

Context attestation for EU digital identity wallets

A public technical overview of Calmproof: what it observes on the phone, what it signs, how a verifier uses it, and why it protects people without watching them.

Document	Version	Published
Public technical overview	1.0	26 October 2026

Schema	Status
Attestation record v1 (draft)	Public, informative

1 Summary

The EU Digital Identity Wallet proves *who* is presenting a credential. It does not show *under what circumstances* the person agreed. Social-engineering fraud lives in that gap: the credential is genuine, the wallet is genuine, the bank is genuine, and the person shares their identity while a caller on the line tells them what to tap.

Calmproof adds one small, signed object to the presentation response: a **context attestation**. In the moments of consent, the wallet reads a short, fixed list of signals on the device, such as an active call, screen sharing, remote control or a request relayed from another device. It derives a risk class from them, protects the person where needed, and seals a record containing only categories. The verifier receives the record inside the response it already checks, bound to the same request, and applies its own policy.

On the device Signals are read and evaluated locally. Raw signals never leave the phone.	Categories only No phone numbers, no call content, no screen content, no location.
Bound to the request Same key, same nonce, same response. Nothing to detach, replay or suppress.	The verifier decides The record is evidence, not a verdict. Each bank applies its own policy.

2 The problem: consent under pressure

Most credit-transfer fraud by value in the EU is no longer a technical break-in. According to the joint EBA/ECB report of December 2025, 74 % of it is the payer being manipulated into authorising the payment themselves. The typical script is an impersonation call: the police, the bank's fraud department, the tax office. Under the EU Payment Services Regulation, banks will carry the loss when a fraudster impersonates them.

From 2027, banks and other regulated services in the EU have to accept the EU Digital Identity Wallet. The wallet ecosystem already secures what it was designed to secure:

- the authenticity of the credential (issuer signature),
- holder binding (the presenting device holds the key),
- freshness (a one-time nonce from the verifier),
- the identity of the verifier and the integrity of the wallet.

None of these checks says anything about the situation the person is in when they tap *Share*. A presentation made calmly at the kitchen table and one made while a stranger on the phone dictates each step look identical to the verifier.

Existing fraud signals either stay on the phone (a warning the person can dismiss), travel through a separate channel that is not bound to the presentation, or depend on a central service watching the user. Calmproof closes the gap inside the presentation itself.

3 Threat model

Actors

Holder	The person presenting a credential from their wallet. Honest, but possibly under pressure, deceived or guided by someone else.
Verifier	A bank, payment service provider, authority or other relying party that requests a presentation and acts on it.
Attacker	A remote fraudster who cannot break the cryptography and does not hold the device, and therefore works through the person: by voice, by remote control, by screen sharing or by relaying a request.

Scenarios in scope

Scenario	What happens	What the record shows
The police call	A caller claims the account is being misused and asks the person to confirm their identity in the wallet, during the call.	<code>voice_call</code>
The remote support	A fake technician installs a remote-control app and guides the person through a presentation.	<code>remote_access</code> , often with <code>voice_call</code>
The relayed QR code	The attacker starts a session in their own browser and sends the QR code to the victim, who scans it with the wallet.	<code>cross_device_no_proximity</code>
The shared screen	A screen-sharing session runs while the person consents; someone else sees or controls the screen.	<code>screen_share</code>
The call that goes on	The presentation itself looked calm; the call starts right after, and the next step is a transfer.	A follow-up record bound to the same request

Out of scope

Calmproof is deliberately narrow. It does not claim to detect:

- **Pressure in the room.** Someone standing next to the person leaves no signal on the device.
- **A fully compromised device.** Device-integrity signals and platform attestation raise the bar and are reported as such, but a record from a compromised operating system cannot be stronger than that operating system.
- **A malicious wallet.** Wallet integrity is the job of wallet certification and wallet attestation in the EU framework.

Saying where the boundary is matters for the verifier's policy: a calm record lowers no other control. It adds evidence where there was none.

4 Architecture

Calmproof has three parts. The **context unit** runs next to the wallet on the device. The **attestation record** travels inside the presentation response. The **verifier policy** runs in the verifier's existing systems.

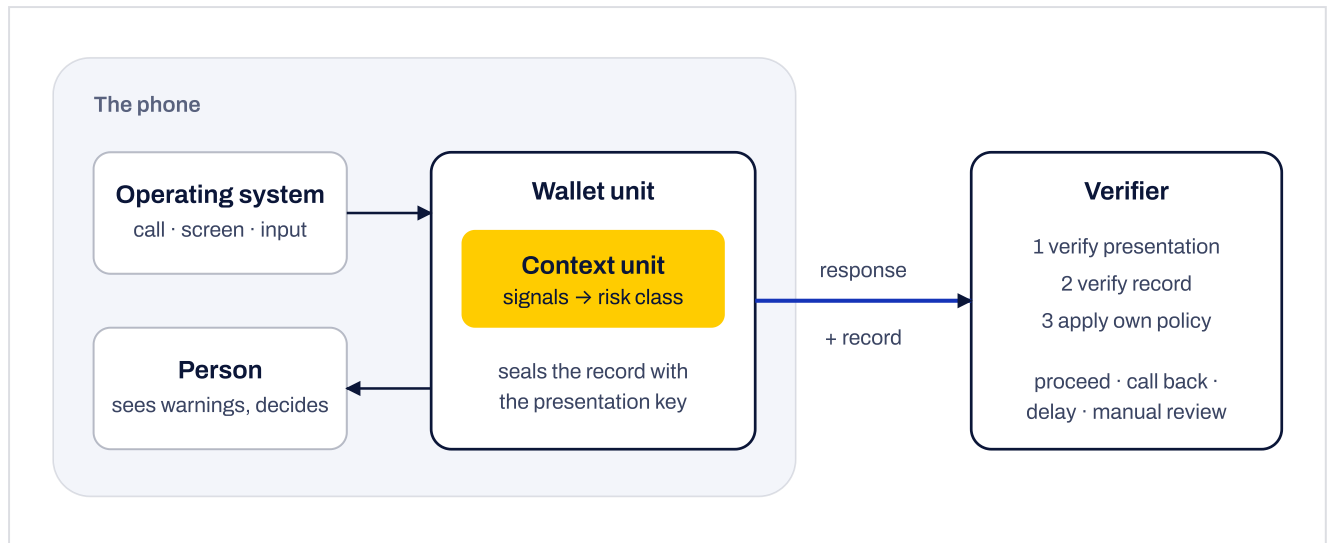


Figure 1. Only the sealed record leaves the phone, inside the presentation response the verifier already receives.

Deployment options

Option	Where the context unit runs	Typical use
Wallet	Inside the wallet application	Wallet providers adding protection for all their users
Verifier app	As an SDK inside the verifier's own app, same-device flow	A bank protecting presentations made from its banking app
Operating system	As a system service used by both	Platform-level integration

In all three options the record has the same format and is verified the same way.

5 The flow, step by step

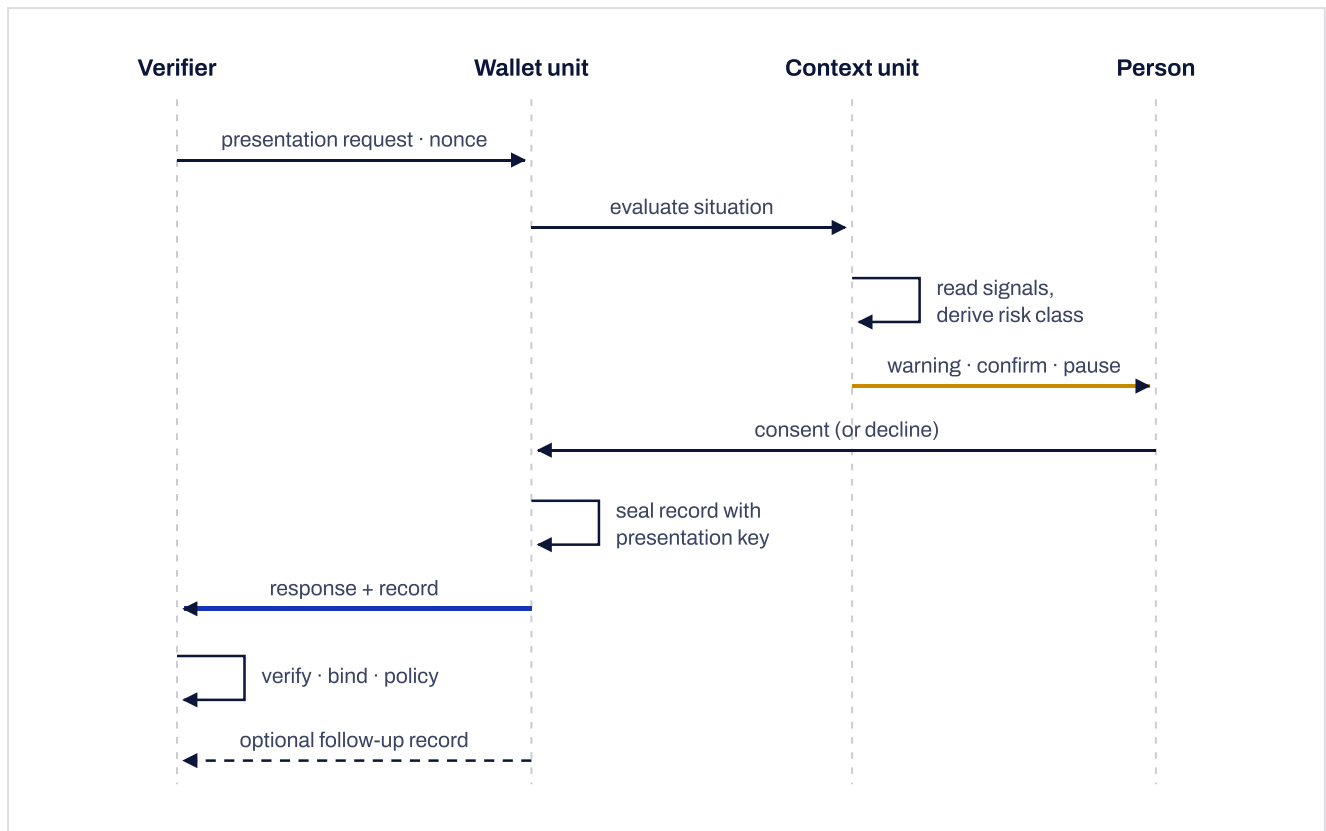


Figure 2. One request, one response. The follow-up record is bound to the same request.

- 1 The verifier asks.** It requests a credential and, optionally, the context attestation, through the standard presentation request with its one-time nonce.
- 2 The wallet looks around.** While the person decides, the context unit reads operating-system signals and derives a risk class from 0 to 3.
- 3 The person is protected.** Depending on the class: a plain warning, a second confirmation, a short pause, or a refusal. The person always sees what is being recorded and can decline the presentation.
- 4 The record is sealed.** Risk class, category flags, timestamp, the hashes of request and presentation and the capability vector are signed with the presentation key and, where available, covered by a platform attestation.
- 5 The verifier decides.** It verifies seal, binding and freshness and applies its own policy.
- 6 If the situation changes.** If a call starts right after a calm presentation, the wallet can send a follow-up record bound to the same request, for example before a transfer is executed.

6 Signal categories

The context unit reads a fixed, short list of signal categories. Each one is evaluated on the device; only the category name can appear in the record.

Category	Meaning
voice_call	A voice or VoIP call is active during consent.
screen_share	The screen is being shared or recorded.
remote_access	A remote-control session is active.
cross_device_no_proximity	The request came from another device without a proof that both devices are close together.
device_integrity	The device or operating system failed an integrity check.
history	Presentation history on this device, for example a verifier the wallet has never seen before.

Observable is not the same as calm. Not every platform and OS version exposes every signal. The record therefore carries a *capability vector* listing the categories this device could observe. A missing flag on a device that cannot observe calls means "unknown", never "no call". Policies stay fair across platforms and a downgrade cannot pass as a calm situation.

7 Risk classes and protection

The context unit condenses the observed signals into one of four risk classes. How signals combine into a class is defined by a versioned rule set; its weights and thresholds are not part of this document.

 Calm No relevant signal observed. The presentation proceeds normally.	 Notice A plain, short hint to the person. One tap to continue.
 Elevated A clear warning and a second confirmation, or a short pause.	 High The wallet recommends refusing; the verifier is told why, in categories.

The warning text is written for the person, not for the verifier: *"Banks and authorities never ask you to share your ID during a call. Pause, hang up, call the number on your card."* It names the situation without blaming anyone and gives one concrete action.

8 The attestation record (schema v1)

The record is a small JSON object. Field names below are from the schema draft v1 and may still change before the specification is published.

Field	Type	Meaning
<code>risk_class</code>	integer 0–3	Risk class derived on the device
<code>flags</code>	array of strings	Category flags from section 6, or "not observable"
<code>issued_at</code>	RFC 3339 timestamp	Time within the consent window
<code>request_hash</code>	string	Hash of the presentation request (nonce, transaction data)
<code>presentation_hash</code>	string	Hash of the presentation the record belongs to
<code>channel</code>	object	Same-device or cross-device, with the proximity-proof result
<code>user_response</code>	string	Outcome of the prompt: confirmed, confirmed after warning, paused, declined
<code>policy_version</code>	string	Version of the rule set used on the device
<code>capabilities</code>	array of strings	Signal categories this device could observe
<code>seal</code>	object	Signature with the presentation key and, where available, a platform attestation over the record's hash

```
{
  "risk_class": 2,
  "flags": ["voice_call"],
  "issued_at": "2027-03-02T10:41:07Z",
  "request_hash": "sha-256:9K_2A5P...",
  "presentation_hash": "sha-256:Qa7mC...",
  "channel": { "mode": "same_device" },
  "user_response": "confirmed_after_warning",
  "policy_version": "1.0.0",
  "capabilities": ["voice_call", "screen_share", "device_integrity"],
  "seal": { "alg": "ES256", "kid": "presentation-key", "platform": "verified" }
}
```

Illustrative example. Hash values are shortened.

9 Verifier processing

A verifier that requested the attestation processes it right after its usual presentation checks:

- 1 Verify the presentation as usual: issuer signature, holder binding, nonce.
- 2 Verify the seal with the same holder key, and the platform attestation where present.
- 3 Check that `request_hash` matches the request it sent, including nonce and transaction data.
- 4 Check that `presentation_hash` matches the presentation received.
- 5 Check that `issued_at` lies within the consent window.
- 6 Read `flags` against `capabilities` : separate "not observed" from "not observable".
- 7 Apply its own policy for the transaction type: proceed, proceed with a call-back, delay, or manual review.
- 8 Keep the record as evidence, under the verifier's own retention rules.

A presentation without a valid record is not rejected by the protocol. It is treated as "no attestation", and the verifier's policy decides which fallback applies.

10 Protocol fit

The context attestation uses the formats and transports of the EU wallet ecosystem. It does not change any credential format.

OpenID for Verifiable Presentations	Requested as a credential type through DCQL or as a <code>transaction_data</code> element; returned inside the same presentation response, signed with the presentation key.
SD-JWT VC and W3C VC	Carried as a claim in the key-binding part, or as an additional credential in the response.
ISO/IEC 18013-5 (mdoc, proximity)	Included in the device-signed data and bound to the session transcript. Proximity proofs for cross-device hand-overs are part of the record.
EUDI Architecture and Reference Framework	A wallet's ability to produce the record can be signalled through the Wallet Unit Attestation.
Digital Credentials API	Same-device browser flows carry the record in the same response as the presentation.

A specification draft for OpenID4VP is available to pilot partners and standards contributors on request.

11 Security properties

Property	How it is achieved
Non-separable	Same key, same nonce, same response as the presentation. There is no second channel to suppress or misroute.
Fresh and bound	Bound to the verifier's one-time request and to the presentation hash. A copied record proves nothing elsewhere.
Tamper-evident	Signed with the presentation key and, where available, covered by a platform attestation over the record's hash. A modified record fails verification.
Downgrade-resistant	The capability vector states what the platform could observe, so a missing flag cannot be passed off as a calm situation.
Relay-aware	Cross-device requests without a proximity proof become a category of their own.
Measured, not promised	Detection per attack scenario, false-positive rate and latency are measured in every pilot against the partner's own test scenarios.

12 Privacy by construction

- **Local evaluation.** Signals are read and evaluated on the device, at the moment of a presentation request.
- **Categories only.** The record never contains phone numbers, contacts, call or screen content, app names or location.
- **Only to the verifier the person chose.** The record travels in the response to the request the person is answering. There is no Calmproof server in the flow, and BuonaLabs receives nothing.
- **Visible and declinable.** The person sees what is recorded and can decline the presentation.
- **Minimal history.** The history category relies on a small local record of past presentations, kept on the device.

More detail for data-protection officers: [Data protection](#).

13 Accessibility and fairness

Accessibility services can look like remote control to a naive detector. Calmproof never counts them alone and never produces a category of their own, so people who depend on screen readers or switch access are not flagged for using them. Warnings are short, plain-language and compatible with screen readers.

14 What this document leaves out

The following are trade secrets and are shared under NDA with pilot partners only:

- scoring weights and thresholds of the rule set,
- the pattern library behind the risk classes,
- per-platform signal reliability and the interfaces used to read each signal,
- results from the attacker lab.

15 Status and roadmap

21 October 2026

German patent application filed (patent pending).

Q1 2027

Pilot programme with banks, payment providers and wallet providers in Germany and Italy: verifier SDK, reference verifier, Android reference context unit, attestation schema v1.

2027

Specification draft for OpenID4VP shared with pilot partners and standards contributors; measured pilot results.

Later

iOS context unit and operating-system integration.

16 Glossary

Context attestation	The signed record of the situation in which a presentation was made.
Context unit	The component on the device that reads signals, derives the risk class and seals the record.
Capability vector	The list of signal categories a given device can observe.
Follow-up record	A later record, bound to the same request, reporting a change in the situation after consent.
Presentation	The holder's response to a verifier's request, containing the requested credential data.
Verifier	The relying party that requests and evaluates a presentation.

17 References

- [1] Regulation (EU) 2024/1183 amending Regulation (EU) No 910/2014 (European Digital Identity Framework).
- [2] OpenID Foundation: OpenID for Verifiable Presentations.
- [3] IETF: SD-JWT-based Verifiable Credentials (SD-JWT VC).
- [4] ISO/IEC 18013-5: Personal identification — ISO-compliant driving licence — Part 5: Mobile driving licence (mDL) application.
- [5] European Commission: EUDI Wallet Architecture and Reference Framework.
- [6] W3C: Verifiable Credentials Data Model; Digital Credentials API.
- [7] EBA and ECB: joint report on payment fraud, December 2025.
- [8] EU Payment Services Regulation, final texts, April 2026.

Calmproof is a product of BuonaLabs UG (haftungsbeschränkt), Lübeck. Patent pending (DE). This document is informative and does not grant any licence.