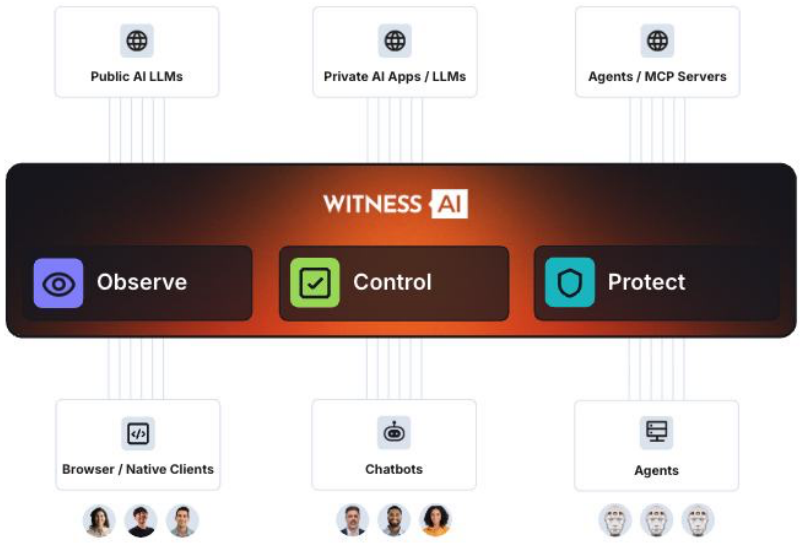
Unified AI Security and Governance Platform

The Confidence Layer for Enterprise AI

Secure AI Interactions Without Hindering Innovation

WitnessAI is the unified AI security and governance platform enterprises trust to protect AI interactions across global deployments. Our platform provides complete network-level visibility across thousands of AI applications including shadow AI, while extending governance to AI agents connecting to external tools and MCP servers.

Unlike basic keyword blocking, we understand employee and agent intent to build context-aware policies that enable productivity while stopping real data leakage and preventing emerging AI threats. Intelligent AI routing delivers effective control by optimizing requests based on risk, cost, or purpose. With single-tenant architecture and customer-controlled encryption, we transform AI security from a bottleneck into a competitive advantage.



Proven Results

4000+ AI applications catalog

350,000+ employees secured

40+ operating across 40+ countries

Millions of daily ai interactions monitored and secured

99.3% true positive model protection guardrails

Unified AI Security and Governance

Observe

Find every AI app, agent, and conversation in your environment

- Catalog thousands of AI applications, including shadow AI
- Capture actual conversations to understand what information is being shared
- Detect agentic activity across AI apps like Claude Desktop, VSCode, and local agents
- Identify public and private MCP servers, categorized by function

Protect

Shield models, apps, chatbots and agents from manipulation

- Block prompt injections and jailbreak attempts before they reach your models or agents
- Filter harmful or off-brand responses before they reach users
- Tokenize sensitive data in real-time while maintaining usability
- Apply bidirectional runtime protection to secure both prompts and responses

Control

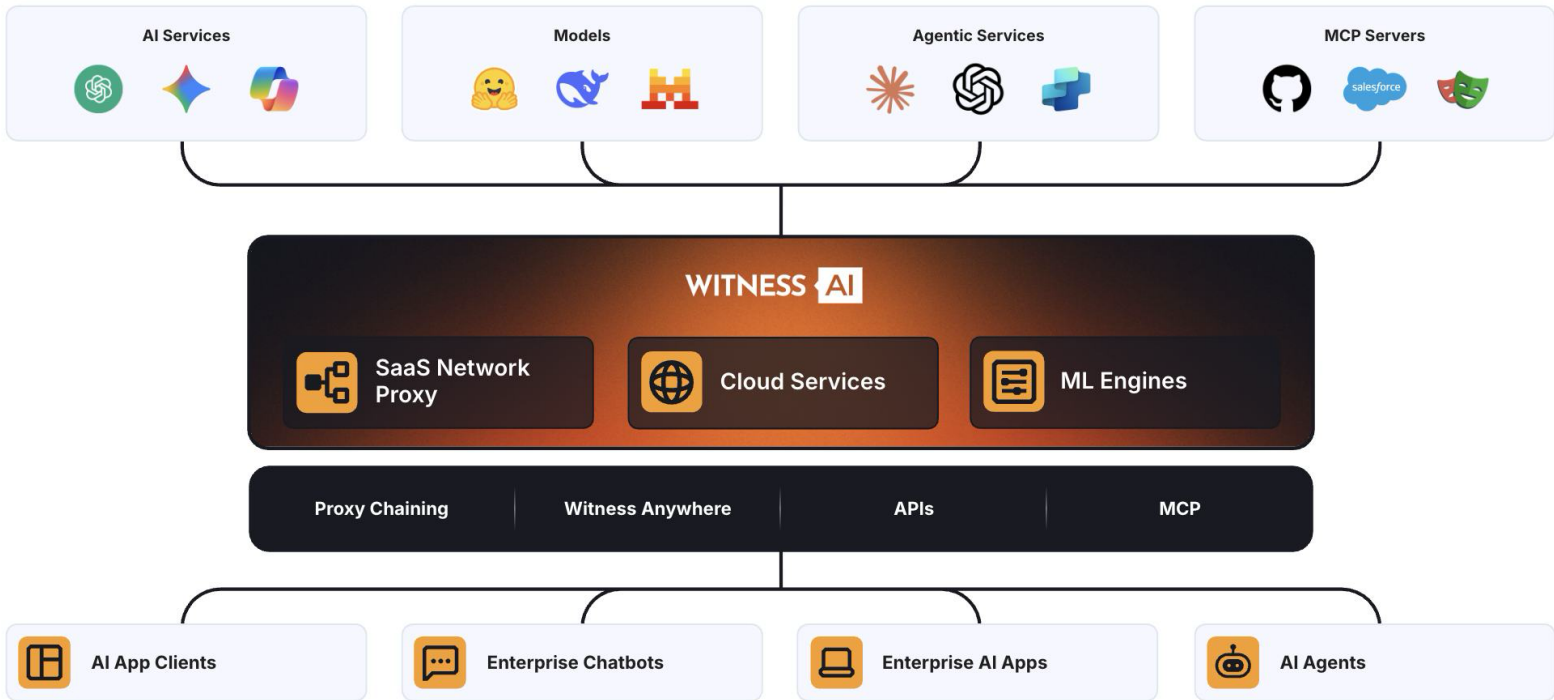
Enforce policies based on user intent and context

- Build policies and controls based on user intention and not hardcoded regex or keywords
- Generate comprehensive audit trails for compliance across human and agent activity
- Route requests intelligently to appropriate models based on risk, cost and purpose
- Attribute every agent to a human identity for accountability

Financial Services

As a company, we knew we needed a way to maintain security and compliance while encouraging our teams to leverage modern approaches with GenAI . We chose WitnessAI because they help us achieve just that with our diverse portfolio, safely.

CISO
InComm Payments



Key Platform Capabilities

Multi-Generational AI Security Platform: WitnessAI has evolved with enterprise AI: from securing human employee interactions, to protecting AI models and applications, to governing autonomous agents. One platform for comprehensive observability, protection, and control across every employee, agent, model, and application.

Network-Level Operations: See 100% of AI activity including thick clients like Windows 11 Copilot and Office 365 without browser extensions or agents. Capture tool calls from autonomous agents and MCP server connections across your entire network.

Intention Based Policy: Enable sophisticated data and IP protection through context-aware policies based on user intentions. Our models analyze conversations and context rather than isolated fields, detecting patterns and risks that static keyword rules cannot adapt to.

Unified Workforce Governance: Govern AI activity of human employees and AI agents from a single console. Attribute every agent action back to its human origin for complete accountability.

Runtime AI Defense: Protect models, applications, and agents with bidirectional inspection that blocks threats like prompt injection before they reach your systems and filters harmful outputs before they reach users. Tokenize sensitive data automatically.

Enterprise-Grade Architecture: Deploy single-tenant infrastructure with complete data sovereignty and customer-controlled encryption. Support executive privacy modes and multi-region deployment to address regulatory requirements across your global enterprise.

Intelligent AI Routing: Route prompts to LLMs based on risk, cost, or purpose. Direct sensitive queries to secure internal models while steering low-risk tasks to cost-effective options, enabling productive and secure workflows.

Enterprise Solutions

Govern Your Human and Digital Workforce: WitnessAI provides unified governance and policy control across human employees and AI agents, with network-level visibility into every AI tool, MCP server, and agent connection in your environment.

Runtime Security for Models, Apps and Agents: WitnessAI delivers multi-layered, bidirectional defense for chatbots, copilots, and autonomous agents against threats traditional security cannot see.

Deployment Options

WitnessAI seamlessly integrates with your existing infrastructure:

Network Connectors: Zscaler, Palo Alto Networks, Netskope

Witness Anywhere: Extend protection to remote workers without endpoint clients

API Integration: Connect to existing analytics and security systems

Global Airline

The ability to see every AI interaction across our global workforce has transformed our security posture. WitnessAI helps us maintain compliance while enabling our teams to leverage AI for competitive advantage.

VP of Cybersecurity
Anonymous (Top 5 Airline)

Secure Your AI Adoption

Book a Demo

About WitnessAI

WitnessAI is the unified AI security and governance platform enterprises trust to govern and protect all AI activity. We provide complete, network-level visibility into every interaction including employees and autonomous agents, even in native apps where legacy tools are blind. WitnessAI transforms security from a bottleneck into the enabler of your AI strategy as the confidence layer for enterprise AI.

Learn more at witness.ai.