Architecture and data flow

Why the architecture is the product

Most AI governance tools sit downstream of the device, on the network or at a gateway, where they can report on traffic but cannot decide what an AI tool does inside an encrypted session, and where the log they keep is one an insider or an attacker can quietly edit. Verillian is built the other way around. The decision is made at the device, before egress, and the evidence is signed there and chained, so the record's integrity is verifiable mathematically, independent of vendor attestation. The sections below show exactly where your data goes, where it never goes, and what makes the record hold up.

The path a request takes

Request and response travel the same axis: device, decide, provider. The agent sits at the device and evaluates your policy before egress. Only permitted content continues to the provider, after any required redaction. A blocked request never leaves the device. Every captured interaction is signed into a hash chain held by your institution, under your keys. The provider receives only what policy allows, and the record of the decision stays with you.

SIGNED CHAIN ENTRY

Each captured interaction is signed and hash-linked to the one before it, so the order and content of decisions cannot be quietly rewritten without breaking the chain. The chain is held by your institution, under your keys, never by Verillian.

Components

Agent

Runs on each user device. It sits in the path of the AI tools that reach a provider, evaluates policy, encrypts content with your organization's key, signs each entry, and uploads to the admin service. It enrolls through your identity provider over OIDC and receives all configuration from the server. Signing keys are generated and held on the device and never leave it.

Admin service

The control plane: users, policy authoring, enrollment, event streams, and the audit viewer. Distributes Ed25519-signed policy bundles to agents. Runs on your infrastructure.

Ingest service

Receives signed records over gRPC, verifies the hash chain on arrival, and stores them. Rejects any record that does not chain to its predecessor. Runs on your infrastructure.

Database (PostgreSQL)	Stores metadata and opaque ciphertext. The content is encrypted under your organization's key, so the database holds no readable content, including to anyone with direct database access. Runs on your infrastructure.

What crosses each boundary

BOUNDARY	WHAT CROSSES
Device to provider	Only the content your policy permits, after any required redaction. Redaction and sensitive-data detection are best-effort and do not guarantee that every sensitive value is caught. A blocked request never leaves the device.
Device to your admin and ingest	Signed audit records: metadata plus content encrypted with your key. The content is opaque ciphertext to the server.
Admin to device	Ed25519-signed policy bundles and configuration. An agent enforces only a policy it can verify the signature on.
Anything to Verillian	Nothing. Verillian the company receives no traffic, no records, and no keys.

What Verillian never receives

- [ ] Your prompts, model responses, or tool calls
- [ ] Your audit record or its contents
- [ ] Your organization's encryption key or any device signing key
- [ ] Any data used to train models, because none is ever sent

Cryptographic properties

These are the properties a security team verifies for itself. None of them depend on trusting Verillian's word; each can be checked against the record and the code.

Content encryption	X25519 envelope encryption with your organization's key is shipped today. The server stores metadata and opaque ciphertext, and decryption happens in your authorized reviewers' browsers. Because the key is yours and never sent, even Verillian cannot read the content.
Tamper-evidence	Each record is SHA-256 hash-chained and Ed25519-signed on the originating device. Any modification, reordering, or deletion of an entry breaks the chain and is detectable on verification. The record is tamper-evident, not tamper-proof: the value is that a change cannot be made without leaving a mathematically detectable mark.
Non-repudiation	Each entry is signed by the originating device's key, which the server never holds, so a signature cannot be forged by the server or any other party. The origin of an entry cannot later be denied.
Policy integrity	Policies are Ed25519-signed by your admin service and verified by each agent before enforcement. An unsigned or revoked policy is rejected, and with no valid policy the agent fails closed.
Licensing	Source-available under the Elastic License 2.0, gated by Ed25519-signed, time-limited license keys verified locally on each start. Works air-gapped, with no outbound call and no phone-home.

How this maps to the controls you answer to

The same architecture is what the framework language asks for. The audit chain is the tamper-evident record behind the NIST 800-53 audit (AU) family, and it is aligned to CJIS Security Policy v6.0 logging with one-year retention; the customer-key encryption and redaction-before-egress support HIPAA data minimization and six-year documentation retention. Verillian is built for these frameworks and aligned to their requirements, not independently certified, and CJIS compliance is validated through your agency's state-agency audit. The control sits at the device, the evidence is the signed chain, and the mapping is what a reviewer checks.

Deployment

The admin, ingest, and database run on your infrastructure, cloud or on-premises, including air-gapped. The agent installs as a code-signed, notarized package on macOS, with Windows support, and deploys through your existing device management such as Jamf, Intune, or Group Policy. Enrollment is through your identity provider over OIDC. You operate the system; you hold your data, your backups, and your key.

> The decision happens at the device, before egress. Only permitted content continues to the provider, and every captured interaction is signed into a chain you hold under your own keys, verifiable independent of anything Verillian attests.

For an architecture review, we will walk the data path, the boundaries, and the cryptographic properties against your framework with your security team, under NDA.

hello@verillian.ai