Sentinel

The evidence gap

The AI governance market has produced a rich vocabulary for what organizations should do. It has produced almost no infrastructure for proving they actually did it.

This is the central failure of the current generation of AI compliance tooling, and it is about to become both commercially and legally consequential. The first wave of governance products gave organizations a language for risk: taxonomies, policy templates, maturity models, and committee charters. Those tools were appropriate to their moment. They told an organization what it intended to do. They did not give it the means to prove it had governed.

That distinction maps directly onto how a regulator thinks. A regulator examining an AI deployment does not primarily want a policy document, which states intent. It wants evidence of what actually happened: which AI actions were evaluated against which policy, when, under what human oversight, and with what result. The question is not "do you have a data-protection policy?" It is "Can you demonstrate that this agent never acted outside its authorized scope across the past twelve months?" Only the second question admits a factual answer grounded in production behavior, and most organizations cannot answer it.

The gap between policy and evidence is not a documentation gap. It is an infrastructure gap, and you cannot close it by writing better policies. You need a runtime: a system present at every AI action that applies governance at execution and produces a verifiable record of every decision. That is what has been missing. That is what Sentinel is.

Why now

The regulatory timeline has made this urgent. The EU AI Act's deployer obligations take effect on August 2, 2026. Article 9 requires a risk management system that is

Qaori.Tech OÜ, Sepapaya 6, 15551 Tallinn, Estonia, Register 17169588.
Qaori UK LIMITED, 66 Paul Street, London EC2A 4NE, UK, Register 16952497.
Qaori AI LLC, 1207 Delaware Ave #1634, Wilmington, DE 19806, United States, Register 61-2310301
Qaori AI Japan, Minato Mirai, Yokohama 220-0012, Japan. (Launch Q3 2026)
Qaori AI Inc., 1207 Delaware Ave #1634, Wilmington, DE 19806, United States, (Registration ongoing)
Qaori AI Tecnologia Brasil Ltda., Rio de Janeiro. (Launch Q4 2026)
Qaori.Coop is a service operated by Qaori Tech OÜ, Estonia, registered as Qaori DAO LLC registered at 108 West 13th Street, Wilmington, New Castle, Delaware, 19801, USA.

© 2026 Qaori.Tech OÜ. All rights reserved. Patents Pending - USPTO Provisional Application No. 64/084,387

continuous and iterative across the system's whole lifecycle. Article 14 requires human oversight capable of actually understanding and overriding the system. Article 26 places these obligations squarely on the organizations deploying high-risk AI. Parallel demands come from the NIST AI Risk Management Framework, ISO 42001, and a widening field of national regimes.

The load-bearing word in Article 9 is continuous. The regulation does not describe a periodic audit. It describes a live operational posture maintained across every model update and every new tool integration. The practical question follows directly: how does an organization produce continuous evidence of continuous monitoring? An annual assessment does not satisfy a continuous requirement. Neither is it a quarterly questionnaire nor a retrospective log review. Only a system that monitors continuously and is present at every action can.

Why the existing approaches fall short

Three approaches dominate the market today, and all three fail the evidence test in structurally similar ways.

Point-in-time assessments are accurate at the moment of assessment and begin aging the instant they end. AI systems are updated, agents are given new access, and configurations change. By the time the report reaches a board or a regulator, it documents a system that no longer exists in its described form.

Self-reported questionnaires measure intent, not behavior. The organization asserts compliance, and the questionnaire records the assertion, with no mechanism to verify it against production. The global trajectory of AI regulation is away from self-attestation and toward verifiable evidence, precisely because assertions are weak governance.

Retrospective log analysis is more sophisticated but still backward-looking. A log tells you what happened after it happened. An agent that exfiltrated sensitive data or executed an irreversible action has already caused the harm by the time analysis finds it. The value of a governance control is highest at the point of decision, before the action is executed. This is the difference between a flight recorder, which explains the crash, and a flight envelope, which prevents the aircraft from leaving safe parameters in the first place.

Each of these is a flight recorder. Sentinel is a flight envelope.

The principle: evidence generated where the action happens

Sentinel separates governance into two planes, and the separation is an architectural commitment rather than a deployment convenience.

The **Sentinel plane** runs inside the client's own infrastructure, their cloud account, their network, and their control. It is present at every AI action, applies the client's governance policy at the moment of decision, and produces a signed, tamper-evident record of every decision it makes. It enforces this by blocking prohibited actions and holding designated categories for human approval before those actions ever reach an external system.

The **Qaori control plane** distributes governance policy to the Sentinel plane and receives signed, batched evidence records back. It never receives the underlying AI actions, the data the agents operated on, or the substance of human approval decisions. Policy travels down. Signed evidence travels up. Raw data crosses the boundary in neither direction.

This matters for three compounding reasons. A governance runtime that observed the client's AI from outside the client's environment would first have to pull the client's sensitive data out of the client's environment to a third party before it could govern it, which is unacceptable for exactly the data categories that most need governing. An outside observer can detect and report, but cannot enforce, because enforcement requires being present at the point of execution. And evidence of your own compliance is more credible when it is generated inside your own environment by a runtime you control than when it is produced for you, on their infrastructure, by the vendor selling you the governance.

What Sentinel is, and is not

Sentinel is a client-side execution and control component. It is present in every AI action, applies policy at the moment of decision, and resolves each action to one of four outcomes: permit, block, escalate, or hold for human approval. Its primary output is a continuous stream of signed governance events, produced at the point of action, stored in the client's environment, and transmitted to the control plane on a schedule set by the client.

Sentinel is not a document generation tool. It does not produce policy templates or compliance reports as its main output.

Sentinel is not an auditing tool. It does not examine completed operations after the fact. It evaluates each action before execution and decides in real time.

Sentinel is not a substitute for governance design. It enforces whatever policy the client configures. It does not decide whether that policy is correct. The governance posture is the client's. The Sentinel's job is to enforce it continuously and to prove that it did.

The evidence belongs to the client

A platform that runs inside your environment and reports back to the vendor is, by definition, working for the vendor. In a governance runtime with visibility into every AI action, the question of whose interests the system serves is not a philosophical nicety. It is a design constraint, and Sentinel answers it at the architectural level rather than the contractual one.

The evidence Sentinel produces is written to a store inside the client's own infrastructure. It belongs to the client because it lives on the client's systems. The client can read it, export it, audit it, and present it to a regulator without asking Qaori for permission or involving Qaori's infrastructure.

What travels to the control plane is a different thing entirely: the signed, abstracted records of governance decisions, carrying the type of decision, the agent's identity, the class of action, the policy version, and a cryptographic signature. These records are not the AI actions themselves, not the data the agents handled, and not the content of what a human approver weighed. The governance record travels. The underlying operation does not.

This is the line that defines the product. The control plane can verify that a governance decision occurred, when, under which policy version, and with what outcome. It cannot see what the agent was substantively trying to do, what data it touched, or what the approver considered. Different clients' records are signed with different keys and carry no shared identifier, so the architecture has no path to correlate one client's activity against another's.

Architecture as commitment, not promise

A policy promise says: we will not do this. An architectural commitment says: we cannot do this without re-engineering the system. Sentinel's separation of client data from vendor visibility is the second kind.

The right question for a data protection officer to ask any vendor is not "do you promise not to use my data?" but "can you show me, structurally, how you are prevented from using it?" For Sentinel, the answer is concrete: the data does not leave your environment because the runtime has no path that routes it outbound, and the evidence is yours because it is signed inside your environment with a key that never leaves it. These are properties of the architecture, not clauses in a contract that a future revision could quietly soften.

The human approval gate follows the same discipline. When an action falls into a category the client has designated as requiring sign-off, an irreversible operation, a high-risk external call, or an action on regulated data, it is held, not executed, until a

registered human approves it. The approver sees what the action intends, which rule triggered the gate, and the agent's identity and scope. They can approve, reject, or defer. The decision itself becomes part of the evidence record. It is a genuine gate, not a post hoc notification.

How does it fit the wider platform

Sentinel delivers value as a standalone runtime, and many clients will deploy it that way. For clients on the wider Qaori platform, it is also the foundation on which the rest of the compliance picture is built. The evidence Sentinel generates is what makes the platform's regulatory mapping, its pre-execution proof to the Veraxis VEIP standard, and its governance scoring real rather than asserted, because each is grounded in evidence produced at the point of action rather than assembled afterward from documents. Sentinel is the source. The rest of the compliance program is downstream of it.

What changes for the buyer

For a **CISO or CTO**, the execution path is no longer a place where AI runs unsupervised at machine speed. Every consequential action is evaluated, stopped if necessary, and recorded with proof within your environment, without your data leaving it to obtain a verdict.

For a **general counsel or compliance officer**, the answer to a regulator is no longer an attestation and a stack of logs. It becomes a continuous, tamper-evident chain of evidence that shows governance was in place before each action, mapped to the specific obligations of the EU AI Act and the regimes you operate under. The continuous monitoring requirement is met by a system that continuously monitors.

For **procurement and the board**, the deployment model is favorable on its own terms: the runtime lives in your infrastructure, the evidence is yours, and the architecture, not a contract, is what protects your data. The reception from industry leaders working on this problem has been enthusiastic for a practical reason. The approach is elegant, it deploys quickly, and it is markedly more cost-effective than assembling a heavyweight governance program from assessments, questionnaires, and after-the-fact reviews that still cannot answer the regulator's core question.

Conclusion

After August 2, 2026, organizations deploying high-risk AI will be asked a question their current tooling cannot answer: Can you demonstrate, with signed and tamper-evident records, that your AI systems operated within policy at every point in

production? A policy document does not answer it. A questionnaire does not answer it. A log does not answer it.

Sentinel answers it by being present at every action, enforcing before execution, and producing the proof inside the environment where the action took place. It is the difference between intending to govern and being able to show that you did.