Visibility, enforcement, evidence

01 Visibility

See what AI is doing. Discover the AI tools in use and log what goes in and out. This is necessary, and almost everyone does it competently. On its own it tells you what happened, after it happened, in a record you have to trust.

02 Enforcement

Stop AI doing something harmful. Block a tool call, redact a prompt, deny a risky action. Most controls do this in some form. It reduces the chance of a bad outcome, but it does not, by itself, prove what was stopped or what was allowed. Where redaction is in play, detection is best-effort, so it lowers exposure rather than guaranteeing every sensitive value is caught.

03 Evidence

Prove what AI did. Produce a record a regulator, auditor, or court can verify mathematically, independent of the vendor's attestation. Signed at the source, tamper-evident, encrypted under keys you hold, and written across the AI tools that reach a provider. This is the job regulated institutions actually have to satisfy, and the one most tools were never built for.

> In regulated environments, "we caught it" is not enough. You have to prove it, and prove the record was not changed.

Why evidence is the hard job, not the last job

Visibility and enforcement are prerequisites. Evidence is the obligation. When a family sues, a federal investigator asks, or a state attorney general subpoenas the record, the question is never "did your dashboard show it." It is "produce the record, and show it was not altered." A log you, an insider, or an attacker could edit is not that record. A record signed on the device and hash-chained, so any change is detectable, is.

The third job is the one the headlines turn on

The cost of getting this wrong is not abstract, and it is concentrated in exactly the sectors that now want AI in the workflow. When the incident lands, the demand is always the same: produce a record that holds up. The figures below are what that demand looks like at scale.

About 243 million

Individuals whose protected health information was exposed in U.S. healthcare breaches that occurred in 2024, driven largely by the single Change Healthcare incident. Source: HHS Office for Civil Rights reports to Congress, as reported by The HIPAA Journal.

Around 62.4 million students	The figure the PowerSchool attacker claimed to have stolen from K-12 districts. PowerSchool did not confirm a total affected; only a subset of individuals had Social Security numbers exfiltrated. Source: BleepingComputer, January 2025.
276 ransomware attacks	On government organizations in the first nine months of 2025, a 41 percent increase over the same period in 2024, with 443,522 records known to have been breached in the confirmed attacks. Source: Comparitech, Government Ransomware Roundup, Q1 to Q3 2025.
$4.44 million	Global average total cost of a data breach, the first decline in five years, credited to faster AI-assisted containment. Healthcare remains the costliest industry at $7.42 million. Source: IBM Cost of a Data Breach Report, 2025.

Each of those becomes a records demand. After a healthcare breach the regulator wants documentation of what was accessed and how. After a student-data breach the parents, the district, and the state attorney general want it. After a ransomware hit on an agency, oversight wants it. Visibility tooling can show a dashboard while the incident is live. None of that survives the moment someone asks you to **prove the record was not changed.**

WHAT EACH JOB LEAVES OPEN

- **Visibility alone:** tells you what happened after it happened, in a record you have to trust.
- **Enforcement alone:** reduces the chance of a bad outcome, but does not prove what was stopped or allowed.
- **A log alone:** is mutable and held by the same system that made it, so it is weak under a determined challenge.

WHAT EVIDENCE ADDS

- **Signed at the source:** each captured interaction is signed by the originating device, so its origin cannot be repudiated later.
- **Tamper-evident:** entries are hash-chained, so any change after the fact breaks every fingerprint that follows it. Alteration is not prevented, it is made detectable.
- **Yours to verify:** content is encrypted under keys you hold, so even Verillian cannot read the record, and the chain can be checked mathematically rather than on the vendor's word.

Where Verillian sits

Verillian treats visibility and enforcement as the floor and builds for evidence. This is runtime governance and evidence for regulated AI. The policy decision is made at the device, before content reaches a provider: allow, redact, or block, and fail-closed when no valid policy applies. Every captured interaction is signed by the originating device and hash-chained, and content is encrypted under keys you hold, so the record is verifiable mathematically, independent of vendor attestation. That maps to the controls your frameworks already require: CJIS Security Policy v6.0 tamper-evident logging with one-year retention, HIPAA data minimization with six-year documentation retention, and the NIST 800-53 audit family. Verillian is built to align with these requirements and is aligned, not independently certified; CJIS is validated through your agency's state-agency audit.

One precision worth stating plainly, because a security buyer will ask. Tamper-evidence means any change to a captured record is detectable, not that capture or deletion is prevented. Completeness holds for every captured interaction, not for interactions a host-controlling insider could keep from being captured at all. That honesty is the point: a control that overstates what it proves is the one that fails when the proof is actually demanded. Verillian claims what the cryptography delivers, signed at the source, hash-chained, and readable only with your key, and says so in those terms.

Audit content is fully parsed for Anthropic, Claude and Claude Code today, with best-effort detection across other providers, and the software is **source-available (Elastic License 2.0)**, so your security team can read exactly what runs on your machines. Visibility tells you. Enforcement stops it. Evidence lets you stand behind it years later.

The model proposes. Verillian decides, and proves it.

To see the signed, hash-chained record built on your own infrastructure, under your own keys, name a sponsor and a technical contact and we will scope a pilot.

hello@verillian.ai