SOLUTION BRIEF

vijil × modulos

End-to-End AI Governance and Enforcement

Translate AI governance into policy-driven guardrails, with verifiable evidence and continuous monitoring for the compliance lifecycle.

Modulos and Vijil Better Together

Goal: Close the loop between organizational governance policies and AI system evaluation, risk reporting, and monitoring for compliant and trusted AI agents.

The Challenge

- Governance teams have the tools to define policies and AI system property compliance requirements, but not to assess AI agents for technical compliance certification
- Manual validation of agents is time-consuming, does not scale, and doesn't replicate production conditions.
- Manual reviews don't yield quantifiable metrics for **proof of accuracy, robustness, and security at a system level**
- **Approaching** EU AI Act deadlines for day-1 compliance and continuous monitoring based on the nature and risks of the application

Impact: Organizations struggle to operationalize trustworthy and compliant AI at enterprise scale.

The Partnership at a Glance

The Modulos and Vijil partnership unifies the AI governance layer with AI agent policy evaluation, enforcement at runtime, and ongoing monitoring.

With the main EU AI Act applications due to go into effect on Aug 2, 2026, the partnership enables covered entities to operationalize at scale obligations for continuous monitoring based on risk.

Governance teams use Modulos to define and specify governance policies, controls, and risk tolerance for AI projects. Security and engineering teams can automatically test and evaluate agents based on these policies and risk tolerance thresholds before deployment, answering the question "is this safe with acceptable risk?" with a quantifiable response and audit-ready documentation through Vijil integration.

Once in production, Vijil's guardrails monitor AI agent behavior and log policy violations and detections, addressing requirements to actively collect and review data on system performance.

Key Benefits

For Security & Risk Leaders

- Closedloop process betweenpolicy definition and policy enforcement
- Quantitative, system-level assessment of agent compliance and resilience
- Reduced manual review burden

For Compliance & Audit Teams

- Automated evidence collection through system-level evaluation
- Framework mapping (e.g., NIST, ISO, EU AI Act) for model and agent evaluation
- System level reporting to address continuous monitoring requirements

For Engineering & ML Teams

- Reliable modelassessment/evaluation
- Built-in guardrails specific to compliance and model properties
- Reduced friction to production

Reference Architecture

Modulos

- Define and document objectives, high-level requirements
- Identify stakeholders
- Establish policies, risk threshold and relevant regulations
- Specify evidence sources

Vijil Diamond

- Ingestpolicies, risk contexts, & AI system categories
- Translate governance parameters into specific, executable test cases
- Feed metrics, logs, & test results back as "objective evidence"

Vijil Dome

- Generate custom guardrail configurations tuned to agent & risks
- Protect AI agent by filtering requests and responses based on evaluation policies
- Capture the full picture of agent behavior in production with built-in telemetry
- Ongoing monitoring for compliance requirements

Typical Flow

1. AI models onboarded into Modulos
2. Policies and controls defined
3. Evidence and telemetry fed into Modulos
4. Automated evaluation against frameworks and custom policies by Vijil
5. Generation of Vijil Trust Score and risk quantification. Mitigation guidance automatically generated
6. Automated reporting and audit readiness

Use Cases

- AI model risk management and certification
- Regulatory readiness (EU AI Act, NIST AI RMF)
- Holistic operationalization of EU AI Act obligations

Contact

Kevin Shinners, Modulos

Tim G. Rudner, Vijil

vijil × modulos