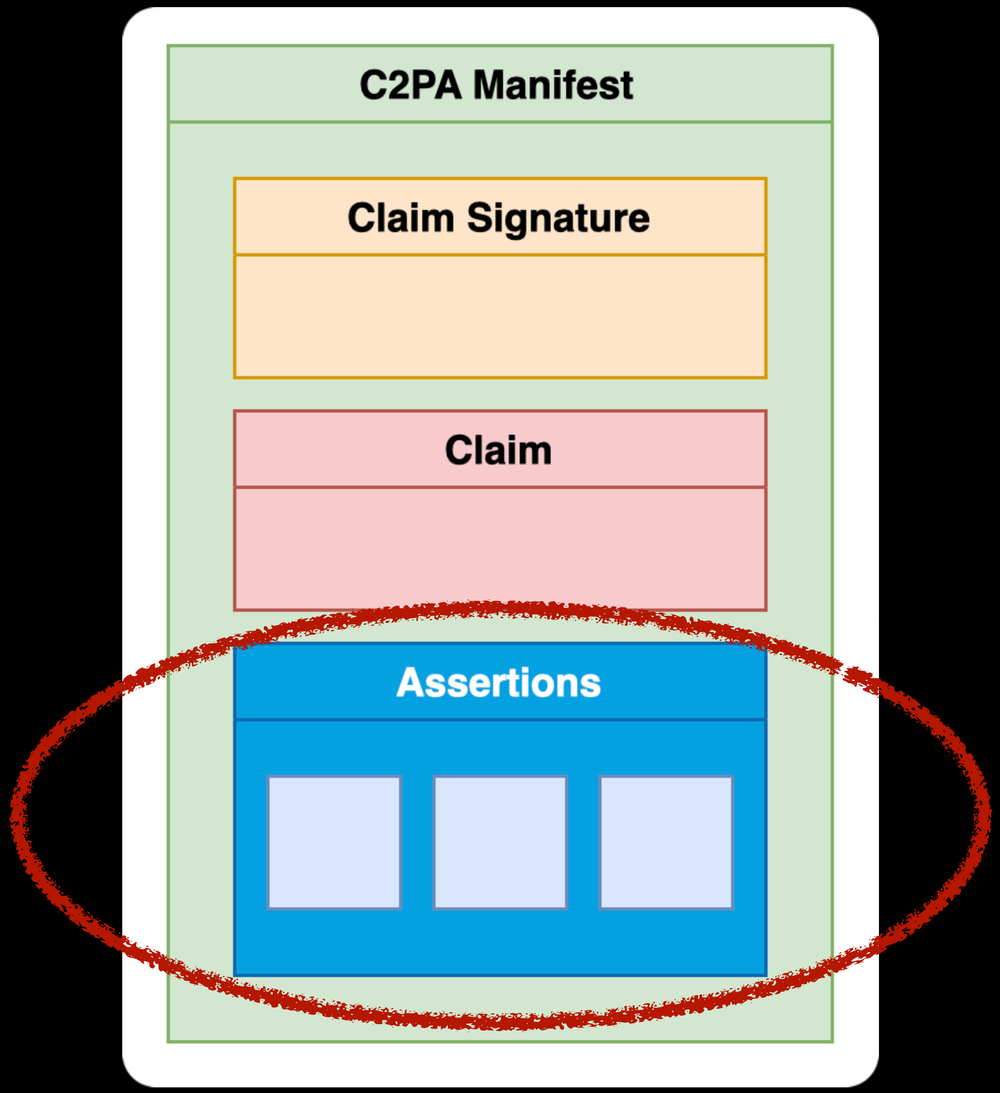
C2PA data model

Assertions

Assertions are opt-in statements that cover areas such as:

- hard binding to asset's binary content
- capture device details
- edit actions
- thumbnail of the content
- other content (ingredients) that were incorporated into this content

This mechanism is **extensible.**

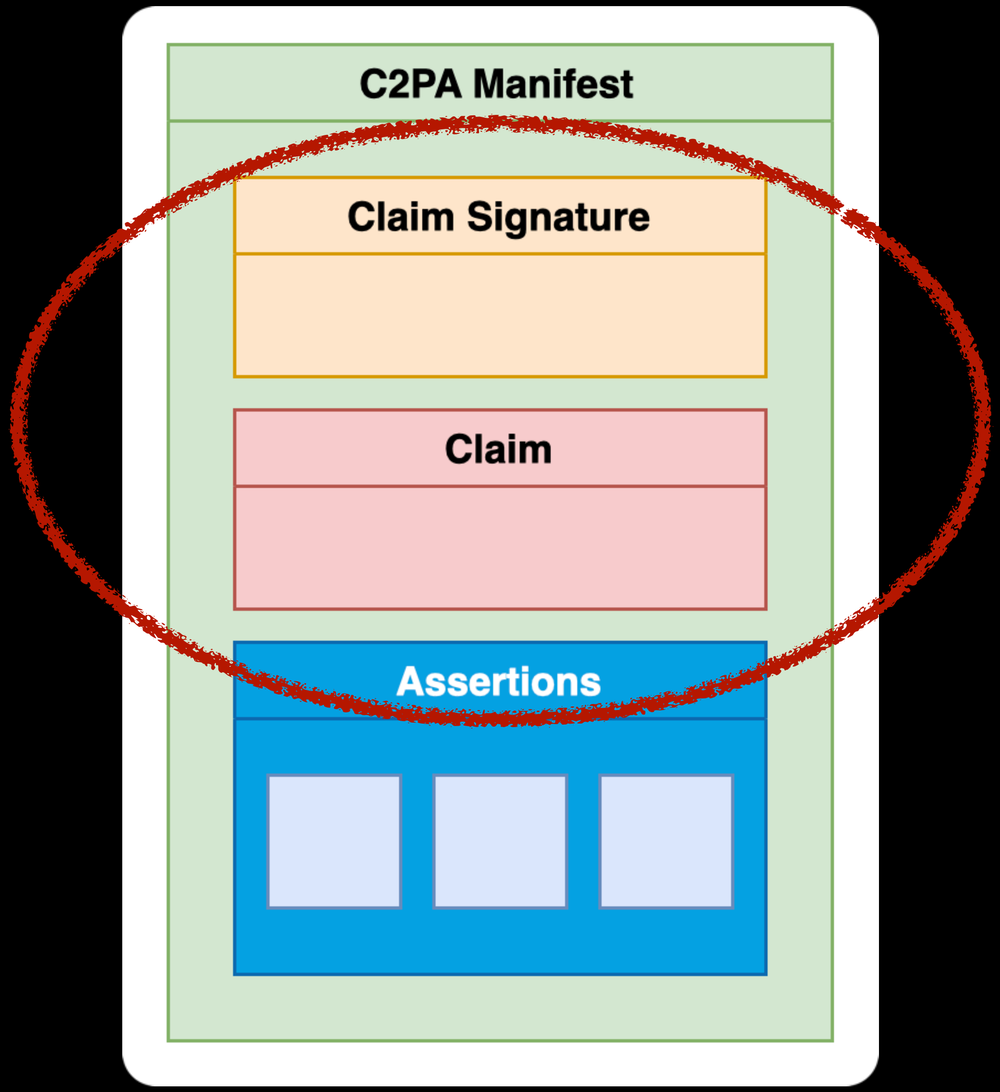


C2PA data model

Claim

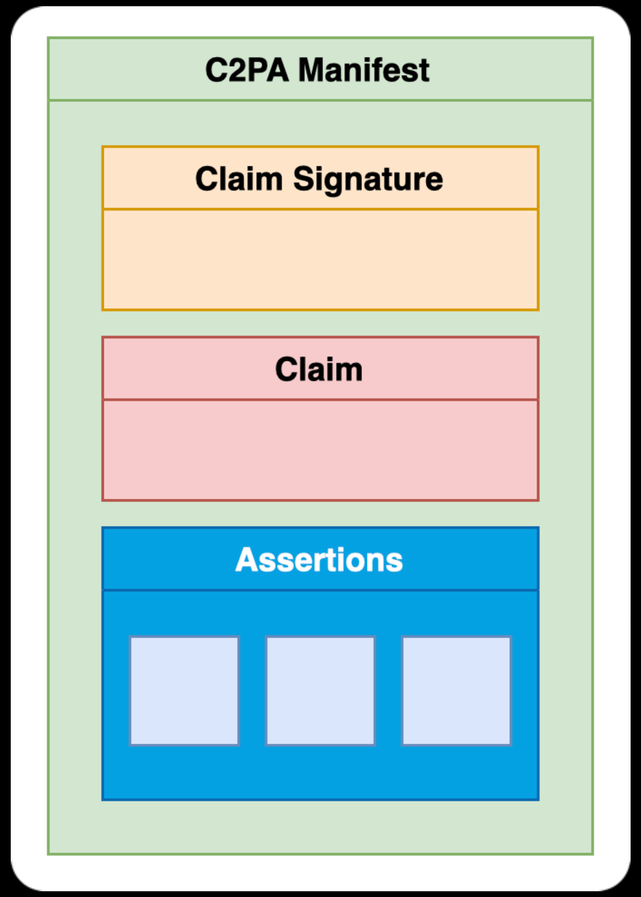
Every C2PA Manifest has exactly one **claim,** which lists the assertions and describes the claim generator (tool that built the Manifest).

This claim is signed using an X.509 certificate, which provides evidence of the tool used and provides tamper evidence if a third party attempts to change the claim after the fact.

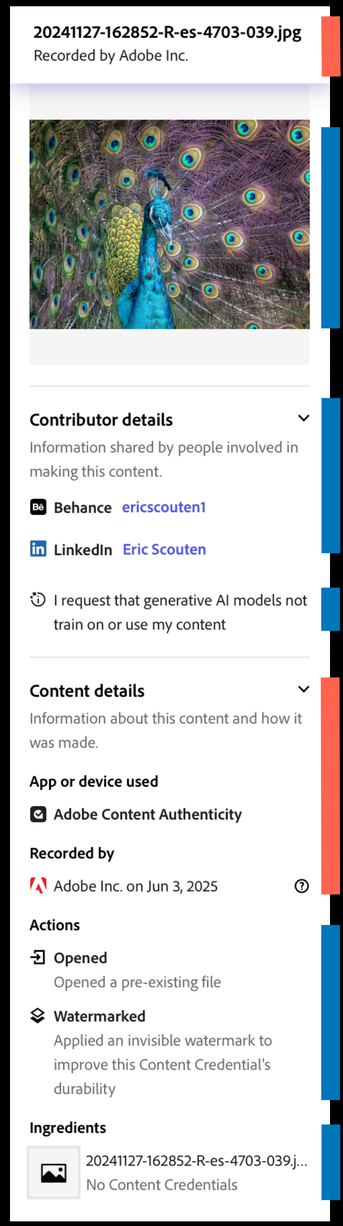


C2PA data model

How we display it



contentauthenticity.adobe.com/inspect



claim generator (C2PA)

thumbnail assertion (C2PA)

identity assertion (CAWG)

training + data mining assertion (CAWG)

claim generator (C2PA)

actions assertion (C2PA)

ingredients assertion (C2PA)

Introducing CAWG

CAWG (Creator Assertions Working Group)

was created in early 2024 to create technical standards to house metadata sourced from individual and organizational content creators.

CAWG became a working group within DIF in March 2025.

What does CAWG do?

Four assertion standards, building on C2PA technical spec:

- **Endorsement** ► Forward permission for CDN-style renditions on C2PA assets
- **Identity** ► Binding digital identity credentials to C2PA assets
- **Metadata** ► Associate user-generated metadata with C2PA assets
- **Training and Data Mining** ► Express permissions regarding AI training and data mining usage

What does CAWG do?

Four assertion standards, building on C2PA technical spec:

- **Endorsement** ► Forward permission for CDN-style renditions on C2PA assets
- **Identity** ► Binding digital identity credentials to C2PA assets
- **Metadata** ► Associate user-generated metadata with C2PA assets
- **Training and Data Mining** ► Express permissions regarding AI training and data mining usage

Identity assertion

is a framework

The actor* described by ... *${credential}*

using a credential issued by ... *${issuer}*

produced the content described by ... *${signer_payload}*

Signed by ... *${credential_holder}*

*actor can be human, non-human, or organization of humans

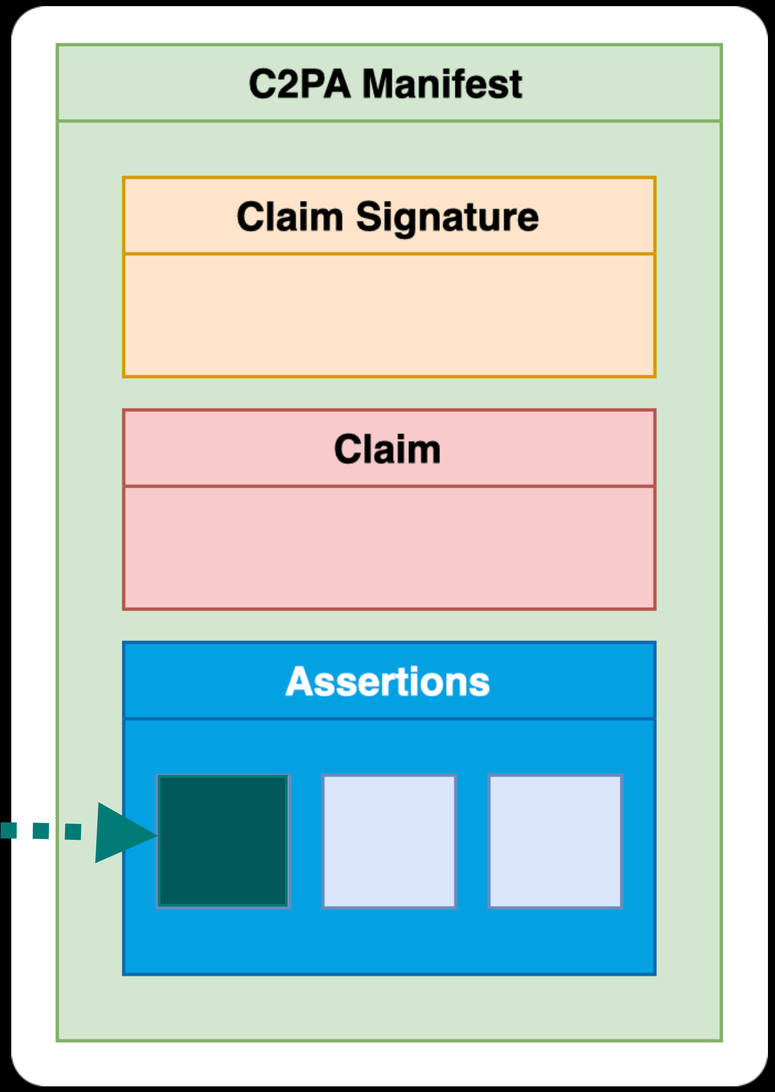
Identity assertion

in the C2PA data model

A **CAWG identity assertion** is typically meant to indicate subject's **authorization of** or **active participation in** production of the asset.

It provides a **tamper-evident binding** between a digital credential and the asset described by the C2PA Manifest *and* potentially other assertions in the same C2PA Manifest.

The actor* described by ...	*${credential}*
using a credential issued by ...	*${issuer}*
produced the content described by ...	*${signer_payload}*
Signed by ...	*${credential holder}*

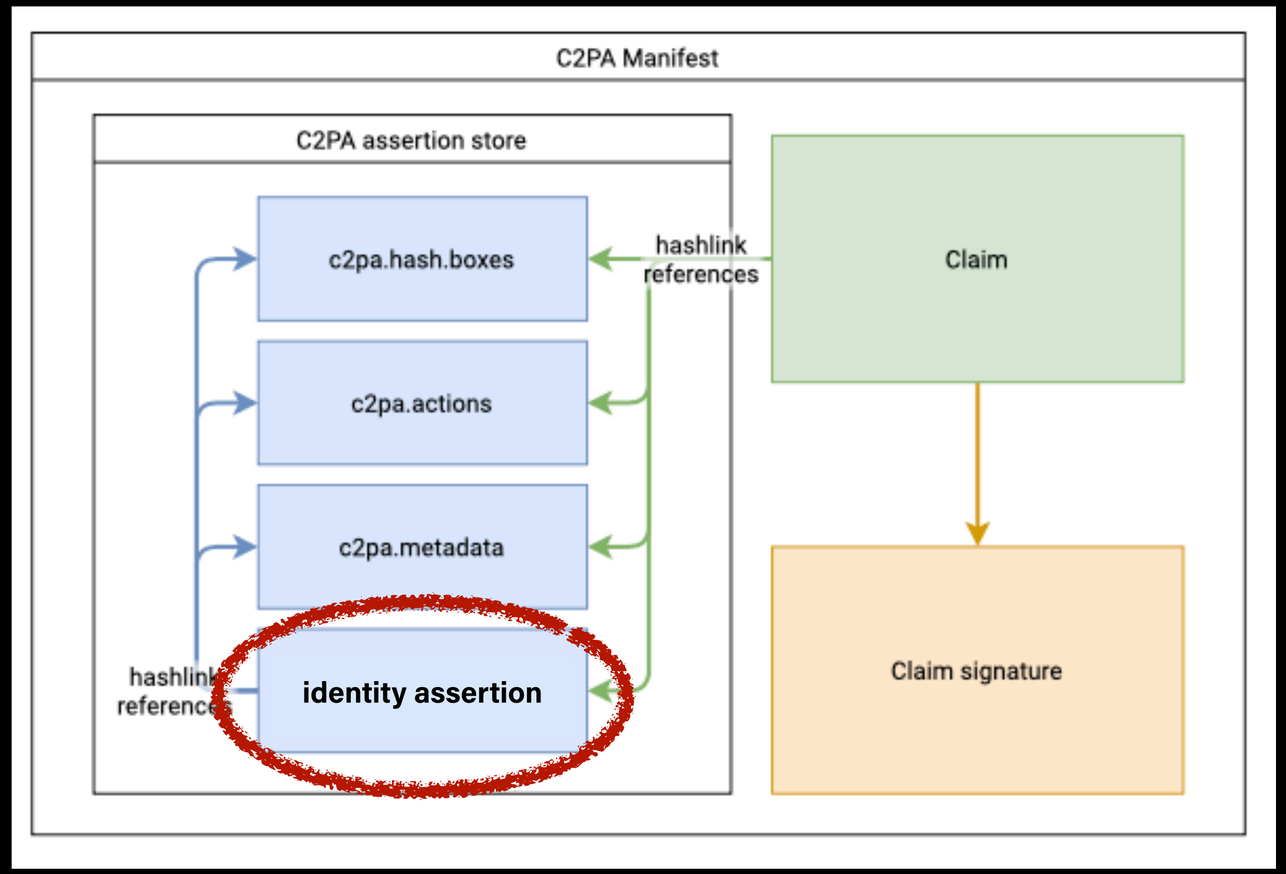


Identity assertion

Overview

Identity assertion allows a credential holder to sign a **signer payload** data structure which contains tamper-evident references to one or more other assertions in the same C2PA Manifest (including hard-binding assertion).

New trust signal separate from C2PA claim generator.



Identity assertion

Two flavors (so far)

- **X.509 certificate**

  Typically used for institutional content creators such as news media. (More discussion in breakout part 2.)

- **Identity claims aggregation**

  Targeted for individual content creators;
  contains links to social media, web site, etc.

- **Extensible**

  More flavors coming in 2026

Identity assertion

Individual content creators

- Instagram
- Twitter
- Other social media
- Web site
- Identity document (mDL or physical drivers license, etc.)

Problem: These credentials can generally be *observed* or *gathered* temporarily, but they generally don’t have autonomous signing capability.

Identity assertion

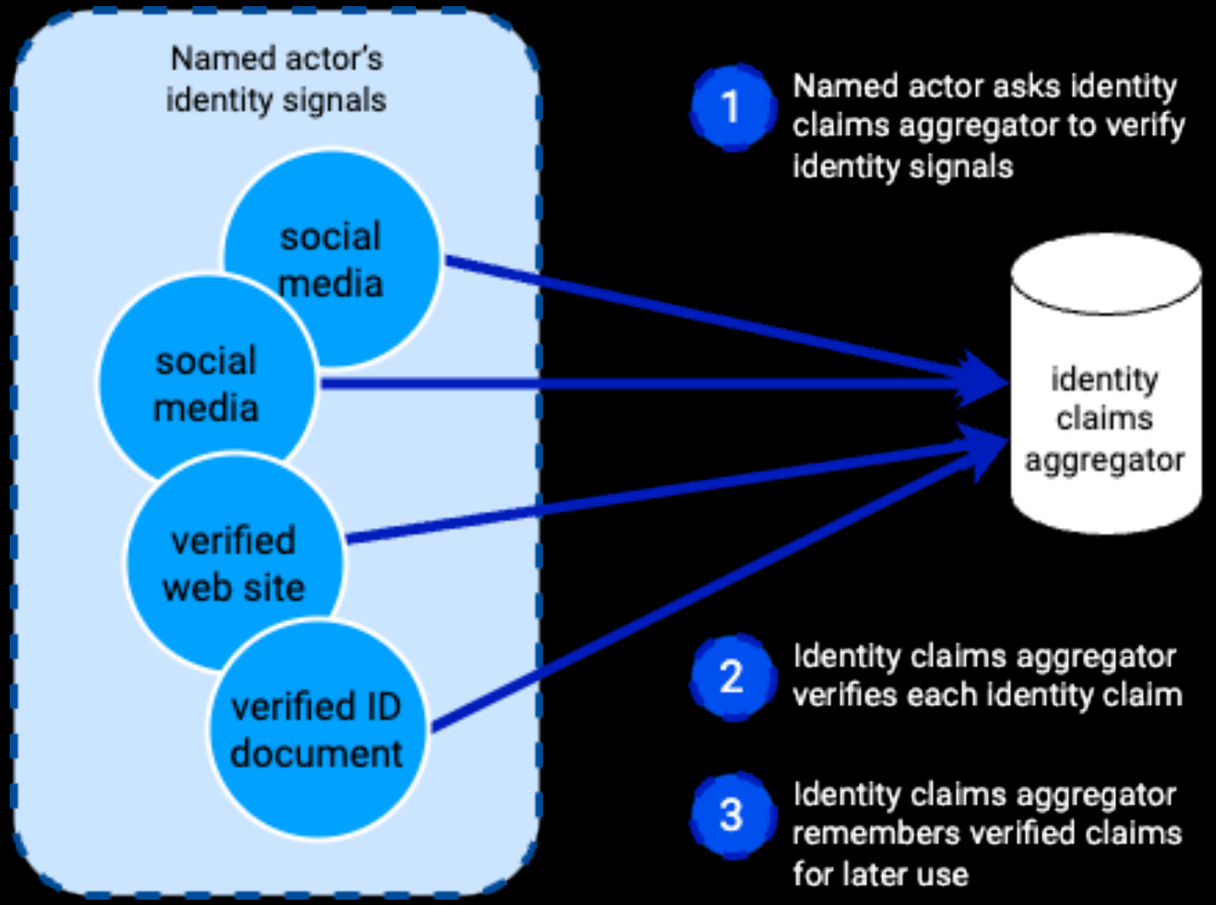
Individual content creators

- Instagram
- Twitter
- Other social media
- Web site
- Identity document (mDL or physical drivers license, etc.)

Solution: Describe how a platform vendor can *aggregate* these identity signals and attest to them on behalf of their customer.

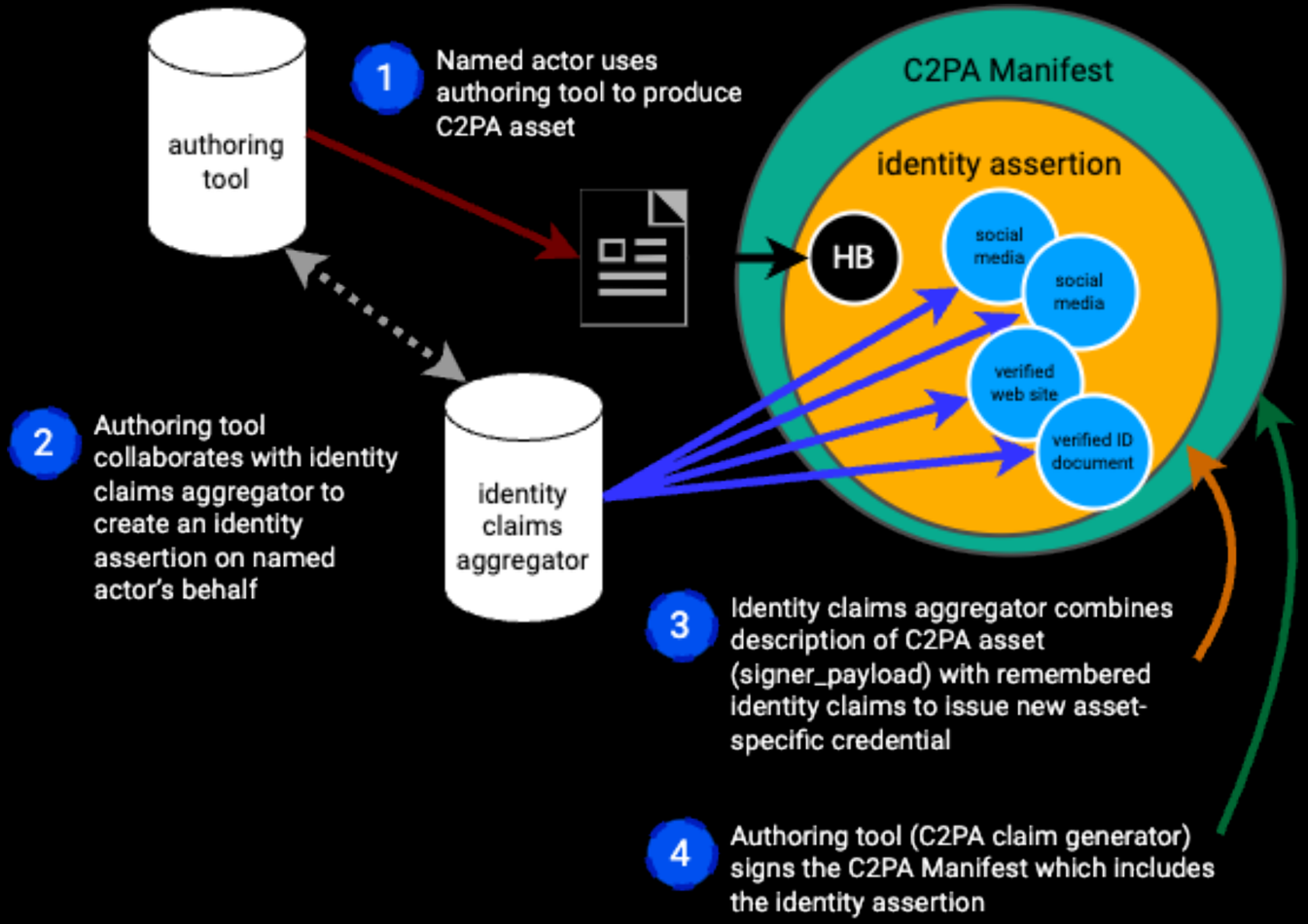
Identity assertion

Individual content creators: Verifying identity attestations



Identity assertion

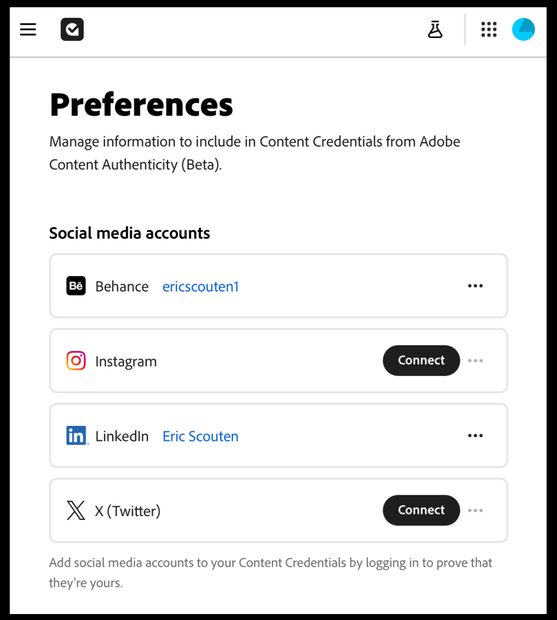
Individual content creators: Creating content



Adobe Content Authenticity

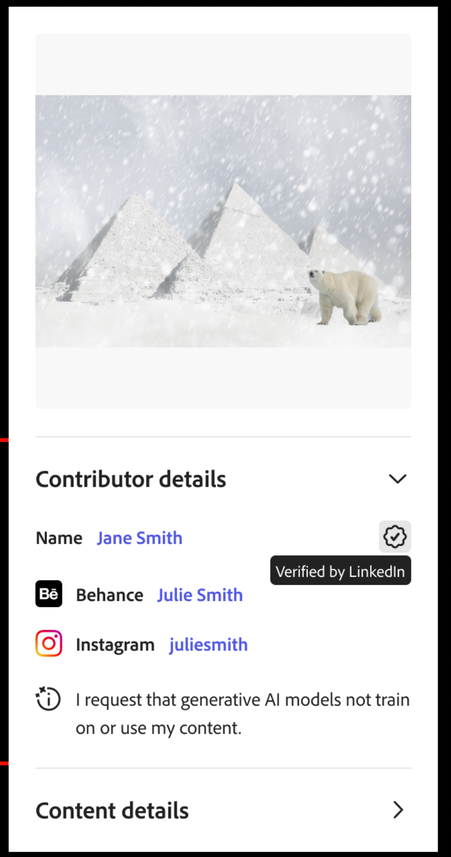
UX for CAWG identity claims aggregation

contentauthenticity.adobe.com/preferences



Data sourced from **CAWG identity** assertion

contentauthenticity.adobe.com/inspect



Identity assertion

Individual content creators

The actor described by ... VC with aggregated ID signals

using a credential issued by ... identity claims aggregator

produced the content described by ... ${signer_payload}

Signed by ... identity claims aggregator

(CAWG identity assertion 1.1 – ratified May 2025)

Caution

Identity claims aggregation is *one way* to provide information about a content creator.

It’s useful as a bridge between the identity signals mentioned before and current credential technology, but it is *not* fundamental to the identity assertion.

Identity assertion

Organizational identity via CAWG X.509

The actor described by ... X.509 certificate

using a credential issued by ... certificate authority

produced the content described by ... ${signer_payload}

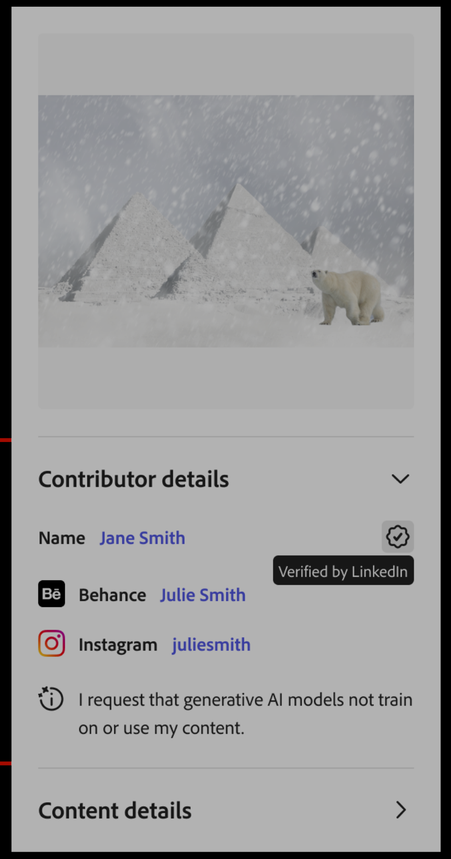
Signed by ... X.509 credential holder

(CAWG identity assertion 1.0 – ratified August 2024)

Adobe Content Authenticity

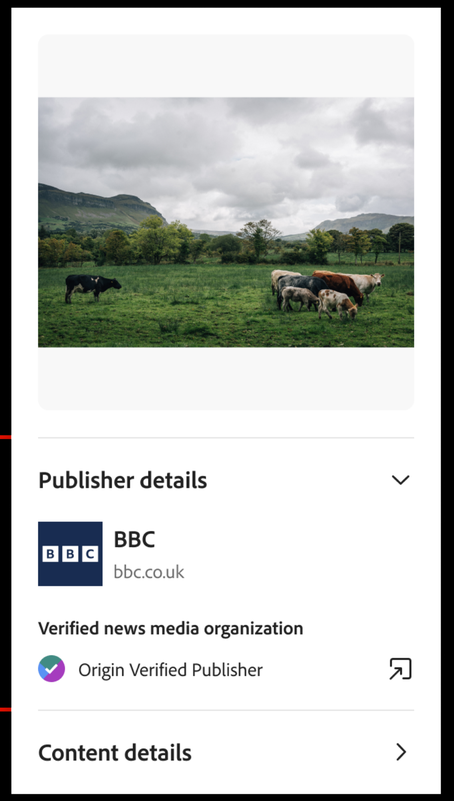
Proposed UX for Origin verified publisher content

contentauthenticity.adobe.com/inspect
(for individual-created content)



Data sourced from **CAWG identity** assertion

contentauthenticity.adobe.com/inspect
(for news media content)



Adobe Content Authenticity

Publisher provides additional metadata

Publisher details

BBC
bbc.co.uk

Verified news media organization

Origin Verified Publisher

Content details

Data sourced from CAWG **metadata** assertion

Publisher details

BBC
bbc.co.uk

Title

Cattle Grazing Beneath Stormy Skies in County Sligo

Description

A group of cows graze in a green pasture surrounded by rolling hills and cloudy skies in a rural landscape, likely in a temperate region.

Credit

Photo by Jane McConnell / Reuters

Published on September 15, 2012

Verified news media organization

Origin Verified Publisher

Content details

Current work

Identity and metadata 1.2

Identity assertion 1.2 – likely very soon

Driven by interest from news media, entertainment, and organizational brands:

- Establish an **interim** trust model for CAWG X.509, largely based on S/MIME certificate infrastructure
- Add guidance for using logos and icons contained in X.509 certificates
- Fix a C2PA compatibility issue, allowing identity assertions to appear in update manifests

Metadata assertion 1.2 – likely very soon

- New guidance for how to use metadata assertion to document involvement of many contributors with particular emphasis on motion picture and recorded music.
- Recommend use of identity assertion to attest to metadata authorship.

Upcoming work

Identity 1.3

Identity assertion 1.3

Proposed goals – still under discussion

- Identity evolution
- Self-control of identity signals
- Privacy preservation
- Broader integration with W3C VCs and VPs

Identity assertion 1.3

Proposed goals – **identity evolution**

- Name changes
- New social media / web site / contact

Allow content creators to provide new information – or remove existing information – about their identity for existing C2PA assets, *even if no information was provided at time of asset creation.*

Identity assertion 1.3

Proposed goals – **self-control**

- Allow content creators to control their own identity signals.
- Allow content creators to use the same identity signals across authoring tools.
- Allow content creators the ability to choose when/if to disclose aspects of their identity.

Identity assertion 1.3

Proposed goals – **privacy preservation**

- Allow content creators to control whether identity signals among C2PA assets are correlatable.
- Avoid unintended identity signals through unintended correlation between identity assertions.
- Allow content creators the ability to choose when/if to disclose aspects of their identity. *(repeat)*

Identity assertion 1.3

Proposed technical approaches – still under discussion

- Identity hooks
- First-person credentials

Identity assertion 1.3

Proposed technical approaches – **identity hooks**

Core idea: **Automatically create an “identity hook” (a public-private key pair or DID) for every C2PA asset created and privately remember the association between private key and asset.**

This allows the content creator to subsequently release information that is associated with that specific asset of the form: “I can prove that I created the specific asset in question and I would now like you to know ___.”

… without inadvertently proving that you created any *other* C2PA asset.

More info: github.com/decentralized-identity/cawg-identity-assertion/issues/216

Identity assertion 1.3

Proposed technical approaches – **first-person credentials**

Core idea: **Establish a self-controlled credential that is based on verifiable relationships.**

Allow content creators to associate reputation-based credentials with the content they create.

Intending to collaborate closely with First Person Project.

Come help us bind content provenance with identity!

CAWG is part of DIF

Meetings are every other Monday at 0800 Pacific / 1100 Eastern / 1500 UTC.

Next meeting: 3 November