Zscaler AI Guard

AT-A-GLANCE

Secure production AI with guardrails that protect against malicious behavior and sensitive data leakage

Why AI Guard?

Enterprises are increasingly building AI assistants and agents to enhance customer experience, drive employee productivity, and streamline business processes. However, deploying AI in production comes with risk, even when extensive red teaming is used to identify vulnerabilities. Once live, AI apps become targets for manipulation. Attackers use prompt injection, jailbreaks, and supply-chain threats to trigger data leakage, harmful outputs, or undesired behaviors. Employees using commercial GenAI apps also pose data leakage risk.

Zscaler AI Guard helps you securely operate AI in production by detecting and blocking threats between users, AI apps, and LLMs, including leading commercial GenAI tools. It enforces guardrails on AI behavior, ensures alignment with internal and external policies, and provides the visibility and control needed to manage risk. AI Guard also monitors and governs usage of apps like ChatGPT, Gemini, and Microsoft 365 Copilot to keep AI secure, trustworthy, and business-aligned.

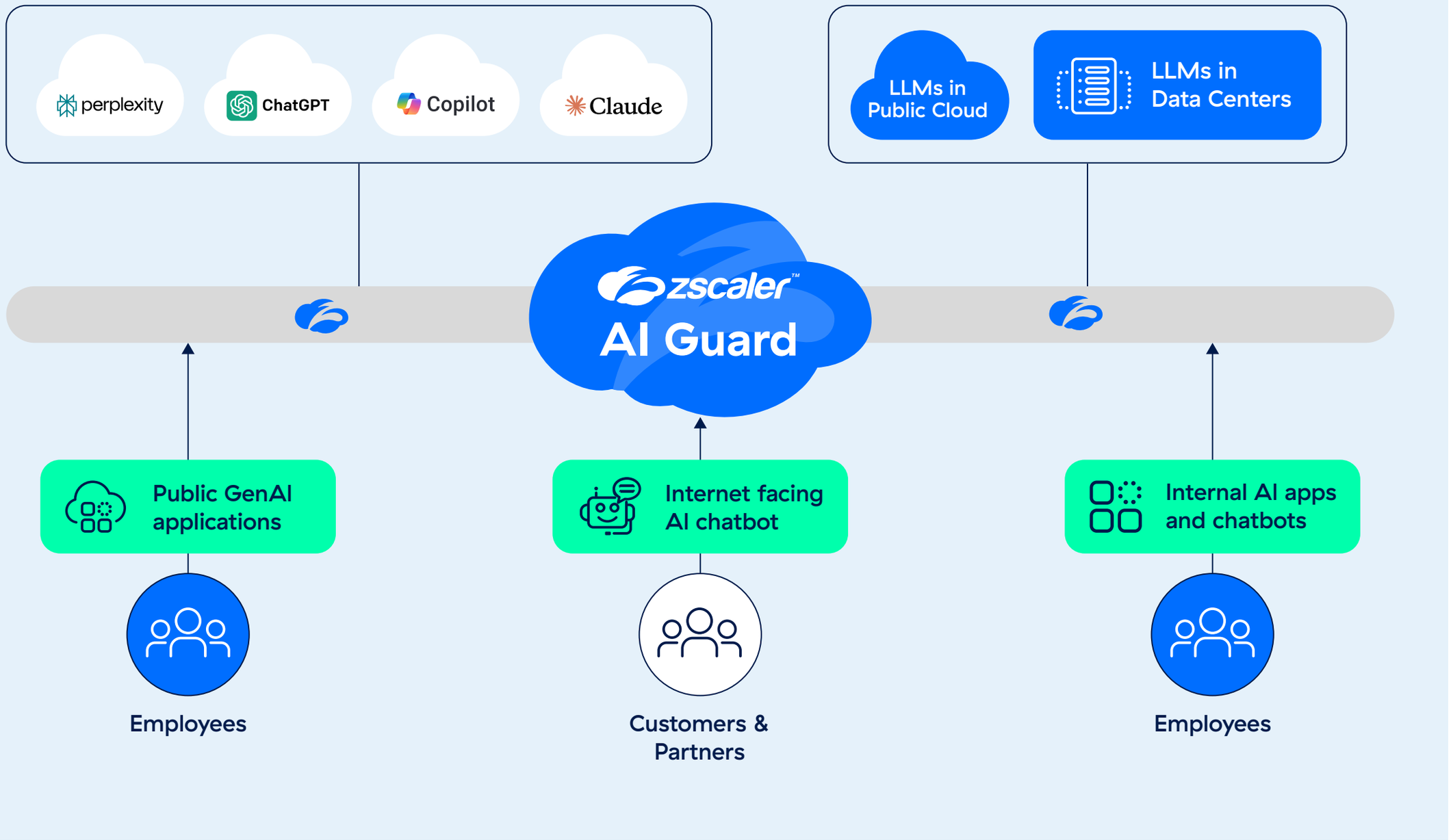
Higher accuracy and lower false positives: AI-powered engines help block toxic or malicious inputs without overfiltering safe content. Severity settings help you set the right priorities.

Broad detection coverage: More than 18 detectors that can be applied on prompts as well as responses to uncover a wide range of attack patterns and exposure points. Support for popular apps such as ChatGPT and Microsoft Copilot.

High fidelity identification for sensitive content: Powerful safeguards that identify and prevent leakage of sensitive data – both for users accessing popular GenAI apps as well as for enterprise AI apps talking to LLMs.

Security for the complete AI lifecycle: Zscaler AI Red Teaming and AI Guard work seamlessly to help identify vulnerabilities and secure AI apps from development through production and operation.

Leverage with cloud-based AI platforms and frameworks: Easily connect to LLMs and use across popular cloud-based AI platforms and frameworks including AWS, Azure, Google Cloud, Langchain and Palantir Webhooks.

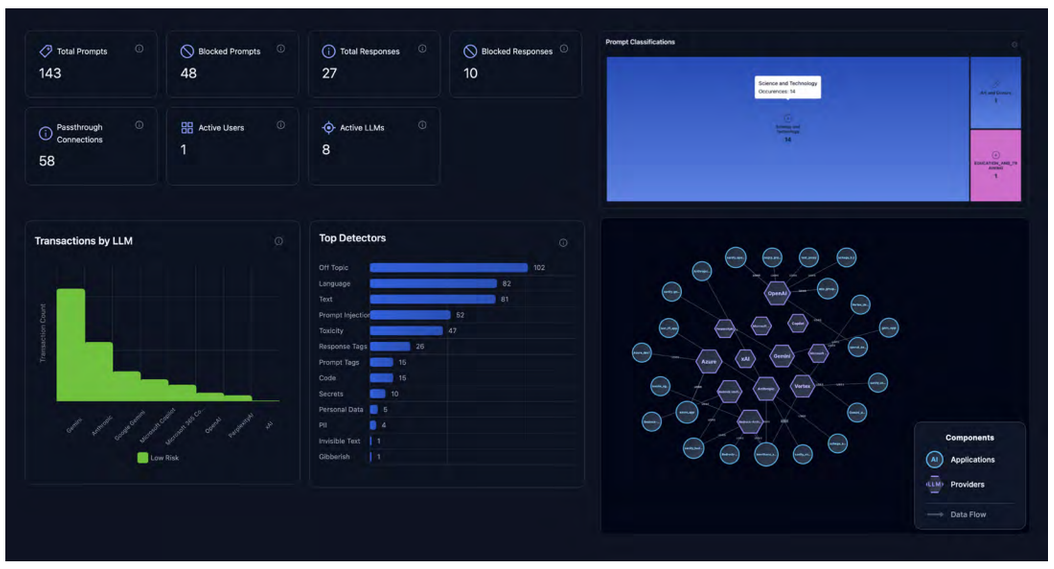


- Reduce Legal & Compliance Risks
- Protect your Brand Reputation
- Automated Incident Response
- Respond to Threats Proactively

Lightning-Fast Guardrails For AI apps

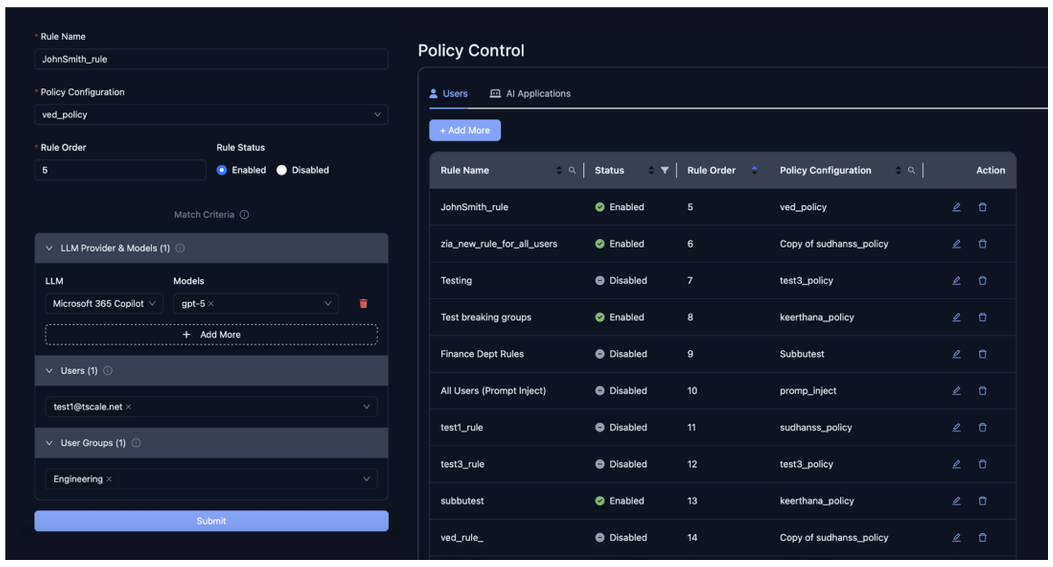
GET ONGOING VISIBILITY AND INSIGHTS

Enable oversight of AI apps – both commercial as well as enterprise, the LLMs they access, their usage, and prompts and responses to discover shadow AI and enforce business alignment.



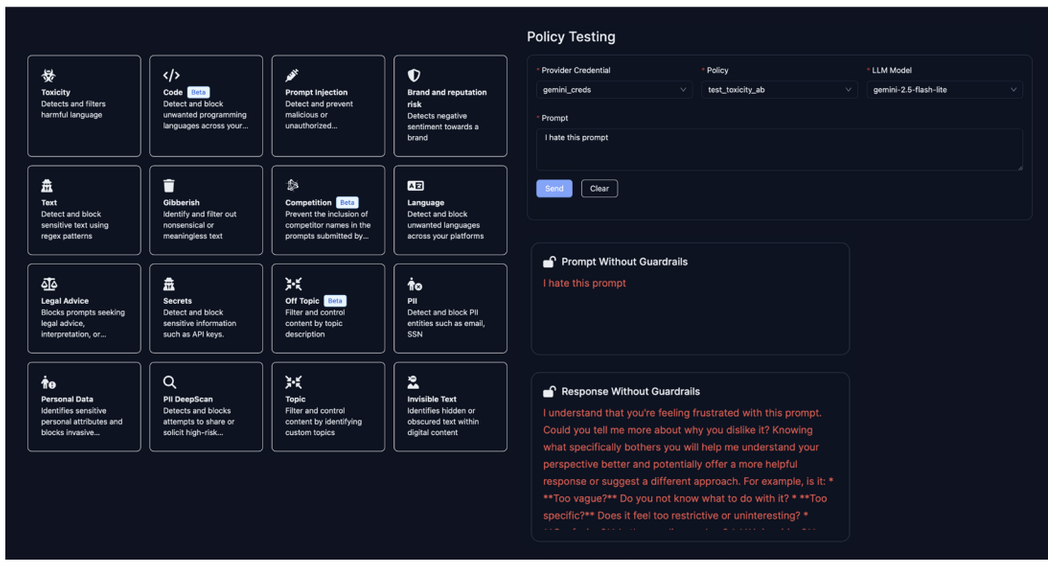
GOVERN THE USE OF POPULAR AI APPS

Control who and how popular AI apps such as ChatGPT, Gemini and Microsoft Copilot are accessed and used. Create and apply policies to specific users or groups of users to ensure compliance and prevent data leakage.



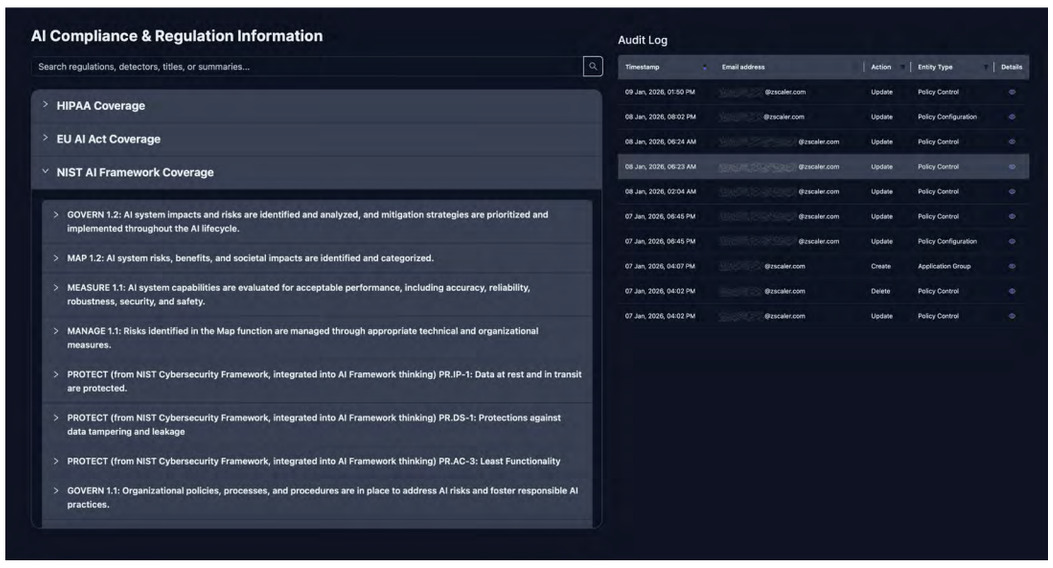
SECURE AI APPS IN PRODUCTION

Protect your enterprise AI apps against adversarial attacks such as prompt injections, embedded code, off-topic discussions, toxicity, and malicious outputs to ensure business alignment and reduce risk. Integrate seamlessly with popular cloud-based AI platforms, SDKs and AI agent frameworks including AWS, Azure, Google Cloud, Langchain Agents, Palantir Webhooks, AWS Boto3, and Google Vertex SDK.



AUDIT AI USAGE, ENSURE PRIVACY, AND ENABLE COMPLIANCE

Maintain logs of users, apps, prompts, responses, policies, and actions to enable compliance with latest standards including the NIST AI Framework and the EU AI Act. Zscaler does not store prompts or responses. All customer data is stored in an Amazon S3 bucket with a customer key to secure access.



Deploy AI Apps with Full Confidence

Protect your AI apps in production by identifying threats and enforcing guardrails

BOOK A DEMO

About Zscaler

Zscaler (NASDAQ: ZS) accelerates digital transformation so customers can be more agile, efficient, resilient, and secure. The Zscaler Zero Trust Exchange™ platform protects thousands of customers from cyberattacks and data loss by securely connecting users, devices, and applications in any location. Distributed across more than 150 data centers globally, the SSE-based Zero Trust Exchange™ is the world's largest in-line cloud security platform. Learn more at **zscaler.com** or follow us on Twitter **@zscaler**.

+1 408.533.0288 **Zscaler, Inc. (HQ) • 120 Holger Way • San Jose, CA 95134** **zscaler.com**

zscaler™

Zero Trust Everywhere