|  |  |
| --- | --- |
| **ITU – Telecommunications Standardization Sector**  STUDY GROUP 21 Question 6  **Video Coding Experts Group (VCEG)**  77th Meeting: 26 June – 4 July 2025, Daejeon, KR | Document VCEG-BY19-v1 |

|  |  |  |  |
| --- | --- | --- | --- |
| Question: | 6/21 (VCEG) | | |
| Source: | Iraj Sodagar Dolby Labs USA | Email: | [irajs@live.com](mailto:irajs@live.com) |
| Title: | **Proposed technical paper on media authentication** | | |
| Purpose: | Proposal | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
| **Abstract:** | This contribution proposes a Technical Report on media authentication in Q6. |

# Introduction

The SG21 recently started an ad hoc group on media authentication. The AHGMMA had two meetings and came to the conclusion that developing a technical report on media authentication would be beneficial before starting any normative work, and suggested such an effort could be conducted as part of Q6. In this contribution, we propose developing such a TR.

# Media authentication in general

As discussed in AHGMMA-DOC1, the media authentication can be achieved by two classes of technologies:

1. Proactive transmission of metadata information with the content, which enables authentication of the content by a consumer. This can be achieved by two means:
   1. Explicit inclusion of the metadata, i.e., easily identifiable, accompanying the content. Examples of such metadata include provenance information by C2PA and JPEG Trust, and Digitally Signed Content SEI messages by JVET. In any of those cases, the metadata provides adequate information for a consumer or a receiver of the content to verify the content through a certificate mechanism issued by the content producer.
   2. Implicit inclusion of the metadata, i.e., including the metadata with the content in a perceptually undetectable way such that a tool can detect and extract the information. Digital watermarking techniques are one class of such methods.
2. Detection-based authentication, in which the content is analysed by the consumer for authenticity and checks whether the content is authentic or not. The two main categories are:
   1. Fingerprinting the content and checking whether the content was previously generated by searching for a matching fingerprint
   2. Content-based detection: detecting whether the content is generated or altered by analysing the characteristics of the content and finding either a mismatch between those characteristics and naturally captured content or finding patterns that AI-generated/altered tools generate.

Figure 1 captures the above four techniques in media authentication.

Find trusted owner from metadata

Processing metadata

Trust

content

Find provenance

Trusted content and info

Trust profiles

Level of trust

*Metadata?*

CA

content

Metadata

Manifests

Find Watermark

Get metadata from cloud

*Watermark?*

Find fingerprint

Find originals

Find metadata

*Fingerprinting?*

FP registry

FP catalogue

AI detection

Mark Generated

Create fingerprint

*AI detection?*

AI detection resources

Altered vs Generated?

*Passive*

*defence*

As is shown in the figure, in general, a public or professional consumer can use these techniques to find out about the authenticity of the content consumes or is about to consume:

1. Workstream 1: Whether there is metadata included in the content, and whether that metadata can authenticate the content? If not, then
2. Workstream 2: Whether there are watermarks in the content, and using the information carried in the watermark, can the content be authenticated? If not, then
3. Workstream 3: Find the fingerprint of the content and check if there is any matching fingerprint in one or more catalogues. If not, then
4. Workstream 4: Use AI detection tools to see if the content is generated or altered. If so, optionally generate a fingerprint and submit it to the fingerprint catalogue.

# Standardisation scope

The goal of standardisation in general is to achieve interoperability to enable different entities in an ecosystem to exchange data with each other and achieve the goals with a reasonable cost and reduced complexity.

In the context of media authentication, we note that:

1. The carriage of metadata with content is in the scope of the standards since interoperability between producers and consumers is needed for delivering such data. While C2PA, JPEG Trust and MPEG are working on the carriage of metadata, there might be gaps in those standardisation efforts. However, the best way to fill the gaps and provide adequate standardisation for workstream 1 is to work with those organisations to complete the solutions.
2. There has been an ongoing debate about whether workstream 2 needs standardisation or not. Standardising a watermarking technique certainly reduces the complexity of the devices the consumer uses to detect the watermark. However, a standard watermark may be more perceptible to tampering. Regardless of those debates, there are benefits to standardising metadata a watermark may carry and the interfaces between the device detecting the watermark and the services that store the corresponding metadata in the cloud. C2PA has taken steps toward this goal. But we believe this may need to be standardised in a bigger context.
3. For workstream 3, to our knowledge, there is no standard that specifies a standard for media fingerprinting. ISO/IEC 24138 (International Standard Content Code (ISCC)) enables decentralised digital content identification and allows the use of any fingerprinting to generate identifiers. MPEG 7 (ISO/IEC 15938) provides a framework that can be used to express multimedia content fingerprints. SMPTE ST 2064 defines a method for generating audio/video fingerprints for the purpose of measuring audio/video synchronisation. Outside of standards, several proprietary algorithms are used by industry for audio, image and video fingerprinting.
4. For workstream 4, as indicated in AHG-MMA-DoC1, there is an ongoing arms race between sophisticated AI generation/alternation tools and detection counterparts. It is expected that tools on both sides become more sophisticated, and therefore, a standard for AI detection seems impractical at this point. However, there might be benefits of a standard output format for AI detection tools, i.e. capturing the result of a detection algorithm in a standard way so it can be used by other AI detection tools. This data, along with the fingerprinting, can be used to mark the content that is suspicious to be generated or altered by AI, to be used for further analysis by other parties.

# Proposal for the new TR on media authentication

Following Clause 3 arguments, we proposed before starting any normative activity, the Q6 produces a technical report covering a general end-to-end framework for media authentication, including prescribing various collaborative scenarios for media authentication covering production, repurposing and consumption of the media, prescribing the best practices that cover the use of various media authentication workstreams using standards as well as nonstandard technologies.