Zero Trust,
Govern, Innovate

Zscaler IT's Formula for Securing AI

WHITEPAPER

Introduction

In today's rapidly evolving digital landscape, Artificial Intelligence (AI) represents a double-edged sword, offering transformative opportunities while introducing unprecedented complexities in security, governance, and risk management. As enterprises embrace AI to drive innovation, agility, and efficiency, they are confronted with a pivotal challenge: How can organizations secure AI to protect sensitive data, ensure compliance, and build trust—both internally and externally?

At Zscaler, this challenge lies at the heart of our IT strategy. As a global leader in cloud security, we have taken a proactive approach to integrating AI into enterprise operations while maintaining strict diligence and governance measures. This commitment reflects not only our leadership in the cybersecurity industry, but also our acknowledgment of the critical importance of securing AI for enterprises across industries. In this whitepaper, we share **Zscaler IT's approach to securing AI adoption and discuss the foundational principles enterprises must adopt to balance the power of AI with the necessity of mitigating risks.**

Facing the Complex Realities of AI

The promise of AI is inherently accompanied by risks that enterprises must address to ensure success. From the **threat of data exposure due to model compromise, to the dangers of bias, misuse, or unregulated outputs, businesses are stepping into a minefield of security, compliance, and governance concerns.** As organizations turn to AI to transform their operations, they must manage these risks without compromising innovation. Achieving this requires proactive strategies and robust frameworks for AI security.

Without comprehensive safeguards, companies face significant risks, including the compromise of intellectual property, the erosion of customer trust, and potential regulatory violations. Enterprises must not only block unauthorized AI usage but also ensure rigorous governance over models and outputs to maintain trust and reliability. As a leader in security innovation, **Zscaler IT has embraced this responsibility, creating a comprehensive AI Security and Governance Framework built on Zero Trust principles and adhering to global standards such as the NIST AI Risk Management Framework.** Our layered approach addresses the unique challenges AI introduces, ensuring sensitive data remains secure and organizational resilience is preserved.

The stakes are high, and mitigating risk in the age of AI demands adopting solutions that enable trust while safeguarding operations. It is only through diligent governance, thoughtfully designed frameworks, and cutting-edge technology that enterprises can confidently navigate the future of AI adoption. At Zscaler, the rigorous, governance-based approach to securing AI is a foundational, mandatory standard for all operations, not merely an optional priority.

AI Governance Framework and Process in Zscaler

Securing AI starts with governance and process, our AI governance is guided by core principles and robust policy documents, including the Generative Artificial Intelligence Use Policy and internal due diligence standards. The fundamental goal of the AI security framework is to safeguard data (both customer and corporate) from issues like model leakage, hallucinations, and misuse.

Core AI Principles and Policy Adoption

Zscaler's approach to AI is driven by four core principles: Accountability, Fairness, Transparency, and Security and Compliance.

Our Generative AI Security Standards document is grounded in the NIST AI Risk Management Framework and industry best practices. Key requirements for all employees and contingent workers include:

- **Restricting certain Public GenAI:** Unsanctioned or High-Risk Nation States AI models and LLMs are blocked. Limited access can be granted only through a Security exception for specific business use cases, such as research and development or innovation.

- **Secured Public AI:** All use of public AI tools must be routed through approved security controls, including DLP policies and AI Guard, to ensure data protection and governance. These measures are designed to establish clear guardrails and monitor public AI interactions in accordance with Zscaler's security requirements.

- **Data Restrictions:** Users must not enter personally identifying information (PII) or material nonpublic information (e.g., acquisition targets, unannounced product roadmaps) into GenAI tools. Users are also prohibited from entering confidential information (e.g., security information/reports, technology and technical information, customer data).

- **Human Review Mandate:** Outputs from GenAI tools are subject to human review. GenAI tools cannot be used to make automated decisions without human intervention or review.

- **HR Tool Prohibitions:** AI/ML tools are explicitly restricted from being used for employment-related purposes, such as hiring evaluations, hiring decisions, or determining terms and conditions of employment, without prior review and approval from P&C and Employment Legal.

Approved Internal AI Platforms

Zscaler maintains a list of Approved GenAI Tools, Platforms & LLMs. The use of AI models, tools, or services associated with sanctioned or high-risk nation-states, including Chinese LLMs like DeepSeek and Qwen, is explicitly prohibited for Product, R&D, or Corporate purposes.

Scope	Approved Platform/Tool	Security & Risk Mitigation
General Employee Usage	**ZChat** (Company approved and hosted on Zscaler dedicated tenant). **ChatGPT** and **Gemini** (Approved Public tools)	Requires ZPA-based posture checks or ZIA SSL inspection, AI Guard and mandatory opt-out of third-party GenAI platform abuse monitoring.
Engineering/ Product Dev.	Microsoft AzureAI, Google Vertex AI, Amazon AWS AI Services, Z-Llama (internal LLM Gateway).	Requires ZIA DLP, Customer Managed Encryption Key (CMEK) for SaaS, ZPA connectivity to "Crown-Jewels" systems (like BitBucket), and running SAST/SCA security tools on generated code.
Software Development	Approved tools can generate suggestions/snippets for code in customer-facing products, or generate code for internal testing or business applications (e.g., payroll automation script).	All generated code is subject to established Zscaler software development lifecycle standards, including human peer review, unit tests, and security vulnerability assessment.

Vendor AI Procurement and Due Diligence Process:

Prior to procuring any vendor AI product or service, we mandated a rigorous internal approval process, involving security, privacy, and legal teams. This is designed to manage the risk associated with external AI usage.

The process begins with the business owner ensuring a signed NDA is in place with the vendor and submitting a completed request form.

The internal security review requires comprehensive documentation from the vendor:

- **Vendor Security Questionnaire (VSQ):** Must be completed in its entirety.
- **SOC 2 Type 2 Report:** A valid report or an in-progress bridge letter is required.
- **Penetration Test Results:** Required within the last two years, demonstrating no outstanding Critical / High vulnerabilities in production, or providing a clear guidance/timeline for mitigation.

Data Security and Model Hosting Standards:

Zscaler evaluates the AI architecture and data flow, using the Five Safes Framework for AI Risk Evaluation and requiring secure data management practices (anonymization, masking, tokenization) for sensitive data.

Acceptance criteria depend on where the AI model is hosted and the type of data processed (Public, Restricted, or Confidential):

- **4th Party Hosted Models (e.g., OpenAI API):** Acceptance for Public or Restricted data highly prefers zero data retention and the vendor opting out of abuse monitoring processes. Confidential data requires the vendor to maintain user and tenant isolation, the 4th party service to operate on a zero data retention basis, and assurance that 4th party provider employees never access Zscaler Confidential data.
- **Model Training Restrictions:** No customer data is ever used to train any internal or external AI. See our post on “Zscaler’s Commitment to Responsible AI” on our external website.
  Zscaler mandates that Zscaler internal data cannot be used to train a vendor’s general model. If Restricted or Confidential internal data is used for training, it must either be heavily de-identified/ aggregated, or the model must be trained and used solely for Zscaler. Crucially, a vendor’s 4th party providers are not allowed to train their models with any Zscaler Data.

Legal and Contractual Safeguards:

The Procurement Legal team embeds language into vendor contracts to severely restrict the use of Zscaler Data. Standard language explicitly prohibits the vendor from accessing, collecting, or using Zscaler Data to:

- **Train, provide, or improve** the vendor's or any third party's AI or ML models.
- **Generate output or provide services** of any kind for itself, its affiliates, or any third party.

If a vendor requires Zscaler Data for AI/ML training, Legal must seek an **exception.** If approved, specific protective language is drafted to limit Zscaler's risk, including restricting use to a **Zscaler-only instance** and ensuring Zscaler owns all intellectual property rights in the Zscaler Data and any generated output.

AI Security Posture: Zscaler Products in Action

After completing all the governance, framework establishment, and vendor onboarding processes, securing AI is the next critical step—achieved through the use of Zscaler's own products. Zscaler's tools and strategies for securing AI evolution revolve around protecting sensitive data, ensuring privacy, and preventing misuse through advanced contextual policies and a modern Zero Trust Architecture.

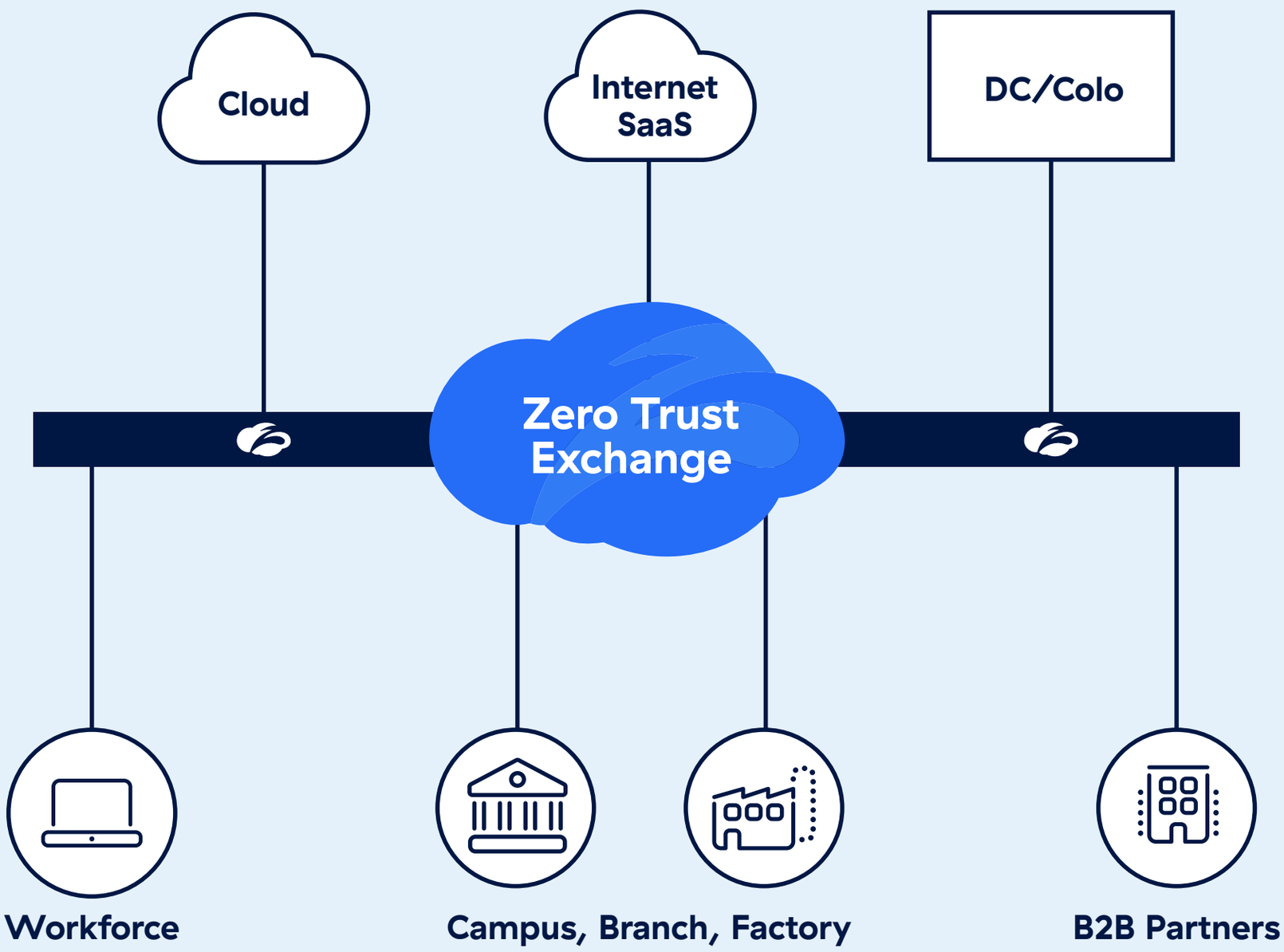

FIGURE 1: Zscaler architecture for networking and security removes complexity and risk

Zero Trust Foundation and Network Security

Zscaler's security posture is built on Zero Trust principles, ensuring connectivity is secure and access is tightly controlled.

- **Zscaler Internet Access (ZIA) for Internet Security:** All users internet traffic is routed to ZIA for policy decisions before access is permitted. Policies include: SSL inspected, Advanced Threat detection, Sandboxing, and DLP policies. Cloud environments are deployed with Cloud Connectors to secure the cloud workloads egress to the internet.

- **Zscaler Private Access (ZPA) for Internal Applications:** All internal applications travel over ZPA, requiring posture checks before access is granted. Cloud environments have deployed app connectors providing seamless secure access to privately hosted cloud applications.

- **Zscaler AI Guard** runtime protection for enterprise-managed AI applications, ensuring both security and compliance in AI operations.

- **Zero Trust Branch (ZTB):** All Zscaler offices have been migrated to a cafe model, eliminating costly firewalls and routing equipment. Using the ZTB appliances in the offices it combines network edge security, SD-WAN, and microsegmentation. Explicitly separating all users, visitors and IOT into individual networks of 1. Eliminating any lateral movement in the branch locations.

Securing Public AI Usage

Public AI platforms (e.g. ChatGPT, Perplexity, CoPiliot, Grok, etc.) are useful to accelerate workflows, solve complex problems, and process data. However, their use can inadvertently expose sensitive data, create pathways for malware, and exacerbate risks related to unsafe usage scenarios. Zscaler addresses these challenges by deploying layered defenses.

Generative AI:

Types of risks

- Access to risky AI apps
- AI misuse – sensitive topics, injection, prompt, harmful prompts
- Confidential data/PII leakage

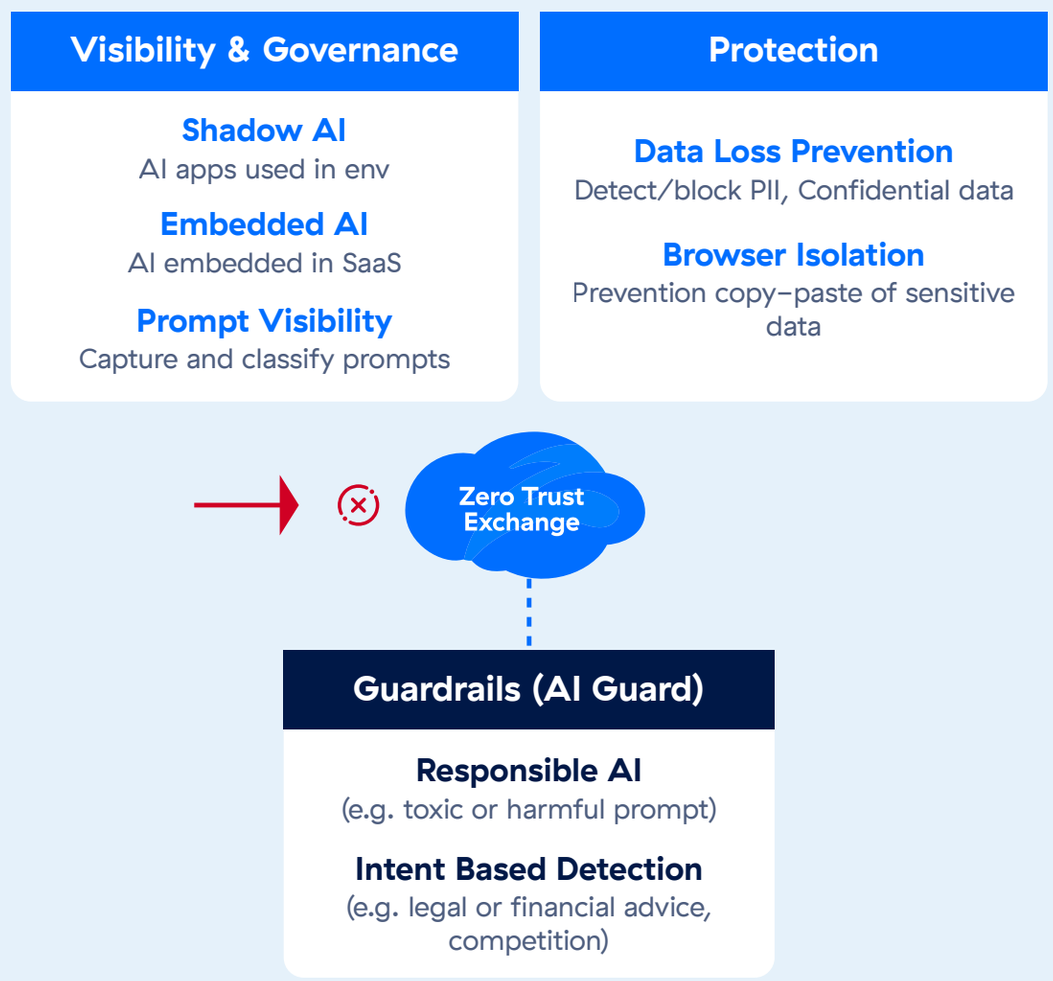

FIGURE 2: Zscaler secures usage of public and private generative AI

- **Blocking Unwanted AI URLs and Shadow AI Cloud Apps**
  Using ZIA's URL filtering and Cloud Application policies are set up to block unauthorized access to certain AI platforms and cloud applications used for generative AI.

- **Isolate AI/ML URL Categories to Restrict Data Sharing**
  Leveraging the AI/ML URL categorization also part of ZIA, we allow access to general AI/ML but in an isolated browser session. All traffic to the category that was not explicitly blocked by the above policy is sent to Zscaler's Zero Trust Browser. Restricting end users from pasting sensitive or confidential data can mitigate potential breaches or AI misuse, but not impeding innovation.

- **Inline DLP Inspection**
  Monitoring all internet traffic including AI interactions in real time to prevent sensitive data from entering prompt workflows or being exfiltrated.

- **Gen AI Security**
  All allowed AI traffic is recorded and analyzed by the GenAI Security product. This report is actively running in ZIA with prompt capture enabled. It provides Corporate Security and the CISO with detailed reports on:

  - Which GenAI tools are being used (allowed or blocked).
  - Who is using these tools.
  - Categorizing prompts

- **AI Guard for Public AI (LLM Guardrails)**
  Prompt and Response Guardrails with AI Guard, adding barriers to scrutinize both prompts sent to allowed public AI platforms and their responses. This includes:

  - Blocking Personally Identifiable Information (PII) from being sent to the LLM.
  - Preventing toxic or inappropriate queries and responses.
  - Blocking prompt injection and jailbreaking attempts to bypass controls, which are continuously updated as new AI-specific attack vectors emerge.

These guardrails ensure that users interact responsibly with AI tools without exposing sensitive corporate data. Also limiting what is returned from the LLM to protect the user and company.

Establishing a Secure Private GenAI platform

Zscaler IT created a private generative AI platform for all employees to securely embrace AI. ZChat's security model is built on Zscaler's zero trust principles, leveraging all the AI and ZPA features of the Zscaler platform combined with Okta for identity and RBAC. ZChat's specific details:

- **Internal Hosting:** Hosting in Zscaler AWS: ZChat is hosted within Amazon Web Services (AWS), utilizing Zscaler's cloud-native architecture. This allows for scalable, resilient deployment and leverages AWS's inherent security features.

- **User Authentication:**
  ZChat is integrated with Okta for authentication. This ensures that only authorized users—provisioned and managed through Okta—can access ZChat.

- **Role-Based Access Controls (RBAC):**
  Access to ZChat is further restricted using role-based access controls in the application. This ensures that users only have access to the features and data appropriate for their role.

- **Communication with Zscaler's OpenAI on Azure:**
  ZChat communicates securely with Zscaler's OpenAI services hosted on Microsoft Azure

- **Access Control via ZPA:**
  Zscaler Private Access provides secure access to private AI applications by segmenting user access using software-defined perimeters. This prevents lateral movement, reduces attack surfaces, and ensures that only authenticated, authorized users can interact with data-intensive AI systems.

- **Device Posture:**
  Mandatory compliance and full management of all devices are required before establishing a connection to private, ZPA-enabled applications.

- **AI Guard for Private AI:**
  Similar to public AI, Zscaler equips Private AI tools with Prompt and Response Guardrails that block activities like PII sharing, inappropriate usage, or exploitation through malicious instructions (e.g., jailbreak attempts). By embedding these protections into private apps, we can enforce strict AI use policies while leveraging AI's full power.

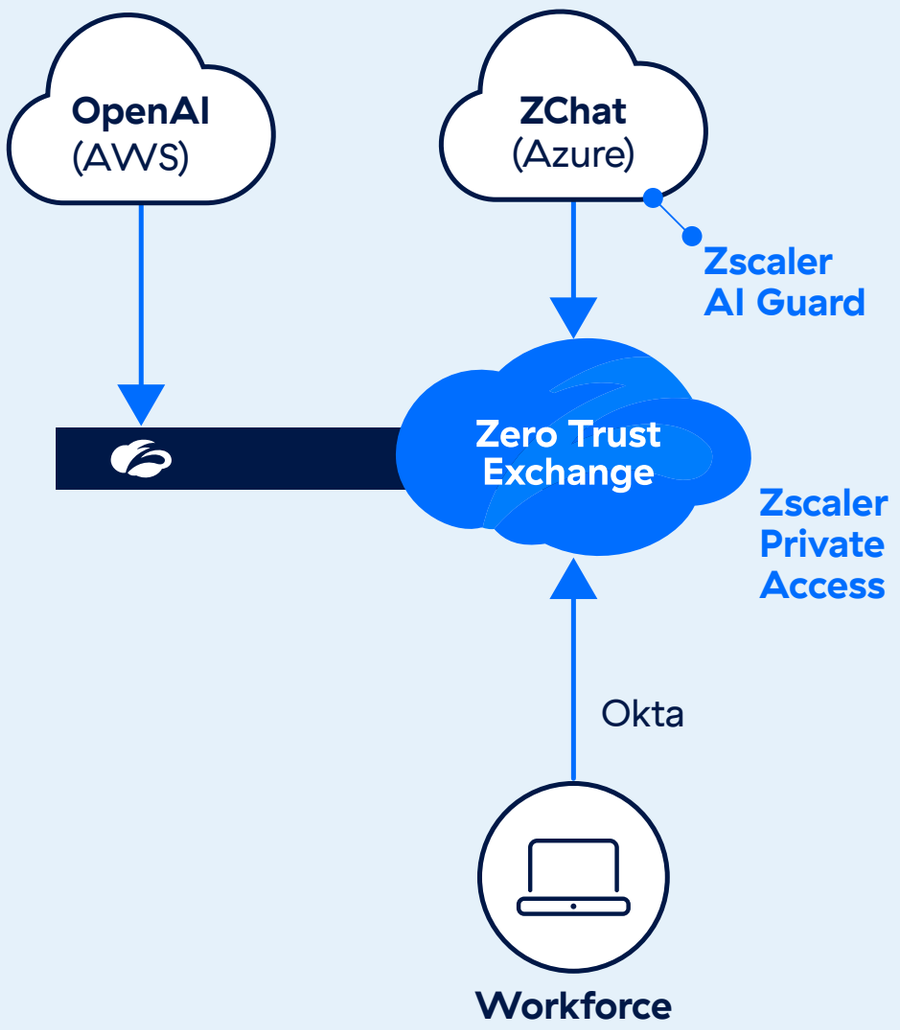

FIGURE 3: Zscaler Private Access (ZPA) provides secure access for employees to the ZChat private application (running in AWS). ZPA also secures the backend communication between ZChat and Zscaler's Open AI services hosted in Microsoft Azure. All private GenAI interactions (prompts & responses) within ZChat are protected by Zscaler AI Guard

AI Secured:
Comprehensive Safeguards for the Enterprise

Security Benefits

- **Zero Trust Architecture:** Zscaler enforces least-privilege, secure-by-design access for all users, devices, and applications, minimizing attack surfaces and preventing lateral movement across environments.

- **Comprehensive Governance:** All AI initiatives are governed by strong internal policies and are aligned with frameworks such as the NIST AI Risk Management Framework.

- **Robust Data Protection:** Sensitive data (e.g., PII, confidential business information) is protected through strict data input and output controls, inline DLP, and browser isolation—ensuring that inadvertent leaks or misuse are prevented.

- **Continuous AI Activity Monitoring:** The Gen AI Security Report provides detailed oversight and reporting capabilities, allowing real-time tracking of tool usage and ensuring compliance with usage policies.

- **Layered AI Guardrails:** Zscaler AI Guard adds prompt and response monitoring—blocking attempts to share sensitive data, stopping jailbreak exploits, and preventing toxic or unauthorized queries both for public GenAI and internal/private AI.

- **Vendor Risk Mitigation:** All third-party AI tools undergo comprehensive due diligence, including security reviews, penetration testing, and contractual/legal data safeguards (e.g., no data used for external AI training).

Employee Productivity Gains

- **Safe Enablement of AI:** By providing secure access to both private and public AI tools. Zscaler allows employees to securely utilize AI for research, coding, and workflow automation—driving innovation without exposure risk.

- **Faster Decision-Making:** Employees can use AI to generate insights, summarize data, and automate tasks with confidence, knowing that guardrails and compliance checks are in place—eliminating bottlenecks.

- **Accelerated Development:** Engineers leverage safe, pre-approved GenAI and code suggestion tools (protected by code review and security checks), shortening the development lifecycle while maintaining high standards.

Business Value Delivered

- **Stronger Trust & Compliance:** By aligning with global standards and proactively implementing AI governance, Zscaler reduces regulatory risks, inspires trust with customers, and protects company reputation.

- **Reduced Cyber Risk Exposure:** Strict segmentation, traffic inspection, and policy enforcement drastically lower the risk of data breaches, compliance violations, and intellectual property loss—translating to cost savings.

- **Greater Agility & Innovation:** Zscaler's secure frameworks empower the organization to explore AI's potential with speed and confidence, supporting differentiated business models and faster time-to-market for new offerings.

- **Leadership & Market Differentiation:** By embracing the use of our own products and transparently documenting success, Zscaler demonstrates leadership in secure AI, setting an example for customers and strengthening its competitive position.

Conclusion: Trusted AI in a Zero Trust World at Scale

Zscaler IT exemplifies how to embrace AI innovation without compromising security, privacy, or compliance. By implementing a comprehensive governance framework, integrating strict policies, and leveraging the power of its own Zero Trust Exchange platform, Zscaler has established itself as a leader in secure AI adoption. Tools such as ZIA with AI prompt capture, AI Guard, and integrated DLP ensure that AI usage is not only scalable but also transparent and rigorously governed. These practices allow Zscaler to mitigate cyber threats, safeguard sensitive data, and adhere to regulatory requirements, all while driving operational efficiency and fostering innovation.

Zscaler's approach to AI security proves that enterprises can confidently harness the transformative potential of AI while maintaining trust and resilience. By accelerating innovation responsibly and setting a benchmark for secure AI governance, Zscaler continues to lead the way in enabling trusted AI at enterprise scale.

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™

Act Fast.
Stay Secure.