ISO/IEC JTC1/SC29/WG1
(ITU-T SG21)

Coding of Still Pictures

JBIG
Joint Bi-level Image Experts Group

JPEG
Joint Photographic Experts Group

TITLE:	JPEG Trust Part 3: Media Asset Watermarking Use Cases and Requirements – v1.0
SOURCE:	WG1
EDITORS:	*Deepayan Bhowmik (deepayan.bhowmik@ncl.ac.uk)* *Frederik Temmermans (frederik.temmermans@vub.be)*
PROJECT:	ISO/IEC AWI 21617-3 (JPEG Trust Part 3: Media Asset Watermarking)
STATUS:	Final
REQUESTED ACTION:	Public distribution

Contact:
ISO/IEC JTC 1/SC 29/WG 1 Convener – Prof. Touradj Ebrahimi
EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 2606, Fax: +41 21 693 7600, E-mail: Touradj.Ebrahimi@epfl.ch

Contents

1 Summary 1
2 Introduction 1
3 Scope 3
4 Definitions 3
5 Use Cases 3
5.1. AI-Generated Content 3
5.2 Trust record recovery 4
5.3 Copyright Assertion 4
5.4 Digital Rights Management (DRM) 5
5.5 Provable Anteriority (timestamp) 5
5.6 Authentication and Verification 5
5.7 Content Tracking and Monitoring 5
5.8 Forensic Watermarking 6
5.9 Brand Protection 6
5.11 Governance (Policymaker) 6
6 Requirements 7
6.1 Watermark, embedding and detection format 7
6.2 Watermark embedding performance 7
6.3 Watermarking robustness performance 7
6.4 Integration with Trust Profiles 7
6.5 Embedding, referencing and asset registration 8

1 Summary

Digital watermarking is considered one of the core elements that provide a mechanism for media asset provenance, integrity, and copyright protection by embedding identifiers into files. An effective system must be robust against attacks (like compression or cropping) and maintain imperceptibility to preserve quality. As part of global standardisation efforts, JPEG Trust Part 3: Media Asset Watermarking (ISO/IEC 21617-3) aims to define frameworks and protocols to bind trust metadata to media asset content. These capabilities support critical use cases from labelling AI-generated media assets to Digital Rights Management (DRM) and source tracing. This document provides descriptions of various use cases and identifies a list of requirements for standardisation.

2 Introduction

Current technologies, especially with the rise of generative AI, have made the modification or synthetic creation of media assets easy for general users. While such developments offer new opportunities, particularly in creative industries, the rapid proliferation of digital media asset content, coupled with the negligible cost of copying and redistribution, has significantly increased the risk of piracy, unauthorised replication and spread of misinformation.

In many application domains, creators may want or need to declare the type of modifications that were performed on the media asset, in contrast to other situations where the intention is to conceal the existence of manipulations. Media modifications are not always negative, as they are increasingly a normal and legal component of the production pipeline. Therefore, governmental organisations may consider new legislation, including the right to opt out of text and data mining, as well as the mandatory declaration of synthetic media and their manipulations. Companies, such as social media platforms and news outlets, may develop mechanisms that clearly detect and annotate manipulated media when it is shared, in support of establishing trustworthiness in media assets.

To facilitate globally interoperable media asset authenticity, integrity and provenance, JPEG (ISO/IEC JTC 1/SC 29/WG 1) developed and published an international standard: **JPEG Trust, Part 1: Core Foundation (ISO/IEC 21617-1:2025)**. The JPEG Trust Part 1 standard defines a media asset as comprising its content, metadata, and a tamper-evident Trust Record. This Trust Record includes one or more Trust Manifests, which contain assertions about the asset's origin, device details, attribution, rights, edits, and related information. From this, Trust Indicators are derived and compiled into a Trust Indicator Set, which is evaluated using a Trust Profile to produce a Trust Report that assesses the asset's trustworthiness.

Digital watermarking, in use for several decades, has been increasingly adopted as a method for embedding information directly into media assets in a way that can be both imperceptible and robust. This technique, as illustrated in Figure 1, establishes a link between the metadata and the content, one that is challenging to disrupt without compromising the intended usage of the media asset itself. Since the inception and rapid rise of generative AI, watermarking has increasingly gained popularity, both within the industry and among policymakers, as a solution to signal whether the media asset is AI-generated or AI-manipulated content. Such watermarking is equally beneficial for media assets generated outside the context of AI (e.g., photographs, edited images). Furthermore, watermarking is also helpful to link other associated content provenance information, such as intellectual property rights, authorship, and ownership.

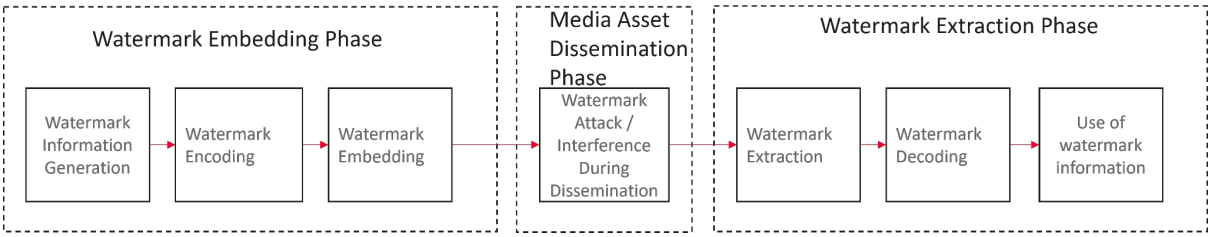

Figure 1: A generic watermarking embedding and extraction workflow in the context of the media distribution/asset dissemination phase.

While growing efforts are noticeable in developing technologies, there is a need to have a standardised way to watermark media assets and link them with other metadata-based information. Therefore, the JPEG standardisation committee (under the auspices of ISO and IEC) has launched an initiative to identify the standardisation needs for facilitating the watermarking of media assets through an in-depth analysis of various usage scenarios.

To facilitate globally interoperable media asset authenticity, JPEG (ISO/IEC JTC 1/SC 29/WG 1), via JPEG Trust, initiated the development of a new international standard: JPEG Trust Part 3: Media Asset Watermarking, with an aim to strengthen the overall media Trust ecosystem which also complement the development of annotation based approach in JPEG Trust Part 1:

Core foundation. This initiative arises from an exploration that began a year ago, focusing on the requirements for addressing aspects of provenance, authenticity, integrity, copyright, and identification of assets and stakeholders.

A media asset watermarking framework aims to facilitate the embedding of an invisible watermark into digital media files and the extraction of this information to ensure media asset provenance, integrity, and copyright protection. Main aspects of such a framework include:

- **Ownership and Authenticity:** Watermarking helps digital assets by embedding tamper-evident content provenance that can be used to establish ownership and authenticity. This is essential for ensuring that the content remains intact.
- **Leak Detection and Provenance Tracing:** Forensic watermarking can identify unauthorised leaks and trace them back to the source. This is particularly useful for sensitive or high-value content, as it helps mitigate risks such as revenue loss, reputational damage, or the protection of rightful ownership.
- **Robustness and Quality:** A robust watermarking framework ensures that the watermark remains intact against various transformations, such as compression, cropping, and scaling, while maintaining the intended use of the original media asset.
- **Compatibility:** A framework aims to ensure compatibility with existing solutions for watermarking, and, where applicable, intellectual property rights management, content provenance, and media asset workflows, allowing seamless integration and efficient asset management, providing the interfaces between the device detecting the watermark and the services that store the corresponding trust record in the cloud.

3 Scope

Part 3 “Media Asset Watermarking” will define the use of watermarking as one of the components of the JPEG Trust framework to support tools and mechanisms for content authenticity, provenance, integrity, labelling, and binding between JPEG Trust metadata and corresponding media assets. Part 3 will focus on various watermarking types (including explicit/visible and implicit/invisible watermarking), watermarking usage scenarios, and watermarking assessment criteria.

4 Definitions

To ensure a correct understanding of the descriptions in this document, this section refers to terms and definitions listed in ISO/IEC JTC 1/SC29/WG1 N101310, REQ “Terms and Definitions for Media Asset Watermarking” and ISO/IEC JTC 1/SC29/WG1 N100972, REQ "Terms and Definitions for JPEG Trust v2.0” and to concepts as they are used in the context of this work.

5 Use Cases

One of the key objectives of JPEG Trust is to better understand topics and use cases that fall under its scope and to analyse their implications, especially from a standardisation point of view. Currently, the JPEG committee has identified the following topics and use cases:

5.1. AI-Generated Content

Watermarking of AI-generated content (AIGC) is a technique used to embed digital marks or indicators into content created by AI. This practice helps identify the origin and authenticity of the content, making it easier to distinguish between increasingly realistic AI-generated content and other media assets. The primary goal is to label the AI-generated content. This is

particularly important in combating misinformation, copyright infringement, and unauthorised use of AI models.

Examples:

- AIGC creation models could include a common model-specific watermark in every image they generate, thereby conforming to the provenance and source model.
- An AIGC system could embed a watermark to any media asset after its creation.
- An external repository can be linked through watermarking, providing rich metadata that includes copyright, ownership, authorship, and additional information.

5.2 Trust record recovery

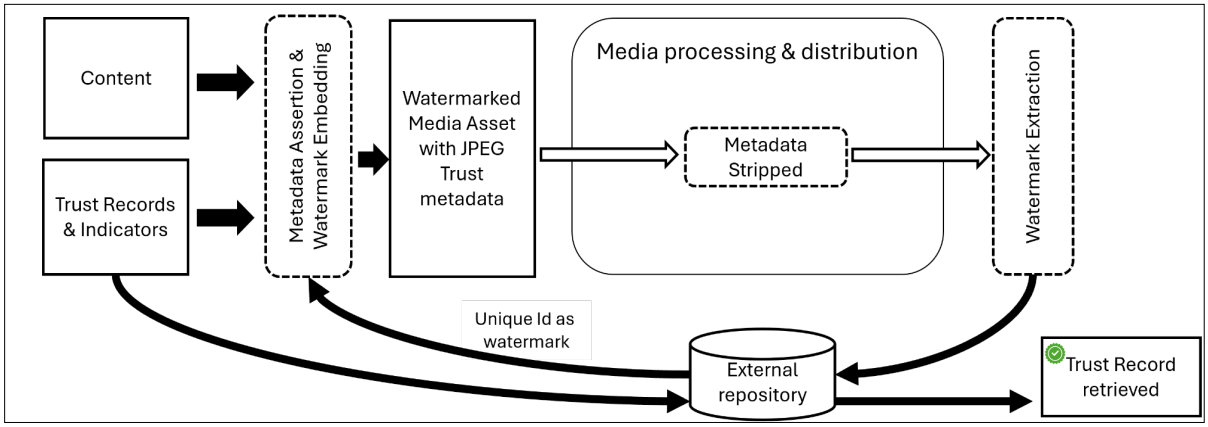

Figure 2: Incorporating media asset watermarking into JPEG Trust ecosystem.

This use case demonstrates the importance of watermarking-based techniques for retrieving the Trust record when metadata is lost. For example, JPEG Trust Part 1 describes how trust records with metadata assertions can be embedded within media assets. However, metadata can be removed during media preprocessing as part of the media distribution ecosystem (e.g., content-sharing platforms) or intentionally for misuse. Such alterations impact the extraction of the Trust Record and the subsequent evaluations by which the trustworthiness of the media asset is based. As a mitigation, a watermarking-based solution can bind the Trust Record with the media asset content in a persistent way, as illustrated in Figure 2.

In this scenario, the Trust Record is embedded within the media asset following JPEG Trust Part 1 and also kept in an external repository. A unique ID for the Trust Record is then used as part of the watermark, which is then embedded into the media asset. Therefore, in the event of complete or partial metadata loss, the watermark can be recovered and queried in the same external repository to retrieve the Trust Record, thereby reestablishing trustworthiness.

5.3 Copyright Assertion

Watermarking is widely used to assert intellectual property rights. By embedding a watermark pointing to rights assertions, creators can assert their authorship, and rights holders can assert their ownership. This is particularly important for photographers, artists, and content creators who share their work online.

Examples:

- Photographers can embed their name or logo into their images in an invisible manner to ensure proper attribution.

- Graphic designers can watermark their digital artwork to identify their creative designs, which can help alleviate unauthorised copying.

5.4 Digital Rights Management (DRM)

Watermarking is an integral part of DRM systems, which protect digital content from unauthorised access and distribution. Watermarks can also enforce usage policies and track the usage of content.

Examples:

- Images/streaming media can include watermarks that restrict copying and sharing, ensuring that only authorised users can access the content.
- Streaming services can embed watermarks in media files to prevent unauthorised distribution and track usage.

5.5 Provable Anteriority (timestamp)

Watermark can be used as technical evidence to support legal validation of ownership/authorship/copyright claims. However, the legal value of a watermark depends on its provable anteriority (timestamp), provable identity of the declarer, and provable identification of the object. Within the JPEG Trust framework, use of Part 1: Core Foundation and the proposed Part 3: Media Watermarking can provide a means to support such a use case, where the watermarking process would be documented, in the Trust Record, as a soft-binding *assertion.* Additionally, information related to the timestamp and identity of the declarer is also recorded alongside the watermark itself. This use case, "Provable", could be a service provided by a Qualified Trust Service Provider.

5.6 Authentication and Verification

Watermarks can be used to verify the authenticity of a media asset. This is crucial in industries where the integrity of information is paramount, such as legal, financial, and academic sectors.

Examples:

- Legal documents with visual content (e.g., passports) can include a digital watermark that verifies their authenticity and enables the detection of tampering.
- Academic certificates often feature watermarks to verify their authenticity and help alleviate forgery.

5.7 Content Tracking and Monitoring

Watermarking allows content owners to track the distribution and usage of their digital assets. This is useful for monitoring how content is shared and ensuring it is used in accordance with licensing agreements.

Examples:

- Media companies can embed unique watermarks in their videos to track where and how the content is being distributed online.
- Stock photo agencies can use watermarks to monitor the usage of their images and ensure compliance with licensing terms.

5.8 Forensic Watermarking

Forensic watermarking involves embedding unique identifiers into digital content to trace the source of unauthorised distribution. This is particularly useful in combating piracy and identifying leaks through media provenance.

Examples:

- Movie studios and media distribution companies can embed forensic watermarks in pre-release copies of films to trace the source of any leaks.

5.9 Brand Protection

Companies use watermarking to protect their brand identity and prevent counterfeiting. By embedding watermarks into product images and promotional materials, brands can ensure the authenticity of their products and maintain their reputation.

Examples:

- Fashion brands can watermark product images to prevent counterfeiters from using their images to sell fake products.
- Luxury goods manufacturers can embed watermarks in product photos to verify the authenticity of their online campaigns.

5.10 Medical Imaging Integrity and Privacy

Medical images like X-rays and MRIs are vital for diagnosis but vulnerable to tampering or misuse. Digital watermarking embeds tamper-evident information into images to provide authenticity. Through the use of encrypted identifiers, it would also be possible to protect patient privacy. This approach strengthens trust in healthcare systems and supports privacy-preserving sharing of medical data.

Examples:

- Inserting encrypted patient identifiers within ultrasound images to enable reliable exchange across hospital systems without exposing sensitive data.
- Watermarking medical images with the content provenance of access and sharing, helping hospitals comply with privacy regulations while preserving image quality.
- Embedding consent declarations invisibly into medical scans so that researchers can confirm usage rights without risk of patient re-identification.

5.11 Governance (Policymaker)

It is not within the purview of policymakers to actually implement watermarks, but rather to describe how watermarks could be implemented by the organisations they govern. The chief goals of watermarking sought by policymakers are to provide transparency and accountability with respect to media assets. Presently, this is most notably focused on AIGC.

For example, in August 2023, China[1] has introduced regulations requiring all AIGC to be clearly marked. To improve transparency and mitigate misuse of AIGC, it proposed that generative AI providers must include: A) Visible labels with explicit watermarks stating the content is AI-generated; B) Invisible/implicit watermarks in media like images, videos, and audio, containing

[1] https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/757583/EPRS_BRI(2023)757583_EN.pdf

provider information, and C) Metadata tags embedded in files to identify them as AI-generated.

6 Requirements

Based on the identified use cases, several JPEG Trust Part 3: Media Asset Watermarking requirements have been identified and organised into the following five main categories:

- Watermark, embedding and detection format
- Watermark embedding performance
- Watermarking robustness performance
- Integration with Trust Profiles and report generation
- Embedding, referencing and asset registration

The sections below list the already identified requirements for each identified category.

6.1 Watermark, embedding and detection format

- R1.1 The standard shall provide means to signal the presence of a watermark.
- R1.2 The standard shall provide information regarding the technique/algorithm used to embed and extract the watermark.
- R1.3 The standard shall provide means to reference an external repository or repositories.

6.2 Watermark embedding performance

- R2.1 The standard shall provide means to evaluate watermark embedding distortion performance for a given media asset and report the evaluation outcome. For example, the user may choose from a suite of objective or subjective quality metrics, such as PSNR, SSIM, or ITU-recommended Subjective Tests.

6.3 Watermarking robustness performance

- R3.1 The standard shall provide means to evaluate watermarking robustness performance against one or more watermarking attacks, such as filtering, compression, cropping or modification and report the outcome. For example:
    - **Signal processing**: JPEG compression, various noises, and image enhancement.
    - **Geometric attacks:** rotation, resize, cropping, scaling, etc.
    - **Removal:** Various filtering, GAN or diffusion-based content modification,
    - **AI-based:** Manipulation or enhancement using deep learning or generative AI, JPEG AI compression etc.

6.4 Integration with Trust Profiles

- R4.1 The standard shall provide mechanisms to create and evaluate Trust Profiles for specific use cases to conform to the target embedding and robustness performance.
- R4.2 The standard shall be compliant with JPEG Trust Part 2: Trust Profiles and Reports.

6.5 Embedding, referencing and asset registration

- R5.1 The standard shall provide means to embed references to externally hosted repositories and services.
- R5.2 The standard shall provide means to register media assets, unique identifiers/watermarks.