Fighting Misinformation with Authenticated C2PA Provenance Metadata

Dr. Nigel Earnshaw
The British Broadcasting Corporation
London, United Kingdom
Nigel.Earnshaw@bbc.co.uk

Jonathan Dupras
CBC/Radio-Canada
Montreal, Quebec, Canada
jonathan.dupras@radio-canada.ca

Bruce MacCormack
Neural Transform
Lunenburg, Nova Scotia, Canada
Bruce@NeuralTransform.com

Abstract – Over the last three years, teams from Microsoft, The New York Times, CBC/Radio-Canada and The BBC have come together as Project Origin. This group has participated as part of a wider community in the standardization of provenance signaling technologies to attach authenticated metadata to media content. The Coalition for Content Provenance and Authenticity (C2PA) specification was the result of these collaborative efforts. This paper will outline the features of the C2PA specification, and the work being undertaken to add this functionality to existing media production workflows to add transparency and counter disinformation and malicious use of synthetic media.

The Growing Threat to the News Ecosystem

The widespread availability of powerful generative media tools has challenged the fundamental “seeing is believing” basis of much of our modern news ecosystem. A need for a higher level of validation will be required for publishers to maintain the trust of our audiences. Established media organizations are at risk of being fooled into using illegitimate content. Brands can be impersonated and have their credibility used to amplify both disinformation and misinformation. The long-established practices for fact checking and defending the veracity of news content are at risk of being overwhelmed by the growing volume of unsubstantiated content in the system.

The emergence of generative AI tools for creating media has led to a call for the development of AI tools to detect this new form of content. This approach will always be faced with a cat and mouse dilemma. The advances in creation will outpace the ability to detect. Worse, detection tools when used with Generative Adversarial Networks will dramatically improve the ability of misinformation to avoid

This paper is excerpted from the Proceedings of the 2023 NAB Broadcast Engineering and Information Technology (BEIT) Conference, © 2023, National Association of Broadcasters, 1 M Street SE, Washington, DC 20003 USA.

detection. It is anticipated that detection tools in general use will have usable life spans measured in weeks.

As an alternative, media provenance uses well understood cryptographic technology to validate the identity of a publisher and the technical integrity of the content. It is a solution that can operate at scale in a mechanized approach. The challenge will be to come to a universally accepted method of using provenance manifests across unrelated organizations and industries.

Building a Coalition

In 2019 teams from Microsoft, the New York Times, CBC/Radio-Canada and the BBC came together as Project Origin to focus on protecting the integrity of content when consumed over any platform. Rather than think about detecting which content might have been 'faked' the more straightforward approach of being able to determine in a cryptographically secure way the organization or individual which claims to have published content was felt to be more stable and practical. In addition, this approach does not create any uncertainty or imply judgment and so enriches the media ecosystem in a positive way.

At the same time another consortium, the Content Authentication Initiative (CAI), led by Adobe, had been considering this challenge from a creator perspective and had made the same observations. Together the combined membership of these two groups, along with ARM, Intel, Truepic and Sony and others, formed an open standard body – the Coalition for Content Provenance and Authenticity (C2PA). The C2PA open technical standard (V1.0) for authenticated provenance metadata was published in January 2022. This has been updated and maintained over the last year [1].

The current membership of C2PA reflects the way provenance data can originate at different points in the media value chain and can, if permitted, be carried along the contribution and distribution networks in a way that preserves the chain of provenance through to the end consumer. Participation in C2PA includes well known and substantial organizations representing camera designers and manufacturers, content infrastructure companies, broadcasters and news organizations, content tool and application vendors, human rights advocates, and social media platforms.

The C2PA Specification - A common core approach

Design Goals

The C2PA process began by articulating its Guiding Principles for the C2PA Designs and Specifications. These include the overarching goals of providing a way for producers and custodians of any given content to assert provenance data in a verifiable manner. A second goal is that the "C2PA specifications should not provide value judgments about whether a given set of provenance data is good or bad, merely whether the assertions included within can be verified as associated with the underlying asset, correctly formed, and free from tampering".

Throughout the whole process much emphasis has been placed on the design goals of ensuring privacy, responsibility, scalability, extensibility and interoperability among others.

Overview of the C2PA technical features

The C2PA technical approach is to authenticate provenance related metadata in the form of a manifest data structure which includes a unique impression of the content bits in the form of a secure cryptographic hash. Having then 'bound' the content impression within the metadata manifest, the manifest is then cryptographically signed using public credentials recognized by validators in the field. In this way a validator can determine the technical integrity of the received manifest and content by validating the signature on the manifest and checking that the received content matches the content impression within the manifest.

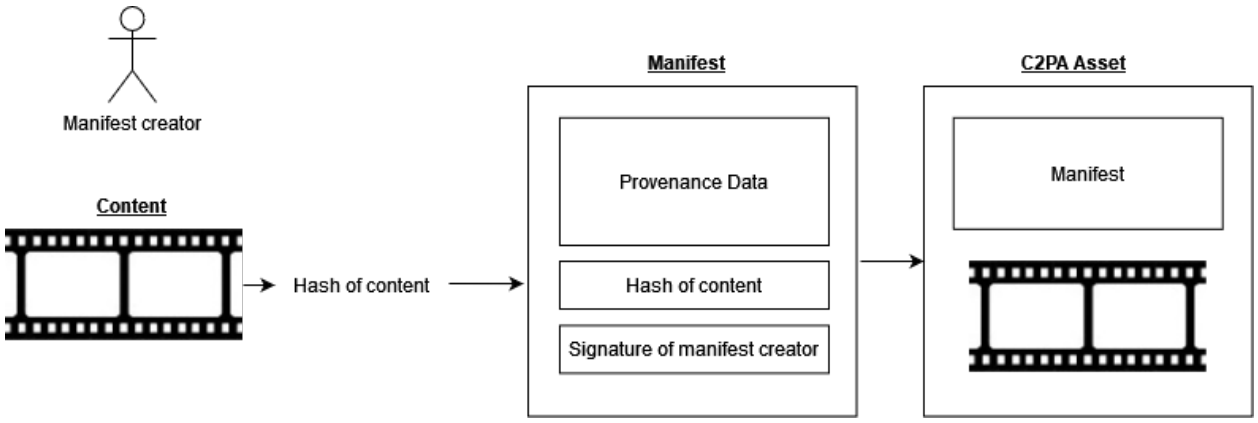

FIGURE 1. SIMPLIFIED SCHEMATIC OF THE C2PA ASSET

The simplified schematic above shows the relationship between the provenance metadata, content manifest and the C2PA asset. The first consideration is the binding between the bits that make up the content and the provenance metadata. In the simplest case as shown above, the cryptographic hash of the content is included in the manifest. The exact parts of the file included within the scope of the hash are carefully chosen according to the specification and signaled within the manifest.

The next consideration is the integrity of the manifest as a whole. This is assured through well-known and established digital signature algorithms being used to protect the critical parts of the manifest to prevent any changes to either the manifest or the content being made without invalidating the overall structure.

Finally, the specification describes how to include the manifest within the well-known container formats commonly used today. These include JPEG, PNG, SVG, PDF amongst others and for audio/video presentations the base media file format (BMFF) [2] is included. Note that being able to insert the manifest into the data file container to form a C2PA asset is a practical consideration only and does not imply any added security. The C2PA specification anticipates the separation of manifest from content and the possibility of services to discover manifests from content alone. When considering validation, no assumption is made that the co-located manifest is the correct manifest authored for the accompanying content.

When a validating client, e.g., a browser validates a C2PA asset as shown in the simple case as above, then the following steps are taken.

1. The validator locates and extracts the manifest from the file format.
2. The validator verifies the data in the manifest is valid according to the signature. This paper discusses the trust model and acquisition of trusted public keys below. Note that some intermediate security credentials such as x.509 certificates are conveyed in the manifest structure to facilitate the conveyance of a signing certificate related to a known trust anchor. Details of signature and hash algorithms are also conveyed within the signed data.
3. The validator independently determines the hash impression of the content using the hash algorithm signaled in the manifest and verifies this matches the hash embedded within the manifest.
4. The validator matches the assertions included in the manifest with the signed representation of these assertions in the manifest claim. This 'claim' structure is described later in this paper.

Once these checks have been performed, and if all are passed, the manifest can be accepted as provenance data about the content as authored by the party who signed the manifest.

Metadata and security information standards

The metadata contained within the manifest is described in the specification and allows for proprietary schema. In addition to the C2PA defined terms, guidance is provided for the inclusion of EXIF data [3], IPTC Photo and Video metadata [4] and the use of Schema.org [5].

C2PA manifests use the JUMBF standard [6] to create the various boxes described in the next section. The schema that define the inner data structures are specified using the CBOR Data Definition Language (CDDL) [7] and represented in the data as Compressed Binary Object Representation (CBOR) [8]. In this way the manifest structure is conveyed in a compact format to reduce the size of the manifest within the file, in line with the design goals.

Security data structures which facilitate signing and transport of credentials are represented using CBOR Object Signing and Encryption standard (COSE) [9]. Presently the only security signing credentials specified are the X.509 certificates [9] which use their customary ASN.1 specification rules and Distinguished Encoding Rules (DER) binary format.

The components of a manifest

This section describes in a little more detail the internal component parts of a manifest. This enables an understanding of how elements within a manifest are related to each other, how manifests within a manifest store relate to each other in a chain of provenance and how updates can be added without breaking an established chain. This section will also describe how endorsements may be used to allow those parties who may have to interrupt the provenance chain demonstrate agreement with a previous signer.

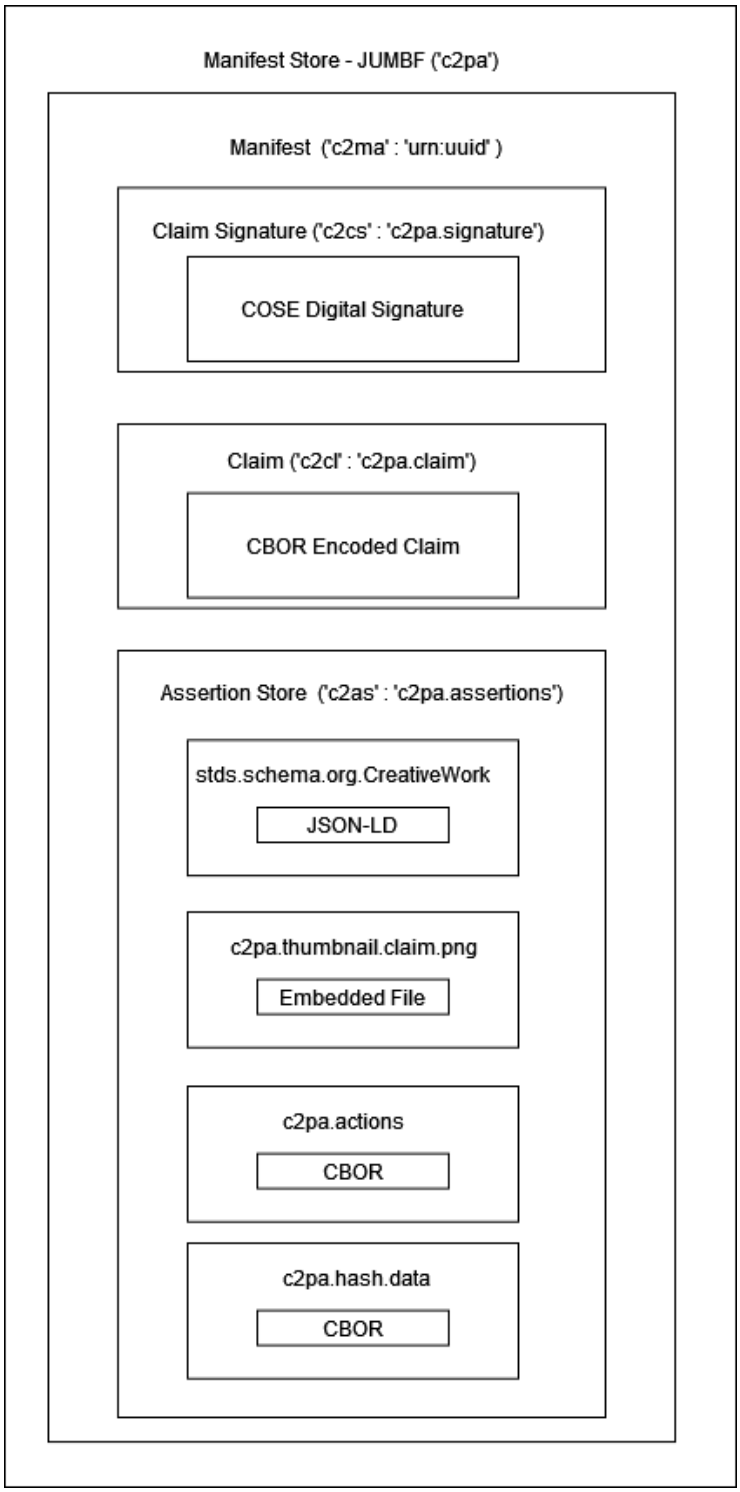

FIGURE 2.THE COMPONENTS OF A MANIFEST

The schema above shows a visualization of the manifest store. This JUMBF box represents the self-contained unit that conveys the C2PA authenticated provenance metadata. All JUMBF boxes are shown with both their 4-character reference identifier and a dot delimited label encoded into the JUMBF box structure.

From the above we see that this outer manifest store JUMBF box contains within it a series of functionally related JUMBF boxes. As the name suggests, the manifest store can contain a number of independently verifiable manifest structures, though only one is shown above for clarity. Note that in the case where more than one manifest structure is included in the store, each will play a specific role in the provenance history of the content and there will be a signaled relationship between them as described later in this paper.

The inner structure of the manifest shows the core C2PA concept of assertions. Assertions are where the creator of the manifest, which may be an editing tool, expresses actions or includes data about the content for consumption by subsequent actors in the value chain. Assertions may be expressed as a predefined metadata schema or be one of the C2PA defined standard assertions. The data for each assertion is held within its own defined structure and may be encoded as CBOR, JSON or Embedded File content type. CBOR is most typical for the native C2PA standard assertions, JSON is typically where an external schema has been adopted and a thumbnail jpeg image would be an example of an assertion using Embedded File content type.

In many cases metadata about assertions can be optionally included, expressing additional information about the assertion such as date and time associated with that assertion. For example, if the assertion expressed an edit, the metadata could indicate the date this was made. This makes for a comprehensive vocabulary of provenance data.

Assertions

The standard C2PA assertions include amongst others:

- Assertion metadata - to further describe assertions.
- Cloud data - serving as a reference to any data hosted on the cloud.
- Thumbnail - allowing inclusion of a thumbnail image.
- Actions - a multivalued assertion to describe typical editing actions.
- Ingredient - to reference an included ingredient manifest.
- Endorsement - to describe the endorsement of some actions by another organization.
- Claim review - to describe a review of a third-party claim.
- Creative Work - a schema to describe a creative work in detail.
- Various bindings - discussed below.

Binding the content to the manifest

As listed above, the standard assertions include a special assertion responsible for capturing the appropriate content binding data which cryptographically ensures the logical relationship between the content data and the manifest. C2PA refers to these techniques of using cryptographically unique representations of content as “hard bindings”. This contrasts with additional data that might assist such a process in a less precise manner termed “soft bindings”. The way in which these hard bindings are formulated depends to some extent on the file format of the content, just as the specification for how to embed a manifest within a file has to be file specific. To this end C2PA has defined how to assemble the content data into an input of a secure cryptographic hash for the case of the base media file format (BMFF) the common box-based format for audio and video presentations. Additionally, the case for

non-BMFF-based file formats, typical for images, is covered through the ‘Data Hash’ assertion. Finally, for those formats which are ‘box like’ but not compatible with BMFF, a ‘General Box Hash’ is specified.

Note that the choice of cryptographic hash algorithm is part of the specification, as are all choices of cryptographic algorithms including signature algorithms. These are currently well known and trusted algorithms used in many applications today and include for example SHA-256 as a cryptographic hash and signatures based upon ECDSA using well known curves.

Focusing in on the BMFF binding which has been designed with audio/video presentations in mind there are currently two techniques fully specified. The first relates to situations in which an entire file can be downloaded before playback begins. In this case the manifest contains a hash of all the data except for specific boxes, or parts thereof, which may safely be subject to change during transit. Additionally, the position of the boxes within the file is included in the hash algorithm input data to prevent manipulation of the overall structure by an adversary.

There is a second technique for forming hard bindings over BMFF file formats for those cases in which the data is likely to be downloaded in a progressive or fragmented way, that is, for cases where playback is expected to begin before the whole file has been downloaded by the client. In these cases, the data representing those parts of the file not relating to the audio/video samples can still be verified before playback. However, the content samples themselves are hashed not as one monolithic block, but as a series of hashes relating to each sample chunk of data held within the BMFF mdat box and addressed through the BMFF format. This leads to a requirement to validate many individual hashes separately. To make this manageable these hashes are prearranged within a Merkle tree structure when preparing the manifest. In this way, only the root of the Merkle tree needs to be included within the BMFF assertion structure of the manifest since this relates in a precise mathematical way to all individual hashes.

Validation of individual chunks can be made by the client on download through the associated timely download of data about that hash in relation to the overall Merkle tree structure. This additional Merkle data, which represents the unique list of siblings of each element in the path between the hash and the row element included in the manifest, allows the client to evaluate and process each individual chunk hash to determine whether this cryptographically aligns with the declared value in static manifest. In this way each chunk can be verified independently of each other and without waiting for all the data.

A simplified example of how to validate content data without associated adjacent data is shown in the schematic below. In this case the manifest contains the representation of the root of a Merkle tree, though the validator does not have access to all the Merkle elements at one time. The missing elements are marked in dotted boxes and arrows. However, the Merkle tree elements represented by the boxes marked A, B and C have been delivered with the data in a format that makes their position within the tree structure known. In the figure below the validator can verify the data by first calculating the hash and then using the data at A, B and C to recalculate the root of the Merkle tree. Matching this against the authenticated root value in the manifest allows the data to be verified.

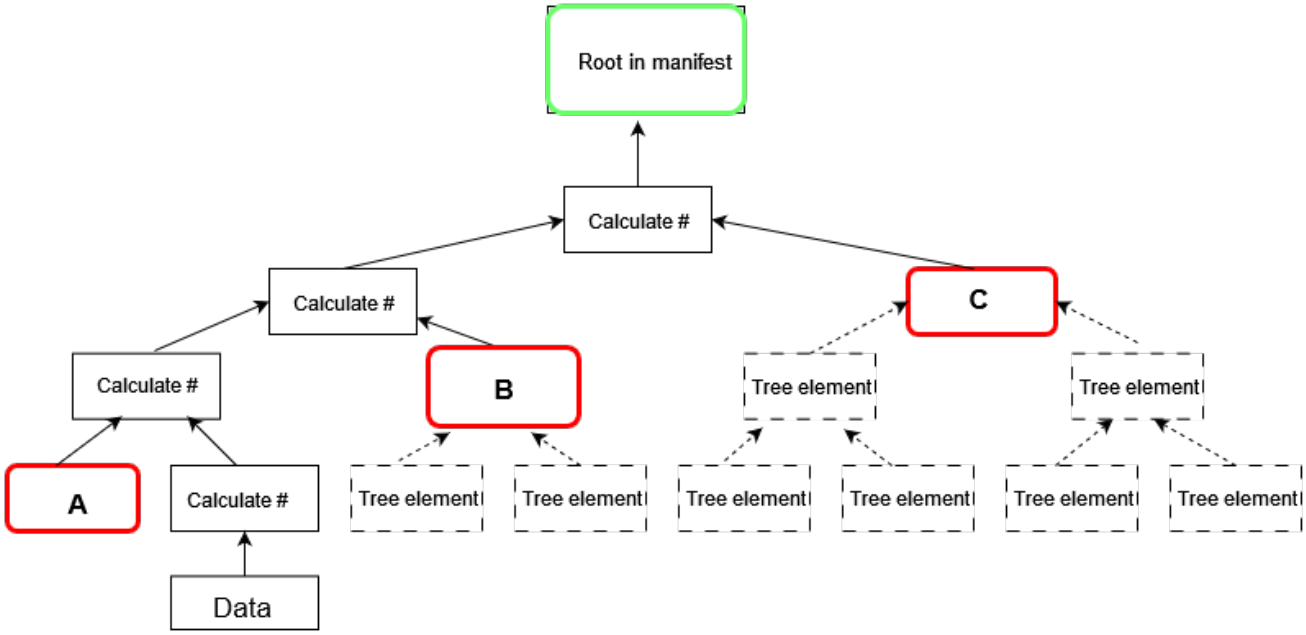

FIGURE 3. DATA CHUNK HASHES ARRANGED INTO A MERKLE TREE STRUCTURE.

Claims

The claim box is also shown in as a component of the manifest structure in Figure 2 above.

The claim is where a cryptographic hash of each assertion is collected together by the claim generator responsible for the manifest. This list of hashes and references can be enriched with other data such as data describing the claim generator itself, a title of the asset, its format, or some housekeeping data indicating the algorithms involved.

The claim box, represented in the manifest as CBOR data, ultimately forms part of the payload of the signature algorithm. Having this indirection between assertion and the signature payload enables assertions to be hosted in the cloud, since it makes it possible to know the manifest has not been tampered with even without the retrieving the cloud assertion. This structure also allows a chain of manifests to be created and verified without requiring access to the underlying content associated with the earlier manifests. A description of a chain of manifests is given in a later section.

Signature

The signature box contains a representation of the signature of the claim payload. The COSE signature data structure used also conveys data about how the signature is formed. All signature algorithms defined use well known public key cryptography and C2PA uses x.509 certificates to allow a public signing key to be cryptographically chained via intermediate certificates to a trust anchor known to the validator. In addition, the data structure can optionally convey an countersignature from a third-party time-stamp authority, witnessing the existence of the manifest at a specific time. Again, there is capacity to convey the PKI credentials associated with the time-stamp authority to a known trust anchor residing in the validator.

Types and special uses of manifests

The most straightforward application of the C2PA standard is as described with a single actor authoring provenance data for content they wish to create or publish. However, a strong use case is when a creator wishes to take already existing content with its own C2PA manifest and modify this or create something based on this, either exclusively or with other content items.

Ingredients. To meet this use case the C2PA manifest store can include one or more component manifest as an 'Ingredient' manifest. These ingredient manifests cannot be altered since that would invalidate their signatures which protect their integrity.

Inclusion of ingredients means that the manifest being authored is now known as the 'active manifest' to designate that this is the manifest at the head of the chain and must be validated first. Note that the inclusion of an ingredient is represented cryptographically by adding a reference into the assertion store of the parent active manifest, thereby cryptographically binding these items together logically.

This concept can be extended again when a creator uses the composite item and includes the active manifest of the original composite item and its associated ingredients as a chain of ingredients in the new items with a new active manifest.

Note that the original content represented by the ingredient manifests is not conveyed with the composite manifest which is bound to the composite content. Therefore, the content bindings assertions in the ingredient manifests cannot be verified. However, due to the indirection of the claim box, the signature of the ingredient manifests can be validated along with the remaining assertion values.

Endorsement. A further feature of manifest signaling is when a party wants to acknowledge or 'endorse' permitted actions of specific downstream parties as they transform the content in specific technical ways.

This can happen for example when a publisher provides a direct content feed to a social media platform or content distribution network for wider distribution. There are use cases where the downstream platform will wish to transcode the incoming content to provide a service to many different client devices with different technical parameters but would not wish to change the meaning of the content in any way. However, with C2PA hard bindings any such change will invalidate the original manifest by changing the content bits.

To handle this case a downstream party can be endorsed to publish, transcode or repackage content by obtaining an authenticated endorsement from the content originator that contains the downstream party public key. In this way the downstream distributer can issue a new active manifest and place the original as an ingredient. The new manifest should include the endorsement assertion to signal to the validator that this content transformation is limited and by arrangement with the originating party.

User experience recommendations

The power of provenance data resides in the signaling of the received authenticated data to the asset consumer after processing by the validator. In this way the asset consumer can be assured of the

source of the content, which may or may not coincide with the appearance, and whether it is ‘as sent’, or has been tampered with. More likely any changes have been recorded along the path and so there is a requirement to consider how this information can be conveyed to the asset consumer and how effectively it can be managed by the content creators. This is a matter for the creators of the validator and content creation tools and will provide a point of discrimination in the market as this technology evolves further.

The C2PA has developed clear recommendations and guidance [11] for implementers of provenance-enabled user experiences (UX). The guiding principles are;

- provide asset creators a means to capture information and history about the content they are creating, and

- provide asset consumers information and history about the content they are experiencing, thereby empowering them to understand where it came from and decide how much to trust it.

These principles and the work done are particularly important so as not to overload the user, To avoid this the C2PA considers 4 levels of disclosure to guide the designs

- Level 1: An indication that C2PA data is present and its cryptographic validation status.

- Level 2: A summary of C2PA data available for a given asset. This level should provide enough information for the particular content, user, and context to allow the consumer to understand to a sufficient degree how the asset came to its current state.

- Level 3: A detailed display of all relevant provenance data. Note that the relevance of certain items over others is contextual and determined by the UX implementer.

- Level 4: For sophisticated, forensic investigatory usage, a tool capable of revealing all the granular detail of signatures and trust signals is recommended.

FIGURE 4. FOUR LEVELS OF DISCLOSURE (FROM C2PA USER EXPERIENCE GUIDANCE DOC)

Trusted Public Key Infrastructure

As described above, the manifest achieves integrity and security through public key cryptography. This in turn assumes the use of Public Key Infrastructure (PKI) to create an ecosystem of appropriately managed private keys and certificates containing the corresponding public keys which can be cryptographically chained in the classic way to a trust anchor known to validators. Validators may contain one of more trust lists of trust anchors depending on which provenance ecosystem they participate in. Validators will have been provisioned with trust lists out of band with respect to online operations. Currently C2PA specifies how to use X.509 certificates to achieve this aim.

C2PA does not mandate or even suggest particular trust lists or public key infrastructure (PKI), and instead takes them and other related configuration as inputs. This is because each application built using the C2PA standard that operates within its own ecosystem will have unique requirements and relationships amongst the participants of that ecosystem. Application implementers may be tempted to take existing trust lists used in other applications, such as those used for validating secure web sites or signed documents and adopt such lists without due consideration. To this end C2PA is developing guidance on the establishing and operating of an application's trust model and identity ecosystem.

Considerations of harms, misuse and abuse

The C2PA has considered harms modelling and analyzing how a socio-technical system might negatively impact users, stakeholders, broader society, or otherwise create or re-enforce structures of injustice, threats to human rights, or disproportionate risks to vulnerable groups globally. This work has been carried out to understand and offer guidance on the creation and use of C2PA in such a way as to prevent accidental harm through oversight or misuse. This work is described in the C2PA Harms Modelling guidance document [12].

Adoption into News Workflows

The interdependence of news publishers requires a broad alignment of approaches at key points in newsroom workflows to allow easy interoperability of media files.

The framework for doing this includes.

1) Digitally signing and securing output when it is published. This can first be done at the organizational level, and ultimately at the show or reporter level.
2) Validate media on ingest. This will aid in sorting valid files from AI generated noise. The value of this function will gradually increase as provenance methods gain wide acceptance and the amount of signed content grows. Eventually, the absence of provenance information will be a signal that the content requires greater skepticism.
3) Adding reputational assertions. The optional ability to embed secured endorsements can assist in adding credibility to media exposed outside of its usual markets.

All of the activities will facilitate mechanized validation of content at scale.

It is also important to note that the inclusion of C2PA based secure meta-data will not break existing workflows, and the absence of this data will not, in itself, be a sign that the information is false. These

two attributes will allow for a gradual introduction of the technology into the established ecosystem. After a period of transition, when C2PA signals are widely adopted, the absence of secure metadata will become a suspicious signal, much the same way the absence of a HTTPS lock on a purported corporate website is today.

The BBC learnings

The BBC is one of the Origin partners who have participated in the technical working groups of C2PA with a focus on media distribution from a broadcaster's perspective. This has involved testing the specification by coding basic manifest generators and coordinating with other members of the C2PA to ensure interoperable content files can be independently created and verified, thereby strengthening the specification and helping bring the concept to life within the organization.

This early work is being further reinforced through work on a more robust and complete local demonstrator, built within our R&D facility. This will help experiment with user experience as we develop further ideas around the presentation and use of this technology and build upon the original work done by our colleagues and partners within the BBC and C2PA members.

Like all new technologies within a fast-moving news technical landscape, there is now a process of education, evangelizing and assessing how the technology can be integrated into the newsroom workflows. This process takes on a number of forms, from looking at the top level strategic threat from synthetic content and how this might be countered, through to the exact nature of the current systems and where we might add C2PA functionality in a robust way to our pipelines to ensure secure use. In our experience this touches not only the engineering capabilities, but also journalistic practices and corporate policy and all these areas are now being developed.

Work is now planned on a formal technical analysis of systems integration whilst at the same time exploring the technology from the point of view of a journalist tool, including understanding how using these signals will affect content ingestion and be used with third party contributions. Further questions are also being considered around how provenance data will engage the audience - not least by making it clearer how journalists select and process the material they use. This is a new and exciting area for news practitioners to explore.

CBC/Radio-Canada learnings

CBC/Radio-Canada is a founding member of Project Origin and an early advocate for the use of the C2PA specification within its news infrastructure. The initial proof of concept used C2PA to securely sign the output emitting from the central distribution hub. This is the point of highest value creation and is compatible with the currently available C2PA tools. This simple intervention allows the signing of the corporation's content, helping to maintain the credibility and reputation of our work. This trial also allowed us to confirm that the addition of C2PA manifests did not disrupt the downstream flow of media.

Starting early 2022, along with Microsoft and Ravnur, the corporation has participated and completed multiple proof of concepts that allowed us to gain much needed experience with the specification along with determining a target architecture on how to ideally deploy the specification to position the corporation to sign our output.

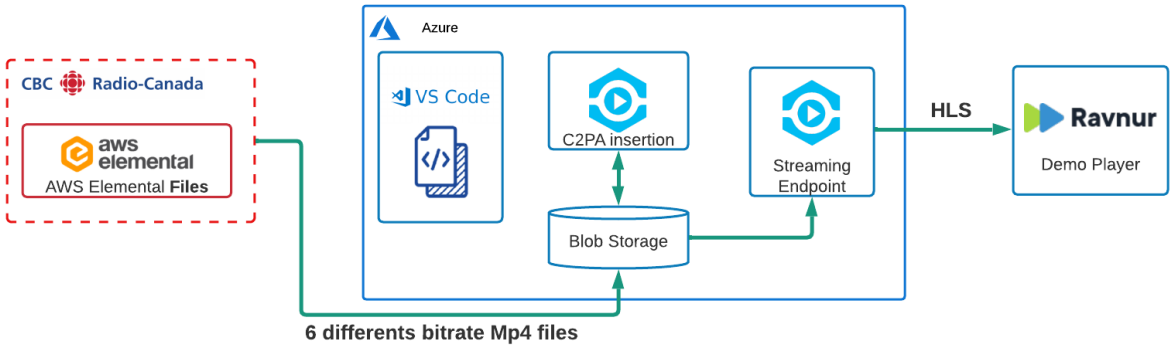

FIGURE 5. SCHEMATIC OF THE CBC/MICROSOFT/RAVNUR PROOF OF CONCEPT OUTPUT SIGNING

Our initial proof of concept was to sign a piece of content and play it back while displaying it, along with provenance in a video player. These trials were quickly successful.

The implementation of these concepts in our current production environment is challenged by the current lack of universal adoption of the C2PA standard. While it is possible to implement the standard now, it would involve adjusting current workflows to send content to service providers who are leading in C2PA adoption. It is hoped that broader adoption of the standard by current suppliers will allow us to avoid this disruption.

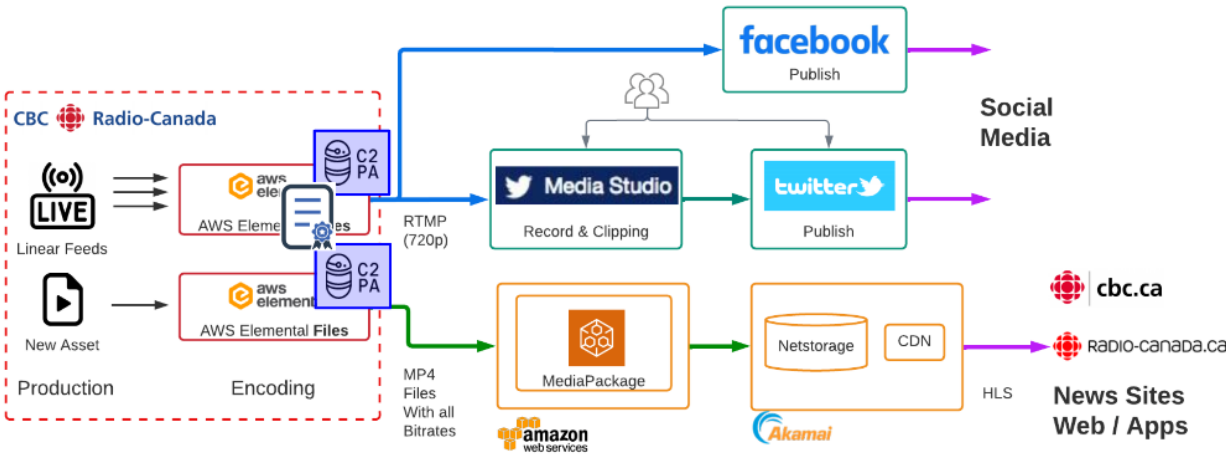

FIGURE 6. IDEAL ARCHITECTURE FOR OUTPUT SIGNING

In a separate trial, CBC/Radio-Canada is participating in an effort to validate photos coming into its news infrastructure. This assists in confirming the veracity of the content coming into the news systems, and will determine how and where the C2PA manifests are accessible in the current workflows.

By validating inputs and signing outputs the corporation is beginning the process of building an end-to-end provenance-based production chain. This is being done to establish the technical experience with provenance tools, and to socialize the use of provenance data within the internal operations of the organization.

This work is also being done to be shared with the wider news and technical communities. Provenance requires broad industry adoption. The corporation will continue to share its learnings with other publishers. In order to signal the importance of media provenance to potential vendors, the corporation has added criteria to its RFP technical compliance grid to emphasize the value of supporting the C2PA specification.

Growing the community

Media Provenance is a viable way to protect the news ecosystem from the threat of powerful synthetic media-based disinformation attacks. A large community of media and technology companies and subject matter experts have come together to define a common approach to the problem. The key to making it work is community adoption of common provenance practices.

Widespread adoption of signed and secured news media will make authentic content stand out in a sea of misinformation and synthetically generated noise.

As a broadcaster you can help by:

1) Adding C2PA secure digital manifests to your output at your point of origin.
2) Ask your vendors what their roadmap is to add C2PA functionality to their products.
3) Add a provenance validation step to your newsroom ingest function.
4) Join the growing community of practice to inform others of the need for updated media provenance processes.

Adding C2PA secure media manifests to our outputs is a necessary step to protect the integrity of the global news ecosystem. The Project Origin team is working to develop best practices for the common use of this technology in newsrooms. We look forward to working with you.

For companies that build media technologies, the Coalition for Content Provenance and Authenticity is a diverse and dedicated cross industry effort. Your participation in furthering the development of the specification or implementing this open standard into your product roadmap is welcome and encouraged.

However you chose to participate, the open secure media provenance community welcomes you to join our efforts.

Acknowledgements

The authors would like to acknowledge the work of the C2PA in creating the technical specification described in this paper. In particular we would like to acknowledge the ongoing work of the C2PA Technical Working Group chaired by Leonard Rosenthol of Adobe and thank them for their permission to present this paper. The authors thank Laura Ellis of the BBC and Kevin Kane of Microsoft for their expert review of this paper and would like to acknowledge the support of the Project Origin members in the production of this paper.

References

[1] *C2PA Specifications*. https://c2pa.org/specifications/specifications/1.2/index.html

[2] *ISO/IEC 14496-12:2020 Information technology — Coding of audio-visual objects — Part 12: ISO base media file format* https://www.iso.org/standard/74428.html

[3] *Exif. Exchangeable image file format for digital still cameras: Exif Version 2.32* https://www.cipa.jp/std/documents/download_e.html?DC-008-Translation-2019-E

[4] *IPTC Metadata standard*. http://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata

[5] Schema.org vocabularies. https://schema.org/

[6] *ISO/IEC 19566-5:2019 Information technologies — JPEG systems — Part 5: JPEG universal metadata box format (JUMBF)* https://www.iso.org/standard/73604.html

[7] Birkholz H, Vigano C, Bormann C. *RFC 8610. Concise Data Definition Language (CDDL): A Notational Convention to Express Concise Binary Object Representation (CBOR) and JSON Data Structures*. June 2019. https://datatracker.ietf.org/doc/html/rfc8610

[8] Bormann C, Hoffman P*. RFC 8949 Concise Binary Object Representation (CBOR).* December 2020. https://www.rfc-editor.org/rfc/rfc8949

[9] Shaad J. *RFC 8152 CBOR Object Signing and Encryption (COSE).* July 2017. https://www.rfc-editor.org/rfc/rfc8152.txt.pdf

[10] Cooper D, Santesson S, Farrell S, Boeyen S, Housley R, Polk W. *RFC 5280. Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile*. May 2008. https://www.rfc-editor.org/rfc/rfc5280

[11] *C2PA User Experience Guidance for Implementers*. https://c2pa.org/specifications/specifications/1.0/ux/UX_Recommendations.html

[12] *C2PA Harms Modelling.* https://c2pa.org/specifications/specifications/1.0/security/Harms_Modelling.html