

**Coalition for
Content Provenance
and Authenticity**

Content Credentials & the C2PA

W3C Media and Entertainment IG August 20, 2024

Andy Parsons
Leonard Rosenthol

Today's agenda

01. Overview of the C2PA

02. Content Credentials

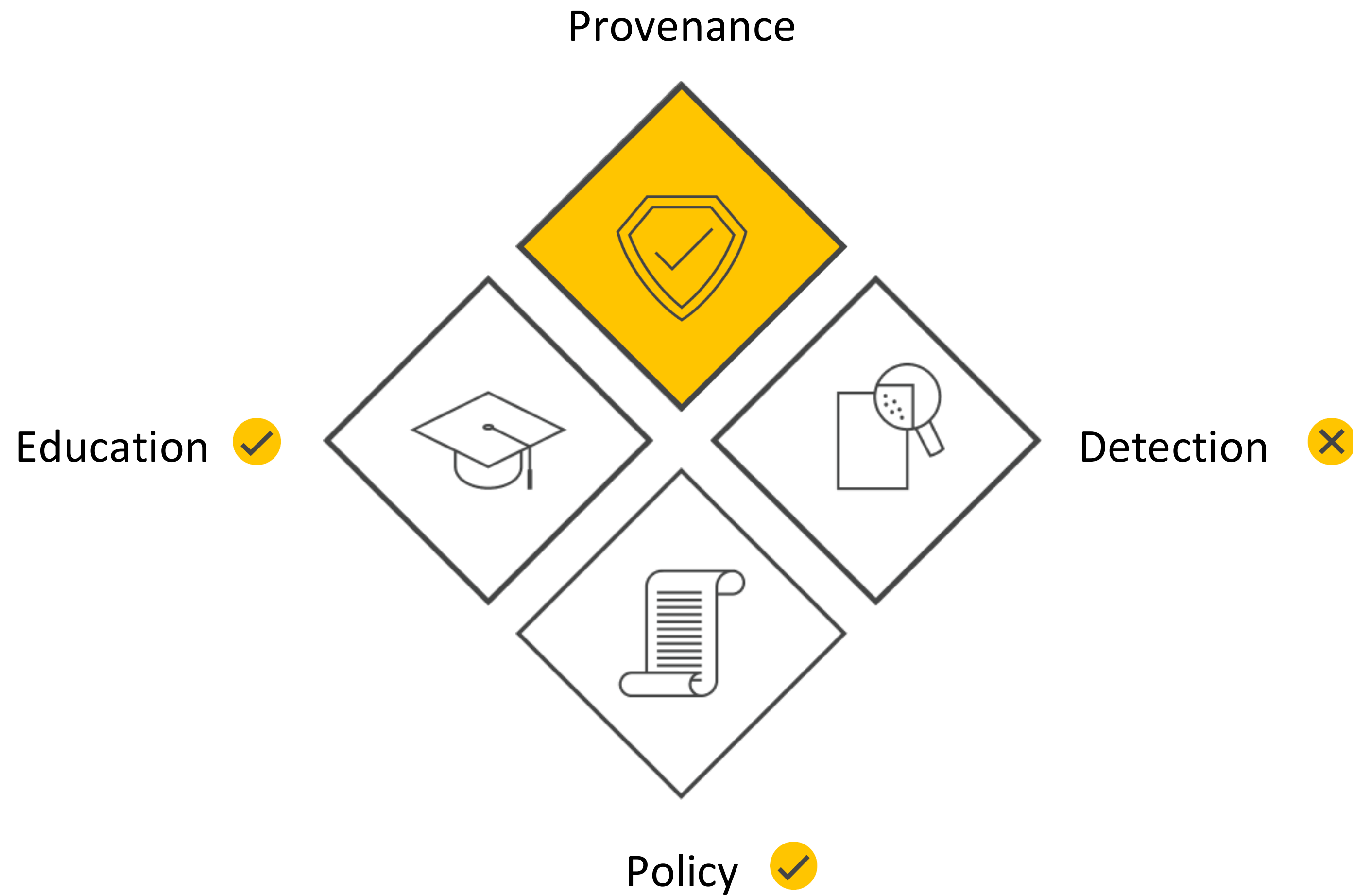
03. The specification

04. Next steps

05. Q & A



Pillars of Transparency and Authenticity

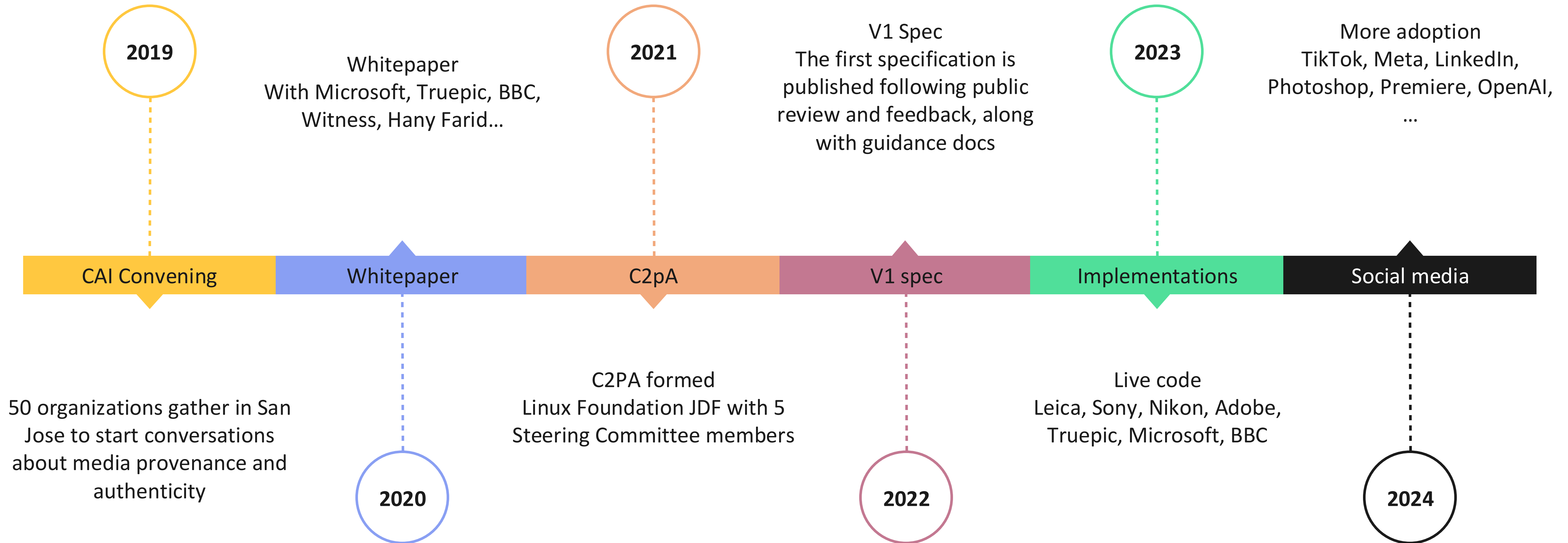


Provenance



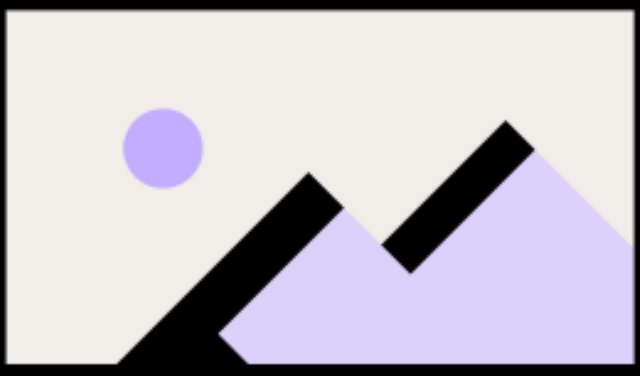
The basic trustworthy facts about the origins of a piece of digital content (image, video, audio recording, document).

A rough Content Credentials timeline






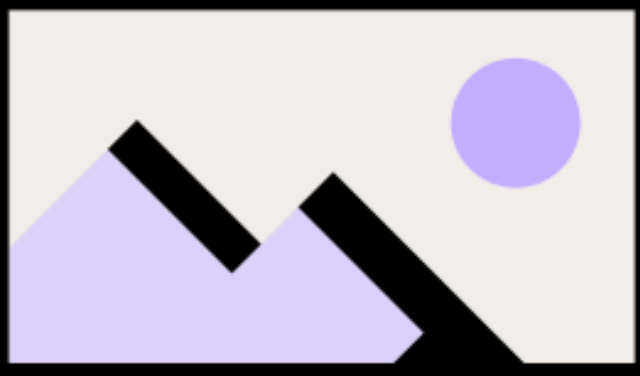
Content Credentials





Creation



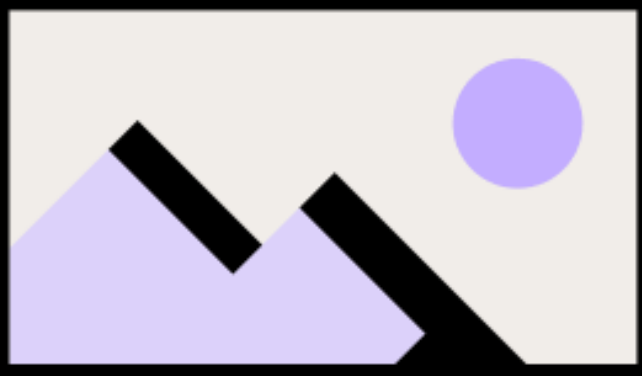
Creation date,
Created with,
Creator






Editing




Edit date,
Edits made,
Edits with,
Editor







Publishing



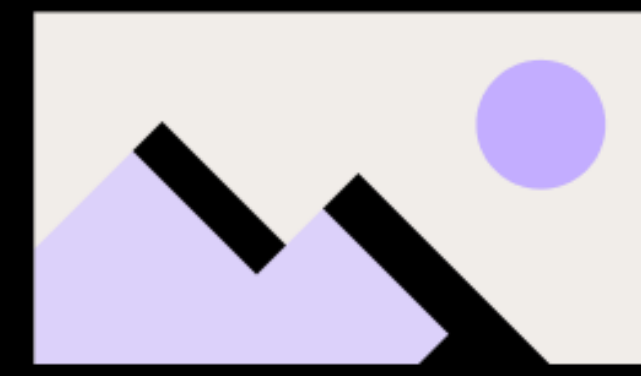
Review,
Publish date,
Published by








Sharing



Optimized for
delivery



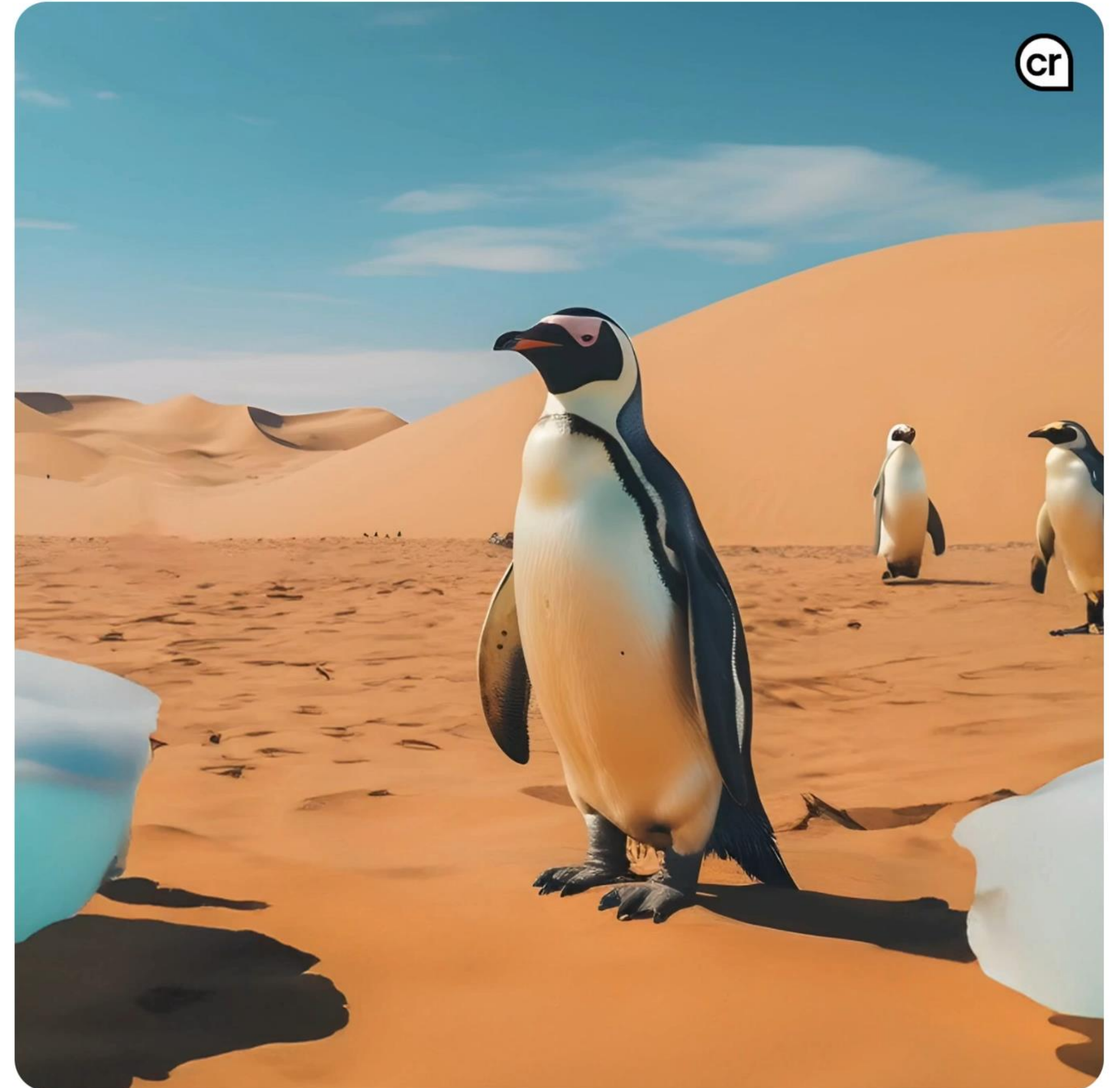
Viewing



Provenance

Wait, where did this image come from?

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



DASH support

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.



The Coalition for Content Provenance and Authenticity (C2PA)

- ▶ Standards organization (SDO)
- ▶ Linux Foundation JDF
- ▶ 9 Steering Committee members
- ▶ V1.0 specification: Jan 2022
V2.1 DRAFT: August 2024



What is the C2PA Specification?



An **open standard** for
storing and accessing
cryptographically
verifiable and tamper-
evident information
whose trustworthiness
can be assessed
based on a
defined trust model.

Specification (v2.0) Available - <https://c2pa.org/specifications/>

The screenshot displays the C2PA Specifications website interface. At the top, there is a header with the C2PA logo, the text "C2PA Specifications", a "Download" button with a dropdown arrow, a search box labeled "Search docs", and a chat icon. Below the header, a navigation sidebar on the left lists "C2PA Specifications" in red, followed by "Technical Specifications" (with a dropdown arrow) containing "Content Credentials" and "Attestations", and "Informative Documents" (with a dropdown arrow) containing "Explainer", "Guidance for Implementers", "User Experience Guidance", "C2PA Security Considerations", "C2PA Harms Modelling", "Guidance for Artificial Intelligence and Machine Learning". The main content area features a breadcrumb "C2PA Specifications" and a version indicator "2.0" with a dropdown arrow. The main heading is "C2PA Specifications". Below it, a paragraph explains that C2PA addresses misleading information through technical standards for certifying media content, mentioning a joint development foundation with Adobe, Arm, Intel, Microsoft, and Truepic. A second paragraph states that the site contains various specifications and documents produced by C2PA. The content is organized into two sections: "Technical Specifications" and "Guidance & Informative Documents", each with a list of links. A "Contents" sidebar on the right lists "Technical Specifications" and "Guidance & Informative Documents". At the bottom, a note indicates that PDF versions of the documents are available via the "Download" button in the page header.

C2PA Specifications

Download ▾ Search docs

C2PA Specifications 2.0 ▾

C2PA Specifications

The Coalition for Content Provenance and Authenticity (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 is a Joint Development Foundation project, formed through an alliance between Adobe, Arm, Intel, Microsoft and Truepic.

This site contains the various specifications and documents produced by the C2PA.

Technical Specifications

- [Content Credentials](#)
- [Attestations](#)

Guidance & Informative Documents

- [Explainer](#)
- [Guidance for Implementers](#)
- [User Experience Guidance](#)
- [Security Considerations](#)
- [Harms Modelling](#)
- [Guidance for Artificial Intelligence and Machine Learning](#)

PDF Versions of these documents are also available via the Download button in the page header.

Contents

- Technical Specifications
- Guidance & Informative Documents

Towards 2.1

- Currently in **public draft review**
 - <https://c2pa.org/specifications/specifications/2.1/index.html>
- Sent to **ISO TC 171/SC 2 for Fast Track** standardization (DIS)
 - WG 13 – meeting to take place in Prague in Nov.
 - Update to final 2.1, then off to publication
- **W3C Review**
 - W3C task force for the express purpose of returning a review of 2.1, the scope of this review should be how C2PA fits into the web and how it makes use of web technologies.

Technical Review of C2PA

(Some of our) Design Goals

- Create only the **minimum required novel technology** by relying on **prior, battle-tested techniques**.
- **Do not require cloud storage** or distributed ledgers/blockchain **but allow for it**.
- **Maintain an audit trail across multiple tools**, from asset creation through all subsequent modification and publication/distribution.
- **Support all standard asset formats** supported by common authoring tools, across media types such as images, videos, audio, and documents.

Core Technologies w/Open Standards

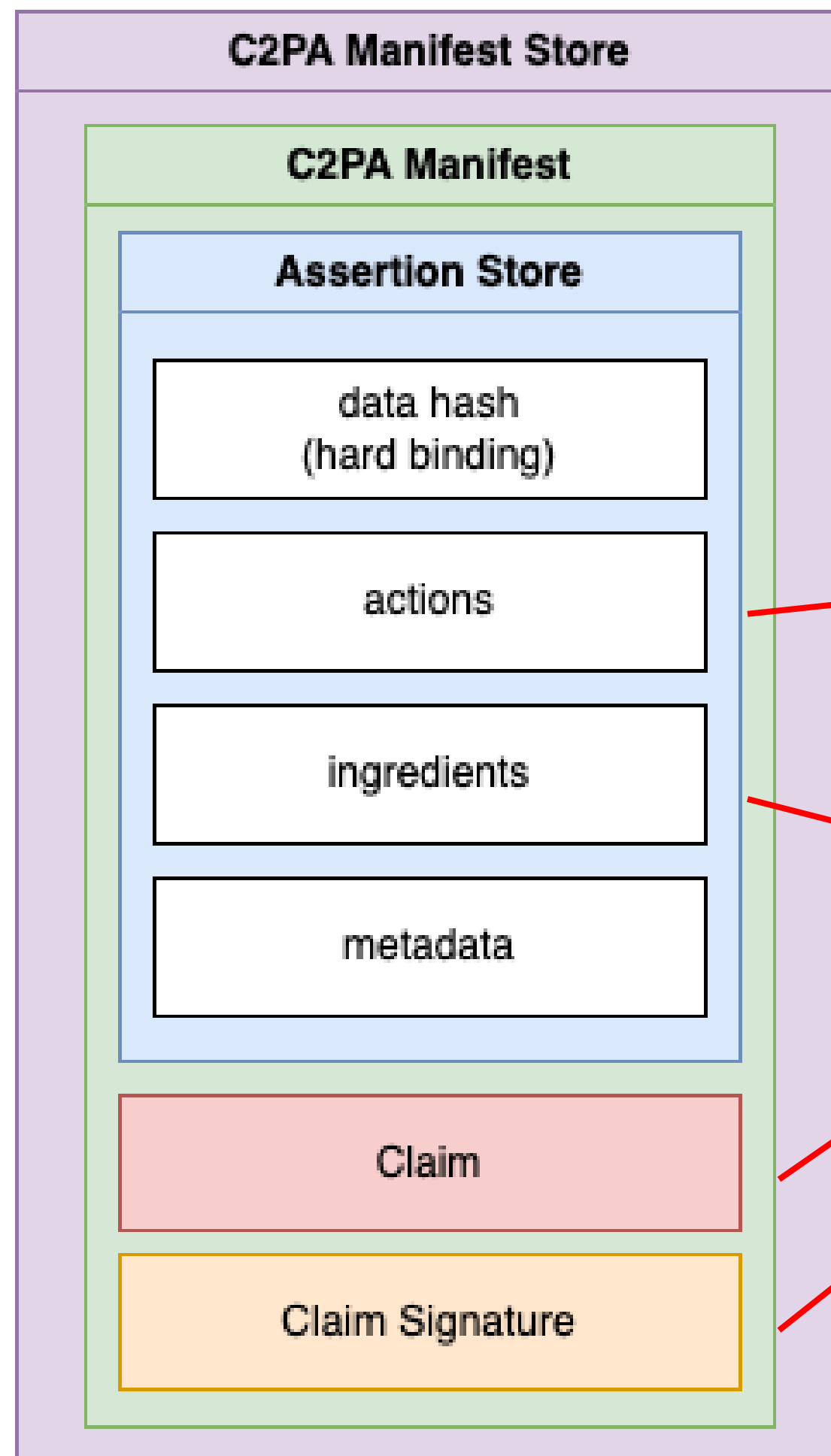
- JPEG universal metadata box format (JUMBF)
- JSON-LD & CBOR
- Cryptographic Message Syntax (CMS)
- CBOR Object Signing and Encryption (COSE)
 - NOTE: working with ETSI on an AdES for COSE, cbAdES



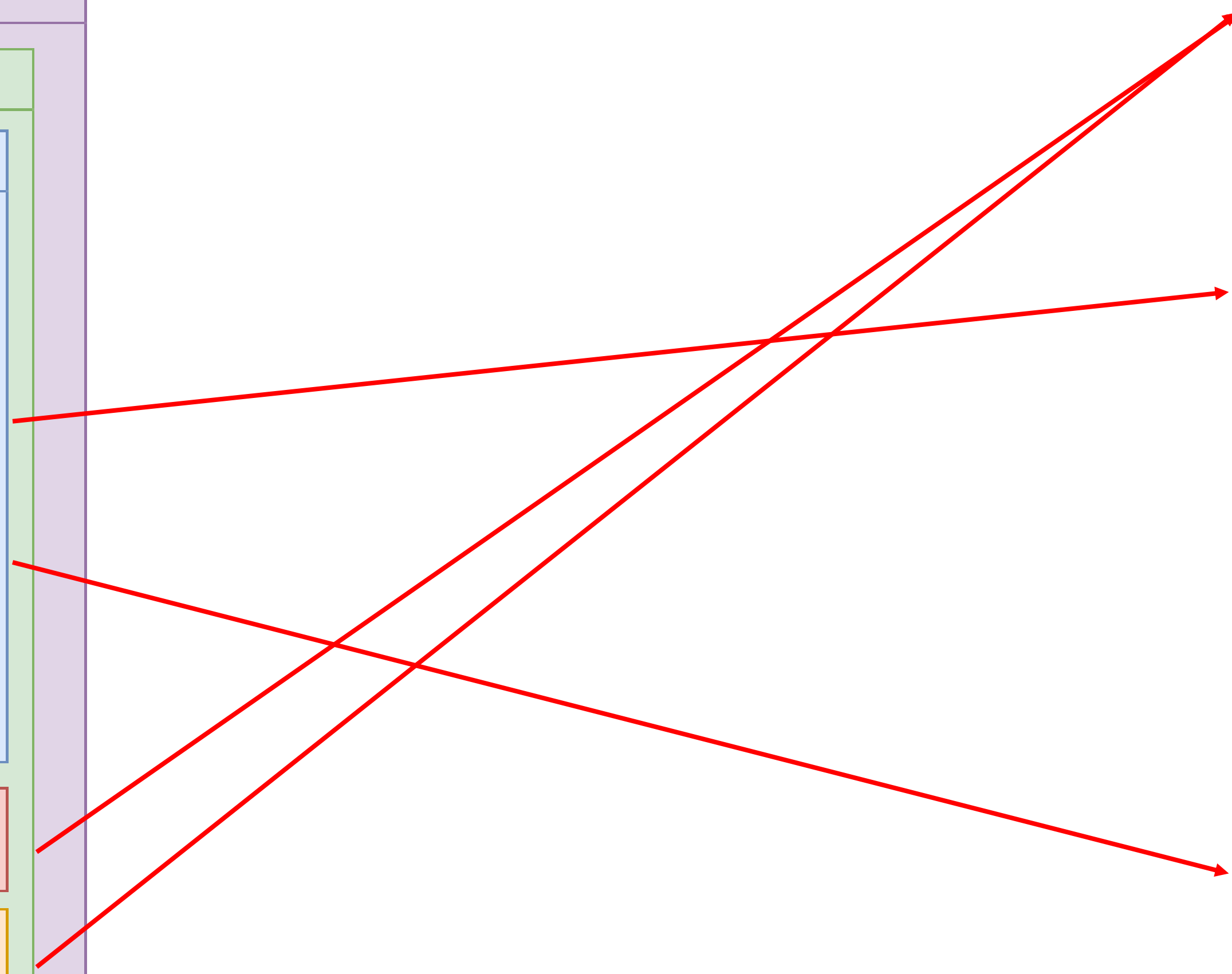
Core Components to C2PA

Content Credentials

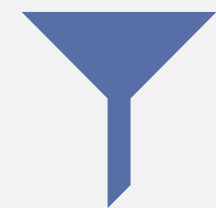
<https://contentcredentials.org/>



A screenshot of a social media post from EditSuite, dated September 8, 2021, at 10:34 AM. The post features a photo of a snowy mountain peak. Below the photo, the text reads "PRODUCED BY John Smith". Underneath, the "EDITS AND ACTIVITY" section lists several actions: "Color adjustments" (Changed tone, saturation, etc.), "Combined assets" (Composited 2 or more assets), and "Size and position adjustments" (Changed size, orientation, direction, or position). The "ASSETS" section shows two small images: one with a gear icon and another with a mountain peak. A "View more" button is located at the bottom of the post.



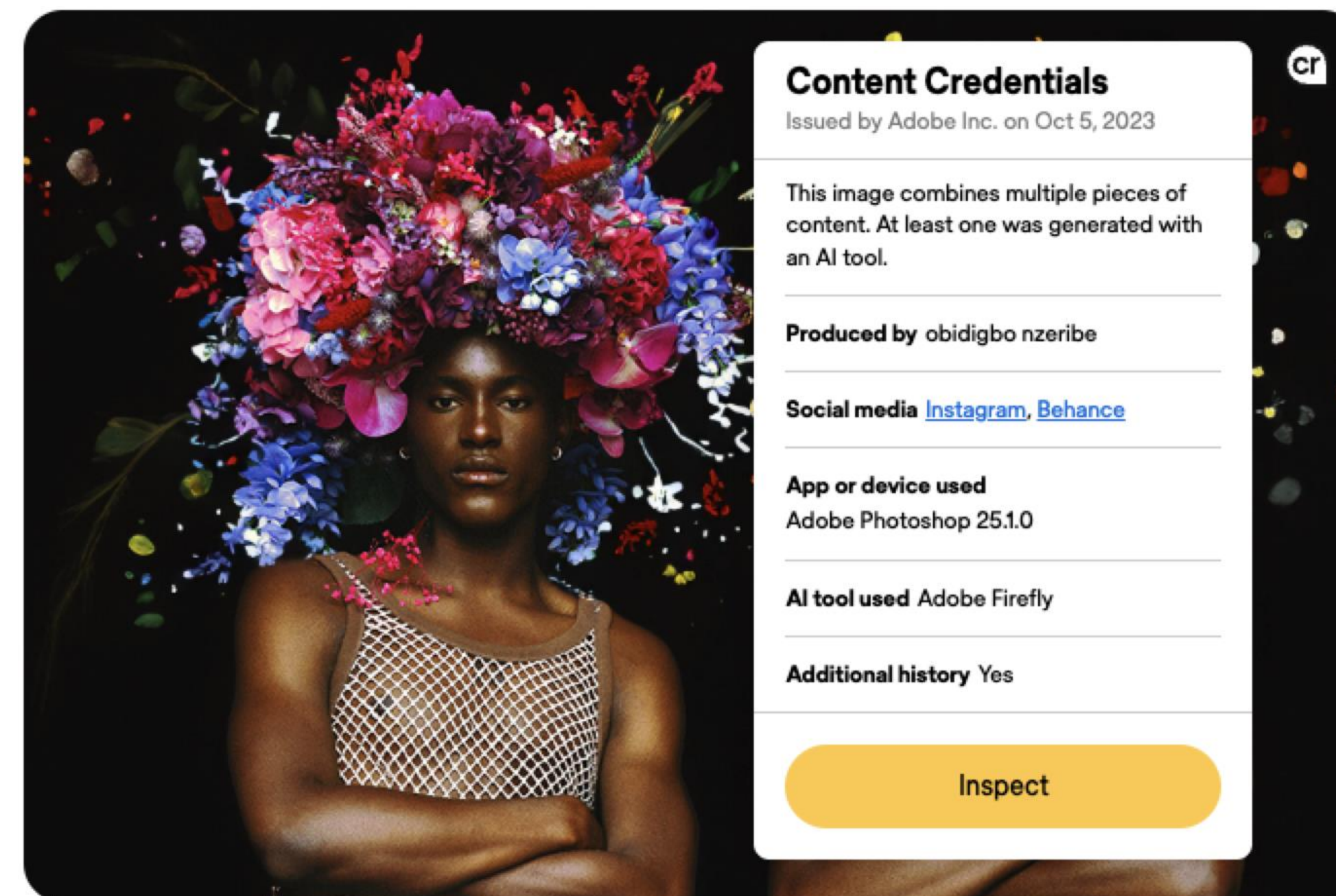
Functionality around Generative AI



Identifying assets & regions of interest
that have been created/modified by AI



Information about the "recipe"



Cryptography - a Trust Model

- Modelled on the **same approach as PDF and the Web**
- Cryptographic foundation
 - Hard bindings between metadata and content (e.g., SHA256, SHA3,...)
 - X.509 certificates for signing
 - Certificate authorities and trust lists
 - Hardware attestation for cameras/mobile devices
- Soft bindings allow recovery of stripped metadata



Trust Lists in 2.0

- C2PA Trust List
 - Official trust list from the C2PA
 - LOTS of details to work out!
 - Will represent the trust anchors for compliant hardware/software
- Additional Trust Lists
 - E.g., Project Origin
- Private/Personal Trust Lists



Well-Formed & C2PA Valid

- Well-Formed Manifest

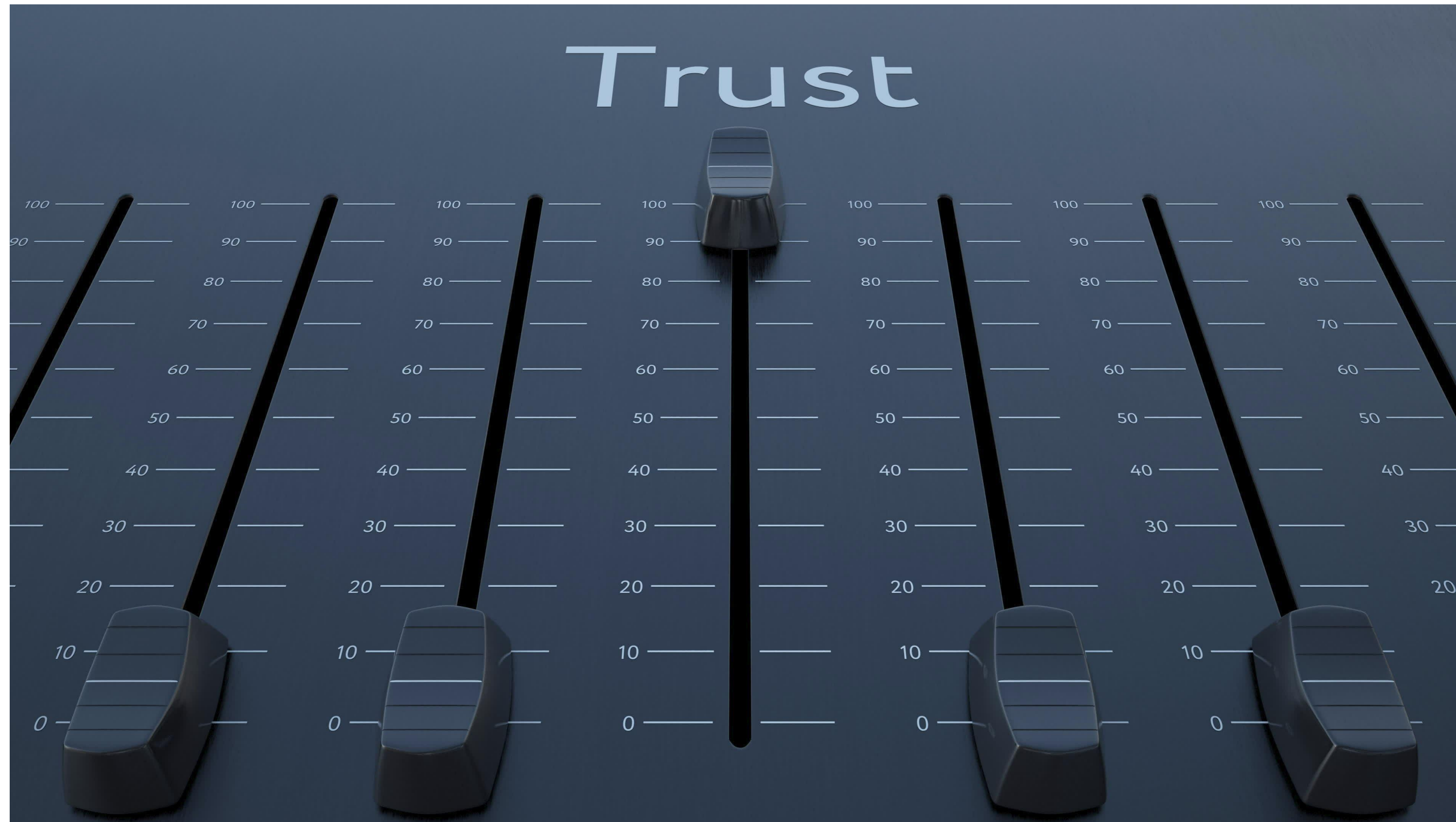
- All the bytes are intact & nothing missing or added
 - All the hashes match.
 - Complete set of assertions
 - Etc.
- Validation section is more complete & includes success codes

- C2PA Valid Manifest

- The signer of the Manifest is on one of the Validator's Trust Lists



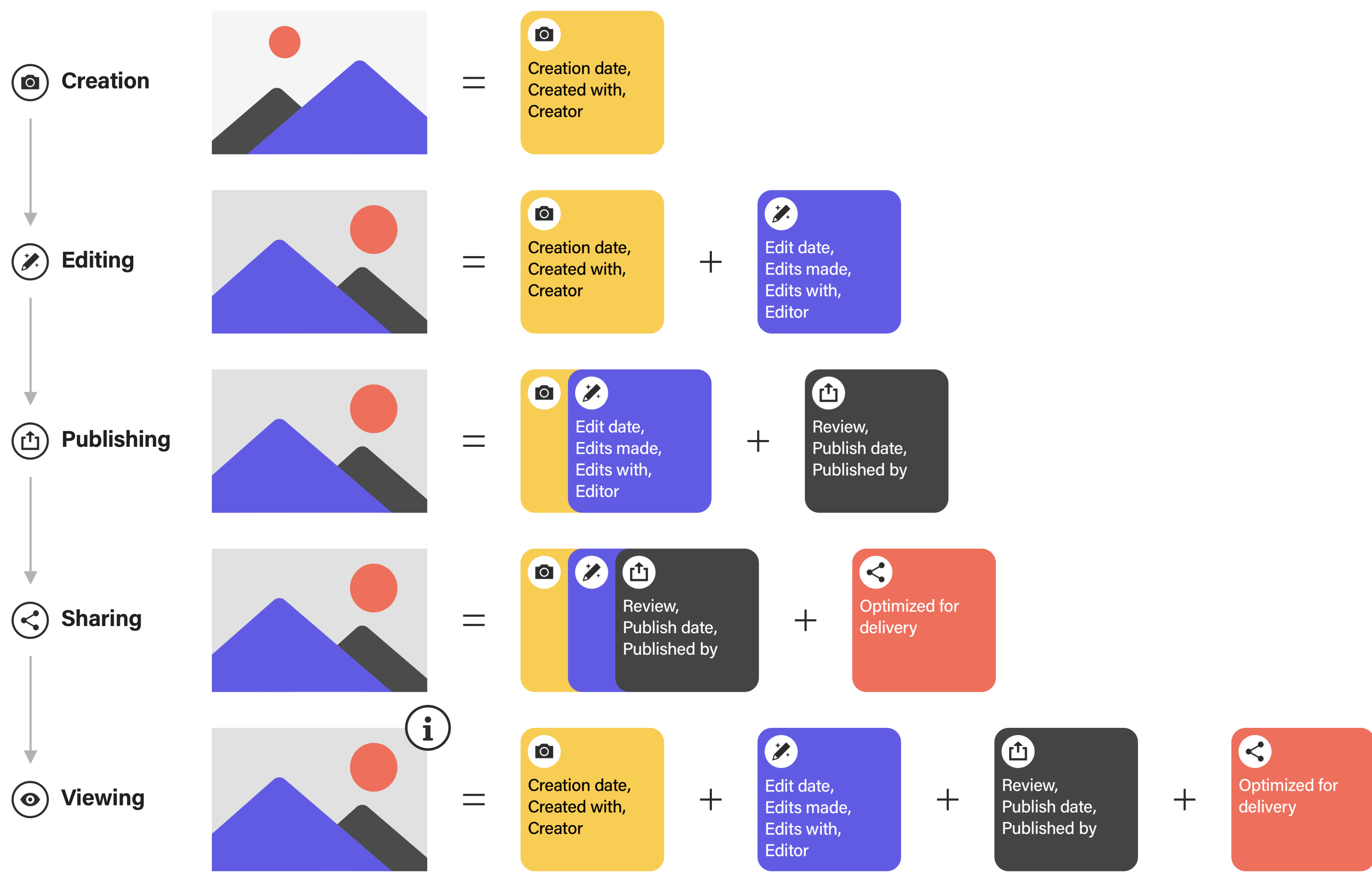
Enhancing Trust with Trust Signals



- Trust isn't binary (true|false)
 - Determined by a human, not a machine!
- Based on a series of “signals” that aid in the decision-making process

Embedded or Referenced

- Content Credentials can be embedded into
 - Images (JPEG, PNG, GIF, WebP, AVIF, HEIC/HEIF, TIFF, DNG, SVG)
 - Videos (MP4, MOV, AVI, BMFF)
 - Audio (FLAC, MP3, WAV, BWF)
 - Documents (PDF, Office, EPUB)
 - Fonts (OTF, TTF)
- They can be stored separately in file systems, the cloud, DLT/Blockchains & referenced by URL, HTTP headers, file system paths and more.



Coming soon (2.1)

Security Improvements

- Stricter & better spelled out validation algorithms
- Mandatory validation of Ingredients
- Improved Timestamping
 - Incl. new “TimeStamp Manifest”
- Improvements in hashing of BMFF and ZIP-based formats



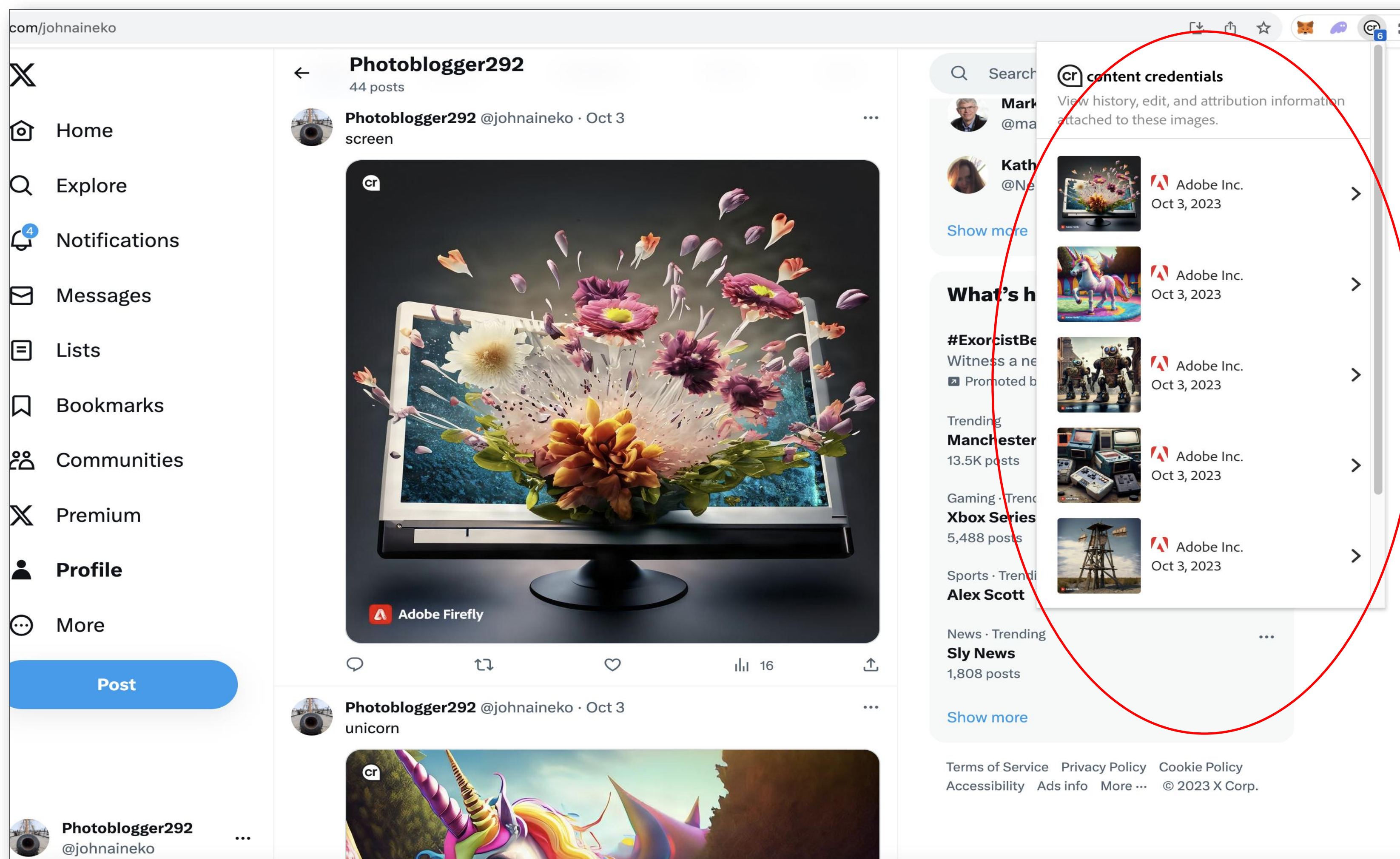
Improvements with Actions

- Required Actions - Created or Opened
- Deeper integration of Actions & Digital Source Types
 - Improves identification of various uses of AI



Durable Content Credentials

Browser detects invisible watermarks and **recovers stripped content credentials** from cloud



Content Credentials
recovered

<https://contentauthenticity.org/blog/durable-content-credentials>

Attestations

- <https://c2pa.org/specifications/specifications/1.4/attestations/attestation.html>
- Attestation is a platform security feature that allows software to prove its identity and the identity and/or security characteristics of the device or platform on which it is running.
- This specification describes how to add attestation signals and measurements to C2PA-marked assets. C2PA seeks feedback on this proposal, especially from implementers.

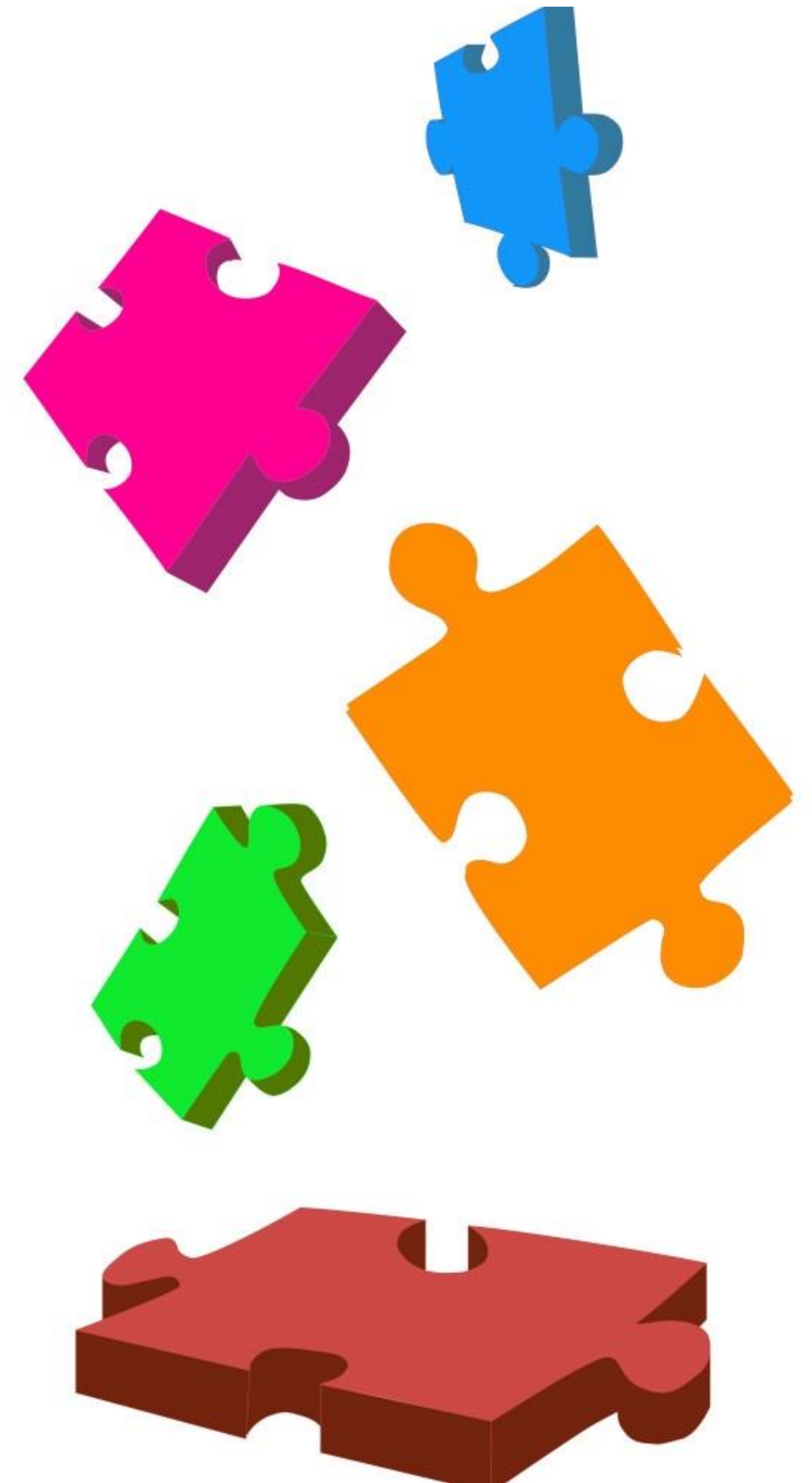
Work beyond the Core Specification

C2PA Task Forces

- AI/ML
- Conformance
- Ledgers
- Live Video
- Text
- Threats and Harms
- User Experience
- Watermark

Extensions to C2PA

- Creator Assertions Working Group (CAWG)
 - Identity
 - Training & Data Mining
 - Metadata
- PDF Association
- SCITT
- IPTC
- Etc..

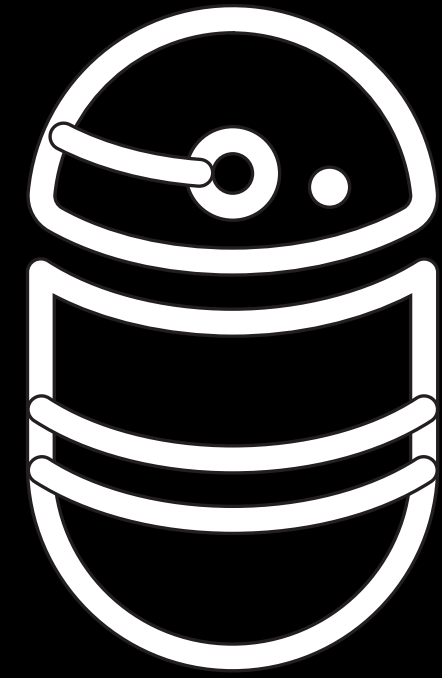


Thank you!



Questions





**Coalition for
Content Provenance
and Authenticity**