TRUSTCAPTURE

Trust infrastructure for digital media

Berlin, Germany · 2026

TrustCapture signs photos and video with hardware-backed cryptography at the moment of capture, registers them in an append-only chain, and gives anyone — a platform, a journalist, a stranger with a phone — a confidence assessment instead of a guess.

The problem

During the Iran–Israel conflict in early 2026, AI-generated battle footage spread across X, Telegram, and WhatsApp faster than anyone could verify it[1]. X's own head of product, Nikita Bier, ended up announcing a blunt fix: creators who post undisclosed AI-generated war footage lose access to ad revenue sharing[2]. Not because X had a good way to catch fakes at the moment of posting — it didn't, and mostly still doesn't. Detection runs behind the content, after millions of people have already seen it and formed an opinion.

This isn't only a war-coverage problem. The same pattern shows up constantly at smaller scale: a renter deciding whether an apartment listing is real, a buyer deciding whether a second-hand item exists as pictured, a driver deciding whether a rental car already had that scratch. The stakes are lower per incident, but the structure is identical — media is cheap to fabricate, cheap to copy, cheap to strip of context, and verification is still slow, manual, and mostly a matter of trusting whoever presented it.

The problem is not carelessness. It's structural. The tools for verification don't move at the speed fabrication does — for anyone, however careful.

The insight

This isn't a new kind of problem. The telegraph collapsed the time between an event and public opinion about it faster than institutions could build filters for truth, and yellow journalism exploited exactly that gap. What closed it wasn't media literacy or better fake-detection — it was infrastructure: wire services like Reuters and AP that created authenticated source provenance.

HTTPS did the same thing for the web. It didn't try to detect malicious sites — it gave legitimate ones a way to prove themselves, and gave everyone else a simple, visual way to check.

TrustCapture applies the same move to visual media, with one adjustment we think matters: trust isn't binary. A rental-car handover, a dating profile photo, an insurance claim, and a piece of legal evidence don't need the same level of assurance, and shouldn't get the same yes/no answer. What they need is a confidence signal built from whatever verifiable evidence actually exists — capture conditions, hardware attestation, registry history, and increasingly, standards like C2PA — that gets stronger as more signals stack up, and stays honest when they don't.

We don't detect fakes. We accumulate evidence, and we say plainly how much of it there is.

How it works

1. Capture

Photo or video is captured live inside the app. Gallery uploads are blocked in the core flow — this is not a filter applied after the fact.

2. Sign

At the moment of capture, before the image touches app memory: a SHA-256 hash of the raw bytes, GPS, timestamp, and device metadata are combined into a manifest and signed with a hardware-backed key — Secure Enclave on iOS, StrongBox on Android flagships. The private key never leaves the chip.

3. Register

The signed manifest is sent to an append-only registry. Records are cryptographically chained; nothing is altered retroactively.

4. Verify

A QR code carrying the verification URL is embedded in the image itself — not just in its metadata — so it survives WhatsApp, screenshots, and re-uploads, all of which strip conventional metadata (including C2PA's) at every hop.

5. Score

The verification page doesn't just say "valid" or "invalid." It shows what evidence actually exists — live capture, hardware signature, GPS/timestamp/device consistency, registry integrity, C2PA or App Attest signals where present — and gives a confidence assessment based on what's there. Missing evidence lowers confidence. It doesn't break the system.

Why this is different from C2PA — and why that's fine

C2PA — the open standard backed by Adobe, Microsoft, Google, and the BBC — is the right foundation, and we build on it rather than compete with it. Google's Pixel 10 now signs every photo by default with hardware-backed keys, offline timestamping, and the highest C2PA conformance rating available[3]. That's good news for us, not bad news: it's exactly the kind of signal our scoring engine is built to ingest.

But C2PA metadata lives inside the file, and the moment that file is shared — WhatsApp, screenshot, re-upload — the metadata is gone. The path most consequential images actually travel (phone → WhatsApp crop → repost → news screenshot) strips provenance at every step. Our QR layer is the part that survives that path, because it's part of the visual, not the file.

The durable value isn't the camera app, and it isn't any single signing technology — those get commoditized by OEMs over time, and that's already starting. The durable value is the registry and scoring engine sitting on top of every signal, old and new, that gets better as the ecosystem adds more of them. We're not betting on one standard winning. We're betting that high-trust decisions will always need something that weighs many signals together.

What's already built

- iOS live capture (Swift), gallery upload hard-blocked in the core flow
- Secure Enclave key generation and ECDSA (ES256) signing, verified live on real hardware via App Attest, DER-encoded
- Full manifest collection: GPS, EXIF, device model, camera/lens metadata, LiDAR depth signal where available
- SHA-256 image hash bound to the manifest

- Backend: Native AOT C# gRPC service, deployed, validating signatures against an append-only PostgreSQL registry
- Production infrastructure with TLS certificate pinning and hardened routing for the API and verification endpoints
- Draggable QR overlay on the capture review screen

The cryptographic core — signing, App Attest validation, registry write — is working end to end on real hardware. This is a working system, not a concept.

Roadmap

TrustCapture is a working MVP moving toward pilot readiness, not a finished platform. In the interest of saying plainly what exists — the same standard we apply to media — here is what's next:

- Public verification website — in progress
- C2PA signal ingestion into the scoring engine
- Android client (StrongBox-backed signing)
- Video manifest support on the backend (client-side already implemented)
- Risk-tiering by use case — rental, legal, insurance
- Admin and dashboard tooling for platform customers

Why now

- C2PA is open, free, and has real institutional weight — Adobe, Microsoft, Google, BBC
- Google Pixel 10 ships hardware-backed C2PA by default — the first mainstream phone to do it, and proof OEMs will build the signing layer themselves[3]
- The EU AI Act's Article 50 transparency rules become mandatory on 2 August 2026, requiring machine-readable marking of AI-generated content and labelling of deepfakes[4]. The European Commission's Code of Practice names digitally signed, timestamped metadata as one of the two required marking layers[5] — which is, functionally, what TrustCapture already does at the point of capture
- LiDAR is already in the install base of iPhone Pro devices as an anti-spoof sensor
- Public trust in visual media crossed a threshold in 2025–2026 that it hadn't crossed before
- Trust is already something people pay for — LinkedIn Premium and Bumble verification proved the willingness

Use cases

Vehicle handover documentation

Peer-to-peer car rental and car-sharing platforms like Turo and Getaround already require timestamped, geotagged photos at pickup and drop-off as evidence for damage disputes[6] — but that timestamp is applied by the app, not by hardware. It can be argued with, and it regularly is: independent reviews score both platforms poorly specifically on how fairly damage disputes are resolved[7], and keyless pickup models remove human inspection from the process entirely[8]. In one documented case, a renter nearly paid $3,000 in disputed damages because she hadn't produced the app-timestamped photos the platform asks for[9] — exactly the failure mode a hardware-signed manifest is built to close.

The fraud vector here is entirely photo-based, and TrustCapture integrates as an API/SDK upgrade to a pickup/drop-off flow platforms already run — not a new consumer app anyone has to be talked into downloading.

High-trust marketplaces

Rental and housing listings, second-hand marketplaces, and other platforms where visual fraud creates real moderation cost. Same registry, same scoring engine, different customer.

Media, insurance, and legal documentation

Anywhere a photo or video is presented as evidence — an insurance claim, a news submission, a condition report — the same confidence assessment applies. Different use cases need different assurance levels, and the scoring model is built to reflect that rather than force a single yes/no answer.

About the founder

TrustCapture is built by a senior backend engineer with more than a decade in distributed systems, event-driven architecture, and production delivery across fintech and e-commerce; a second-time founder. The current MVP — backend and iOS — was built end to end, largely solo, with systems architecture and security-adjacent infrastructure as the core strength.

Even a person who cares and tries their best can still be deceived — because the infrastructure for verification doesn't exist yet. That is what we are building.

Contact

Web: trustcapture.app · Berlin, Germany

Sources

[1] Fake, AI-generated images and videos of the Iran war are spreading on social media — CNN Politics — https://www.cnn.com/2026/03/11/politics/fake-ai-images-videos-iran-war

[2] X to crack down on AI-generated war footage — The Times of Israel — https://www.timesofisrael.com/liveblog_entry/x-to-crack-down-on-ai-generated-war-footage/

[3] How Pixel and Android are bringing a new level of trust to your images with C2PA Content Credentials — Google — https://blog.google/security/pixel-android-trusted-images-c2pa-content-credentials/

[4] Commission publishes Code of Practice on marking and labelling AI-generated content — European Commission — https://digital-strategy.ec.europa.eu/en/news/commission-publishes-code-practice-marking-and-labelling-ai-generated-content

[5] Taking the EU AI Act to Practice: The Final Transparency Code of Practice — Bird & Bird — https://www.twobirds.com/en/insights/2026/taking-the-eu-ai-act-to-practice-the-final-transparency-code-of-practice

[6] Checking in a guest and checking out — Turo Help Center — https://help.turo.com/en_us/checking-in-a-guest-H1gsHEg4c

[7] Turo Terms of Service Review — ToS Watchdog — https://terms.law/ToS-Watchdog/car-sharing/turo/

[8] Getaround Terms of Service Review — ToS Watchdog — https://terms.law/ToS-Watchdog/car-sharing/getaround/

[9] Renter disputes suspicious damage claims to a Turo car — syndicated consumer-advice column — https://tucson.com/news/nation-world/business/article_91f9414e-a99b-11ee-9733-dfb24ac0075d.html

Historical analogies (telegraph-era wire services, HTTPS) and named consumer precedents (LinkedIn Premium, Bumble verification) are widely documented and not separately footnoted here.