vijil dome

Immune System for AI Agents

Why Vijil

Vijil helps you ship AI agents to enterprises 4x faster while lowering operational risks. Vijil provides a layer of trust between agent development frameworks and agent runtime platforms to make AI agents more reliable, secure, and safe for enterprises from development to operations, and back.

Why You

You balance innovation and risk: You're the senior leader at a fast-moving company who wants to use AI agents in production quickly. You want to see evidence of reliability, security, and safety throughout development and operations.

You use special agents: Your AI team is building or buying custom domain-specific agents using (open or closed) LLMs with access to confidential data and restricted tools.

What Dome Does

- Enforces custom policies on the behavior of AI agents
- Monitors agents at runtime for adherence to policy
- Continuously adapts to new attacks and vulnerabilities

Why Dome is Better

- **Comprehensive**: Covers adversarial inputs, prompt injections, jailbreaks, PII leakage, toxicity, stereotypes, and violations of policies, standards, and regulations
- **Customizable**: Automatically generates guardrails based on evaluation results customized to agent scope, user personas, and org policies
- **Accurate**: Leading prompt-injection detector model
- **Transparent**: Observability compliant with OTEL
- **Fast**: 17 ms on Groq; less than 100 ms on GPUs

Deploy Dome in Minutes

- Load open source Python library into an agent built with any leading framework (`pip install vijil-dome`)
- Run containerized application on Kubernetes anywhere
- On marketplaces at AWS, DigitalOcean, Google Cloud

> Google is pleased to collaborate with Vijil on tools to help enterprises customize Gemma and other open models for trust as well as task.
> - Manvinder Singh, Director of Product, **Google**