

Do AI-generated images on Wikimedia Commons carry provenance marks?

EasyxLab · EasyByte Hub S. Coop. Mad. · Study S6

 Working draft Published 2026-10-03

Status: working draft, not peer-reviewed

EasyxLab, study S6 · a research lab by EasyByte · October 2026

Cite as: EasyxLab (2026). Do AI-generated images on Wikimedia Commons carry provenance marks? Study S6. EasyByte Hub S. Coop. Mad. <https://github.com/easybytehub/easyxlab>

Abstract

Wikimedia Commons stores uploaded files byte for byte. It therefore shows which AI provenance marks reach a public archive when the host does not strip them. We enumerated every file Commons labels as AI-generated: the `Category:AI-generated images` tree (530 categories) and `{{PD-algorithm}}`, 8,971 files in total. We then inspected 2,084 originals from a month-stratified sample with ai-mark-lint, downloading, checking and deleting each one.

Two populations are reported. The *main* population is the category tree plus `{{PD-algorithm}}` files with AI evidence in their categories (1,585 measured files); some of these images are AI-modified rather than AI-generated. The *union* adds the other `{{PD-algorithm}}` files (2,084 measured).

Marks in the main population. The share of files with a machine-readable AI mark rose sharply in **June 2026**: from 25.8% of January–May uploads (n = 532) to 50.2% of June–July uploads (n = 313). That is +24.4 percentage points; a bootstrap that resamples upload days gives +14.7 to +33.6. August–September (n = 461) shows 43.4%, or 53.3% without one large unmarked batch of coats of arms.

Before and after 2 August. The comparison of January–July with August–September, when Art. 50(2) of the EU AI Act became applicable, gives +8.4 pp (day-cluster bootstrap +0.2 to +17). It depends on that batch: without it, the difference is +17.8 pp. Neither comparison can be attributed to the AI Act.

Kinds of mark. Almost every mark is a C2PA manifest, mostly from OpenAI and Google. The IPTC `DigitalSourceType` alone marks 1.5% of main files. A readable Chinese AIGC label appears on 1 file of the union, and 3 more files carry a double-encoded label.

Validation. We found a defect in our own validator: ai-mark-lint 0.1.0 rejected 302 of the 625 readable manifests (union), 187 of them because no extended-key-usage list was configured. The defect is fixed in 0.1.1. With the corrected rule, 117 manifests (18.7%) fail, for these reasons:

- certificates that have expired in manifests without a time-stamp (85);
- content changed after signing (32);
- C2PA first-action rule violations (26);
- Microsoft Paint manifests (12).

With the official C2PA Trust List, 452 of 624 manifests (72%) are valid and trusted.

1. Why this matters now

Art. 50(2) of the EU AI Act, applicable since 2 August 2026, requires providers of generative systems to mark outputs "in a machine-readable format and detectable as artificially generated or manipulated". California's AI Transparency Act became operative the same day. In practice the marks are:

- a C2PA manifest declaring an IPTC `digitalSourceType` such as `trainedAlgorithmicMedia` ;
- the same IPTC term in XMP;
- China's `TC260:AIGC` label (GB 45438-2025).

Marks reach the public only if generators write them and publishers keep them. Our study S3 found that no common re-encoding pipeline kept embedded C2PA. X strips C2PA on upload (Zewde et al., arXiv:2604.25370). Commons keeps the original: "Content credentials should not be stripped however they will only be in the original file, not any thumbnails" (Bawolff, Commons Village pump, 30 December 2025). For all 2,084 measured files, the SHA-1 of the download matched the SHA-1 the API reports for the stored original.

2. Prior work

Search. We ran it on 2026-10-02 and again on 2026-10-03 after review, with `scripts/prior_work.py` (`--extra` for the second pass):

- **arXiv API.** First pass: `all:C2PA` (16 results), `all:"content credentials"` (4), `all:"Wikimedia Commons"` AND `all:"AI-generated"` (0), `abs:provenance` AND `abs:"AI-generated images"` AND `abs:metadata` (0). Second pass: `abs:"AI Act"` AND `abs:watermarking` (8), `abs:"machine-readable"` AND `abs:"AI-generated"` (7), `abs:"AI Act"` AND `abs:labelling` (12), `abs:watermarking` AND `abs:adoption` AND `abs:generators` (62).
- **Crossref**, two queries.
- **Semantic Scholar Graph API**, three queries. Several attempts returned HTTP 429. The successful ones returned 1 result for "C2PA content credentials in the wild" and 48 for "AI-generated images Wikimedia Commons".
- **GitHub** via `gh search repos` , four queries, 0 results each.
- **Commons search API** over the project, help and template namespaces: `C2PA` (101), `"Content Credentials"` (5), `"DigitalSourceType"` (9), `"trainedAlgorithmicMedia"` (5). We also read the wikitext of `Commons:AI-generated media` , `Commons:Signs of AI media` and two Village pump archives.
- **Phabricator** Conduit refused anonymous search ("Session key is not present").
- **Specifications and code:** C2PA 2.2, the C2PA conformance trust list, the IPTC vocabulary, and the `c2pa-rs v0.91.0` source.

The first-pass queries missed the closest empirical work, because its abstract does not say "C2PA". The second pass found it.

Closest results.

- **Generator side.** Rijsbosch, van Dijck and Kollnig (arXiv:2503.18156, March 2025) audited image generators against the AI Act: "we find that only a minority number of AI image generators currently implement adequate watermarking (38%) and deep fake labelling (18%) practices". They measured what generators emit; we measure what reaches a public archive.
- **Platform side.** Rijsbosch et al. (arXiv:2609.38571, September 2026) audited labelling on four platforms: "In controlled uploads of AI-generated content carrying standard AI provenance signals, platforms labelled only 61% of uploads, and commonly strip those signals after uploading." Zewde et al. (arXiv:2604.25370) report that X strips C2PA "systematically".
- **Commons.** On Village pump/Technical in September 2025, Bawolff wrote "I believe chatgpt uses C2PA metadata which we currently don't support" (19 September). He sampled "about 10 recent ones at random" from `Category:PD-algorithm`, and "only [one] had it" (20 September); Omphalographer listed more files with C2PA in reply. Bawolff also observed what we call class F: "a lot of the chatgpt files give invalid when i use the web app or c2patool (error: first action must be created or opened), but work fine with exiftool" (03:26, 20 September 2025). On 1 October 2025 he announced that MediaWiki would display XMP `Iptc4xmpExt:DigitalSourceType`, adding "Unfortunately i don't think any AI image generators actually add that".
- **Commons guidance.** `Commons:Signs of AI media` advises to "Scan for C2PA information". Deletion requests cite C2PA and IPTC values as evidence (for example `File:Physicien du 19ème siècle.png`, August 2026). Ckoerner opened Phabricator task T387075 on C2PA support on 22 February 2025.
- **Specification analyses.** arXiv:2604.24890 (security of C2PA) and arXiv:2603.02378 (provenance-watermark "Integrity Clash") analyse the specification, not files in the wild.

What this study adds. Generator audits and platform audits exist; so do anecdotal Commons checks, including an earlier sighting of the first-action failure. What we did not find is:

1. population-level rates, with intervals that allow for batch uploads, of each kind of mark on a public archive that keeps originals, month by month across the date the AI Act became applicable;
2. validation of every manifest under several trust and ECU configurations and three validator builds, with a cause for every failure;
3. evidence that an ECU-less configuration of `c2pa-python`, the one our own tool used, rejects about six in ten otherwise-acceptable manifests.

We do not claim to be the first to see C2PA, or its failures, on Commons.

3. Method (summary; METHOD.md has the details)

Population. Read from the Commons Action API on 2026-10-02. It has two parts:

- the `Category:AI-generated images` tree, breadth-first to depth 6. That is 530 categories and 6,014 files. The code does not list subcategories at depth 6; the reviewer checked the five depth-6 categories and none has subcategories.
- files transcluding `{{PD-algorithm}}` : 7,143.

The union holds 8,971 files, 7,090 of them JPEG, PNG or WebP. The **main** population is the tree plus the `{{PD-algorithm}}` files whose own categories match an AI pattern (`scripts/fetch_meta.py`).

"Main" is broader than "AI-generated":

- 149 of the 1,585 measured main files carry a category suggesting AI *modification* of an existing image (retouched, upscaled, colourised, restored);
- the 85 admitted `{{PD-algorithm}}` files include a few screenshots of AI products.

The union adds the remaining `{{PD-algorithm}}` files, which mix AI images with fractals and diagrams.

Sample. Stratified by the month of the current file version: a census of January–September 2026, 40 per month for 2024–2025, and 100 files before 2024.

Files were processed in a shuffled order by rounds. The agent fixed the time budget at about 23:25 CEST on 2 October, after it had looked at interim results, and the run stopped at 00:15. The measured files are therefore the first k of each month's random order: a simple random subsample, given a stop rule that does not depend on outcomes. Ours depends on them only weakly, through download time and file size.

- Coverage: 72% of January–July 2026 files and 70% of August–September 2026 files; 360 files from 2024–2025 and 25 from before 2024.
- Seven September files returned HTTP 404 (deleted after enumeration). They are treated as missing at random.
- Two files whose certificate our own extraction step could not parse are included. An earlier draft had wrongly excluded them.

Estimation.

- Period estimates are weighted by month.
- Intervals are Wilson intervals on the effective sample size, without a finite-population correction, because the claims are about the upload process, not about the files.
- Uploads come in batches. Comparisons therefore also report a **day-cluster bootstrap**: upload days are resampled within each month (2,000 replicates). Newcombe intervals, which treat files as independent, are given for reference.

Measurement. For each file:

1. download it and check its SHA-1;
2. run `ai-mark-lint FILE --format json --jurisdiction none` (version 0.1.0 from its release wheel, with `c2pa-python 0.38.0` / `c2pa-rs 0.91.0`);
3. for C2PA files, re-read the manifest under six trust and ECU configurations, with `c2pa-python 0.10.0` and with `c2patool 0.27.22`, and extract the certificate chain;
4. record the results and delete the file.

No step fetched remote manifests or used OCSP.

Definitions. A file carries an *AI mark* if any of these holds:

- it has a readable C2PA manifest declaring an AI `digitalSourceType` (in its own actions or its ingredients');
- its XMP carries an IPTC `DigitalSourceType` with such a value;
- it has a readable AIGC label.

Validity is reported under the corrected rule of ai-mark-lint 0.1.1, emulated from the recorded outputs (§4.4). The 0.1.0 figures are kept as the finding.

4. Results

4.1 How many AI images carry a mark

Main population, weighted by month; 95% Wilson intervals (files treated as independent):

UPLOAD PERIOD	N	ANY AI MARK	C2PA DECLARING AI	OF WHICH VALID (0.1.1)	TRUSTED (OFFICIAL LIST)	IPTC AI	NO MARK
2024	122	17.4% (11.3–25.7)	16.8%	15.8%	0.0%	0.6%	82.6%
2025	133	29.3% (22.2–37.6)	24.8%	3.9%	0.0%	4.5%	70.7%
2026 Jan–May	532	25.8% (22.2–29.7)	25.0%	15.5%	14.8%	2.3%	74.2%
2026 Jun–Jul	313	50.2% (44.6–55.7)	49.2%	45.4%	41.6%	1.3%	49.8%
2026 Aug–Sep	461	43.4% (39.1–47.9)	41.0%	38.0%	37.8%	2.6%	56.6%

The 2025 C2PA marks are mostly 2025 ChatGPT manifests: their certificates have since expired, and they carry no time-stamp (§4.4).

Mark profile of the 1,585 measured main files (unweighted):

PROFILE	SHARE
no AI mark	63.6%
C2PA, AI declared, valid (0.1.1)	26.7%
C2PA, AI declared, failing (0.1.1)	6.7%
IPTC <code>DigitalSourceType</code> only	1.5%
C2PA without an AI declaration	1.6%

AIGC.

- One file in the union (none in the main population) carries a readable AIGC label.
- Three more (one in the main population) carry a label encoded as a JSON string that contains the JSON object. ai-mark-lint 0.1.0 and 0.1.1 did not parse it.
- ai-mark-lint 0.1.2 parses it and flags the encoding as non-conformant.

4.2 The June jump, and 2 August

Main population, share with an AI mark by month of upload in 2026:

MONTH	SHARE (N)	WITHOUT COATS OF ARMS
January	20.6% (68)	22.2%
February	22.8% (57)	24.0%
March	25.4% (114)	26.6%
April	28.6% (98)	27.4%
May	27.2% (195)	26.4%
June	51.7% (145)	50.4%
July	48.8% (168)	51.0%
August	32.6% (224)	49.2%
September	53.6% (237)	55.7%

The coats-of-arms batch is large. In August, 96 of the 224 sampled main files belong to the `AI-generated coats of arms` branch, and only 10 of those 96 are marked. Re-dating files by first upload, as the review did, would move 8 of them back to July.

Comparisons (main population; differences in percentage points):

COMPARISON	ESTIMATE	NEWCOMBE 95% (INDEPENDENT FILES)	DAY-CLUSTER BOOTSTRAP 95%
Jun–Jul vs Jan–May	+24.4	+17.6, +31.0	+14.7, +33.6
Aug–Sep vs Jun–Jul	–6.7	–13.7, +0.4	–16.7, +4.6
Aug–Sep vs Jan–Jul	+8.4	+3.0, +13.9	+0.2, +16.8
same, without coats of arms	+17.8	+11.6, +24.0	+9.2, +27.0
Aug–Sep vs Jun–Jul, without coats of arms	+2.6	–5.2, +10.4	–8.0, +14.1

The reviewer's independent bootstrap of the January–July vs August–September comparison gave –0.9 to +17.8. That comparison is borderline once clustering is allowed for, and it depends on one batch. The robust finding is the June jump. August–September is not higher than June–July. Over the union (n = 2,084), the pattern is the same: +24.7 pp for the June jump, +8.0 pp for January–July vs August–September.

Decomposition of the June jump. By the source of the mark, weighted share of main files marked, January–May → June–July:

SOURCE OF THE MARK	JAN–MAY	JUN–JUL	CHANGE
OpenAI-signed	14.4%	28.8%	+14.4 pp
Google-signed	8.8%	14.1%	+5.3 pp
Microsoft-signed	1.0%	5.4%	+4.4 pp
other sources	–	–	+0.2 pp

Microsoft's share comes mostly from 9 Bing Image Creator files uploaded on 8 June. From May to June alone, OpenAI-signed files went from 16.4% to 25.5%, Google from 7.7% to 11.7% and Microsoft from 2.1% to 11.0%.

OpenAI's "Media Service API" manifests were already present in April (7 files) and May (31). The jump therefore did not start with them. It coincides with more OpenAI- and Google-signed uploads and with one Microsoft batch. Our data cannot tell whether people changed tools or tools changed behaviour (80% of main files name no generator).

On the AI Act. There is no step in August. A provider complying in anticipation of 2 August could, however, have changed its products before that date, so the data say nothing about the Act in either direction. We also count calendar months: the 16 main files uploaded on 1 August (3 marked) are in "August–September".

4.3 By generator and by signer

Selected generator rows (main population, generator inferred from category names):

GENERATOR CATEGORY	N	SHARE WITH ANY AI MARK
Microsoft Bing / Designer / Copilot	20	55.0% (all 11 manifests valid under 0.1.1)
Google (Gemini / Imagen / "Nano Banana")	86	48.8%
OpenAI GPT-4o / gpt-image / ChatGPT	88	45.5%
DALL-E	51	21.6%
Midjourney	12	16.7% (IPTC only)
Stable Diffusion	14	7.1%
Grok	12	0%
no generator category	1,272	34.7%

Two September 2026 files without a Grok category (one in the main population) carry a valid manifest declaring AI from **xAI Grok Imagine**. It is signed with a self-signed "LOCAL USE ONLY" certificate, so no trust list can trust it. The 12 Grok-category files carry no mark; they may predate C2PA in Grok Imagine or have passed through other tools.

Signers (union, 627 files with a manifest; main counts in `data/summary.md` §6):

SIGNER	MANIFESTS	VALID (0.1.1)	FAILING (0.1.1)	TRUSTED (OFFICIAL LIST)
OpenAI	403	327	76	317
Google	149	138	11	135
Microsoft	44	25	19	0
Canva	11	0	11	0
Adobe	9	9	0	0
Samsung	4	4	0	0
Anthropic	2	2	0	0
xAI (self- signed)	2	2	0	0
ByteDance	1	1	0	0

Notes on the table:

- Two more manifests are remote references to Adobe's manifest store (`cai-manifests.adobe.com`), which an offline validator cannot fetch.
- Microsoft's 19 failures are 12 "certificate params incorrect" plus the update-manifest errors from **Microsoft Paint**, an image editor, and 7 hash or claim errors.
- Canva's 11 manifests are expired and have no time-stamp.

4.4 Why manifests fail validation

All counts in this section are for the 625 readable manifests of the union (553 in the main population).

The finding: ai-mark-lint 0.1.0 rejected 302 of them (267 main). We reconstructed what each version would report from the recorded outputs:

VALIDATOR	TRUSTED	VALID, UNTRUSTED	FAILING
ai-mark-lint 0.1.0 (no trust list, no ECU list)	0	323	302
ai-mark-lint 0.1.1 rule, no trust list	0	508	117
0.1.1 rule with the official C2PA list + TSA list (624 files)	452	53	119
strict C2PA §15.8.2 with the official lists + ECU list (624 files)	452	53	119
c2patool 0.27.22, defaults	–	–	118
c2pa-python 0.10.0, defaults	–	–	105 of 296 it can read (cannot read 329, 320 of them OpenAI)

Two configurations cover 624 files, because one follow-up download returned 404. Two Google manifests come back `Invalid` with no failure code under the official-list configurations, which accounts for the 117 → 119 difference.

Of the 302 failures under 0.1.0, 185 disappear under 0.1.1. The causes (a file can have several):

CLASS	FILES (MAIN)	VERDICT
D. <code>signingCredential.invalid</code> , "certificate missing required EKU"	187 (161)	validator defect, fixed in 0.1.1
B. certificate expired, no time-stamp	85 (77)	real: §15.8.2 requires rejection
A. content changed after signing (hash mismatch)	32 (31)	real
F. first action not <code>c2pa.created</code> / <code>c2pa.opened</code> (§18.14.2)	26 (24)	real; 24 ChatGPT, 1 Sora, 1 Microsoft Designer
G. <code>manifest.update.invalid</code> / <code>claim.malformed</code>	16 (16)	real
E. "certificate params incorrect" (Microsoft Paint)	12 (12)	real
C. expired; time-stamp from a TSA on the official C2PA TSA list	9 (9)	configuration-dependent
B'. expired; time-stamp from a TSA on no list we used (Microsoft)	4 (4)	configuration-dependent

Class D: the defect. In c2pa-rs 0.91.0, `has_allowed_eku` always accepts `emailProtection` , `timeStamping` and `OCSPSigning` . Any other EKU is accepted only if it is in the configured list. The library's default file lists `documentSigning` , C2PA claim signing (`1.3.6.1.4.1.62558.2.1`) and Microsoft's C2PA OID. A context built from `Settings` without `trust_config` behaves as if that list were empty. The library's own tests load an EKU list (`settings/mod.rs` , `#[cfg(test)]`), so they do not exercise this path; c2patool does not show it either.

C2PA 2.2 §14.5.1 says EKUs outside the configured list "shall not cause the certificate to be rejected". Rejecting the claim-signing certificates is therefore non-conformant: that covers OpenAI's 153, issued by Trufo, which is on the official list, and Anthropic's 2. Samsung's `documentSigning` certificates (4) are covered by the "should allow" in §14.4.1. Microsoft's vendor-only EKU (28) depends on the validator's list.

ai-mark-lint 0.1.1 (released 2026-10-03) passes c2pa-rs's default EKU list.

Classes C and B': configuration-dependent, not artefacts. C2PA 2.2 §15.8.2 says of a TSA certificate that cannot be chained to a trust list: "the validator shall issue a timestamp.untrusted informational code and ignore the time-stamp". Without a valid time-stamp, "the C2PA Manifest is valid if the current time at validation is within the validity period of the signer's certificate ... If it is not, the C2PA Manifest shall be rejected with a failure code of `claimSignature.outsideValidity`."

- Class C is rejected without a TSA trust list and valid with the official C2PA TSA list (Google's TSA is on it).

- Class B' is rejected under every list we have.

ai-mark-lint 0.1.1 has no TSA trust list of its own. When the time-stamp is present but untrusted, it reports an expired certificate as a trust question (C2PA-003, inconclusive) rather than as valid or invalid. This is a documented deviation from §15.8.2's strict rejection, consistent with how it treats

`signingCredential.untrusted`. It does not check that the certificate was valid at the untrusted time. In this sample, every class C and B' certificate was valid at its TSA time.

Class B: expired, no time-stamp. These are the 2024–2025 ChatGPT manifests with certificates issued by Truepic, and Canva's. No configuration can validate them today. That does not show tampering, but they no longer work as a verifiable credential.

4.5 The `{{PD-algorithm}}` files without AI evidence

Of the 499 measured union files outside the main population, 14% carry an AI mark, almost all of them C2PA. The category tree therefore misses some AI uploads. The comparisons are unchanged on the union (§4.2).

5. What this means

- **For readers of Commons.** Between 43% and 53% of the AI images uploaded since June 2026 carry a machine-readable declaration (main population, depending on month and batch). Many do not, and the absence of a mark says nothing about whether an image is synthetic.
- **For people who build validators.** An "invalid" rate means nothing without the ECU list and the signer and TSA trust lists. Publish the configuration with the number. Our own tool got it wrong (§4.4).
- **For providers.** Manifests without a time-stamp stop validating when the certificate expires (85 manifests here). A self-signed certificate (Grok Imagine) can never be trusted.
- **For Commons.** Showing whether a file has C2PA, as discussed on the Village pump in 2025, would surface a declaration on 43–53% of new AI uploads (main population). Showing whether it is valid needs the configuration above.

6. Competing interests

ai-mark-lint, the instrument of this study, is EasyxLab's own open-source tool. This study found two defects in it:

- the ECU defect, which made 0.1.0 report 185 manifests as invalid that the corrected rule accepts (§4.4);
- the double-encoded AIGC label it did not parse.

The fixes shipped as 0.1.1 (2026-10-03) and 0.1.2. A study that reports problems in its own tool has an interest in showing that the tool is now right. We therefore give the 0.1.0 figures next to the corrected ones, and every corrected figure is computed by `scripts/analyze.py` from the recorded validator outputs. They are not recomputed by the new version of the tool. ai-mark-lint is free and open source.

7. Limitations

- **Self-selection.** People who upload AI images to Commons are not a random sample of AI users. Commons policy discourages many AI uploads.
- **Removed or never present.** We cannot distinguish a mark that was removed before upload from one that never existed.
- **Invisible watermarks.** We did not examine invisible watermarks such as SynthID.
- **Revocation.** We did not check revocation (no OCSP or CRL).
- **Survivorship.** Deleted files are absent, and exposure to deletion grows with age: January uploads have had nine months, September uploads days. Deletion requests now cite C2PA metadata as evidence, so survival may depend on the mark. This confounds the month-by-month comparison, and we did not bound it.
- **Dates.** We use the current file version. The review found that 84 of 1,305 measured 2026 main files (6.4%) have several versions, and that dating them by first upload would move 23 files between periods, 8 of them from August to July.
- **Population.** It relies on volunteer categorisation. "Main" includes AI-modified images and a few screenshots of AI products (§3).
- **Clustering.** Batch uploads make the files dependent. Newcombe intervals understate uncertainty; use the bootstrap.
- **Partial census and timing.** About 70% of 2026 uploads were measured. The time budget was set after interim looks at the results.
- **Generators.** Generator labels come from categories; 80% of main files name none.

8. Automation and review

An AI agent (an LLM-based coding agent supervised by EasyByte) did the following:

- designed and ran the study;
- wrote the scripts;
- traced the validation failures in the c2pa-rs source;
- drafted this paper.

A second, independent LLM-based agent then reviewed it adversarially. That agent recomputed every figure with its own code, re-validated 13 files with ai-mark-lint 0.1.0 and 0.1.1, checked the quotes against their sources, and checked the sampling. The changes it required are applied in this draft.

`scripts/analyze.py` produces every number in this paper and the README from `data/` .
`scripts/check_headlines.py` asserts each headline figure against `data/` and its presence in the text.

9. Data, ethics and reproducibility

`data/per_file.csv` has one row per inspected file. It gives:

- the Commons title, page URL and month;
- the marks;
- the signer and failure codes under each configuration;
- the 0.1.0 and 0.1.1 verdicts.

Privacy.

- No uploader name is requested or stored.
- Categories named after users are published only as placeholders. Raw category lists stay local.
- Samsung device-attestation certificates are reduced to their issuer.

Images and requests. No image was kept; each was deleted after inspection, with at most one on disk at a time. Nothing was edited on Commons. Requests were serial, at most one per second, with `maxlag` and a descriptive User-Agent (METHOD §6).

Code: Apache-2.0. Data and text: CC BY 4.0.

References

- Regulation (EU) 2024/1689 (AI Act), Art. 50(2).
- C2PA Technical Specification 2.2: §14.4.1, §14.5.1, §15.8.2, §18.14.2 (spec.c2pa.org, retrieved 2026-10-03). C2PA conformance trust list, `c2pa-org/conformance-public @ 3573be5` (list issued 2026-08-05).
- contentauth/c2pa-rs v0.91.0: `certificate_trust_policy.rs`, `certificate_profile.rs`, `valid_eku_oids.cfg`, `settings/mod.rs`.
- Rijsbosch B., van Dijck G., Kollnig K. (2025). Adoption of Watermarking for Generative AI Systems in Practice and Implications under the new EU AI Act. arXiv:2503.18156.
- Rijsbosch B., Bekavac L., Tari H., van Dijck G., Kollnig K. (2026). Drowning in AI Slop: How Social Media Platforms (Do Not) Label AI and Deepfake Content under EU law. arXiv:2609.38571.
- Zewde K. et al. (2026). GPT-Image-2 in the Wild. arXiv:2604.25370.
- Verifying Provenance of Digital Media: Why the C2PA Specifications Fall Short (2026). arXiv:2604.24890.
- Authenticated Contradictions from Desynchronized Provenance and Watermarking (2026). arXiv:2603.02378.
- Wikimedia Commons: `Commons:Village pump/Technical/Archive/2025/09`, `Commons:Village pump/Archive/2025/12`, `Commons:Signs of AI media`, `Commons:AI-generated media`; Phabricator T387075.
- EasyxLab (2026). Do AI provenance marks survive common image, video and audio pipelines? Study S3. EasyByte Hub S. Coop. Mad. <https://github.com/easybytehub/easyxlab>
- ai-mark-lint 0.1.0–0.1.2, <https://github.com/easybytehub/ai-mark-lint>.

Cite this study

CITATION

EasyxLab (2026). Do AI-generated images on Wikimedia Commons carry provenance marks? Study S6. EasyByte Hub S. Coop. Mad.
<https://github.com/easybytehub/easyxlab/tree/main/studies/s6-commons-ai-marks>

BIBTEX

```
@techreport{easyxlab_s6,  
  title      = {Do AI-generated images on Wikimedia Commons carry provenance  
marks?},  
  author     = {{EasyxLab}},  
  institution = {EasyByte Hub S. Coop. Mad.},  
  number     = {S6},  
  year       = {2026},  
  url        = {https://github.com/easybytehub/easyxlab/tree/main/studies/s6-  
commons-ai-marks}  
}
```