EINPRESSWIRE

# ZeroTrusted.ai Launches Revolutionary Anti-AI Tool for Enhancing Security and Privacy in Large Language Models

*ZeroTrusted.ai launches groundbreaking Anti-AI tool, setting a new standard in AI security. It addresses LLM vulnerabilities, ensuring privacy, and compliance.*

ZeroTrusted.ai logo

KISSIMMEE, FLORIDA, USA, February 23, 2024 /EINPresswire.com/ -- ZeroTrusted.ai, a leading Generative AI (GenAI) security firm, today announced the release of its groundbreaking anti-AI tool designed to comprehensively address security and privacy concerns associated with Large Language Models (LLMs). This innovative tool establishes a new standard in AI security by providing unparalleled protection against injection attacks, hallucinations, compliance violations, plagiarism, copyright violations, and other vulnerabilities.

The newly launched ZeroTrusted.ai AI anti-AI tool utilizes advanced algorithms to automatically identify and mitigate issues related to the following:

Anonymize LLM usage: Maintain prompt anonymity when utilizing LLMs to ensure privacy.

Compliance Validation: Check for instances of compliance breaches and sensitive data such as Personally Identifiable Information (PII), Payment Card Industry (PCI) data, etc. Rectify any issues by appropriately sanitizing the data.

Detect Plagiarism and Copyright Infringements: Ensure content generated by LLMs is original and respects intellectual property rights.

Improve Accuracy: Enhance accuracy by optimizing prompts and validating results.

Prevent Hallucinations: Verify that LLM outputs are not fabricated.

Injection Attacks: Safeguard against malicious inputs designed to exploit model vulnerabilities.

Prompt Injection, Data Poisoning, and Bias: Address issues that compromise the integrity and impartiality of LLM outputs.

Adversarial Attacks and Leakage: Fortify defenses against attempts to deceive models or extract confidential information.

Secure LLM Interaction: Maintain privacy by implementing end-to-end encryption for data exchanged between user prompts and proprietary LLMs.

Insecure Code and Model Theft: Prevent unauthorized access to proprietary model code and intellectual property.

Excessive Agency and Inadequate Sandboxing: Limit unintended model behaviors and ensure robust containment.

Insufficient Access Controls and Overreliance: Establish stringent access measures and promote balanced utilization of AI capabilities.

Prompt Disclosure and Supply Chain Vulnerabilities: Secure the entire AI development pipeline from external threats and internal lapses.

SSRF Vulnerabilities: Shield against Server-Side Request Forgery attacks that target internal systems.

ZeroTrusted.ai's guardrail tool is engineered for organizations deploying LLMs across various sectors, including technology, finance, healthcare, and government, offering a critical layer of security that promotes trust and reliability in AI applications.

"AI technologies are transforming industries, but with great power comes great responsibility," said Waylon Krush, CEO of ZeroTrusted.ai. "Our anti-AI tool addresses the urgent need for comprehensive security solutions that protect against a wide array of threats. By securing LLMs against these vulnerabilities, we empower organizations to leverage AI with confidence, ensuring their innovations are both safe and compliant with the highest standards of privacy and security."

The launch of this anti-AI tool marks a significant milestone in ZeroTrusted.ai's mission to advance AI governance and application methodologies that prioritize security, compliance, and operational excellence. With its state-of-the-art features, the tool is poised to become an essential asset for any entity looking to navigate the complex landscape of AI with assurance and integrity.

"Our anti-AI tool is integrated directly into our Zero Trust Policy Server that significantly enhances our customer experience on ensuring the accuracy and real zero trust of all 3rd party and proprietary LLMs," said Femi Fashakin, CTO of ZeroTrusted.ai.

Meet Waylon and Femi at the upcoming Orlando Code Camp 2024, organized by the Orlando .NET User Group (ONETUG) and hosted at the Sanford/Lake Mary campus of Seminole State College. They will be speaking on Zero Trust Architecture: Principles and Best Practices on Saturday, February 24, 2024, from 2:30 PM - 3:20 PM at Seminole State College, Sanford, Florida - UP 2208.

For more information about ZeroTrusted.ai and its new Anti-AI tool, please visit https://www.zerotrusted.ai/

About ZeroTrusted.ai

ZeroTrusted.ai is a pioneer in cybersecurity and Generative AI security, dedicated to developing secure, innovative technologies that address the most pressing challenges in AI applications. With a team of leading experts in cybersecurity, AI research, and ethical computing, ZeroTrusted.ai is at the forefront of creating solutions that ensure the responsible and effective use of artificial intelligence.

Sharon Lam
ZeroTrusted.ai
email us here
Visit us on social media:
Facebook
Twitter
LinkedIn
Instagram

This press release can be viewed online at: https://www.einpresswire.com/article/690635364

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.
© 1995-2024 Newsmatics Inc. All Right Reserved.