World **Privacy** Forum

September 2025

Privacy, Identity and Trust in C2PA

A Technical Review and Analysis of the C2PA Digital Media Provenance Framework

Kate Kaye Pam Dixon

Privacy, Identity and Trust in C2PA
A Technical Review and Analysis of the C2PA Digital Media Provenance Framework

World Privacy Forum
www.worldprivacyforum.org

EBook/Digital: ISBN: 978-0-9914500-3-9
Publication Date: September 2025
Nothing in this material constitutes legal advice.

This report is available free of charge at: https://worldprivacyforum.org/posts/privacy-identity-and-trust-in-c2pa
Updates to the report will be made at : https://worldprivacyforum.org/posts/privacy-identity-and-trust-in-c2pa

About This Report

This report is a technical review and analysis of the C2PA technical framework that connects digital media content such as images, video, audio and documents to data about the origins of and changes made to that content. C2PA is designed to undergird media infrastructures with detailed, automated, encoded and shareable data and trust signals about media and its creators. In its analysis of C2PA, this report considers and discusses C2PA use cases and interactions with data privacy, identity and trust in digital information ecosystems. C2PA is an abbreviation for the Coalition for Content Provenance and Authenticity, which developed the framework.

Brief Summary

Today, C2PA can be used to trace the edits made to digital photos, to manage media content in publishing pipelines, or to indicate whether certain Al platforms were utilized to alter or produce an image or video. Although C2PA is not built to hunt for Al deepfake clues or to fact-check information in content, it is intended to signal trustworthiness, not unlike provenance documentation indicating the authenticity of an oil painting or how an ancient artifact changed hands over time.

As such C2PA has on occasion been described as a content labeling system. However, it has additional capacities that reach beyond content labeling. C2PA is also designed to undergird digital media infrastructures with granular, automated, shareable, machine-readable data and trust signals about digital content and its creators. The encoded data it generates is meant to be readily ingested, analyzed, and exchanged by any and all systems that support C2PA, from tiny open-source software tools to massive cloud content delivery networks.

C2PA adoption is underway and accelerating. Understanding how this complex technical framework introduces new considerations for data use, data exchange, and system interoperability; how it affects people's privacy and identity, and how it facilitates measures of trustworthiness in an ever-morphing information ecosystem is vitally important, particularly at this early stage. The in-depth research this report provides aims to foster that understanding.

Key technical components of C2PA and considerations this report discusses include:

C2PA at Its Molecular Metadata Level: Think of metadata —the data about the media itself — as the molecules of C2PA. Products built using C2PA such as cameras or media management software automatically generate, ingest, trace and interpret a wide variety of C2PA metadata including geographic data marking the precise locations where media was created and details about content creators. Though technical hurdles can get in the way, the C2PA framework is designed to compile an ongoing chain of content provenance metadata that travels along with a piece of content wherever it goes.

Privacy and Identity in C2PA: This report reviews technical methods for asserting identity in the C2PA data workflow and connecting commercial and government identity systems to the C2PA workflow. Identity information and data flows carry privacy considerations, a connection well-documented across privacy and technical literature. Privacy, identifiable information and connections to identity systems are important considerations when it comes to understanding C2PA. Some C2PA users want to use the framework to assert the rights of content creators and owners by ensuring their identities are always attached to their content. However, there is disagreement on this point of use as not all creators or owners of digital media want to be identified. The C2PA specifications enable redaction of some sensitive metadata and call on C2PA-based systems to allow creators and publishers to opt in to its use. However, actual implementation could override some C2PA privacy guidance and mechanisms. C2PA's own Harms Modeling documentation recognizes the privacy and civil liberties threats posed by C2PA and states that loss of control over personal information and enforced suppression of speech are possible through use of C2PA.

C2PA's Trust Model: C2PA is not intended to directly determine the trustworthiness of content, but it

incorporates a technical trust model just the same. The C2PA trust model is a process by which certified authorities decide whether or not trust signals are valid or should be marked as fully trusted. A conformance program for those authorities was opened in June 2025[1] and has not been reviewed for this report.[2] C2PA metadata is meant to be used as signals for measuring content trustworthiness; indeed, the very absence of C2PA metadata can negatively affect C2PA-based interpretations of trust.

1 *Conformance, Conforming products*, Coalition for Content Provenance and Authenticity, Linux Foundation Projects, https://c2pa.org/conformance/. For program details see: *C2PA Conformance Program*, *C2pA technical working group conformance task force, Version 0.1, 2025-06-022*, GitHub (docs - archive - current - schemas - C2PA Conformance Program.pdf), https://github.com/c2pa-org/conformance-public/blob/main/docs/current/C2PA Conformance Program.pdf . *See also*: GitHub (docs- archive - archive_2025-06-13_15-38-22 readme.md) https://github.com/c2pa-org/conformance-public/tree/main/docs/archive .

2 *Conformance, Conforming products*, Coalition for Content Provenance and Authenticity, Linux Foundation Projects, https://c2pa.org/conformance/.

About the Authors

Kate Kaye contributed the primary research and writing for this report and conducted all interviews.

Pam Dixon contributed substantive editing and analysis, and research and analysis pertaining to governance, privacy, and identity. She also contributed to the methodology.

John Emerson designed the cover, the timeline illustration, and the report's online and eBook versions.

(Bios listed alphabetically)

Pam Dixon is the founder and executive director of the World Privacy Forum, a respected nonprofit, non-partisan, public interest research group. An author and researcher, she has written influential studies in the area of identity, AI, health, and complex data ecosystems and their governance for more than 20 years. Dixon has worked extensively on data governance and privacy across multiple jurisdictions, including the US, India, Africa, Asia, the EU, and additional jurisdictions. Recently, she completed the most comprehensive to date research charting data governance laws, treaties, and conventions globally. Also recently, she convened and chaired a workshop and roundtable of the data protection authorities of Africa to learn about the status of identity ecosystems and data protection across the regions of Africa. Her field research on India's Aadhaar identity ecosystem, peer-reviewed and published in Nature Springer, was cited in India's landmark Aadhaar Privacy Supreme Court opinion. Dixon has served as the co-chair of the UN Statistics Data Governance and Legal Frameworks working group, and is an advisor to the WHO's Health Data Collaborative. At OECD, Dixon is a member of the OECD.AI Network of Experts and serves in multiple expert groups, including the AI Futures group. In prior work at OECD, Dixon was part of the original AI expert group that crafted the OECD AI Principles, which were ratified in 2019. Dixon has presented her work on complex data ecosystems governance to the The National Academies of Sciences, Engineering, and Medicine, the Royal Academies of Science, and the Mongolian Academies of Science. She is the author of nine books and numerous studies and articles, and she serves on the editorial board of the Journal of Technology Science, a Harvard-based publication. Dixon was named one of the most influential global experts in digital identity in 2021. Dixon received the Electronic Frontier Foundation Pioneer Award in 2021 for her ongoing oeuvre of groundbreaking research regarding privacy and data ecosystems. Dixon has forthcoming peer-reviewed work on complex privacy governance.

John Emerson is a graphic designer, writer, and programmer based in New York City. He has designed web sites, printed materials and motion graphics for leading media companies as well as local and international non-profit organizations including Amnesty International USA, Human Rights Watch, the Committee to Protect Journalists, and the United Nations.

Kate Kaye joined World Privacy Forum as its deputy director in February 2023. In her role she focuses on national and international work on AI and machine learning including AI governance tools, digital identity ecosystems, health data ecosystems, and WPF's ongoing work on data governance. As part of her research on AI governance tools and use of those tools, Kate is the co-author of World Privacy Forum's December 2023 report, *Risky Analysis: Assessing and Improving AI Governance Tools*. She is also the author of *Uncovering Areas for AI Governance Tools Refinement through Real-World Use Case Analysis from Canada, Chile and Singapore,* which was published in July 2025 in the Proceedings of Machine Learning Research as part of the Fourth European Workshop on Algorithmic Fairness. Kate served as a paper reviewer for the Annual Conference on Neural Information Processing Systems (NeurIPS) in 2024 and 2025. Kate speaks often at events addressing data and privacy related topics, and is also host, editor and producer of WPF's *Privacy on the Ground* podcast. Before joining WPF, Kate worked as a journalist for more than 20 years, reporting on data-centric algorithmic technical systems, how they affect people, and on policy regulating tech including for MIT Technology Review, NPR, Protocol, Bloomberg CityLab, OneZero, WSJ and Fast Company. Kate has won several journalism awards including First Place Society of Professional Journalists NW Excellence in Journalism Awards for Technology in 2019 and 2022. Kate was a member of the UN's Hive Data Advisory Board in 2017.

About the World Privacy Forum

The World Privacy Forum is a respected NGO and non-partisan public interest research group focused on conducting research and analysis in the area of privacy and complex data ecosystems and their governance, including in the areas of identity, AI, health, and others. WPF works extensively on privacy and governance across multiple jurisdictions, including the US, India, Africa, Asia, the EU, and additional jurisdictions. For more than 20 years WPF has written in-depth, influential studies, including groundbreaking research regarding AI, including *The Scoring of America,* an early and influential report on machine learning and consumer scores, and most recently, *Risky Analysis — Assessing and Improving AI Governance Tools: An international review of AI Governance Tools and suggestions for pathways forward,* a report that has been cited by multiple governments. WPF conducted extensive work on systemic medical identity theft, bringing the issue to public attention for the first time, and on India's Aadhaar identity ecosystem —and *A Failure to do No Harm*, peer-reviewed work which was cited in the landmark Aadhaar Privacy Opinion of the Indian Supreme Court. WPF has co-chaired the UN Statistics Data Governance and Legal Frameworks working group, and is an advisor to the board of the WHO Health Data Collaborative. At OECD, WPF researchers participate in the OECD.AI AI Expert Groups, among other activities. WPF participated as a member of the first core group of AI experts that collaborated to write the OECD Recommendation on Artificial Intelligence, now widely viewed as the leading normative principles regarding AI. WPF research on complex data ecosystems governance has been presented at the National Academies of Science, the Mongolian Academies of Science, and the Royal Academies of Science. World Privacy Forum: https://www.worldprivacyforum.org.[3]

3 World Privacy Forum's home page includes information about our activities, as well as numerous data governance and privacy research, data visualizations, podcasts, and other resources. https://www.worldprivacyforum.org.

TABLE OF CONTENTS

Part I: Background and Introduction

An Overview of C2PA

C2PA is an emerging set of technical specifications for an open, standardized method enabling creation, storage, distribution and traceability of an array of detailed information about how content was created, how it has been changed, when, where, how, and by whom; this information is generally categorized as content provenance information. Several prototypes and early products and services have already been built using it, some of which are discussed in this report.

The Coalition for Content Provenance and Authenticity (C2PA)[4] developed the C2PA technical specifications[5] — sometimes referred to by their non-technical name, Content Credentials —as a way to generate and trace digital content provenance and authenticity information. C2PA has garnered increasing attention from media outlets[6] and even lawmakers,[7] particularly as use of generative AI systems[8] has exacerbated AI-enabled deepfakes and disinformation.

C2PA is sometimes referred to as a method for distinguishing between AI-generated synthetic media and non-synthetic media in the broader context of "labeling" content[9] to help determine whether or not it is trustworthy. It is used to indicate use of generative AI tools in the creation of images[10] and videos,[11] and in conjunction with digital watermarks.[12] C2PA assists in reviewing whether previous versions of content exist and facilitates comparisons among versions of content. And it is used in complementary systems that aim to determine whether images were used by a generative AI system or to train AI models.[13]

However, AI detection does not fall within C2PA's scope of content provenance[14] and at its technical core,

---

4 Coalition for Content Provenance and Authenticity, Home page, https://c2pa.org.

5 C2PA, *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html Note: This report uses C2PA Technical Specifications 2.1 as its primary technical reference. A subsequent version, 2.2, was published in May 2025. Version 2.2 does not alter or invalidate any technical information presented in this report.

6 Tate Ryan-Mosley, *An internet protocol called C2PA adds a "nutrition label" to images, video, and audio*, MIT Technology Review, (July 28, 2023), https://www.technologyreview.com/2023/07/28/1076843/cryptography-ai-labeling-problem-c2pa-provenance/.

7 Justin Hendrix, *Transcript of the US Senate Subcommittee Hearing on "Protecting Consumers from Artificial Intelligence Enabled Fraud and Scams*," Tech Policy Press (November 20, 2024), https://www.techpolicy.press/transcript-us-senate-subcommittee-hearing-on-protecting-consumers-from-artificial-intelligence-enabled-fraud-and-scams/ See also: United States Senate Commerce Committee on Commerce, Science, and Transportation Subcommittee Hearing on"*Protecting Consumers from Artificial Intelligence Enabled Fraud and Scams"* U.S .Senate, https://www.commerce.senate.gov/2024/11/protecting-consumers-from-artificial-intelligence-enabled-fraud-and-scams.

8 *Generative Artificial Intelligence,* National Institutes of Technology, Security Technology Resource Center, Glossary, https://csrc.nist.gov/glossary/term/generative_artificial_intelligence.

9 Content labels, sometimes referred to as "nutrition labels" are consumer-facing presentations overviewing history or digital content that are displayed when people experience that content. See: C2PA, https://c2pa.org/

10 OpenAI, C2PA in DALL·E 3, Privacy and policies, Policy FAQ, https://help.openai.com/en/articles/8912793-c2pa-in-dall-e-3.

11 Kylie Robinson, Emma Roth, and Richard Lawler, OpenAI has finally released Sora, The Verge, December 9, 2024, https://www.theverge.com/2024/12/9/24317092/openai-sora-text-to-video-ai-launch.

12 Digimarc, Digimarc Brings Digital Watermarking to the C2PA 2.1 Standard, Digimarc Newsroom, (October 8, 2024), https://www.digimarc.com/press-releases/2024/10/08/digimarc-brings-digital-watermarking-c2pa-21-standard.

13 Kar Balan, Alex Black, Andrew Gilbert, Simon Jenni, Andy Parsons, John Collomosse, *DECORAIT - DECentralized Opt-in/out Registry for AI Training*, In Proceedings of the 20th ACM SIGGRAPH European Conference on Visual Media Production (CVMP '23). Association for Computing Machinery, New York, NY, USA, Article 4, 1–10, 2023, https://doi.org/10.1145/3626495.3626506

14 Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe,

C2PA is not designed in itself to detect AI clues such as artifacts, inconsistent noise patterns and other inconsistencies,[15] [16] to hunt for deepfakes or AI-generated misinformation, or to sniff out hidden signs that AI may have been used to create content. Instead, C2PA is designed to generate, trace and recognize intentionally-inserted content provenance information.[17] [18] Some who have implemented C2PA or evaluated it from a policy perspective[19] suggest the goals of the framework have morphed over time in an effort to solve an expanding and expansive set of problems.

It is also important to recognize several broad use cases that are emerging regarding C2PA. C2PA does form the technical underpinnings of tools that display content labels for viewing by everyday humans. Additionally, behind-the-scenes, those technical underpinnings produce and share more granular, machine-readable data about content and its creators, data that is accessible to any system that has adopted C2PA or recognizes C2PA signals. This creates a very broad potential ecosystem for C2PA.

C2PA also has been used as a technical framework for business or legal-related use cases, such as to trace, manage and authenticate media throughout media industry pipelines,[20] or to assess the history and credibility of a photo or digital content. And it has been used to help demonstrate authenticity of photos documenting election

December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. In the December 4, 2024 interview for this report, Rosenthol told WPF that AI detection would not be added as a C2PA capability. However, because of C2PA's open design, implementers of C2PA could create new types of assertions reflecting AI detection. See the *Under the Hood* section for more detail on assertions.

15 Jaron Schneider, *Cameras, Content Authenticity, and the Evolving Fight Against AI Image*s, PetaPixel January 2, 2024, https://petapixel.com/2024/01/02/cameras-content-authenticity-and-the-evolving-fight-against-ai-images/. "The original premise behind content authenticity was not to fight AI — the CAI's initiative and provenance standard predate the proliferation of AI images — but rather as a way for publications to certify that an image has not been altered after it was captured."

16 Many supporters and drafters of C2PA and related frameworks see it as part of a larger mix of tools enabling secure metadata, watermarks, fingerprinting, and other tools for tracking provenance that will be needed to help combat AI-enabled deepfakes, misinformation and disinformation. See International Telecommunication Union (ITU), *Detecting deepfakes and generative AI: Report on standards for AI watermarking and multimedia authenticity workshop: The need for standards collaboration on AI and multimedia authenticity*, 2024, https://www.itu.int/hub/publication/t-ai4g-ai4good-2024-7/.

17 See Truepic, *Letter sent to Senator Mark Warner*, (March 22, 2024), https://www.warner.senate.gov/public/_cache/files/b/4/b460c5d1-f28b-40f9-8c0c-375fbee55566/B7524C0B87E869C5EB46BA1304F414F3.truepic-response-to-senator-warner-5.22.24.pdf. "The solutions we build help enterprises securely adding trust signals (i.e. provenance via C2PA Content Credentials upstream in the tech ecosystem. We also offer tools for programmatically reading and displaying C2PA Content Credentials on existing media files, but we do not consider this to be detection in the traditional sense. Instead, we would refer to this as C2PA ingestion; detecting the presence of C2PA Content Credentials on media files so that platforms can extend that transparency for the benefit of end users."

18 Since its launch in 2021, the Coalition for Content Provenance and Authenticity has referred to C2PA as "technical standards for certifying the source and history (or provenance) of digital content" to address "the prevalence of misleading information online." See *News*, Coalition for Content Provenance and Authenticity, https://c2pa.org/post/.

19 Dean W. Ball, *Deepfakes and the Art of the Possible, The C2PA standard is deeply flawed, but it may be fixable,* Hyperdimensional, May 30, 2024 https://www.hyperdimensional.co/p/deepfakes-and-the-art-of-the-possible . The author of the article noted that C2PA is "trying to do too many things at once."

20 AWS Architecture Center, *AWS Innovation with Sinclair, AWS Innovation Ambassadors* June 2024, https://aws.amazon.com/podcasts/innovation-ambassadors/ia-podcast-ep-47-sinclair/?podcast-list-wide.sort-by=item.additionalFields.EpisodeNum&podcast-list-wide.sort-order=desc.

related events [21] or events in conflict zones,[22] and as a standard for asserting creator attribution,[23] data or content use preferences, and intellectual property rights.[24] [25]

C2PA has a technical Trust Model,[26] however, C2PA is not designed to fact-check or vet the quality or veracity of information carrying its metadata.

Though borne out of earlier efforts, C2PA was launched in 2021 and the development is ongoing. See Figure 1 for a brief timeline. Appendix A contains a list of citations underpinning Figure 1 that provide source material about C2PA's development.

C2PA Timeline

2018

Project Origin Launches

Project Origin, a content provenance system and precursor to C2PA is founded.[T1]

NOVEMBER 2019

Content Authenticity Initiative Launches

Content Authenticity Initiative (CAI), a precursor to C2PA, launches to develop an "industry standard" for digital content attribution.[T2] CAI is fully funded by Adobe and continues to operate a C2PA membership organization and C2PA technical community Discord server.[T3]

FEBRUARY 2021

C2PA Is Created by Some Founders of Project Origin & CAI

Some founders of Project Origin and CAI join to establish the Coalition for Content Provenance and Authenticity (C2PA) as a Linux affiliated Joint Development Foundation. They refer to Project Origin and CAI as efforts influencing C2PA workflows, requirements, and best practices.[T4]

---

21 Ingo Boltz, *Content Credentialed Media in Election Observation Missions – First Lessons Learned*, Electoral Integrity Project, September 26, 2024, https://www.electoralintegrityproject.com/eip-blog/2024/9/20/content-credentialed-media-in-election-observation-missions-first-lessons-learned.

22 *Synthetic Media Framework Case Study: How Truepic used disclosures to help authenticate cultural heritage imagery in conflict zones*, Truepic, submitted as a case study to the Partnership on AI's Synthetic Media Framework, https://partnershiponai.org/wp-content/uploads/2024/11/case-study-truepic.pdf. See also: P*artnership on AI unveils new case studies from supporters of synthetic media framework: Meta, Microsoft, Thorn, and Truepic*, Truepic press release, Truepic, November 19, 2024. https://www.truepic.com/blog/partnership-on-ai-unveils-new-case-studies-from-supporters-of-synthetic-media-framework-meta-microsoft-thorn-and-truepic.

23 *See What Are Content Credentials?*, Adobe, Content Credentials, (October 14, 2024), https://helpx.adobe.com/creative-cloud/help/content-credentials.html.

24 *Content integrity: ensuring media authenticity*, Truepic Blog, Truepic, https://www.truepic.com/blog/content-integrity.

25 Ingo Boltz, *Content Credentialed Media in Election Observation Missions – First Lessons Learned*, Electoral Integrity Project, September 26, 2024, https://www.electoralintegrityproject.com/eip-blog/2024/9/20/content-credentialed-media-in-election-observation-missions-first-lessons-learned. "Corporations have deployed the technology to combat piracy of digital media. The Coalition for Content Provenance and Authenticity (C2PA), currently the most visible actor in the digital provenance space, is a corporate initiative originally founded with that objective."

26 "Trust Model," as used in this report refers specifically to the phrase used in the C2PA technical specifications that refer to the process "which is concerned with trust in a signer's identity." According to the specification, "...the consumer (who is not specified in the trust model), [uses] the identity of the signer, along with other trust signals, to decide whether the assertions made about an asset are true." C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html . See 14,Trust Model. See also: the *Technical Components of C2PA's Trust Model* section of this report for more details on this topic.

2021 – 2024

New Members Join the C2PA Steering Committee

Several Steering Committee members join C2PA in this time.[T5] Steering Committee membership costs $27,000 and is the only membership tier allowing for the right to approve final C2PA deliverables such as specifications for public publications.[T6]

SEPTEMBER 2021

First Draft of C2PA Technical Specs Is Published [T7]

JANUARY 2022

C2PA 1.0 Technical Specs Are Published

A C2PA press release refers to its use "to ensure transparency, understanding, and trust" and to address "deceptive content, such as deepfakes generated by artificial intelligence or more traditionally manipulated media." [T8]

OCTOBER 2023

Content Credentials Icon Launches

A C2PA press release calls the symbol an "icon of transparency" and a "mark that will provide creators, marketers and consumers around the world with the signal of trust-worthy digital content." ContentCredentials.org is also launched as a related "online hub."[T9]

JANUARY 2024

C2PA 2.0 Tech Specs Remove Identity

When version 2.0 of C2PA specs were published in January 2024, its drafters called the removal of references related to human and organization identity or identifiable humans "a significant departure from previous versions."[T10]

FEBRUARY 2024

Identity Proposal Discussed at First CAWG Meeting

One month after identity was removed from the core C2PA specs, a new group known as the Creator Assertions Working Group or CAWG was established, in part to serve as a home for identity related technical specifications development related to C2PA.[T11]

SEPTEMBER 2024

CAWG Identity Specs 1.0 for C2PA Ratified

The first version of an identity assertion for attaching human identifiers to content via C2PA is published. CAWG Identity assertions are expected to be components of the C2PA workflow.[T12]

OCTOBER 2024

C2PA Considered as an ISO Standard

The C2PA 2.1 specification is under review by the International Organization for Standardization TC 171/SC 2 committee with a chance to become an ISO Standard.[T13]

Figure 1: Timeline noting highlights of C2PA's development from 2018 to October 2024. See Appendix A for timeline citations.

Methodology

This report surveys C2PA in detail to assess whether — and if so how — the C2PA framework relates to privacy, identity, and more broadly, digital trust. The research for this report sought to determine the answers to the following research questions related to the focal areas of privacy, identity, and digital trust measurement as well as those related to the baselines of C2PA. Focal C2PA research questions:

1) Does C2PA relate to privacy and data protection? If so, how?
2) Does C2PA relate to the area of identity? If so, how?
3) Does C2PA relate to digital trust measurement? If so, how?

Additional research questions regarding the baselines for C2PA:

4) What are the stated goals of C2PA?
5) What are the capabilities of C2PA in relation to data generation, data sharing, and data use, including downstream data use?
6) How does C2PA interoperate with other frameworks or systems?
7) What are the existing use cases for C2PA?

These research questions were chosen in order to understand how C2PA is designed and how its design relates to data use, privacy, identity, and trust. To assess these research questions, the report utilized the following sources:

- C2PA's technical specifications [27] and related technical documentation, such as developer documentation, technical academic papers and journal articles, technical working papers, and additional relevant technical material.
- Public technical discussions among C2PA developers and those building products and services with C2PA, [28] were also examined.
- Interviews with experts who have worked closely with C2PA in various capacities, including development and implementation, also informed this research.

The research for this report began in October 2023. The technical research for this report concluded May 2025. The policy and governance research for this report concluded June 2025. This publication complements World Privacy Forum's ongoing research in relation to data governance and privacy, AI governance,[29] [30] [31] [32] identity ecosystems, and broadly, the governance of complex digital ecosystems.

---

27 C2PA, *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html Note: This report uses C2PA Technical Specifications 2.1 as its primary technical reference. A subsequent version, 2.2, was published in May 2025 after research and drafting of this report was complete. Version 2.2 does not alter or invalidate any technical information presented in this report.

28 c2pa-standard, Discord, Content Authenticity Initiative, https://discord.com/channels/983153151341371422/1129423756964659201/1199859509468868669.

29 Kate Kaye and Pam Dixon, *Risky Analysis: Assessing and Improving AI Governance Tools*, World Privacy Forum, December 2023, https://www.worldprivacyforum.org/2023/12/new-report-risky-analysis-assessing-and-improving-ai-governance-tools/.

30 Kate Kaye, *AI Governance on the Ground: Chile's Social Security and Medical Insurance Agency Grapples with Balancing New Responsible AI Criteria and Vendor Cost*, World Privacy Forum, November 1, 2024, https://www.worldprivacyforum.org/2024/11/ai-governance-on-the-ground-chiles-social-security-and-medical-insurance-agency-grapples-with-balancing-new-responsible-ai-criteria-and-vendor-cost/.

31 Kate Kaye, *AI Governance on the Ground: Canada's Algorithmic Impact Assessment Process and Algorithm has evolved*, World Privacy Forum, August 14, 2024, https://www.worldprivacyforum.org/2024/08/ai-governance-on-the-ground-series-canada/.

32 Kate Kaye, *Uncovering Areas for AI Governance Tools Refinement through Real-World Use Case Analysis from Canada, Chile and Singapore*: Proceedings of Fourth European Workshop on Algorithmic Fairness in Proceedings of Machine Learning Research, 294:135-151 July 2025, https://proceedings.mlr.press/v294/kaye25a.html.

Findings

Part II of this report discusses the findings of the research in detail. In summary, the research for the focal areas of the report — which queries the privacy, identity, and digital trust aspects of the system — indicates the following:

1.) The C2PA framework does relate to privacy and data protection. It allows for redaction or removal of certain types of data including data that may be considered sensitive such as location data. However, there are limits to what types of data can be redacted, and as stated in the C2PA Harms Modelling documentation,[33] redacted information "may still be accessible" in some cases and "inadvertent disclosure of sensitive information" can occur. C2PA also is intended to allow content creators and others in the media ecosystem to opt in or control whether certain provenance data such as personal identifiers or location information is included, but the C2PA Harms Modelling documentation also states that use of C2PA-enabled tools "may result in human rights violations," and acknowledges "the possibility of malicious actors, including potentially state actors, misusing or abusing the system."[34]

2.) The C2PA framework does relate to identity, including regarding digital identity assertions, a type of C2PA assertion that enables named humans or organizations to prove control over a digital identity and use that identity to assert their role in relation to a particular media file or piece of content.[35]

3.) The C2PA framework does relate to digital trust measurement. It enables a technical "Trust Model" process that uses C2PA metadata as signals for gauging the trustworthiness of media. The C2PA Trust Model involves use of trust lists[36] of certified authorities that determine whether or not trust signals should be marked as valid or fully trusted. The absence of C2PA metadata can negatively affect C2PA-based interpretations of trust.

See Part II: Discussion, for a detailed discussion of the focal area research questions and the baseline research questions.

---

33 *C2PA Harms Modelling*, C2PA Specifications, https://c2pa.org/specifications/specifications/1.4/security/Harms_Modelling.html.

34 "Privacy and security considerations / The following potential harms may still occur in specific implementations of the C2PA specifications:

- Inadvertent disclosure of information: C2PA claim generators may automatically add, or require to add, information to manifests that may be sensitive. Specification-compliant implementations may intentionally or unintentionally hide this feature. User experience guidance has been published alongside the specs to offer a clear acknowledgement of creator consent before a C2PA implementation can begin accumulating data.
- Redacted (deleted) information from Content Credentials may still be accessible: If a soft binding lookup is enabled or required by manifest stores, then previous versions of a manifest with sensitive information may be located.
- The use of C2PA-enabled tools and services in adverse legal or political situations may result in human rights violations: The C2PA specifications include features to protect the privacy of users, but this does not preclude the possibility of malicious actors, including potentially state actors, misusing or abusing the system." Text quoted from: *C2PA Harms Modelling*, C2PA Specifications, https://c2pa.org/specifications/specifications/1.4/security/Harms_Modelling.html.

35 "Identity Assertion The *C2PA technical specification* allows *actors* in a workflow to make cryptographically signed *assertions* about the produced *C2PA asset*. This signature is issued by the vendor whose software or hardware was used to create the C2PA assertions and the C2PA claim, which is why it is called the *C2PA claim generator*. This specification describes a *C2PA assertion* referred to here as the *identity assertion* that can be added to a *C2PA Manifest* to enable a *credential holder* to prove control over a digital identity and to use that identity to document the *named actor's* role(s) in the *C2PA asset 's* lifecycle." Text quoted from: *Identity Assertion*, 1.2-draft, DIF, Creator Assertions Working Group, Version 1.2 Draft from 3 July 2025. https://cawg.io/identity/1.2-draft/ . In the original text, the italicized words and terms link to the specifications. See https://cawg.io/identity/1.2-draft/ for the text with links. See also: *Content Credentials*, C2PA Technical Specification, C2PA, https://spec.c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html .

36 A C2PA Trust List is "A C2PA-managed list of X.509 certificate trust anchors that issue certificates to hardware & software signers that use them to sign claims." C2PA, *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html See 2.3. Core Aspects of C2PA / 2.3.15. C2PA Trust List.

Part II: Discussion

Part II is a detailed discussion based on the analysis of the technical research, original source documents, and source interviews as applied to the research queries articulated in the methodology.

Under the Hood: What C2PA Does and How It Does It

C2PA is an open standard [37] designed to build each component of an ongoing chain of content provenance information throughout the lifecycle of a piece of content, wherever it travels and however it is altered. It is designed to attach encoded, machine-automated or human-generated metadata (i.e. the data about the content itself) detailing the origins and modifications made to digital image, photo, video, audio, font or document files.

The C2PA process involves several layers of statements called "assertions," and validations of those assertions. As illustrated in Figures 2, 3, and 4, assertions in C2PA-based metadata reflect a variety of elements related to the content such as the make and model of a camera or camera lens used to produce a photograph, details of the types of editing modifications made to a video, information about the design software used to produce or alter the content, in addition to information about its creator, [38] the creator's location, time of content creation or modification, and more. While some types of C2PA assertions can be redacted, other types cannot.[39]

Assertions in C2PA metadata inside photographic content include descriptions such as "Digital capture sampled from real life" or "Human-edited media." When it comes to an AI-generated or AI-altered video, assertions include descriptions such as "Edited using Generative AI," "Algorithmically-altered media" or "Pure algorithmic media."[40] Because C2PA is an open standard, implementers can add their own new or custom types of C2PA-based assertions that others can adopt.

C2PA uses a two-step cryptographic hashing[41] and signing process as a validation method that promises resistance to metadata manipulation, tampering or erroneously fraudulent authentication of forged content.[42] The process is intended to ensure that C2PA metadata belongs with a particular content file and to validate modifications made to the content.

C2PA's cryptographic hashing and signing components:

- A cryptographic hashing process referred to as Hard Bindings uniquely identifies a piece of content or portions of that content.

---

37 Natalie R. Ortiz, *Federal Data Management: Issues and Challenges in the Use of Data Standards*, Congressional Research Service, April 29, 2024, https://sgp.fas.org/crs/misc/R48053.pdf. See p.13, Open Data Standards.

38 Adobe, *Connect Accounts for Creator Attribution*, Photoshop, May 23, 2023, https://helpx.adobe.com/photoshop/using/connect-accounts.html. See also See What Are Content Credentials?, Adobe, Content Credentials, (October 14, 2024), https://helpx.adobe.com/creative-cloud/help/content-credentials.html.

39 See the *Privacy in C2PA* section of this report for details on redaction in C2PA.

40 C2PA incorporates standard NewsCodes and digital source types from International Press Telecommunications Council. See: *NewsCodes Scheme (Controlled Vocabulary), Information Technology for News*, IPTC, https://cv.iptc.org/newscodes/digitalsourcetype/.

41 *Cryptographic hash function,* National Institutes of Technology, Security Technology Resource Center, Glossary, https://csrc.nist.gov/glossary/term/cryptographic_hash_function.

42 Neal Krawetz, a media forensics researcher, has extensively evaluated and tested vulnerabilities of C2PA. For example, see Neal Krawetz, *C2PA from the Attacker's Perspective*, The Hacker Factor Blog, May 9, 2024, https://www.hackerfactor.com/blog/index.php?/archives/1031-C2PA-from-the-Attackers-Perspective.html, and Neal Krawetz, *C2PA and Authenticated Disinformation,* The Hacker Factor Blog, October 15, 2024, https://www.hackerfactor.com/blog/index.php?/archives/1046-C2PA-and-Authenticated-Disinformation.html.

- Statements called Assertions reflect various elements of the content, including aspects of its origin, and actions taken in the modification of the content.
- Signers such as hardware and software makers, publishers or app providers create Assertions and sign Claims that encompass all the Assertions about a content file at a given time.
- Trusted certificate authorities issue certificates to Claim Signers to indicate Signer trustworthiness.
- Claims are cryptographically hashed and signed automatically by a signing tool, producing a Claim Signature.
- Assertions, Claims and Claim Signatures are bundled into a Manifest, or a Manifest Store that is embedded inside or attached externally to the associated content file.
- Validators assess and validate the Manifest and Claim Signatures associated with the content file, checking that their values (or checksums[43]) match only those associated with that specific content file, and checking whether the data is well-formed, valid, trusted, or unknown.

C2PA uses a two-step signing approach:

- Step One uses checksums to assess whether existing metadata associated with the content, or claims covering the content, match what is in the manifest. If the metadata or visual content has been altered, the checksums won't match.
- Step Two uses cryptographic signatures to assess whether checksums have been altered.

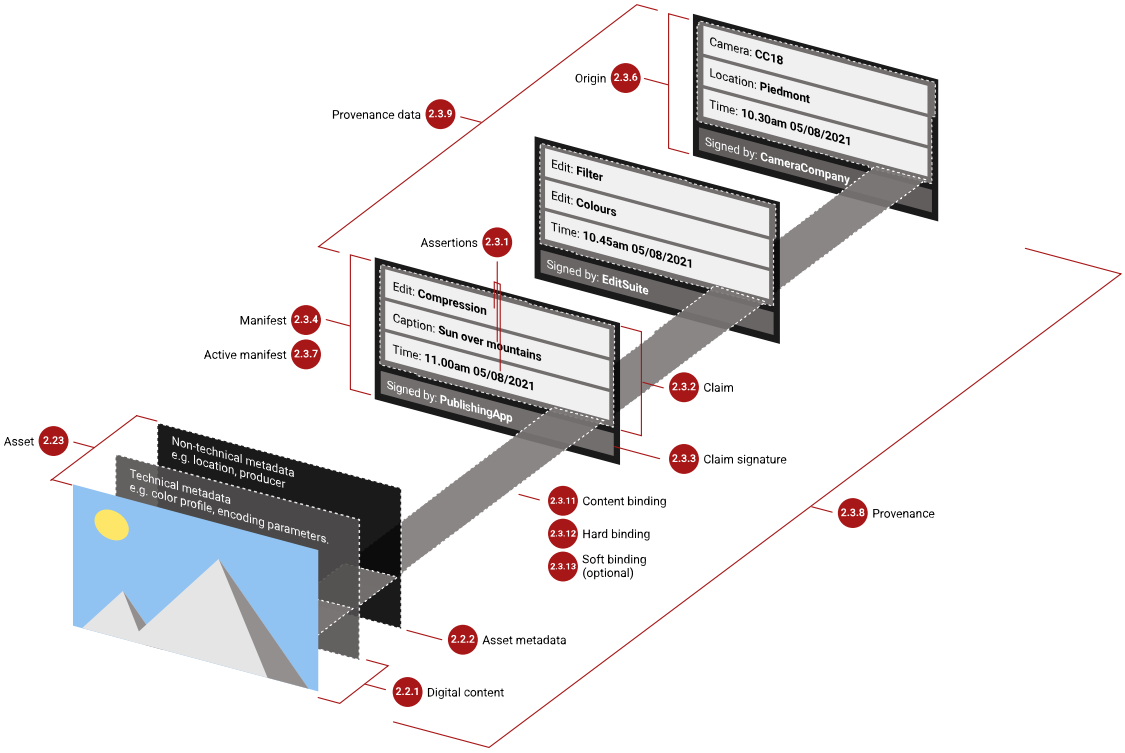


Figure 2: Image of elements of C2PA from the C2PA Technical Specifications (C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html)

43 *Cryptographic checksum*, National Institutes of Technology, Security Technology Resource Center, Glossary, https://csrc.nist.gov/glossary/term/cryptographic_checksum

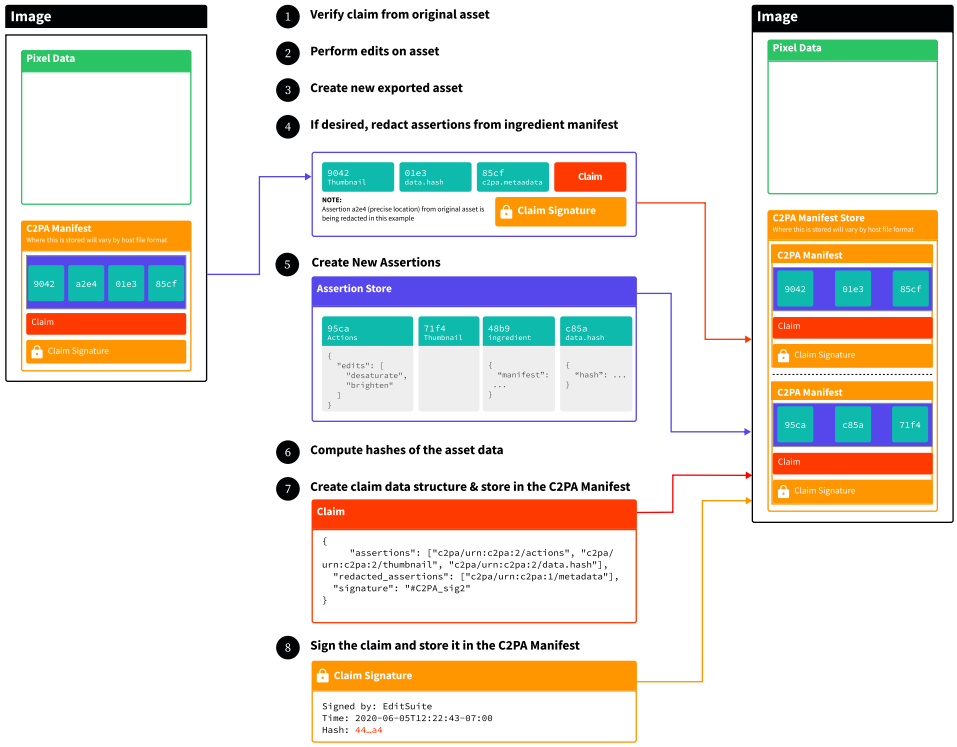


Figure 3: Image from C2PA Technical Specifications (C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html (https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html)

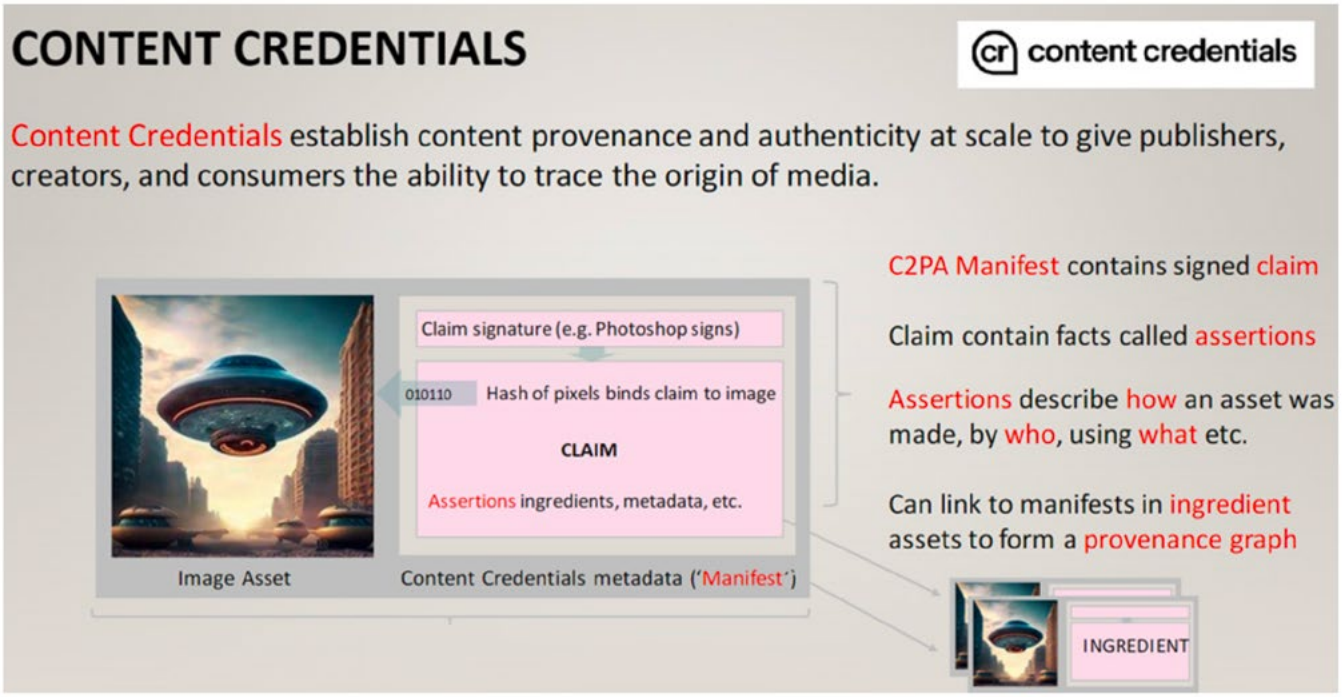


Figure 4: This image is from a presentation given at the International Telecommunication Union (ITU)'s 2024 Standards for AI Watermarking and Multimedia Authenticity Workshop. See International Telecommunication Union (ITU), Detecting deepfakes and generative AI: Report on standards for AI watermarking and multimedia authenticity workshop: The need for standards collaboration on AI and multimedia authenticity, 2024, https://www.itu.int/hub/publication/t-ai4g-ai4good-2024-7/.

Validation States

The absence of C2PA metadata or lack of recognition of provenance metadata also carries weight in the C2PA process. If validated signatures are not present, C2PA design considers this a break in the provenance chain. For instance, if a content creator such as a photographer uses a camera or photo editing software that does not feature C2PA capabilities, the provenance chain weakens or breaks. Removal of C2PA metadata also breaks the content provenance chain. A broken content provenance chain is used by C2PA as a signal of possibly invalid or untrusted alterations to the content.

C2PA assigns one of three validation states —well-formed, valid, or trusted —to a manifest as an indication of its level of trustworthiness. C2PA validators also can also label a manifest as "unknown" if they encounter a manifest that uses constructs from a version of C2PA that they do not support.

The C2PA Metadata Attached to Content and How It Is Used

The C2PA process is designed to generate —often automatically —a wide array of metadata.[44] For example, C2PA technical specs allow for the insertion of metadata that directly and persistently attaches specific entities, people, and their identities to content, showing precise locations of people and entities involved in content creation and modifications, as well as dates and times of content creation and modification.

One way that C2PA metadata is used today is in consumer-facing presentations overviewing content history that are displayed when people experience that content. These presentations are often referred to as content labels. However, it is important to recognize the extent to which C2PA metadata is accessible behind the scenes beyond its use as a labeling system intended for humans to experience. Many entities involved in creating and distributing media or in other parts of the C2PA validation process have access to far more C2PA metadata than everyday content consumers typically access. By design, any entity or external, interoperable system that plays a role in the C2PA content provenance chain is intended to have access to every bit of granular C2PA metadata, barring redacted data. Entities with access to C2PA metadata can be expected to access and use it in automatic, machine-readable form.

These are some examples of the types of metadata produced using tools that adopt C2PA:

- Geographic and temporal metadata indicating GPS latitude, longitude, altitude, dates and timestamps,
- Software, hardware or device related metadata indicating camera make or model, or lens make or model,
- Audio, image and photo related metadata indicating changes in playback speed of a video or audio track, areas of image crops or deletions, or color adjustments,
- 3D depthmap[45] related metadata providing 3D descriptions of a scene captured by a camera and representing the distance or depth information for each pixel in the scene,
- Metadata indicating other types of changes made to content categorized as "editorial transformations" or "non-editorial transformations,"
- Metadata indicating whether content creation or modification involved use of AI or algorithmic systems,thmic systems,[46] and

---

44 Note that while C2PA technical specs distinguish between metadata and other types of data generated using C2PA-based processes, for the sake of relative simplicity this report refers to all types of C2PA-related data attached to content as "metadata."

45 *Depth Maps: How software encodes 3D space*, Looking Glass, Jan. 19, 2023, https://blog.lookingglassfactory.com/depth-maps-how-software-encodes-3d-space/

46 Telephone interview with Carl Seibert, metadata expert, December 18, 2024 by author. According to Seibert, there are "…all kinds of ambiguity and conflicting opinions" regarding interpretations of standardized metadata codes. Interpretations of C2PA labels and whether or how to apply them, such as in relation to whether AI was used to alter an image in a trivial or non-

- Metadata reflecting inclusion of watermarks such as those used to reflect copyright.

C2PA System Interoperability

At this early stage of C2PA adoption, early prototypes and implementations offer some examples of C2PA interoperability, indicating the types of systems that communicate, interconnect and share C2PA related data. A wide variety of systems [47] already interoperate in the burgeoning C2PA ecosystem including camera hardware and software, a mobile phone chipset and platform, content provenance labeling systems, a content delivery network, generative AI platforms, social media platforms, search and digital advertising platforms, publishing and media management systems, business documents software, audio software, digital watermarks, data governance tools, and an entertainment and athlete talent identification system.

According to a member of an AI company technical group who has implemented C2PA, the common practice of removing metadata from media files when they are uploaded or shared is the primary obstacle to C2PA interoperability.[48]

External C2PA Metadata Repositories

In addition to embedding C2PA metadata inside content files, it can be stored externally in a manifest repository, opening up additional points of data connectivity and system interoperability.

For instance, a broadcast media company that enabled C2PA throughout its media content production and distribution processes stored C2PA metadata externally in a "sidecar"[49] outside the content file. The company said in the future it could store C2PA metadata inside a content file in a "media wrapper."

Another early prototype that incorporates C2PA involves interoperability with external content registries.[50] The prototype —a process for connecting, storing and accessing C2PA manifest metadata externally —employs an external registry for metadata storage and large-scale retrieval of content and its metadata included in C2PA manifests. The system stores manifests within a distributed InterPlanetary distributed File System (IPFS), making them searchable and referenced by a hashed URL.

And, a large content delivery network[51] automatically stores all transformed images incorporating C2PA

---

trivial way, will vary depending on interpretation by entities implementing C2PA.

47 More information about these applications of C2PA can be found in the *Early Examples of C2PA in Prototypes and Products* section of this report.

48 Virtual interview with Michael Lampe, a member of OpenAI's technical group, by the author, May 6, 2025. According to Michael Lampe, "At the end of the day the metadata question is the key blocker to being interoperable because if the data is not transferred from place A to place B then there's no question of interoperability even [being feasible] because you don't have a thing you could even make a decision on."

49 AWS Architecture Center, *AWS Innovation with Sinclair, AWS Innovation Ambassadors* June 2024, https://aws.amazon.com/podcasts/innovation-ambassadors/ia-podcast-ep-47-sinclair/?podcast-list-wide.sort-by=item.additionalFields.EpisodeNum&podcast-list-wide.sort-order=desc.

50 Kar Balan, Alex Black, Andrew Gilbert, Simon Jenni, Andy Parsons, John Collomosse, *DECORAIT - DECentralized Opt-in/out Registry for AI Training, In Proceedings of the 20th ACM SIGGRAPH European Conference on Visual Media Production (CVMP '23)*. Association for Computing Machinery, New York, NY, USA, Article 4, 1–10, 2023, https://doi.org/10.1145/3626495.3626506 . See also Frances Liddell, Ella Tallyn, Evan Morgan, Kar Balan, Martin Disley, Theodore Koterwas, Billy Dixon, Caterina Moruzzi, John Collomosse, and Chris Elsden, *ORAgen: Exploring the Design of Attribution through Media Tokenisation*. In Designing Interactive Systems Conference (DIS Companion '24), July 01–05, 2024, IT University of Copenhagen, Denmark. ACM, New York, NY, USA, https://doi.org/10.1145/3656156.3663693.

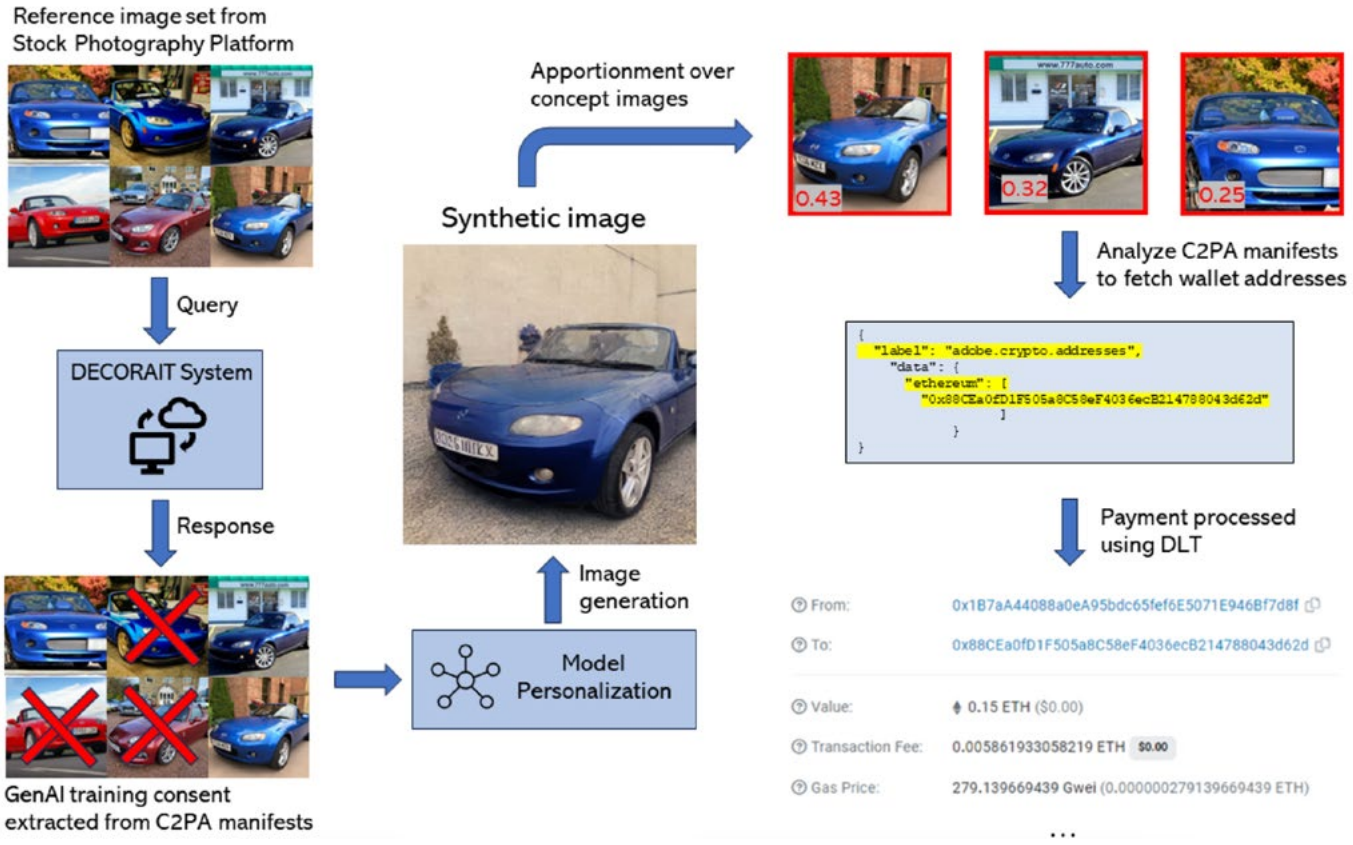
51 Web Technology Surveys, *Historical trends in the usage statistics of reverse proxy services for websites*, May 2024, https://w3techs.com/technologies/history_overview/proxy/all.

metadata on its global network.[52]

Digital Wallets and NFT Connections

One type of digital design software allows content creators to include their digital wallet information in C2PA-based metadata by connecting to external wallet and NFT — or *non fungible token* —systems.[53] The process involves using the design software to log into digital wallet and NFT accounts to add a crypto wallet address to C2PA metadata and mint an image as an NFT for sale.

The aforementioned early prototype[54] enabling metadata storage in an external repository also involves a process that extracts digital wallet information from C2PA metadata[55] to provide "royalty-like payments" or "rewards" to image creators and owners who consent to use of their images in AI model training. See an illustration of this process in Figure 5.



52 Jen Tse, Cloudflare becomes the first major content delivery network to implement Content Credentials, February 3, 2025, https://contentauthenticity.org/blog/cloudflare-becomes-the-first-major-content-delivery-network-to-implement-content-credentials.

53 Adobe, *How to use Content Credentials (Beta) in Photoshop for NFTs*, Creative Cloud, Discover, https://creativecloud.adobe.com/discover/article/how-to-use-content-credentials-beta-in-photoshop-for-nfts.

54 Kar Balan, Alex Black, Andrew Gilbert, Simon Jenni, Andy Parsons, John Collomosse, *DECORAIT - DECentralized Opt-in/out Registry for AI Training, In Proceedings of the 20th ACM SIGGRAPH European Conference on Visual Media Production* (CVMP '23). Association for Computing Machinery, New York, NY, USA, Article 4, 1–10, 2023, https://doi.org/10.1145/3626495.3626506 .

55 The prototype incorporates a training-mining assertion included in an early version of C2PA, version 1.3. Since then, technical specifications for assertions related to content creator preferences regarding use of content for AI model training have moved from C2PA to CAWG. See Creator Assertions Working Group, https://cawg.io. (Note that the research in this report is based on C2PA version 2.1.)

Figure 5: Image from Kar Balan, Alex Black, Andrew Gilbert, Simon Jenni, Andy Parsons, John Collomosse, DECORAIT - DECentralized Opt-in/out Registry for AI Training, In Proceedings of the 20th ACM SIGGRAPH European Conference on Visual Media Production (CVMP '23) https://doi.org/10.1145/3626495.3626506 .

Watermarking Systems

C2PA interoperates with watermarking technology that is intended to ensure original metadata is accessible even if it is stripped from the content file or manipulated to insert fraudulent information about creators, content origins or edits. In one example, a digital watermark containing a reference to the C2PA manifest is added to a content file. [56] The watermark system detects the watermark on the content, verifying a match to the externally-stored manifest.

Data Processing and Media Pipeline Connections

Systems that enable computational data processing —including for training, testing, and deploying machine learning and AI models —also have been used in the C2PA validation process, to create C2PA manifests, and for C2PA manifest storage.[57]

In the aforementioned broadcast media company prototype, the company used an external data processing and storage system to incorporate C2PA in its own media production and content management systems, and to facilitate adoption of C2PA among other entities it works with including for media production or media asset management, video editing, and video format conversion.[58] A prototyping engineer who worked on the project suggested that a benefit was having C2PA data available in a file for use downstream.[59]

Identity System Connections

Regarding C2PA interoperability with identity ecosystems,[60] a variety of private third-party and government identity systems could be used in conjunction with identity assertions expected to be components of the C2PA workflow.[61]

---

56 Digimarc, *Your Digital Assets Are At Risk,* Digimarc Validate, See video https://www.digimarc.com/products/digital-content-authentication.

57 AWS Solutions Library, *Guidance for Media Provenance with C2PA on AWS,* Guidance, July 23, 2024, https://aws.amazon.com/solutions/guidance/media-provenance-with-c2pa-on-aws/, also see https://github.com/aws-solutions-library-samples/guidance-for-media-provenance-with-c2pa-on-aws/blob/main/README.md.

58 NewscastStudio, *Sinclair partners with Embrace for media orchestration in the cloud,* March 27, 2024, https://www.newscaststudio.com/2024/03/27/sinclair-partners-with-embrace-for-media-orchestration-in-the-cloud/. See more about the project below in the Early Examples section.

59 AWS Architecture Center, *AWS Innovation with Sinclair*, AWS Innovation Ambassadors, June 2024, https://aws.amazon.com/podcasts/innovation-ambassadors/ia-podcast-ep-47-sinclair/?podcast-list-wide.sort-by=item.additionalFields.EpisodeNum&podcast-list-wide.sort-order=desc.

60 Identity systems and ecosystems are well-understood and studied at this point. However, they are quite complex. This complexity reaches across the domains of legal, policy, technical, data flows, sectors, and jurisdictions. C2PA and CAWG activities relating to identity will likely be interfacing / intersecting with multiple aspects of existing identity standards, ecosystems, and policy, governmental, and other structures. Due to the nuanced and multifaceted interactions of the underlying technical and legal structures of modern ID ecosystems today, particularly legal or foundational identity systems, there can be tension points unique to each country or ID system that need to be resolved with bespoke policy and technical design and guidance consisting of a variety of intersecting guardrails that can range from implementing identity-specific regulations, technical protections, Memorandums of Understanding (MOUs), to requirements for robust audits, Data Protection Impact Assessments by Data Protection Authorities, and public consultations, among other practices.

61 Virtual interview with Andy Parsons, Senior Director Content Authenticity at Adobe, May 5, 2025 by author. In the interview, Parsons referenced various identity systems and technical identity related projects including decentralized wallets and identity related work at Decentralized Identity Foundation, Apple and Google wallets, Open Wallet Foundation, and Estonia's government-issued digital ID system, and said, "Those can probably all be embraced and will be embraced by CAWG for signing

In one recent example, a C2PA-based system allowing media creators to attach identity information to their media files works in conjunction with a social media platform identity verification system[62] that can incorporate multiple third-party identity service systems. [63]

In addition, some C2PA implementations involve experimentation with connecting information from external identity systems to C2PA-based systems. For instance, C2PA is used in a proof of concept for a system that creates, documents and measures consent-based digital replicas of "notable" legal and natural public figures such as actors and athletes; the system embeds cryptographic metadata in a watermark or fingerprint as a C2PA Content Credential for the digital replica identity in a multi-party workflow and detects C2PA metadata.[64]

Identity in C2PA is discussed in further detail under the heading *Identity in C2PA* in this report.

C2PA Data Storage, Access, Control and Durability

It matters where content provenance metadata lives. The storage location of metadata affects how it is controlled, how it can be accessed, by whom, and under what circumstances. The crafters of C2PA prefer that content provenance information be "durable." That calls for C2PA metadata to be maintained and accessible throughout the lifecycle of a content file and never detached from the content file no matter how it is altered, or where it travels. The durability approach calls for implementers to keep the accumulating elements of an ongoing C2PA chain of content provenance in multiple places or multiple forms, both inside and outside the content file in the following ways:

- Embedding C2PA metadata locally inside a content file,
- Storing C2PA metadata externally for backup, [65] [66]
- Providing a second form of backup using watermarks or fingerprinting, a computational process that uniquely codes content, allowing it to be matched to recover the content with intact C2PA metadata from an external databaseernal database[67] or manifest storage repository. [68]

---

purposes." Also see this report's *Identity in C2PA* section which provides details on CAWG or Creator Assertions Working Group, a C2PA-based framework used to assert identity.

62 Andy Parsons, *Adobe Content Authenticity, now in public beta, helps creators secure attribution*, Adobe Blog, News, April 24, 2025, https://blog.adobe.com/en/publish/2025/04/24/adobe-content-authenticity-now-public-beta-helps-creators-secure-attribution. Also see Oscar Rodriguez, LinkedIn, Building Trust in the Digital Age: LinkedIn's New Verified on LinkedIn Service, April 24, 2025, https://www.linkedin.com/pulse/building-trust-digital-age-linkedins-new-verified-oscar-rodriguez-nhdre/.

63 LinkedIn's verification system partners with identity verification services including CLEAR, DigiLocker, Microsoft Entra Verified ID and Persona. See LinkedIn, *Verifications on your LinkedIn profile,* Help, https://www.linkedin.com/help/linkedin/answer/a1359065. See also LinkedIn, *Workplace verification via Microsoft Entra Verified ID*, Help, January 2025, https://www.linkedin.com/help/linkedin/answer/a1453679.

64 HAND Labs, Digital Replicas and Talent ID: Provenance, Verification and New Automated Workflows, 2024, https://handidentity.com/hand-labs/.

65 A broadcast media company conducted a prototype project involving external storage of C2PA metadata in a "sidecar" outside the content file, and planned for storage of C2PA metadata inside the content file in a "media wrapper" in future implementations. See AWS Architecture Center, *AWS Innovation with Sinclair,* AWS Innovation Ambassadors June 2024, https://aws.amazon.com/podcasts/innovation-ambassadors/ia-podcast-ep-47-sinclair/?podcast-list-wide.sort-by=item.additionalFields.EpisodeNum&podcast-list-wide.sort-order=desc.

66 A large content delivery network that has adopted C2PA for embedding and signing metadata associated with image origin and transformations can automatically cache or store C2PA-enabled images managed by its customers on its global network. See Cloudflare, *Cloudflare Docs, Get Started*, https://developers.cloudflare.com/images/get-started/#enable-transformations.

67 *Content Authenticity and Image Fingerprinting: Q&A with John Collomosse*, News, Adobe Research, April 4, 2023, https://research.adobe.com/news/content-authenticity-and-image-fingerprinting-qa-with-john-collomosse/

68 *C2PA Soft Binding API, Technical Specifications,* C2PA, https://c2pa.org/specifications/specifications/2.2/softbinding/

A key reason C2PA designers want implementers to enable multiple storage methods is for C2PA durability: The content provenance chain can be weakened or broken if C2PA metadata is removed, which is of particular risk through the common practice of metadata stripping. If metadata is stripped when content files are published or distributed,[69] the provenance chain —and therefore the promise of tracing content provenance throughout the digital media ecosystem —is disrupted. If C2PA metadata is not attached to a media file, C2PA validators might label the validation state [70] of content metadata as "unknown" or merely as "well-formed" but not as fully "valid" or "trusted."

C2PA as a New Data Source to Store and Use

The desire to store C2PA metadata in multiple places to protect against metadata loss or manipulation gives external C2PA metadata storage services important connections to C2PA metadata.[71]

Because C2PA has no central host, there is no universal, centralized repository of all C2PA metadata. However, as noted earlier, the design of the framework calls for any and every entity that touches the technical back-end of a content file to have access to every piece of C2PA metadata associated with the file.

Few external storage repositories for C2PA metadata exist today, so many C2PA manifests may only be embedded in media files rather than backed up externally.

Identity in C2PA

Some artists, content creators, media and entertainment brands, businesses, and government agencies[72] [73] want

---

Decoupled.html .

69 *Social Media Sites Photo Metadata Test Results 2019,* International Press Telecommunications Council (IPTC) 2019, https://iptc.org/standards/photo-metadata/social-media-sites-photo-metadata-test-results-2019/. See also: *State of image metadata in 2018*, Imatag, May 11, 2018, https://www.imatag.com/blog/state-of-image-metadata-in-2018 . See also: *What Are Content Credentials?*, Adobe, Content Credentials, October 14, 2024, https://helpx.adobe.com/creative-cloud/help/content-credentials.html .

70 C2PA assigns one of three validation states -- well-formed, valid or trusted -- to a manifest to indicate its level of trustworthiness, in part as a way to address potential adverse impact on creators or content missing C2PA-validated provenance information. Validators assess and validate the Manifest and Claim Signatures associated with the content file, checking whether the data is well-formed, valid or trusted. Validators also can choose to mark a manifest as having "unknown" provenance if it uses constructs from a version of the specification that the validator does not support. See: *Getting started with content credentials, CAI open source SDK*, Content Authenticity Initiative, https://opensource.contentauthenticity.org/docs/getting-started/ . See also: *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, Linux Foundation Projects, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html . See 14.3., Validation states.

71 Adobe software tools use the company's public, persistent cloud service to store C2PA-based metadata, to enable recovery of metadata that can be stripped when media files are uploaded and shared, and to reduce file sizes. See: Getting Started with Content Credentials, Adobe CAI open source SDK, See: Storing a Manifest in the Cloud, https://opensource.contentauthenticity.org/docs/getting-started/#storing-a-manifest-in-the-cloud. Also see Adobe Firefly, Content Credentials Overview, Adobe Help Center, February 12, 2025, https://helpx.adobe.com/firefly/get-set-up/learn-the-basics/content-credentials-overview.html.

72 *Content Credentials: Strengthening multimedia integrity in the generative AI era*, Joint publication: Australian Signals Directorate's Australian Cyber Security Centre; Canadian Centre for Cyber Security, Communications Security Establishment; United Kingdom National Cyber Security Centre; National Security Agency, United States of America; Document U/OO/109191-25/PP-25-0336l. January 2025 V. 1.0, https://media.defense.gov/2025/Jan/29/2003634788/-1/-1/0/CSI-CONTENT-CREDENTIALS.PDF .

Government of Canada, *Joint guidance on content credentials and strengthening multimedia integrity in the generative artificial intelligence era*, Canadian Centre for Cyber Security, January 29, 2025, https://www.cyber.gc.ca/en/news-events/joint-guidance-content-credentials-and-strengthening-multimedia-trust-generative-artificial-intelligence-era .

73 Implementation of data standards for U.S. Federal data, including challenges associated with data standards management and

to use C2PA to mark, trace, and protect their content, to reflect intellectual property rights, to indicate AI training data-related consent or restrictions, or to help distinguish their content as trustworthy. Those types of C2PA users may want to incorporate their identities[74] into C2PA provenance metadata.[75]

For instance, while working on a proof-of-concept for a product feature using C2PA, an engineer at an image and video management platform company suggested that media organizations would want to connect C2PA metadata assertions about media to their identities and reputation.[76]

Indeed, some media brands or news outlets believe the identity of the media brand, company or media creator is the most important signal in C2PA metadata establishing trustworthiness.[77] In one recent example, a C2PA-based system allows media creators to attach identity information to their media files and works in conjunction with a social media platform identity verification system to ensure proper attribution for their work.[78] Another early implementation of C2PA aimed at content creators and owners spotlights use of "name, social media handles, or other identity-related information" in C2PA-based Content Credentials associated with content management software.[79]

Shifting Identity Outside Core C2PA Specifications

Still, in part because C2PA design goals aim to avoid identity exposure of content creators such as marginalized or at-risk content creators, enabling identity signals in C2PA is not without controversy. Assertions reflecting

---

data governance is a well-researched area. See Natalie R. Ortiz, *Federal Data Management: Issues and Challenges in the Use of Data Standards*, Congressional Research Service, April 29, 2024, https://sgp.fas.org/crs/misc/R48053.pdf.

74 Regarding a precise definition of the term "identity" within the C2PA technical specifications and the CAWG technical specifications reviewed for the report, none of the documents reviewed were found to explicitly define the term "identity" as a noun. As noted later in this section of the report, *Identity in C2PA*, assertions reflecting "identified humans" were included in early versions of C2PA technical specifications but that changed in January 2024 when identity-related assertions were removed, as reflected in the 2.0 version of C2PA published at that time. All versions published before C2PA 2.0 (1.0, 1.1, 1.3, and 1.4), include a description of "Identity in Assertions," which seems to be the closest to a direct definition of the term "identity" as a noun in the C2PA universe. The brief section states, "Some assertions (such as stds.schema-org.CreativeWork) allow a person's identity to be associated in a defined way with the asset. This identity is purely scoped via the definition in each assertion and does not imply any larger involvement or responsibility for any assertion made in the claim, or the asset itself. All assertions, as stated below, are made by a signer." C2PA, *C2PA Technical Specifications 1.4, Technical Specifications, Content Credentials*, C2PA, https://spec.c2pa.org/specifications/specifications/1.4/specs/C2PA_Specification.html#_identity_in_assertions. See15.5. *Identity In Assertions*.

75 Some entities promoting C2PA recognize the goals of artists and companies that want to add provenance information to their media assets, and as a result, emphasize the benefits of embedding identity information in content. See: *What Are Content Credentials?*, Adobe, Content Credentials, October 14, 2024, https://helpx.adobe.com/creative-cloud/help/content-credentials.html. From the document: "Content Credentials allow you to attach your identity and contact information to your work, giving people more ways to find and connect with you when they encounter your content online."

76 Eric Portis, senior developer experience engineer at Cloudinary confirmed to WPF in May 2025 that he wrote the following statement on the C2PA related discussion thread on the Content Authenticity Initiative-hosted Discord server: "So, if I'm an organization who is invested in putting verifiable facts out into the world, one of the most powerful tools I have is my identity and reputation, which I would like to tie to those assertions." See: eric.portis, *c2pa-standard, Discord, Content Authenticity Initiative*, January 24, 2024, https://discord.com/channels/983153151341371422/1129423756964659201/1199859509468868669.

77 Virtual interview with Andy Parsons, Senior Director Content Authenticity at Adobe, May 5, 2025 by author. In the interview, Parsons noted that "...the most important trust signal" for one international media company's constituents is the company's identity.

78 Andy Parsons, Adobe Content Authenticity, now in public beta, helps creators secure attribution, Adobe Blog, News, April 24, 2025, https://blog.adobe.com/en/publish/2025/04/24/adobe-content-authenticity-now-public-beta-helps-creators-secure-attribution. Also see Oscar Rodriguez, LinkedIn, Building Trust in the Digital Age: LinkedIn's New Verified on LinkedIn Service, April 24, 2025, https://www.linkedin.com/pulse/building-trust-digital-age-linkedins-new-verified-oscar-rodriguez-nhdre/.

79 Adobe Experience Manager, *Content Credentials, Documentation*, AEM as a Cloud Service, (October 7, 2024), https://experienceleague.adobe.com/en/docs/experience-manager-cloud-service/content/assets/assets-view/content-credentials.

"identified humans" were included in early versions of C2PA. That changed in January 2024 when identity-related assertions[80] were removed, as reflected in the 2.0 version of C2PA published at that time. That 2.0 version removed "identified humans" from assertion metadata, and reduced use of the term "actor," which no longer represented humans and organizations. In the technical specifications document, C2PA drafters called this a "philosophical change" and "a significant departure from previous versions."[81]

The drafters of C2PA recognize that some of its implementers and users will want to include human or business-related identifiers[82] in content provenance metadata. Just a month after identity was removed from the core C2PA specifications, a new group known as the Creator Assertions Working Group or CAWG[83] was established,[84] in part to serve as a home for identity related technical specifications development related to C2PA.[85]

The decision to remove identity factors from C2PA and move them to CAWG was a decision intended to distinguish metadata reflecting how a media file was made from metadata reflecting who or what entity created or produced it, and to ensure that what C2PA drafters call "core technical provenance" was distinct from identity related matters.[86]

CAWG became a working group within the Decentralized Identity Foundation (DIF) in March 2025.[87]

---

80 The term "identity assertions" is used in an introductory section related to Identity Assertions in the CAWG technical specifications which states, "This specification describes a *C2PA assertion* referred to here as the identity assertion that can be added to a *C2PA Manifest* to enable a *credential holder* to prove control over a digital identity and to use that identity to document the *named actor's* role(s) in the *C2PA asset's* lifecycle." (*Creator Assertions Working Group, Identity Assertion 1.1*, Decentralized Identity Foundation, https://cawg.io/identity/1.1/ See Identity Assertion.) Digital identity can be applied to an entity. An existing international standard, ISO/TC 68, *Enabling Secure Digital Identity for Organizations,* focuses on decentralized digital identity verifications for entities and provides important context as to the normative standards that exist in this area of ensuring integrity across the full identity chain of trust. In particular, see Global Legal Entity Identifier Foundation (GLEIF), which manages the Global LEI system that is working on provenance of verifiable legal entity identifiers, or LEIs: https://www.gleif.org/en/about/this-is-gleif and https://www.gleif.org/en/organizational-identity/overview. See also ISO 17442-2:2020, *Financial services - Legal Entity Identifier* https://www.iso.org/standard/79917.html. The ISO standards cited here are primarily related to financial crime and corporate fraud. These are cited as the Global Legal Entity Identifier is an important and existing model that has intersections with entity identity.

81 "This version represents a significant departure from previous versions. It no longer has any references to actors as humans or organizations, they can only be hardware or software entities." As quoted in: *C2PA Technical Specifications 2.0, Technical Specifications, Content Credentials*, C2PA https://spec.c2pa.org/specifications/specifications/2.0/specs/C2PA_Specification.html. See Version History 2.0 - January 2024.

82 An important ISO standard that addresses human and non-human actors in identity systems is ISO/IEC 27001:2022, *Information security, cybersecurity and privacy protection — Information security management systems — Requirements, Annex A Control 5.16- Strengthening Identity Management.* https://www.iso.org/standard/27001 . This standard articulates how "human" and "non-human" entities are defined and how they should be managed regarding information and security governance.

83 Creator Assertions Working Group, Decentralized Identity Foundation, https://cawg.io.

84 The first CAWG meeting was held on February 20, 2024. See 20 February 2024, Creator Assertions Working Group, Decentralized Identity Foundation, Meeting Notes, https://cawg.io/meeting-notes/2024-02-20/.

85 See *Getting Started with Content Credentials*, CAI open source SDK, Identity, Content Authenticity Initiative, https://opensource.contentauthenticity.org/docs/getting-started/#identity . From the document: "The Creator Assertions Working Group (CAWG) is developing a technical specification for an identity assertion for use in the C2PA ecosystem." Document cited was viewed in 2025. A changelog for the SDK is available on the C2PA GitHub: https://github.com/contentauth/c2pa-rs/blob/main/CHANGELOG.md.

86 Virtual interview with Andy Parsons, Senior Director Content Authenticity at Adobe, May 5, 2025 by author. In the interview, Parsons said that it was "quite a process to establish CAWG [and] get it to be part of the Decentralized Identity Foundation."

87 The CAWG in DIF is chaired by Eric Scouten, Identity Standards Architect at Adobe. See Decentralized Identity Foundation, *Welcoming Creator Assertions Working Group to DIF*, March 4, 2025, https://blog.identity.foundation/welcome-cawg/.

How Identity Is Verified and Linked to Content In CAWG

CAWG technical specifications document an identity assertion,[88] a type of C2PA assertion that can be added to a C2PA manifest[89] to enable a Named Actor (that is, a named human, organization of humans, or non-human software or hardware device) to prove control over a digital identity and then use that identity to assert their role in relation to a particular content file.[90]

A variety of private third-party and government identity systems[91] could be used in conjunction with identity assertions created by CAWG, which when implemented will be components of the C2PA workflow.[92]

---

88 In C2PA and CAWG technical specifications there are several instances of the term "identity" used as a qualifier. One such example is the term "identity assertion" used in the CAWG specifications. "Identity assertion" is a well-known term of art in the broader identity field with a standardized definition. Specifically, "identity assertion" is an ISO-defined term (ISO 24760). There are many ISO standards that address identity, and there are multiple terms of art utilized in the CAWG specifications.

89 "Once the signature is obtained, the identity assertion can be created and added to the C2PA Manifest's assertion store, and then referenced in the C2PA claim." As cited in: Decentralized Identity Foundation, Creator Assertions Working Group, https://cawg.io. See Identity Assertion 1.1, See also 6.2. Creating the assertion.

90 Decentralized Identity Foundation, Creator Assertions Working Group, https://cawg.io. See Identity Assertion 1.1.

91 In order to situate the CAWG identity discussion in its broader context, this note provides additional background information regarding identity ecosystems. "Identity" encompasses different types of identity which operate in and transverse across extremely complex socio-technical-legal ecosystems which can and do differ. For example, there is "legal identity" (sometimes also called "foundational identity"), there is also "functional identity," and a third category that can be characterized as "digital foundational identity." An example of a classic legal identity system is, for example, a government that issues a unique and verified national ID, either digitally or otherwise, to be used for general purposes and typically across sectoral domains. Functional ID, on the other hand, is focused on a single sector, or is designed as an ID to be used for a limited activity. An example of a functional ID system is a voter ID card or digital system that is used only for voting. Digital forms of foundational ID are another category, and have diverse models. Some "born digital" foundational IDs are centralized in structure, some are federated and some are decentralized. Tokenization can be present in many forms of digital ID ecosystems. An example of this is the current-day version of India's Aadhaar biometric ID ecosystem, which at 1.4 billion - plus enrollees is arguably the largest "born digital" foundational ID in the world. Aadhaar is a centralized system, but with federation and tokenization that has been added in, among other privacy-preserving measures. The CAWG framework analyzed in this report fits into the broad category of purely digital ID, which also includes what the World Bank has described as "open-market" digital ID, a rapidly evolving category. In this general category, private sector entities might utilize IDs that are derived (called derived digital identity) from a verifiable government or other ID. The open market digital ID ecosystems can and often do differ widely. For further information on these topics, see: United Nations Definition of Legal identity as found in *Vital Statistics and Identity Management Systems: Communication for Development,* UN Department of Economic and Social Affairs, Statistics Division, (UN Legal Identity Expert Group) 2022. See (2.) Definitions, Legal Identity, Para. 16. https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/crvs/crvs-Idm-E.pdf . For a complete overview of digital identity, issues, and challenges, see Dr. Joseph Atick, *Digital Identity Toolkit,* World Bank Group. https://documents1.worldbank.org/curated/en/147961468203357928/pdf/912490WP0Digit00Box385330B00PUBLIC0.pdf . For a discussion of new models and frameworks for digital foundational ID, see *Practitioner's Guide*, ID4D, World Bank Group, https://id4d.worldbank.org/guide/ . See also *Digital Identity Roadmap Guide*, International Telecommunication Union, 2018, http://handle.itu.int/11.1002/pub/81215cb9-en. Note: Appendix 1 of the ITU Guide lists relevant ISO and ITU standards, which define terms for ID ecosystems. Regarding Aadhaar specifically, much has been written about this ecosystem. The post-Indian Supreme Court ruling version of Aadhaar has been greatly improved. A peer-reviewed, field researched evaluation of the pre-2018 Aadhaar and its problems provides a useful discussion of some of the early challenges in the Aadhaar system: Pam Dixon, *A Failure to Do No Harm: India's Aadhaar biometric ID program and its inability to protect privacy in relation to measures in Europe and the U.S.*, Springer Nature, Health Technology. DOI 10.1007/s12553-017-0202-6. http://rdcu.be/tsWv. Open Access via Harvard-Based Technology Science: https://techscience.org/a/2017082901/. For a short, astute analysis of the Landmark Aadhaar Judgment of 2018 and its ongoing importance, see: Anumodan Tiwari, UILS, Chandigarh University, *The landmark Aadhaar Judgment (2018): A delicate balance between empowerment and privacy*, Lawful Legal Law Journal, 19 December 2024. https://lawfullegal.in/the-landmark-aadhaar-judgment-2018-a-delicate-balance-between-empowerment-and-privacy/ . Aadhaar's complex and advanced digital ID system and how it changed and improved its handling of mandatory linking, commercial uses, and approaches to privacy after the Supreme Court ruling forms a benchmark for governance of digital identity ecosystems and continues to be a significant part of the global digital ID discussion.

92 Virtual interview with Andy Parsons, Senior Director Content Authenticity at Adobe, May 5, 2025 by author. In the interview, Parsons referenced various identity systems and technical identity related projects including decentralized wallets and

The CAWG identity assertion process may involve third-party intermediaries called Identity Claims Aggregators[93] that connect verified identity to content and act as trusted parties for both named actors and the variety of entities that use identity assertions in relation to the C2PA process.

Steps in the CAWG Identity Process [94] *(as illustrated in Figure 6)*

- Identity Claims Aggregators collect identity claims of named actors from various identity providers such as social media sites and ID verification vendors.
- Identity Claims Aggregators verify each identity claim.
- Identity Claims Aggregators access or store verified identity claims for later use.
- When a named actor creates content using a tool that has adopted C2PA, the Identity Claims Aggregator creates a unique content file-specific credential called an Identity Claims Aggregation that directly attaches that named actor's identity to that specific content file. See Figure 6 for an illustration of this process.
- The Identity Claims Aggregation created by the Identity Claims Aggregator is based on Verifiable Credentials, a Worldwide Web Consortium (W3C) identity method.entity method.[95]
- The Identity Claims Aggregation must also show the date and time when it was validated.was validated.[96]
- Identity Claims Aggregators can gather multiple identity claims from various identity providers and attach them to content produced by the same actor.
- The process requires that the collection of information related to the named content actor —known as the Verified Identities property —must be included in the identity assertion.ity assertion.[97] See Figure 7 for an example of the encoded Verified Identities property referencing names and identifier information.

identity related work at Decentralized Identity Foundation, Apple and Google wallets, Open Wallet Foundation, and Estonia's government-issued digital ID system, and said, "Those can probably all be embraced and will be embraced by CAWG for signing purposes."

93 "To facilitate the use of such identity signals, the named actor may use the services of a third-party intermediary that they trust to gather these signals and to restate them on their behalf. We call this intermediary an identity claims aggregator. It performs two important roles: 1.) It collects and verifies identity attestation claims from various identity providers such as social media sites and ID verification vendors. 2.) When the named actor creates content, it creates a unique asset-specific credential binding the identity attestation claims collected earlier to the specific C2PA asset being described." Text quoted from: Decentralized Identity Foundation, Creator Assertions Working Group, https://cawg.io. See Identity Assertion 1.1 See also 8.1. Identity claims aggregation, 8.1.1. Identity claims aggregation conceptual overview.

94 These steps describe certain types of CAWG identity assertion processes that involve third-party intermediaries called Identity Claims Aggregators. Decentralized Identity Foundation, Creator Assertions Working Group, https://cawg.io. See Identity Assertion 1.1 See also 8.1. Identity claims aggregation, 8.1.1. Identity claims aggregation conceptual overview.

95 *Verifiable Credentials Overview, Worldwide Web Consortium, standards and drafts,* November 8, 2024, https://www.w3.org/TR/2024/NOTE-vc-overview-20241108/. Normative references in the CAWG specifications also indicate that they could incorporate W3C Decentralized Identifiers. See Decentralized Identifiers (DIDs) v1.0, Worldwide Web Consortium, standards and drafts, (July 19, 2022) https://www.w3.org/TR/did-core/.

96 Decentralized Identity Foundation, Creator Assertions Working Group, https://cawg.io. See Identity Assertion 1.1 See also 8.1.2. Identity claims aggregation technical description, 8.1.2.4. Validity.

97 Decentralized Identity Foundation, Creator Assertions Working Group, https://cawg.io. See Identity Assertion 1.1 See also 8.1.2. Identity claims aggregation technical description, 8.1.2.5. Verified identities.

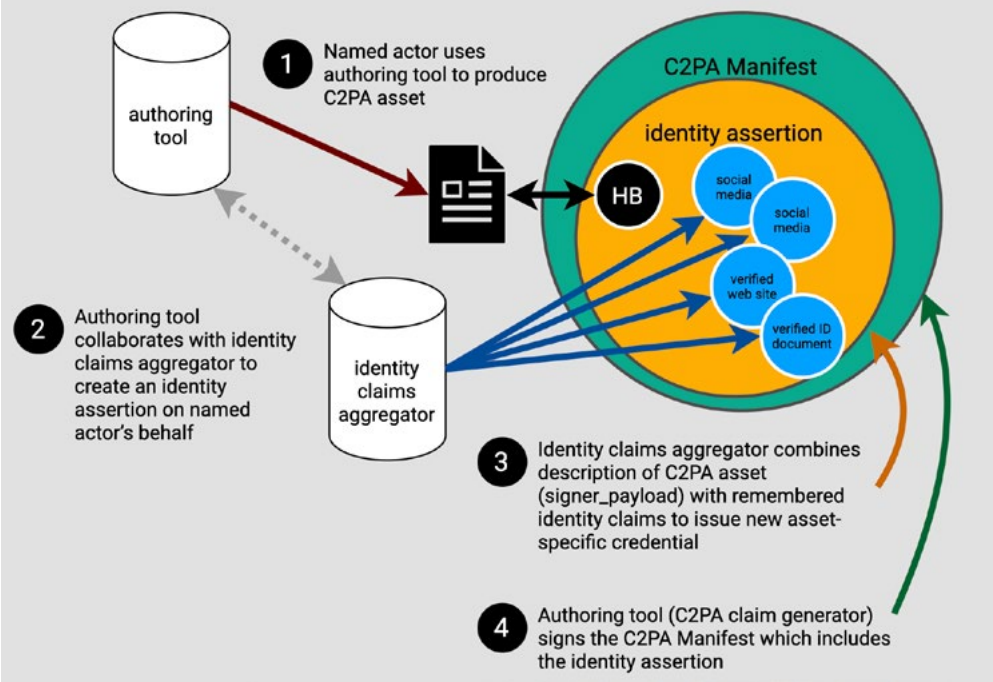


Figure 6: CAWG Identity Claims. Source: Identity Assertion, Creator Assertions Working Group, 1.1-draft, https://cawg.io/identity/1.1-draft/.

```
"credentialSubject": {
  ...
  "verifiedIdentities": [
    {
      "name": "First-Name Last-Name",
      "type": "cawg.document_verification",
      "provider": {
        "id": "https://example-id-verifier.com",
        "name": "Example ID Verifier",
      },
      "verifiedAt": "2024-07-26T22:30:15Z"
    },
    {
      "type": "cawg.web_site",
      "uri": "named-actor-site.example",
      "verifiedAt": "2024-09-25T22:13:35Z"
    },
    {
      "type": "cawg.affiliation",
      "provider": {
        "id": "https://example-affiliated-organization.com",
        "name": "Example Affiliated Organization",
      },
      "verifiedAt": "2024-07-26T22:29:57Z"
```

Figure 7: Image from CAWG technical specifications 1.1, Identity Assertion, section 8.1.2.5, "Verified Identities." This figure shows an example of the encoded Verified Identities property referencing names and identifier information. Available at https://cawg.io/

Verified Identities

The identity assertion specifications list several types of Verified Identities that entities interpreting C2PA identity assertions should accept, including government-issued identity documents such as drive's licenses,[98] professional organization-related credentials, as well as social media account or digital wallet related identifiers. Verified Identities information such as names in government identity documents or unique alphanumeric identifiers associated with social media or digital wallet identifiers must match names or identifiers in the original identity verification source.

Addressing Identity-Related Impacts in C2PA

C2PA's own Harms Modelling documentation[99] discusses potential harms of identity systems used in conjunction with C2PA, noting, "In some countries, governments may issue digital certificates to all of its citizens. These certificates could be potentially used to sign C2PA manifests. If government control and surveillance is not regulated, or if there are laws meant to attach journalistic identity to media posted online, these certificates may be used to enforce suppression of speech or to persecute journalists if required by claim generators that do not guarantee privacy and confidentiality."

The documentation also states, "Similarly, certain C2PA claim generators may allow content creators, including civic, community and indepente (sic) media, to sign manifests with their personal certificates associated with their IDs. Although guidance to allow for anonymity and pseudonymity has been issued, specification-compliant tools may sell information to third-parties, or not follow user experience guidance meant to empower users to retain control of their information."

Observer organizations[100] [101] have also highlighted potential harms for media creators including marginalized or vulnerable creators and people sharing or accessing content attached to C2PA metadata. An election technical expert who tested C2PA to produce digitally credentialed media in an election observation mission said wider adoption of C2PA increases "the threat of persecution through metadata, which can contain information identifying a specific smartphone and its owner."[102]

AI Training Consent and Controls in CAWG

CAWG has also created technical specs for other types of assertions including an assertion intended to give content creators and owners some control over use of their content in relation to AI. The Training and Data

---

98 In an October CAWG meeting, possible "inclusion of credentials such as driver's licenses," was discussed as an example of "verified presentations" in identity assertions. See 7 October 2024, Creator Assertions Working Group, *Meeting Notes*, Decentralized Identity Foundation, https://creator-assertions.github.io/meeting-notes/2024-10-07/. Also see the Identity Claims Aggregator workflow diagram in Figure 6 of this report indicating possible use of a "verified ID document" in addition to other types of identity providers including social media.

99 C2PA Harms Modelling, *C2PA Specifications*, C2PA, https://c2pa.org/specifications/specifications/1.4/security/Harms_Modelling.html.

100 Jacobo Castellanos, *WITNESS and the C2PA Harms and Misuse Assessment Process Confronting Potential Harms Early in the Process of Developing Authenticity and Provenance Infrastructure*, WITNESS, December 2, 2021, https://blog.witness.org/2021/12/witness-and-the-c2pa-harms-and-misuse-assessment-process/.

101 Ingo Boltz, *Content Credentialed Media in Election Observation Missions – First Lessons Learned*, Electoral Integrity Project, September 26, 2024, https://www.electoralintegrityproject.com/eip-blog/2024/9/20/content-credentialed-media-in-election-observation-missions-first-lessons-learned.

102 Ingo Boltz, *Content Credentialed Media in Election Observation Missions – First Lessons Learned*, Electoral Integrity Project, September 26, 2024, https://www.electoralintegrityproject.com/eip-blog/2024/9/20/content-credentialed-media-in-election-observation-missions-first-lessons-learned.

Mining Assertion [103] allows a human actor to state whether or not use of content with C2PA metadata is allowed as part of a data mining, AI or machine learning training workflow, and whether allowed use is conditional or constrained.

The assertion specs refer to four pre-defined purposes or "entries" related to AI use [104]:

- cawg.da*ta_mining addresses* use of text or data extracted from content for determining patterns, trends, and correlations.
- cawg.a*i_inference addresses* use of content as input to a trained AI/ML model to infer a result.
- cawg.ai_*generative_training addresses* use of content as training data to an AI/ML model that could generate other content.
- cawg.a*i_training addresses* use of content as data to train non-generative AI/ML models, such for classification or object detection.

Privacy in C2PA

Along with external evaluators,[105] C2PA's own Harms Modeling documentation[106] recognizes the serious privacy threats posed by C2PA. It states the possibility that C2PA Claim Generators – the device hardware or software systems that produce statements or assertions about the origins and modifications to a piece of content – "may automatically add, or require to add, information to manifests that may be sensitive."[107] And it recognizes C2PA's potential for "inadvertent disclosure of information," accessibility of redacted information, and misuse or abuse of C2PA to build products that violate human rights.[108]

Even though the designers of C2PA cannot make demands about how the quasi-standard actually is used in practice, and they can't themselves regulate or enforce proper C2PA implementations,[109] [110] they can make strong suggestions. C2PA's "Guiding Principles" emphasize use of C2PA in ways that respect privacy, personal control of data and avoid potential abuse and misuse.[111]

---

103 *Training and Data Mining Assertion 1.1,* Decentralized Identity Foundation, Creator Assertions Working Group, https://cawg.io/training-and-data-mining/1.1/ .

104 *Training and Data Mining Assertion 1.1*, Decentralized Identity Foundation, Creator Assertions Working Group, https://cawg.io/training-and-data-mining/1.1/. See 3. Assertion definition.

105 Bilva Chandra, Jesse Dunietz, and Kathleen Roberts, *Reducing Risks Posed by Synthetic Content An Overview of Technical Approaches to Digital Content Transparency*, NIST, NIST Trustworthy and Responsible AI, November 2024, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-4.pdf.

106 *C2PA Harms Modelling,* C2PA Specifications, https://c2pa.org/specifications/specifications/1.4/security/Harms_Modelling.html.

107 *C2PA Harms Modelling*, C2PA Specifications, https://c2pa.org/specifications/specifications/1.4/security/Harms_Modelling.html. See 6.1. General considerations for content creators/Privacy and security considerations.

108 *C2PA Harms Modelling*, C2PA Specifications, https://c2pa.org/specifications/specifications/1.4/security/Harms_Modelling.html. See 6.1. General considerations for content creators/Privacy and security considerations.

109 See Bilva Chandra, Jesse Dunietz, and Kathleen Roberts, *Reducing Risks Posed by Synthetic Content An Overview of Technical Approaches to Digital Content Transparency*, NIST, NIST Trustworthy and Responsible AI, November 2024, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-4.pdf. "[C2PA] recommendations are not enforceable, and malicious or negligent tool providers could claim to promote transparency via metadata recording solutions but end up exposing user information without consent." See 3.1.2.4 Privacy Considerations for Metadata Recording, p. 19.

110 Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe, December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. In a May 2, 2025 follow up email for this research, Rosenthol stated, "If someone wants to mandate its [C2PA's] uses or how it is implemented – they can. That said, the C2PA recommends that they not put such specific details in policy – but instead, focus on mandat[ing] the use of the underlying approach, 'content provenance.'"

111 *Guiding Principles for C2PA Designs and Specifications*, Coalition for Content Provenance and Authenticity, https://c2pa.org/

C2PA's Opt-in Goals

An opt-in approach is a part of C2PA's privacy mitigation and is an important aspect of this effort. The specifications call for implementers to allow content creators and publishers to control whether certain provenance data such as personal identifiers or location information is included. In addition, C2PA's user experience guidance[112] states that C2PA-based products are required to provide disclosure of data collection and obtain "clear acknowledgement of creator consent before a C2PA implementation can begin accumulating data."

Redaction in C2PA

Another primary method of privacy mitigation and control incorporated in C2PA is data redaction; it allows for certain assertions to be removed from a manifest. As an example, the specs state that "a metadata assertion containing both location data and camera information which needs to have the location data redacted could be done through an updated manifest with a new metadata assertion containing only the camera information."[113]

However, there are limits to redaction in C2PA. Action-oriented assertions cannot be redacted, because according to the specs, they are "essential information in understanding the history of an asset."[114] These *c2pa.actions assertions* include content alterations such as color changes or [image] crops, editorial content deletions, assertions noting whether an asset was published, if a content component or ingredient was removed, or if an invisible watermark is inserted into content.

Also, similar to redactions in government records made public, C2PA calls for inclusion of assertions that indicate that the redaction occurred. And, the type of information redacted also will be known assuming the assertion label is maintained, which the specs state "enables both humans and machines to apply rules to determine if the removal was acceptable."[115]

It is unclear whether or how entities that adopt C2PA will enable C2PA metadata removal from their systems. A large content delivery network [116] that adopted C2PA [117] said that if its setting used to automatically embed and attach C2PA metadata to an image is disabled, any existing Content Credentials, i.e. C2PA metadata, will always be discarded.

How the C2PA-related Identity Assertion from CAWG Addresses Privacy

As noted in the Identity section earlier in this report, one of C2PA's most pressing privacy issues —identity — was moved to the Creator Assertions Working Group, or CAWG. So, it is important to review how that group has addressed privacy in relation to identifiers and identity systems.

---

principles.

112 *C2PA User Experience Guidance for Implementers*, Coalition for Content Provenance and Authenticity, C2PA Specifications, https://c2pa.org/specifications/specifications/2.0/ux/UX_Recommendations.html. See 8.1. Opting in, privacy and data collection.

113 C2PA, *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html . See 18.16 Metadata/18.16.4. Redaction of c2pa.metadata.

114 *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html . See 6. Assertions/6.8. Redaction of Assertions.

115 *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html . See 6. Assertions/6.8. Redaction of Assertions

116 *Historical trends in the usage statistics of reverse proxy services for websites*, Web Technology Surveys, May 2024, https://w3techs.com/technologies/history_overview/proxy/all.

117 Will Allen, *Preserving content provenance by integrating Content Credentials into Cloudflare Images*, Cloudflare, The Cloudflare Blog, February 3, 2025, https://blog.cloudflare.com/preserve-content-credentials-with-cloudflare-images.

CAWG specifications acknowledge privacy risks associated with identity assertions such as exposure of identifiable information through insecure C2PA-enabled cameras or devices, and fraudulent identity spoofing. They state:

"An attacker could attempt to extract a valid identity assertion out of one C2PA asset and embed it in a new C2PA asset of an attacker's choosing without causing a validation error. This scenario could allow an attacker to falsely attribute the new C2PA asset to a victim's identity without their consent or knowledge. If the attacker's C2PA asset is controversial or illegal, then falsely attributing it to a victim's identity could result in severe consequences for the victim. The system must ensure that an attacker can not apply an existing identity assertion to a different C2PA asset."[118]

Also, as noted above, CAWG's identity assertion specs[119] incorporate W3C Verifiable Credentials and also reference W3C Decentralized Identifiers. W3C Verifiable Credentials documentation acknowledges their privacy risks,[120] including risks of exposure of personally identifiable information used in Verifiable Credentials, risks of identity exposure enabled through correlations of Verifiable Credential related signatures and identifiers, as well as risks of identity exposure through use of "long-lived" identifiers across more than one web domain. Specs for W3C Decentralized Identifiers also recognize the privacy risks of identifier correlation.[121]

When the technical research for this report concluded, the most recent draft of CAWG Identity Assertion documentation referred to redaction only in reference to redaction of C2PA assertions,[122] and in two use case examples that mention identity assertion redaction as a way to protect the privacy of a fictitious human rights defender and fictitious designer.[123]

The Technical Components of C2PA's Trust Model

C2PA technical specifications state that C2PA should not provide value judgements regarding whether provenance data is good or bad.[124] And the method is not intended to directly determine whether or not a content creator or entity involved in producing or altering content is "trustworthy."

The basis of gauging trust in C2PA is a "trust model" reliant on a process of generating and validating trust signals, comprised of the components described above in the *Under the Hood* Section of this report. At the close of the research phase for this report, the trust afforded to each of those components was determined by the drafters and early adopters of C2PA. Find an illustration of the C2PA Trust Model in Figure 8.

---

118 *Creator Assertions Working Group, Identity Assertion 1.1*, Decentralized Identity Foundation, https://cawg.io/identity/1.1/. See 3. Assertion definition. See 9.4 Threats to trust model/9.4.1. Replay attacks.

119 *Identity Assertion, 1.1-draf*t, Creator Assertions Working Group, Decentralized Identity Foundation, https://cawg.io.

120 *Verifiable Credentials Data Model v2.0*, Worldwide Web Consortium, standards and drafts, October 19, 2024, https://www.w3.org/TR/2024/NOTE-vc-overview-20241108/.

121 *Decentralized Identifiers (DIDs) v1.0*, Worldwide Web Consortium, standards and drafts, July 19, 2022, https://www.w3.org/TR/did-core/, See 10. Privacy Considerations and 10.2 DID Correlation Risks.

122 *Creator Assertions Working Group, Identity Assertion 1.1,* Decentralized Identity Foundation, https://cawg.io/identity/1.1/. See 3.1.3. C2PA claim and 6.3. Interaction with data hash assertion. For dates research concluded for this report, see methodology.

123 Creator Assertions Working Group, Identity Assertion 1.1, Decentralized Identity Foundation, https://cawg.io/identity/1.1/. See 1.5. Use cases and examples/1.5.2. Enhance the evidentiary value of critical footage and 1.5.4. Brand protection in digital marketing.

124 *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, Linux Foundation Projects, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html . See 1.2. Scope.

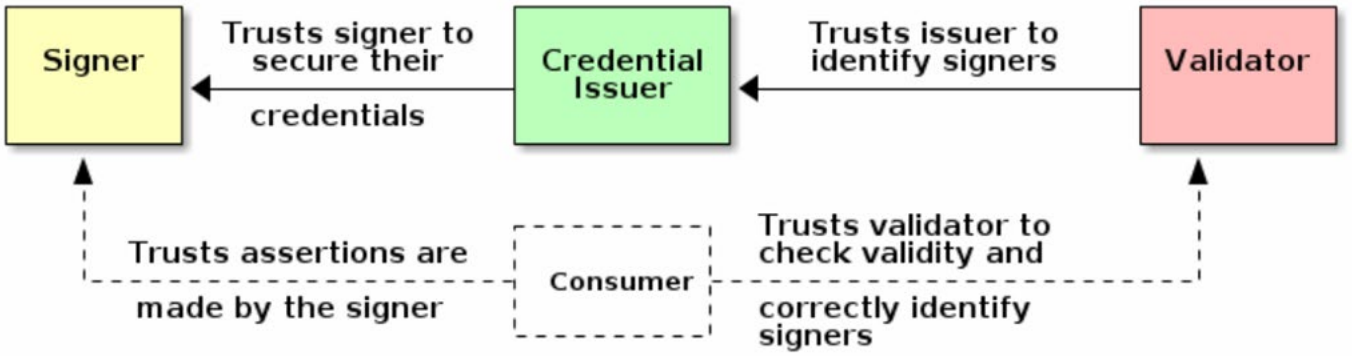


Figure 8: C2PA Trust Model. Source: C2PA, C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html.

Trust Lists in C2PA

C2PA Trust Lists[125] are important components of the trust model. At the time of research for this report, these are lists of certificate authorities that have established certifications made specifically for C2PA, allowing them to add metadata to a content file and validate claim signing certificates. Entities on the lists —currently in a "temporary" form —include camera makers, software providers, AI companies and media platforms.[126]

One early C2PA adopter only extracts and displays C2PA metadata to people viewing images if the content manifest is signed by a Certification Authority included in the C2PA Trust List.[127] When a large[128] content delivery network providing image storage, management and delivery services adopted C2PA, the process required addition of the network to the C2PA Trust List.[129]

The criteria used for inclusion on these lists, or who the arbiters deciding which entities are considered known, and by extension, trustworthy have not been made public. A "conformance program"[130] is in development that involves testing of entities seeking approval as known, trusted certificate authorities.[131] A C2PA Interim Verify Certificate Intake Request form[132] asks organizations interested in inclusion on the C2PA-managed trust list for a

125 A C2PA Trust List is "A C2PA-managed list of X.509 certificate trust anchors that issue certificates to hardware & software signers that use them to sign claims, as quoted in: C2PA, *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html See 2.3. Core Aspects of C2PA / 2.3.15. *C2PA Trust List.*

126 *Verify tool known certificate list,* Content Authenticity Initiative, https://opensource.contentauthenticity.org/docs/verify-known-cert-list/. See methodology heading of this report for dates of research.

127 Google Search Central, *Image Metadata in Google Images, Documentation*, See How C2PA metadata can appear in Google Search results, https://developers.google.com/search/docs/appearance/structured-data/image-license-metadata#c2pa-metadata.

128 *Historical trends in the usage statistics of reverse proxy services for websites,* W3Techs, Web Technology Surveys, May 2024, https://w3techs.com/technologies/history_overview/proxy/all.

129 "In order for Cloudflare to append the Content Credentials with any transformations, we needed to have a publicly available end-entity certificate and join this Trust List." Text quoted from: Will Allen, *Preserving content provenance by integrating Content Credentials into Cloudflare Images,* Cloudflare, The Cloudflare Blog, February 3, 2025, https://blog.cloudflare.com/preserve-content-credentials-with-cloudflare-images.

130 *Conformance,* Linux Foundation Projects, C2PA, https://c2pa.org/conformance/.

131 Virtual interview with Andy Parsons, Senior Director Content Authenticity at Adobe, May 5, 2025 by author. Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe, December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. Both Parsons and Rosenthol stated in separate interviews with the author that a conformance program for the C2PA-managed Trust List was in development. Rosenthol in his comments noted that entities included on the C2PA-managed Trust List have passed conformance tests.

132 *C2PA Interim Verify Certificate Intake Request,* Creator Assertions Working Group, 07 Oct. 2024, https://airtable.com/

name and application description, along with certificate related details.

As of July 2025, very little information regarding the C2PA conformance program testing criteria, or which entities or people manage the C2PA trust list, conformance tests or the testing application process is public. A member of an AI company's technical group who has implemented C2PA but is not on a working group related to the conformance program said the program was intended to establish criteria for tiered trust categories.[133]

Absence of C2PA Signals

The C2PA process is designed to allow validators to assess and label the validation state of content metadata, assigning one of three validation states to a manifest: well-formed, valid or trusted. A fully trusted manifest is one signed by an authority listed on a Trust List.[134] [135] C2PA validators can also label a manifest as unknown if they encounter a manifest that uses constructs from an unsupported version of C2PA. Those label choices affect whether or not a content provenance chain is considered weakened or broken, which is an indication of untrustworthiness.

Thus, it's not just the presence of valid trust signals that can affect C2PA-based trust measures, but also their absence. A lack of a signed C2PA claim is an indication of untrustworthiness. For instance, according to User Experience Guidance[136] associated with the C2PA specs, when content "is signed by an untrusted entity" it results in a break in the provenance metadata, including in the information displayed to someone viewing content. In an example shown in the guidance and also featured in Figure 9, in place of specific information reflecting that step in the content provenance chain, metadata indicates an "untrusted edit" and states that "something unknown or untrusted happened to the manifest;" people viewing content credentials based on that metadata see a step in the provenance chain marked as "invalid."[137]

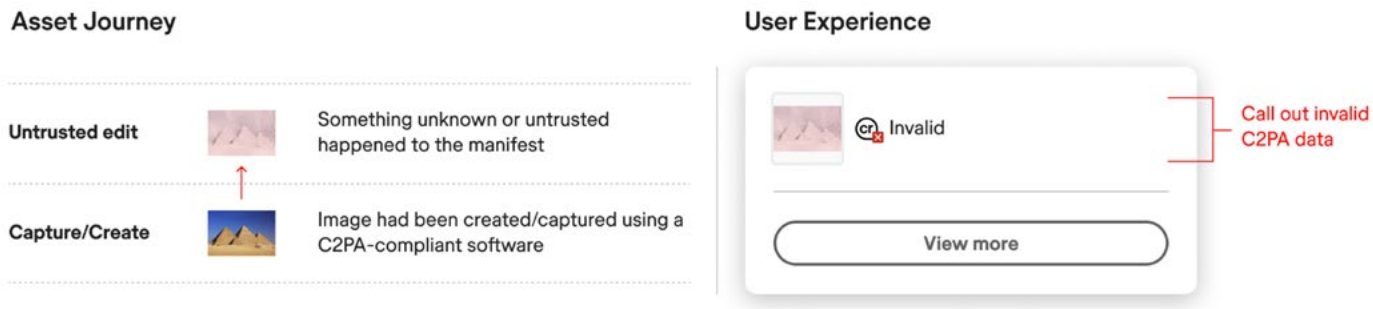


Figure 9: Image from C2PA User Experience Guidance for Implementers illustrating how C2PA metadata reflecting an "invalid state" such as in "instances when C2PA-enabled content has been maliciously edited to tamper with C2PA data, or is signed by an untrusted entity" might be reflected in a content label. See C2PA Specifications, C2PA User Experience Guidance for Implementers, C2PA, https://c2pa.org/specifications/specifications/2.0/ux/UX_Recommendations.html.

appouw5x68pZLwKKY/pagwdnUZ8TQ8t7Zjy/form.

133 Virtual interview with Michael Lampe, a member of OpenAI's technical group, by the author, May 6, 2025. According to Michael Lampe, the conformance program in development "...establishes criteria for being issued a certificate of a certain level of trust. The conformance piece is meant to help bucket people or bucket organizations, producers of content tools into different trust categories."

134 *C2PA Specifications, C2PA User Experience Guidance for Implementers, C2PA*, https://c2pa.org/specifications/specifications/2.0/ux/UX_Recommendations.html. See 5.6. Invalid states.

135 *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, C2PA , https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html . See 14.3. Validation states. Also see 15.7. Validate the Signature.

136 *C2PA Specifications, C2PA User Experience Guidance for Implementers*, C2PA, https://c2pa.org/specifications/specifications/2.0/ux/UX_Recommendations.html.

137 *C2PA Specifications, C2PA User Experience Guidance for Implementers*, C2PA, https://c2pa.org/specifications/specifications/2.0/ux/UX_Recommendations.html. See 5.6. Invalid states.

One camera maker said it plans to offer C2PA functions in certain cameras only to select media outlets who would need to enable C2PA functions available through a separate upgrade license.[138]

Some worry that C2PA's trust model and implementations could penalize content that lacks C2PA metadata, or includes a broken provenance chain or unrecognizable signals —and by extension, creators of that content.

An electoral researcher who described use of C2PA to produce digitally credentialed media in an election observation mission said C2PA-based warnings regarding the "unknown" origin of content that resulted from certificates that were not "whitelisted" would confuse media consumers and cause distrust.[139]

Validation states in C2PA are used in part as a way to address adverse impact on creators and their content and reflect a lack of fully-validated provenance information. For instance, even if some content provenance metadata is not fully "trusted" it can still be considered "well-formed" or "valid."[140]

The C2PA process does not measure trustworthiness of its data, but rather is designed to measure trustworthiness of the Signer of the data. It is up to the end evaluator or "consumer" of the C2PA provenance information to gauge trustworthiness of the metadata provided by the Signer.[141]

C2PA specifications also incorporate quantifiable measures called Review Ratings.[142] While they are not mandatory components of C2PA, the ratings allow claim generators to determine a rating to gauge the quality of an assertion, or lack thereof. The ratings are expressed as integer values where 1 represents the worst quality assertion and 5 represents the best quality.

The CAWG Trust Model

CAWG's identity assertion technical specifications also have a Trust Model;[143] it incorporates trust-related decisions directly related to identified content creators or other named actors (named humans, organizations of humans or non-human software or hardware devices). It calls for named actors to be labeled as:

- Trusted - when a unique content file-specific credential can be verified
- Well-formed - when a unique content file-specific credential cannot be verified but has not been considered invalid either
- Revoked - when the credential used for signing the identity assertion had been deemed invalid at the time the assertion was created

---

138 *Sony Announces Second-Generation Flagship Alpha 1 II,* Sony, November 19, 2024, https://alphauniverse.com/stories/sony-announces-secondgeneration-flagship-alpha-1-ii/.

139 Ingo Boltz, *Content Credentialed Media in Election Observation Missions – First Lessons Learned,* Electoral Integrity Project, September 26, 2024, https://www.electoralintegrityproject.com/eip-blog/2024/9/20/content-credentialed-media-in-election-observation-missions-first-lessons-learned.

140 Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe, December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. Rosenthol described the use of validation states in December 2024 and May 2025 interviews for this report.

141 Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe, December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. The statements in the text cited were confirmed by Leonard Rosenthol in a May 2, 2025 follow up email interview for this report.

142 *C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials*, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html . See 18.3.3. Review Ratings.

143 *Identity Assertion, 1.1-draf*t, Creator Assertions Working Group, Decentralized Identity Foundation, https://cawg.io. See 9. Trust Model

C2PA's Technical Hurdles

There are some technical caveats that limit the accuracy and effectiveness of C2PA. This section provide an overview of these caveats.

Metadata Stripping and Removal through File Modifications

When photos, images, videos, and other content files are uploaded to social media platforms or shared via various systems, they are often stripped of their metadata.[144] Multiple implementers of C2PA have acknowledged metadata stripping as an obstacle to its efficacy and robustness.[145] [146]

Rather than storing C2PA metadata externally and connecting it to content, C2PA drafters prefer storage of C2PA metadata inside content files "because it keeps the provenance with the asset."[147] C2PA circumvents metadata stripping through technical support for external storage of metadata outside the content file itself, in a manifest repository sometimes referred to as a "sidecar."[148]

As a related workaround to the metadata stripping problem, C2PA enables use of third-party watermarks,[149] Non-fungible Tokens,[150] and fingerprinting for provenance metadata storage and recovery. In the C2PA context, fingerprinting is a computational process that uniquely codes content, allowing it to be matched to recover the content with intact C2PA metadata from an external database.[151] [152]

---

144 *Social Media Sites Photo Metadata Test Results 2019*, International Press Telecommunications Council (IPTC), 2019, https://iptc.org/standards/photo-metadata/social-media-sites-photo-metadata-test-results-2019/. See also: *State of image metadata in 2018*, Imatag, May 11, 2018, https://www.imatag.com/blog/state-of-image-metadata-in-2018. See also: *What Are Content Credentials?*, Adobe, Content Credentials, (October 14, 2024), https://helpx.adobe.com/creative-cloud/help/content-credentials.html.

145 Virtual interview with Michael Lampe, a member of OpenAI's technical group, by the author, May 6, 2025. According to Lampe, the common practice of metadata stripping is the biggest obstacle preventing C2PA from working as designed. He said that when OpenAI first implemented C2PA, "...what we found was that the ecosystem wasn't mature enough to not strip this metadata," but added, "We are starting to see major online platforms actually adopt it and that's a positive signal in its own right." See also: *C2PA in DALL·E 3, Privacy and policies,* Policy FAQ, OpenAI, https://help.openai.com/en/articles/8912793-c2pa-in-dall-e-3.

146 *Integration Guide, Truepic Enterprise C2PA Tools,* Truepic, https://lens.truepic.dev/docs/display-library.

147 Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe, December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. According to a May 2, 2025 email follow up with Rosenthol, rather than storing C2PA metadata externally and connecting it to content, storage of C2PA metadata inside content files is preferred "because it keeps the provenance with the asset."

148 *AWS Architecture Center, AWS Innovation with Sinclair,* AWS Innovation Ambassadors June 2024, Amazon, https://aws.amazon.com/podcasts/innovation-ambassadors/ia-podcast-ep-47-sinclair/?podcast-list-wide.sort-by=item.additionalFields.EpisodeNum&podcast-list-wide.sort-order=desc.

149 *Your Digital Assets Are At Risk, Digimarc Validate,* Digimarc, See video https://www.digimarc.com/products/digital-content-authentication. See also Digimarc, *Digimarc Brings Digital Watermarking to the C2PA 2.1 Standard,* Digimarc Newsroom, (October 8, 2024), https://www.digimarc.com/press-releases/2024/10/08/digimarc-brings-digital-watermarking-c2pa-21-standard.

150 Kar Balan, Shruti Agarwal, Simon Jenni, Andy Parsons, Andrew Gilbert, John Collomosse, *EKILA: Synthetic Media Provenance and Attribution for Generative Art, in 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (CVPRW), Vancouver, BC, Canada, 2023, pp. 913-922, https://openaccess.thecvf.com/content/CVPR2023W/WMF/papers/Balan_EKILA_Synthetic_Media_Provenance_and_Attribution_for_Generative_Art_CVPRW_2023_paper.pdf, See also doi: 10.1109/CVPRW59228.2023.00098.

151 Andy Parsons, *Durable Content Credentials,* Content Authenticity Initiative, April 8, 2024, https://contentauthenticity.org/blog/durable-content-credentials.

152 C2PA Technical Specifications 2.1, Technical Specifications, Content Credentials, C2PA, https://c2pa.org/specifications/

Durability and Its Caveats

However, C2PA's technical approaches to "durability" including through metadata stripping circumventions have their own questions and caveats:

- Depending on the fingerprinting process employed, connecting content files to C2PA metadata stored externally using fingerprinting could be negatively affected by image perturbations resulting from repeated image use, uploads, downloads and compressions including noise and changes to image resolution, quality and format.ty and format.[153]
- Questions and concerns remain regarding hosting and control of fingerprinting computation and fingerprint storage.print storage.[154] [155]
- Questions remain regarding the cost ding the cost [156] and energy use required to enable C2PA,[157] such as energy use related to fingerprinting computation.

There also are other factors affecting C2PA metadata generation, storage and maintenance:

- Manifest file size, processing time,ocessing time,[158] frequency of use, and need for access will be considerations for hardware, software and app makers building C2PA-enabled products and systems.[159]
- Some file types — such as .txt files — will not support embedded C2PA information.A information.[160]

Forgeries and Other Validation Vulnerabilities

specifications/2.1/specs/C2PA_Specification.html . See 2.4. Additional Terms/2.4.1. Fingerprint.

153 Kar Balan, Alex Black, Andrew Gilbert, Simon Jenni, Andy Parsons, John Collomosse, *DECORAIT - DECentralized Opt-in/out Registry for AI Training, In Proceedings of the 20th ACM SIGGRAPH European Conference on Visual Media Production* (CVMP '23). Association for Computing Machinery, New York, NY, USA, Article 4, 1–10, 2023, https://doi.org/10.1145/3626495.3626506 see 5.1 Experimental Setup.

154 Andy Parsons, *Durable Content Credentials,* Content Authenticity Initiative, April 8, 2024, https://contentauthenticity.org/blog/durable-content-credentials.

155 K. Balan, S. Agarwal, S. Jenni, A. Parsons, A. Gilbert, and J. Collomosse, EKILA: Synthetic Media Provenance and Attribution for Generative Art, In Proc. CVPR Workshop on Media Forensics, 2023, https://arxiv.org/pdf/2304.04639.

156 Kar Balan, Alex Black, Andrew Gilbert, Simon Jenni, Andy Parsons, John Collomosse, *DECORAIT - DECentralized Opt-in/out Registry for AI Training, In Proceedings of the 20th ACM SIGGRAPH European Conference on Visual Media Production (*CVMP '23). Association for Computing Machinery, New York, NY, USA, Article 4, 1–10, 2023, https://doi.org/10.1145/3626495.3626506 see 5.4 Evaluating Cost.

157 Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe, December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. The statements in the text cited were confirmed by Leonard Rosenthol in a May 2, 2025 follow up email interview for this report. Rosenthol noted that the C2PA Technical Working Group has no specific information regarding computational processing related energy use that may be required to implement C2PA because each implementation (such as implementations for smaller devices such as cameras or smartphones, PCs, servers or blockchain-related integrations) can have different processing requirements.

158 Email interview with Ingo Boltz, an expert in election technology and digital threats to democracy by the author May 5, 2025. Boltz, who has tested C2PA-enabled apps intended to help prove the authenticity of photos, confirmed that camera devices and apps that automatically generate cryptographically signed images at the point of capture using C2PA may add steps and slow down the sometimes high-pressure act of capturing photojournalistic images. In particular, he mentioned added steps related to processing for cryptographic signing, steps that protect against manipulation of C2PA metadata, GPS calibration, and file library functions.

159 Kaushal Rathi, Sathyanarayana Sampath Kumar, Mandanna A N, *Insights into Coalition for Content Provenance and Authenticity (C2PA)*, Infosys Tech Compass, https://www.infosys.com/iki/techcompass/content-provenance-authenticity.html.

160 *C2PA Technical Specifications 2.1*, Technical Specifications, Content Credentials, C2PA, https://c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html . See 11.4. External Manifest.

Multiple researchers who have analyzed or tested C2PA including some who have been interviewed for this report have expressed concern that C2PA does not deter bad actors, is vulnerable to hacking,[161] facilitates authentication of malicious deepfakes or forgeries,[162] [163] and lays groundwork for anti-consumer behavior.[164] They have pointed to use of C2PA to cryptographically sign authentication and provenance information in forged content including images of credit card numbers, drivers' licenses or pregnancy tests used in paternity payment situations,[165] or to sign inaccurate or confusing timestamps or geographic metadata.

One media company announced use of its video content verification process incorporating C2PA-based Content Credentials, and later issued a correction stating that the "verified" video had actually been misleadingly edited by a social media user to include elements of an older, unrelated video.[166]

Specs vs Implementations

A key C2PA drafter interviewed for this research[167] said that the majority of tests showing fraudulent or forged content that had been authenticated using C2PA reflected implementations rather than the technical specs themselves, and suggested they could have been based on older versions of the technical specs. The drafter emphasized distinctions between C2PA technical specs and practical implementations of C2PA, suggesting that the two should be evaluated separately.[168]

Trust Model Nuances and Limitations

There are also important nuances to C2PA's Trust Model and content provenance validation process. The C2PA process does not measure trustworthiness of its data, but rather is designed to measure trustworthiness of the Signer of the data. It is up to the end evaluator or "consumer" of the C2PA provenance information to gauge trustworthiness of the metadata provided by the Signer.[169]

---

161 Martin Thomson, *C2PA Is Not Going To Fix Our Misinformation Problem*, Low Entropy, https://lowentropy.net/posts/c2pa .

162 Kaushal Rathi, Sathyanarayana Sampath Kumar, Mandanna A N, *Insights into Coalition for Content Provenance and Authenticity (C2PA)*, Infosys Tech Compass, https://www.infosys.com/iki/techcompass/content-provenance-authenticity.html.

163 Adam Zeloof, *Falsified Photos: Fooling Adobe's Cryptographically-Signed Metadata,* Hackaday, November 30, 2023, https://hackaday.com/2023/11/30/falsified-photos-fooling-adobes-cryptographically-signed-metadata

164 Martin Thomson, *C2PA Is Not Going To Fix Our Misinformation Problem*, Low Entropy, https://lowentropy.net/posts/c2pa.

165 Neal Krawetz, *C2PA's Butterfly Effect,* The Hacker Factor Blog, November 16, 2023, https://www.hackerfactor.com/blog/index.php?/archives/1010-C2PAs-Butterfly-Effect.html. Also see Neal Krawetz, *C2PA from the Attacker's Perspective*, The Hacker Factor Blog, May 9, 2024, https://www.hackerfactor.com/blog/index.php?/archives/1031-C2PA-from-the-Attackers-Perspective.html, and Neal Krawetz, *C2PA and Authenticated Disinformation*, The Hacker Factor Blog, October 15, 2024, https://www.hackerfactor.com/blog/index.php?/archives/1046-C2PA-and-Authenticated-Disinformation.html.

166 BBC's correction stated, "...it has since been drawn to our attention that some of this was misleadingly edited in from an older unrelated video by a social media user. In line with our commitment to transparency, we have reviewed the video and muted the relevant passage of audio. We have also updated our commentary on the video to reflect those changes." *Transparency tool launched by BBC Verify,* BBC News, March 4, 2024, https://www.bbc.com/news/av/world-68474465. See also: Henri Astier and Gianluca Avagnina, *Haiti violence: Haiti gangs demand PM resign after mass jailbreak,* BBC News, March 4, 2024, https://www.bbc.com/news/world-latin-america-68462851.

167 Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe, December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. The statements in the text cited were confirmed by Leonard Rosenthol in a May 2, 2025 follow up email interview for this report.

168 Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe, December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. The statements in the text cited were confirmed by Leonard Rosenthol in a May 2, 2025 follow up email interview for this report.

169 Virtual interview with Leonard Rosenthol, Chair of the C2PA Technical Working Group and Senior Principal Scientist at Adobe, December 4, 2024 with email follow up in December 2024, January 2025, and May 2025 by author. The statements in the text

It's also worth reiterating that despite C2PA's technical Trust Model process, it is not designed to fact-check or vet the quality or veracity of media or information carrying its metadata.

Other Technical Limitations

Engineers[170] who have implemented C2PA have highlighted other impediments to C2PA's effectiveness:

- The inability to determine how C2PA assertions relate to or reference other assertions;
- The need for non-trivial engineering effort to integrate C2PA in content creation and distribution workflows, including code writing and deployment resources;
- The need for long-term commitment and resources and participants to maintain C2PA and support its infrastructure; and
- A lack of widespread adoption and ecosystem support.

Early Examples of C2PA in Prototypes and Products

Implementations and integrations of C2PA are just beginning to be tested or released, including from members of the C2PA Steering Committee.[171] The following examples of early tests and emerging products illustrate C2PA use for revealing content provenance information to everyday people, as well as for distributing granular metadata throughout the digital information ecosystem.

Content Provenance Labeling Systems

C2PA is at work inside an app "powered by the C2PA standard and the Content Authenticity Initiative,"[172] the growing coalition behind C2PA.[173] As illustrated in Figure 10, the app allows creators to include C2PA metadata in their content, and displays of some of that information in human-readable form when content viewers click a "cr" icon visible in the content. As seen in other examples below, the "cr" icon is used in a variety of software and systems.

cited were confirmed by Leonard Rosenthol in a May 2, 2025 follow up email interview for this report.

170 See: Patrick O'Connor and Jonathan Solomon, *An easy-to-use solution for C2PA workloads,* (October 2024), https://summit.smpte.org/2024/session/2477681/an-easy-to-use-solution-for-c2pa-workloads and https://assets.swoogo.com/uploads/4655929-673f7ac0813b7.pdf. See also: AWS Architecture Center, *AWS Innovation with Sinclair, AWS Innovation Ambassadors June 2024*, Amazon, https://aws.amazon.com/podcasts/innovation-ambassadors/ia-podcast-ep-47-sinclair/?podcast-list-wide.sort-by=item.additionalFields.EpisodeNum&podcast-list-wide.sort-order=desc.

171 *Membership*, Coalition for Content Provenance and Authenticity, https://c2pa.org/membership/

172 Virtual interview with Andy Parsons, Senior Director Content Authenticity at Adobe, May 5, 2025 by author. In the interview Parsons noted that CAI is 100% funded by Adobe.

173 Content Credentials, https://contentcredentials.org.

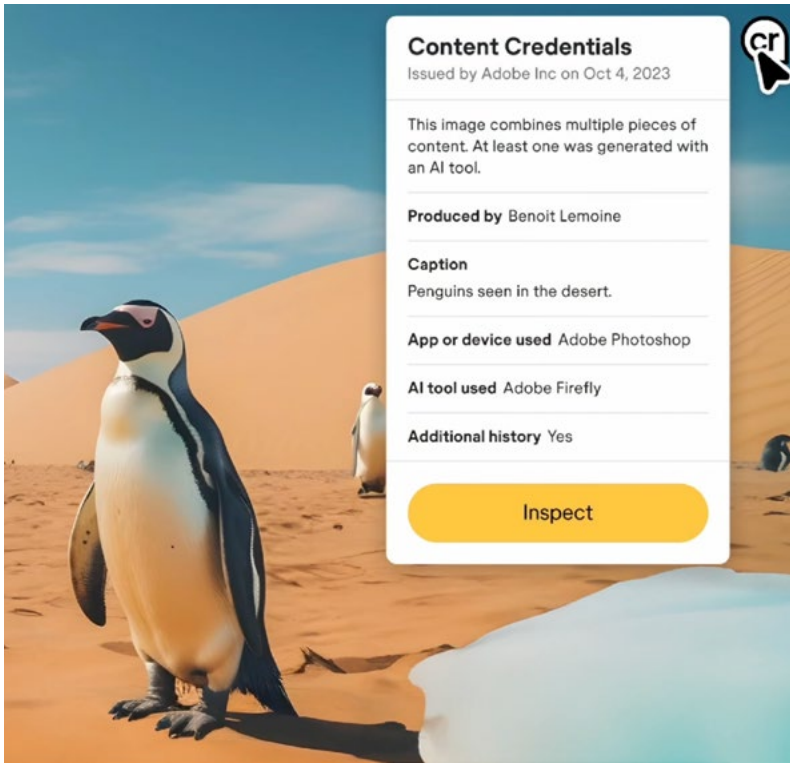


Figure 10: Image from the ContentCredentials.org homepage; this illustration highlights the "Content Credentials" box, which includes the content labeling for the image. (https://contentcredentials.org)

Generative AI Platforms

Generative AI systems automatically insert C2PA metadata in images and videos[174] those systems help produce or modify[175] including to indicate use of those systems in content produced with them.[176] For example, one generative AI system automatically applies C2PA-based metadata to media files in which 100% of the pixels are generated with the generative AI system, including information related to date, apps and devices used and edits made to the media file.[177]

Social Media Platforms

Some social media platforms used in business[178] and consumer settings have begun to incorporate C2PA in their systems as a means of disclosing the presence of AI-generated media to users,[179] and as part of their content

174 Kylie Robinson, Emma Roth, and Richard Lawler, *OpenAI has finally released Sora,* The Verge, December 9, 2024, https://www.theverge.com/2024/12/9/24317092/openai-sora-text-to-video-ai-launch.

175 *OpenAI, C2PA in DALL·E 3, Privacy and policies, Policy FAQ,* OpenAI, https://help.openai.com/en/articles/8912793-c2pa-in-dall-e-3

176 *What Are Content Credentials?*, Adobe, Content Credentials, October 14, 2024, https://helpx.adobe.com/creative-cloud/help/content-credentials.html.

177 *Adobe Firefly, Content Credentials Overview*, Adobe Help Center, February 12, 2025, https://helpx.adobe.com/firefly/get-set-up/learn-the-basics/content-credentials-overview.html.

178 Patrick Corrigan, *LinkedIn Adopts C2PA Standard*, May 15, 2024, https://www.linkedin.com/pulse/linkedin-adopts-c2pa-standard-patrick-corrigan-kwldf/.

179 *About AI Generated Content, Support, Creating Videos, See How is AI-generated content labeled on TikTok?* TikTok, https://support.tiktok.com/en/using-tiktok/creating-videos/ai-generated-content#3 Also see TikTok, Partnering with our industry to advance AI transparency and literacy, Newsroom, May 9, 2024, https://newsroom.tiktok.com/en-us/partnering-with-our-

moderation practices.[180] For instance, a video platform offers expanded video descriptions when video creators use cameras, devices, software or apps that support C2PA.[181]

Search Platforms

Search platforms with image search capabilities can extract metadata derived from C2PA such as information about how an image was created or whether it was edited with AI tools. Image search systems can then display some of that information when people use image search tools.[182]

Digital Advertising Platforms

Many digital and social platform advertising systems also integrate C2PA metadata, or plan to do so.[183] One early C2PA adopter uses C2PA signals to inform the enforcement of advertising policies,[184] and only extracts and displays C2PA metadata if the content manifest is signed by a Certification Authority included in the C2PA Trust List.[185]

Media Creator Identity, and Media Use Preference Systems

A digital media design software system allows media creators to attach C2PA-based identity information including verified names and social media account information to their media files as well as to assert preferences regarding usage of their media files for generative AI related purposes.[186] The identity[187] and use preference information is intended to remain attached to the media files if they are shared in other digital platforms such as social media platforms.

industry-to-advance-ai-transparency-and-literacy.

180 Email interview with Maurice Turner, Global Public Policy at TikTok, by author May 6, 2025. Turner noted in the interview that TikTok uses C2PA to automatically detect and label AI-generated content as part of its content moderation system.

181 *YouTube Help, Building trust on YouTube: 'Captured with a camera' disclosure,* YouTube Policies, YouTube, https://support.google.com/youtube/answer/15446725?hl=en.

182 *Image Metadata in Google Images, Documentation, See How C2PA metadata can appear in Google Search results*, Google Search Central, https://developers.google.com/search/docs/appearance/structured-data/image-license-metadata#c2pa-metadata.

183 Patrick Corrigan, *LinkedIn Adopts C2PA Standard,* May 15, 2024, https://www.linkedin.com/pulse/linkedin-adopts-c2pa-standard-patrick-corrigan-kwldf/.

184 *How we're increasing transparency for gen AI content with the C2PA,* The Keyword, Google, https://blog.google/technology/ai/google-gen-ai-content-transparency-c2pa/.

185 *Image Metadata in Google Images, Documentation, See How C2PA metadata can appear in Google Search results*, Google, *Google Search Central,* https://developers.google.com/search/docs/appearance/structured-data/image-license-metadata#c2pa-metadata..

186 *Andy Parsons, Adobe Content Authenticity, now in public beta, helps creators secure attribution,* Adobe Blog, News, Adobe, April 24, 2025, https://blog.adobe.com/en/publish/2025/04/24/adobe-content-authenticity-now-public-beta-helps-creators-secure-attribution .

187 See: Oscar Rodriguez, *Building Trust in the Digital Age: LinkedIn's New Verified on LinkedIn Service*, LinkedIn, April 24, 2025, https://www.linkedin.com/pulse/building-trust-digital-age-linkedins-new-verified-oscar-rodriguez-nhdre/. See also *Verifications on your LinkedIn profile,* Help, LinkedIn, https://www.linkedin.com/help/linkedin/answer/a1359065 .

Camera Hardware and Software

Camera makers[188] [189] [190] are among early adopters of C2PA, implementing C2PA in physical camera software and related apps to automatically embed C2PA metadata inside photo files.

One camera maker embeds C2PA metadata including 3D depth information into the captured image as an in-camera digital signature,[191] and said it plans to offer C2PA functions in certain cameras only to select media outlets who would need to enable C2PA functions available through a separate upgrade license.

Some camera software providers have integrated C2PA into their own products for generating, signing and sealing provenance metadata in camera or mobile photos[192] and video, and offer extensive technical documentation for C2PA integration.[193]

Mobile Phone Chipsets and Platforms

A mobile chipset provider integrated a C2PA-based authentication system into its mobile platform.[194]

News and Broadcast Publishers

A variety of digital media, print and broadcast media outlets[195] have tested or implemented C2PA, including displaying C2PA metadata in videos and other content.[196] An international news outlet worked with a camera maker and an academic research lab to pilot use of a camera prototype to embed C2PA metadata inside photos.[197] The process involved digitally assigning photos and their corresponding time, date, and location with unique identifiers, cryptographically signing them, registering the photos into a public blockchain, updating the information to reflect subsequent modifications, then distributing the provenance metadata-enhanced photos using C2PA.

One media company announced use of its video content verification process incorporating C2PA-based Content

---

188 *Fuji Film, Lightweight, High Speed: Fujifilm Introduces FUJIFILM GFX100S II Mirrorless Digital Camera*, News, Fuji, (May 16, 2024), https://www.fujifilm.com/us/en/news/digital-cameras/fujifilm-introduces-gfx100sii-mirrorless-digital-camera.

189 *Content Credentials,* Leica Camera, https://leica-camera.com/en-US/photography/content-credentials.

190 *Nikon Z6III firmware update to feature content verification,* Nikon, Technology & Know-how, Nikon,October 14, 2024, https://www.nikon.co.uk/en_GB/learn-and-explore/magazine/gear/nikon-z6iii-firmware-update-to-feature-content-verification.

191 *Sony Announces Second-Generation Flagship Alpha 1 II*, Sony, November 19, 2024, https://alphauniverse.com/stories/sony-announces-secondgeneration-flagship-alpha-1-ii/.

192 *Press Release: Camera Bits Introduces its Solution for Protecting Provenance of C2PA Signed Photos in Effort to Help Combat Fake Imagery,* Blog, Camera Bits, May 6, 2024, https://home.camerabits.com/2024/05/06/press-release-camera-bits-introduces-its-solution-for-protecting-provenance-of-c2pa-signed-photos-in-effort-to-help-combat-fake-imagery/.

193 *Integration Guide, Truepic Enterprise C2PA Tools,* Truepic, https://lens.truepic.dev/docs/display-library.

194 *The OnQ Team, Trust what you see: How Truepic authenticates images and videos in the age of deepfakes,* OnQ Blog, Qualcomm, November 25, 2024, https://www.qualcomm.com/news/onq/2024/11/trust-what-you-see-how-truepic-authenticates-images-and-videos.

195 Transparency tool launched by BBC Verify, BBC News, March 4, 2024, https://www.bbc.com/news/av/world-68474465. See also: Henri Astier and Gianluca Avagnina, Haiti violence: Haiti gangs demand PM resign after mass jailbreak, BBC News, March 4, 2024, https://www.bbc.com/news/world-latin-america-68462851.

196 Charlie Halford, *Mark the good stuff: Content provenance and the fight against disinformation*, BBC, Research and Development, March 4, 2024, https://www.bbc.co.uk/rd/blog/2024-03-c2pa-verification-news-journalism-credentials.

197 *Reuters new proof of concept employs authentication system to securely capture, store and verify photographs*, August 30, 2023, https://www.thomsonreuters.com/en/press-releases/20Thomson Reuters23/august/reuters-new-proof-of-concept-employs-authentication-system-to-securely-capture-store-and-verify-photographs.html.

Credentials, and later issued a correction stating that the "verified" video had actually been misleadingly edited by a social media user to include elements of an older, unrelated video.[198]

Media Management and Delivery Systems

A broadcasting company that owns local news and international media outlets worked with a data services and software company to enable C2PA throughout its media content production and distribution processes, helping to ensure that C2PA metadata is generated and maintained at every step of the content provenance chain.[199] The broadcaster adopted C2PA to track use of AI in content editing and production, and to declare its use in content, as well as for reasons unrelated to AI, including keeping track of intellectual property, and consistently tracking all content metadata including provenance metadata.

The prototype project involved external storage of C2PA metadata in a "sidecar" outside the content file. Developers said future implementation could entail storage of C2PA metadata inside the content file in a "media wrapper."

A large[200] content delivery network adopted C2PA,[201] allowing those using the image content storage and management system such as website publishers to preserve C2PA metadata that has been cryptographically signed by the content delivery network and embed it in the content.[202] The network also automatically updates C2PA metadata if the image has been resized or otherwise transformed, appending the additional information to the image. The network automatically caches all transformed images on its global network,[203] and uses another external entity to conduct end-entity certificates used for C2PA manifests.[204]

An image and video management platform also supports adding signed C2PA metadata to specify whether transformations altered the pixels or merely optimized them for delivery of those media files.[205]

Business Document Software

An organization that provides an open standard protocol for photos and documents has adopted C2PA for smart

---

198 BBC's correction stated, "...it has since been drawn to our attention that some of this was misleadingly edited in from an older unrelated video by a social media user. In line with our commitment to transparency, we have reviewed the video and muted the relevant passage of audio. We have also updated our commentary on the video to reflect those changes." *Transparency tool launched by BBC Verify,* BBC News, March 4, 2024, https://www.bbc.com/news/av/world-68474465. See also: Henri Astier and Gianluca Avagnina, *Haiti violence: Haiti gangs demand PM resign after mass jailbreak*, BBC News, March 4, 2024, https://www.bbc.com/news/world-latin-america-68462851.

199 *AWS Innovation with Sinclair, AWS Innovation Ambassadors,* AWS Architecture Center, June 2024, https://aws.amazon.com/podcasts/innovation-ambassadors/ia-podcast-ep-47-sinclair/?podcast-list-wide.sort-by=item.additionalFields.EpisodeNum&podcast-list-wide.sort-order=desc.

200 *Historical trends in the usage statistics of reverse proxy services for websites*, Web Technology Surveys, May 2024, https://w3techs.com/technologies/history_overview/proxy/all.

201 Will Allen, *Preserving content provenance by integrating Content Credentials into Cloudflare Images,* Cloudflare, The Cloudflare Blog, February 3, 2025, https://blog.cloudflare.com/preserve-content-credentials-with-cloudflare-images

202 *Preserve Content Credentials, Cloudflare Docs, Cloudflare Images*, Cloudflare, https://developers.cloudflare.com/images/transform-images/preserve-content-credentials/.

203 *Getting Started, Cloudflare Docs, Cloudflare Images, See Enable Transformations on Your Zone*, Cloudflare, https://developers.cloudflare.com/images/get-started/#enable-transformations.

204 Will Allen, *Preserving content provenance by integrating Content Credentials into Cloudflare Images,* Cloudflare, The Cloudflare Blog, February 3, 2025, https://blog.cloudflare.com/preserve-content-credentials-with-cloudflare-images.

205 Eric Portis, *Combating Fake Visuals: Cloudinary's New C2PA Standard Implementation*, Cloudinary, Blog, July 2, 2024, https://cloudinary.com/blog/c2pa-standard-implementation.

phone cameras and document scanning applications.[206]

Audio Software

An open-source multimedia authentication and verification system[207] uses C2PA in its audio recorder software.[208]

Digital Watermarks

A digital watermark provider allows recovery of original C2PA metadata via its digital watermarks, even if it is stripped from the content file or manipulated.[209] The system adds a digital watermark containing a reference to the C2PA manifest to the content file; the watermark system detects the watermark on the content, verifying a match to the externally-stored manifest.

Data Governance Tools

A data provenance system records the provenance of images that already include C2PA metadata, and offers C2PA manifest repository services.[210]

Entertainment and Athlete Talent Identification Systems

A Proof of Concept for a system that creates, documents and measures consent-based digital replicas of "notable" legal and natural public figures such as actors and athletes embeds cryptographic metadata in a watermark or fingerprint as a C2PA Content Credential for the digital replica identity in a multi-party workflow and detects C2PA metadata.[211]

---

206 *TWAIN Working Group and the Coalition for Content Provenance and Authenticity (C2PA) enter into a Liaison Partnership Agreement*, TWAIN Working Group, August 27, 2024, https://www.prweb.com/releases/twain-working-group-and-the-coalition-for-content-provenance-and-authenticity-c2pa-enter-into-a-liaison-partnership-agreement-302229575.html.

207 ProofMode was developed by Guardian Project, open tech research and product design organization Okthanks, and WITNESS, an organization that helps people use video and technology to protect and defend human rights that has evaluated potential harms of C2PA. See: Jacobo Castellanos, *WITNESS and the C2PA Harms and Misuse Assessment Process: Confronting Potential Harms Early in the Process of Developing Authenticity and Provenance Infrastructure*, WITNESS, December 2, 2021, https://blog.witness.org/2021/12/witness-and-the-c2pa-harms-and-misuse-assessment-process/.

208 *ProofMode, Adding Content Credentials (C2PA) to Audio Recordings Using SimpleC2PA*, January 25, 2024, https://proofmode.org/blog/audio-c2pa.

209 *Digimarc Brings Digital Watermarking to the C2PA 2.1 Standard, Digimarc Newsroom*, Digimarc, (October 8, 2024), https://www.digimarc.com/press-releases/2024/10/08/digimarc-brings-digital-watermarking-c2pa-21-standard.

210 *Increase long-term trust in C2PA media provenance with DataTrails*, DataTrails Support, https://support.datatrails.ai/hc/en-gb/articles/12546557959058-Increase-long-term-trust-in-C2PA-media-provenance-with-DataTrails.

211 *Digital Replicas and Talent ID: Provenance, Verification and New Automated Workflows*, HAND Labs, 2024, https://handidentity.com/hand-labs/.

Appendix A: C2PA Timeline Reference Citations

This appendix lists the references that specifically support the information in Figure 1, the timeline of C2PA's development.

T1 *Project Origin: Protecting Trusted Media,* Home Page. https://www.originproject.info. See also Laura Ellis, *Project Origin: Securing Trust in Media,* BBC, Trusted News Initiative, https://www.bbc.com/beyondfakenews/trusted-news-initiative/project-origin-securing-trust-in-media. See also: Nils Martin Silvola, *Media City Bergen blir del av Project Origin – skal bekjempe falsk informasjon* (Media City Bergen becomes part of Project Origin - to combat false information), Journalisten, April 11, 2023, https://www.journalisten.no/media-city-bergen-blir-del-av-project-origin-skal-bekjempe-falsk-informasjon/567142.

T2 Adobe Communications Team, *Introducing the Content Authenticity Initiative,* Adobe Blog, November 4, 2019, https://blog.adobe.com/en/publish/2019/11/04/content-authenticity-initiative. See also Leonard Rosenthol, et al. *The Content Authenticity Initiative Setting the Standard for Digital Content Attribution*, Adobe, August 2020, https://acrobat.adobe.com/link/track?uri=urn%3Aaaid%3Ascds%3AUS%3A2c6361d5-b8da-4aca-89bd-1ed66cd22d19&viewer%21megaVerb=group-discover.

T3 Virtual interview with Andy Parsons, Senior Director Content Authenticity at Adobe, May 5, 2025 by author. In the interview Parsons said that CAI is 100% funded by Adobe. See also Discord, Content Authenticity Initiative, https://discord.com/channels/983153151341371422/@home.

T4 *C2PA Founding Press Release*, News, Coalition for Content Provenance and Authenticity, February 22, 2021, https://c2pa.org/post/c2pa_initial_pr/. See also Eric Horvitz, *A promising step forward on disinformation*, Microsoft on the Issues, Microsoft, February 22, 2021, https://blogs.microsoft.com/on-the-issues/2021/02/22/deepfakes-disinformation-c2pa-origin-cai/. See also *Microsoft, Technology and media entities join forces to create standards group aimed at building trust in online content*, Microsoft News Center, Microsoft, February 22, 2021, https://news.microsoft.com/2021/02/22/technology-and-media-entities-join-forces-to-create-standards-group-aimed-at-building-trust-in-online-content/. See also "*Going Beyond Source Code in 2021: Joint Development Foundation and Open Standards Efforts*, Blog, The Linux Foundation, November 16, 2021, https://www.linuxfoundation.org/blog/blog/going-beyond-source-code-in-2021-joint-development-foundation-and-open-standards-efforts. The foundational governance documents of C2PA are are articulated in *Background and Membership Agreement Package Instructions*. See:
Linux Foundation, Joint Development Foundation, *Background and Membership Instructions,* https://cdn.platform.linuxfoundation.org/agreements/c2pa-fund.pdf . See also note T6.

T5 *Twitter Joins C2PA,* News, Coalition for Content Provenance and Authenticity, May 13, 2021, https://c2pa.org/post/twitter_pr/. See also: *Sony Corporation Joins C2PA as Steering Committee Member,* News, Coalition for Content Provenance and Authenticity, March 16, 2022, https://c2pa.org/post/sony_pr/. See also: *C2PA welcomes first advertising holding company to steering committee*, News, Coalition for Content Provenance and Authenticity, June 5, 2023, https://c2pa.org/post/publicis_pr/. See also: *Amazon Joins the C2PA Steering Committee*, News, Coalition for Content Provenance and Authenticity, September 12, 2024, https://c2pa.org/post/amazon_pr/ . See also: *Google to join C2PA to help increase transparency around digital content*, News, Coalition for Content Provenance and Authenticity, February 8, 2024, https://c2pa.org/post/google_pr/ . See also: *Meta Joins the C2PA Steering Committee*, News, Coalition for Content Provenance and Authenticity, September 5, 2024, https://c2pa.org/post/meta_pr/. See also: *OpenAI Joins C2PA Steering Committee*, News, Coalition for Content Provenance and Authenticity, May 7, 2024, https://c2pa.org/post/openai_pr/ . X (formerly Twitter) is not currently listed among C2PA Steering Committee Members on the C2PA website. As of 25 July, 2025, the C2PA website listed the following C2PA Steering Committee Members: Adobe, Amazon, BBC, Google, Intel, Meta, Microsoft, OpenAI, Publicis Groupe, Sony, and Truepic. See Linux Foundation Projects, *C2PA Coalition for Content Provenance and Authenticity*, https://c2pa.org/. See also: Irene Tham, *Fighting deepfakes with content "nutrition labels,"* The Straits Times, (Singapore) June 2, 2025. https://www.straitstimes.com/opinion/fighting-deepfakes-with-content-nutrition-labels .

T6 The foundational charter for C2PA is published by the Linux Foundation's Joint Development Foundation Projects, LLC. C2PA is a project of the Joint Development Foundation, or JDF. The JDF has extensive governance policies and legal agreements in place, which are articulated in the *Background and Membership Agreement Package Instructions.* See:
Linux Foundation, Joint Development Foundation, *Background and Membership Instructions, https://cdn.platform.linuxfoundation.org/agreements/c2pa-fund.pdf* . From the Agreement document: "The Steering Committee is the body that is responsible for governing the Project, including approving Final Deliverables." See Linux Foundation, Joint Development Foundation, Project Charter, 6. Organization., 6.1. Steering Committee. https://cdn.platform.linuxfoundation.org/agreements/c2pa-fund.pdf . "No sooner than 30 days after a Draft Deliverable has been designated as a Working Group Approved Deliverable, the Project Chairperson or their designee will present that Working Group Approved Draft Deliverable to the Steering Committee for Approval. Upon Approval by the Steering Committee, that Draft Deliverable will be designated a Final Deliverable as of the date of such Steering Committee Approval." See Linux Foundation, Joint Development Foundation, Project Charter, Appendix A Traditional Mode Governance, 4. Deliverable Development Process., 4.3. Final Approval, https://cdn.platform.linuxfoundation.org/agreements/c2pa-fund.pdf . "Changes to a Working Group Charter must be approved by the Steering Committee." Linux Foundation, Joint Development Foundation, Project Charter, 9. Working Groups., https://cdn.

platform.linuxfoundation.org/agreements/c2pa-fund.pdf.

For additional details on the governance, fee structure, and other policies regarding C2PA, see, for example in the document: Membership Agreement, p. 1; Project Charter section, pages 1-7 (the charter includes the governance and legal terms for the Project and its Steering Committees); Working Group Charter section, pages 1-3; Traditional Mode Governance section, Appendix A page 1 -3; Appendix C contains the non-member agreement, page 1, and the Project Sponsorship Agreement, pages 1-2, among additional information.

See also (archived at the *Wayback Machine*): *C2PA Introduction Deck, Membership Level Rights*, Coalition for Content Provenance and Authenticity, https://web.archive.org/web/20250214060431/https://c2pa.org/files/C2PA_Introduction_Deck.pdf . Adobe does not provide any funding to the C2PA Joint Development Foundation project beyond its $27,000 Steering Committee fee. In an mail interview with Andrew Cha, Senior Public Relations Manager at Adobe, July 25, 2025 by the author, Cha confirmed this fact.

T7 *C2PA Releases Draft Spec*, News, Coalition for Content Provenance and Authenticity, September 1, 2021, https://c2pa.org/post/draft_1_pr/.

T8 *C2PA Releases Specification of World's First Industry Standard for Content Provenance*, News, Coalition for Content Provenance and Authenticity, January 26, 2022, https://c2pa.org/post/release_1_pr/.

T9 Andy Parsons, *Adobe Max 2023: Milestone wave of Content Credentials adoption with industry partners Microsoft, Leica Camera, Nikon, Publicis Group, and more*, Adobe Blog, News, October 10, 2023, https://blog.adobe.com/en/publish/2023/10/10/new-content-credentials-icon-transparency. See also: *Introducing Official Content Credentials Icon*, News, Coalition for Content Provenance and Authenticity, https://c2pa.org/post/publicis_pr/.

T10 *Content Credentials: C2PA Technical Specifications,* January 2024, Coalition for Content Provenance and Authenticity, https://c2pa.org/specifications/specifications/2.0/specs/C2PA_Specification.html. See: *Version History 2.0 - January 2024*. From the Version 2.0 document: "This version represents a significant departure from previous versions. It no longer has any references to actors as humans or organizations, they can only be hardware or software entities."

T11 *Meeting Notes 20 February 2024*, Creator Assertions Working Group, https://cawg.io/meeting-notes/2024-02-20/.

T12 Eric Scouten, *Identity Assertion*, Creator Assertions Working Group, September 9, 2024, https://cawg.io/identity/1.0/.

T13 See: *The C2PA and Embassy of France Collaborate to Advance Authenticity and Transparency in Digital Content*, News, Coalition for Content Provenance and Authenticity, October 2, 2024, https://c2pa.org/post/embasssyoffrance_pr/. From the document: "The C2PA's specification is currently being examined by the organization's TC 171/SC 2 committee and is expected to become a global standard in early 2025." See also: *ISO/TC171/SC2, Document file formats, EDMS systems and authenticity information,* International Organization for Standardization, https://www.iso.org/committee/53674.html.

Report permalink: https://worldprivacyforum.org/posts/privacy-identity-and-trust-in-c2pa
Version history: 1.0