aws marketplace

GENERATIVE AI SOLUTION SPOTLIGHT

Safeguard your generative AI initiatives at scale with AIShield GuArdIan on AWS

Establish effective guardrails for safe and secure AI adoption

Organizations can use generative AI to streamline mundane processes and free employees for higher-value tasks. However, it also poses potential security concerns, such as data breaches, intellectual property (IP) infringement, and brand, reputation, and policy-related risks. Addressing these issues is essential to safe and responsible AI adoption.

AIShield GuArdIan, an enterprise security solution on Amazon Web Services (AWS), acts as middleware, providing guardrail/firewall capabilities to secure the input and output stages of generative AI. The solution helps ensure security and compliance with your organization's requirements, allowing you to confidently scale your AI initiatives while mitigating safety, security, and compliance risks.

AIShield GuArdIan overview

AIShield GuArdIan provides comprehensive safeguards to help protect your AI projects. As a cybersecurity middleware, it strengthens your generative AI and large language model (LLM) applications against a wide range of cyber threats, providing an added layer of protection.

AIShield GuArdIan offers robust protection against prompt injections, jailbreaks, and sensitive information disclosure. It acts as a secure bridge between user applications and LLMs, analyzing inputs and outputs, enforcing role-based policies, and ensuring compliance with legal and policy standards.

In collaboration with:

AIShield
Powered by Bosch

Securing LLM applications at scale for a hi-tech company

An electronics and technology multinational corporation was looking to widely implement LLMs and generative AI. However, the corporation wanted to ensure adequate security for LLM usage as they scaled. It needed a security solution that would support key security features and offer flexibility for configurations.

The company decided to leverage AIShield GuArdIan to help safeguard its applications, including a chatbot, document summarization, and image moderation. The corporation used the solution to mitigate risks such as data privacy, copyright, and IP loss issues. The enterprise was equipped with 75+ controls across content access, privacy protection, content validation, and SIEM integration. As a result, it was able to achieve a 79 percent reduction in time from proof of concept (PoC) to production scaling.

© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key components

Admin console: Simplifies policy configuration with an intuitive interface for managing security across AI deployments.

Embedded simulation setup: Allows real-time testing of policies in simulated scenarios before deployment, ensuring robust policy validation.

Observability dashboard: Offers comprehensive traffic monitoring and insights into policy violations, including key metrics like the most violated policies.

Operational modes: Operates in monitoring mode for passive logging or blocking mode to actively prevent violations and control data flow.

Advanced policy enforcement: Leverages statistical methods, regular expressions (regex), custom machine learning and deep learning (ML/DL) models, and tailored generative AI in an ensemble approach for multi-layered security.

Key features

Multi-modal guardrail capability
Features over 75 configurable security controls applicable to both text and image-based I/O across use cases and retrieval-augmented generation (RAG)-based applications.

Real-time threat intelligence and protection
Provides real-time monitoring and mitigation of potential threats, enabling threat-informed defenses for ML models and guardrails for LLMs.

Advanced observability
Offers enhanced reporting and is seamlessly integrated with security information and event management (SIEM) tools for comprehensive visibility.

Enterprise-ready integration
Integrates seamlessly with existing enterprise tools, including SIEM platforms and IT service management (ITSM) tools. Offers robust logging, monitoring, user authentication, and management features, making it ideal for organizations with complex IT infrastructures.

LLM-agnostic
Delivers flexibility and broad compatibility across various foundation models as a model-agnostic guardrail solution.

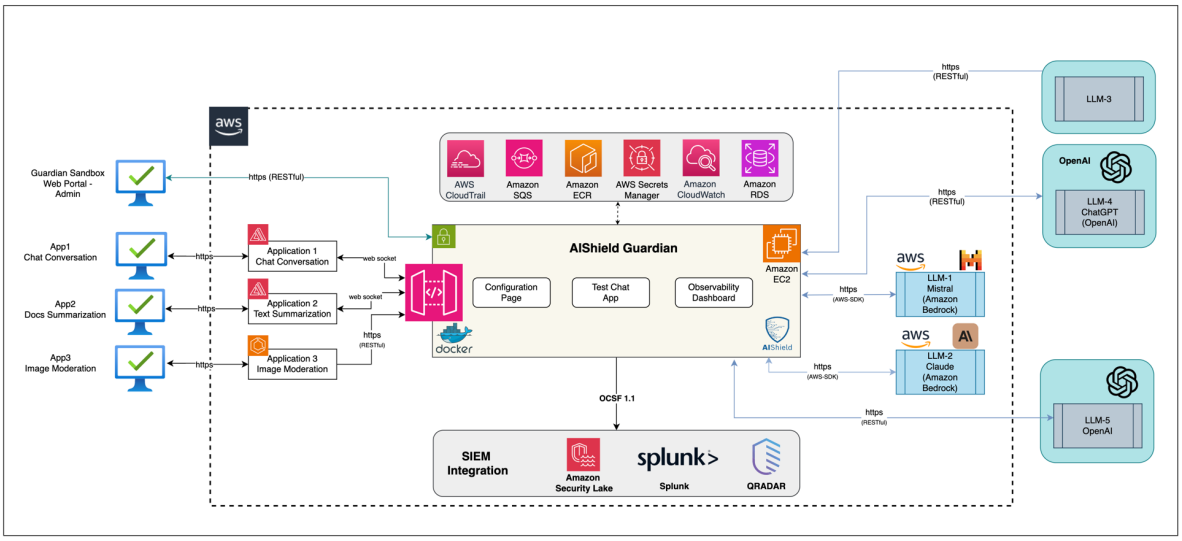
Fully customizable
Meets the unique security and compliance needs of different industries.

Boosting IT support efficiency with RAG and AIShield GuArdIan

A large manufacturing and automotive multinational corporation's IT support team struggled to efficiently address a wide range of technical queries due to the evolving nature of IT problems and solutions. Its traditional support systems struggled to stay updated with the latest information, resulting in delays in query resolution and increased risks to data privacy.

To solve for this challenge, the company implemented AIShield GuArdIan with a RAG-based LLM platform to enhance the effectiveness of its support systems. The corporation's IT support chatbot is now connected to a current external knowledge base, enabling it to deliver accurate and timely solutions. AIShield GuArdIan, acting as middleware, enforces strict data privacy and compliance measures, ensuring all responses are aligned with the company's security standards. The guardrail demonstrated an effectiveness of 97 percent in mitigating attack attempts around sensitive information disclosure and prompt injection/jailbreak attacks.

© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.



- Reduce time from PoC to production of generative AI workloads.
- Achieve efficiency improvements by scaling generative AI usage across the enterprise with security confidence.
- Mitigate and respond to any potential security threats quickly.
- Demonstrate governance and accelerate compliance for generative AI deployments with a single pane of control and detailed insights into policies and violation.

Featured AWS products

Amazon Bedrock
The easiest way to build and scale generative AI applications with foundation models.

Amazon SageMaker
Build, train, and deploy ML models for any use case.

Amazon Security Lake
Automatically centralize security data from various sources into a data lake in your account.

Amazon Virtual Private Cloud (Amazon VPC)
Create and manage your own virtual network for AWS resources.

AWS PrivateLink
Establish connectivity between VPCs and AWS services without exposing data to the internet.

Learn more ›

Enhance generative AI risk management with AIShield GuArdIan on AWS.
See how you can strengthen the security and privacy of your LLMs.

About AIShield

AIShield is a Gartner-recognized leader in AI security, delivering cybersecurity products and services. Backed by 45+ patents, AIShield's enterprise-ready unified AI security platform SecureAIx protects enterprise AI/ML models, applications, and workloads across various stages of development and operation (MLOps/LLMOps). The platform offers a suite of advanced security testing and defense technology to AI/ML teams, facilitating AI risk mitigation, accelerated compliance and time-to-market, effective governance and the protection of brand and intellectual property.

AIShield proudly serves more than 40 customers and has actively participated in developing guidelines and best practices with MITRE, NIST, and ETSI for assessing the security and safety of AI systems globally. AIShield's strategic collaborations and technology integrations with industry-leading players such as AWS, F5, IBM, Dell, Hewlett Packard Enterprise, Deloitte, Dekra, Wipro, and Databricks position it as an integral part of the AI/ML ecosystem, advancing trust, security, and robustness of AI systems.

© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.