Guardrail Sentinel

Guard the Gates of Your Language Model.

Setting a new standard for LLM security.

The Problem: Prompt Injection Risks

LLMs are vulnerable

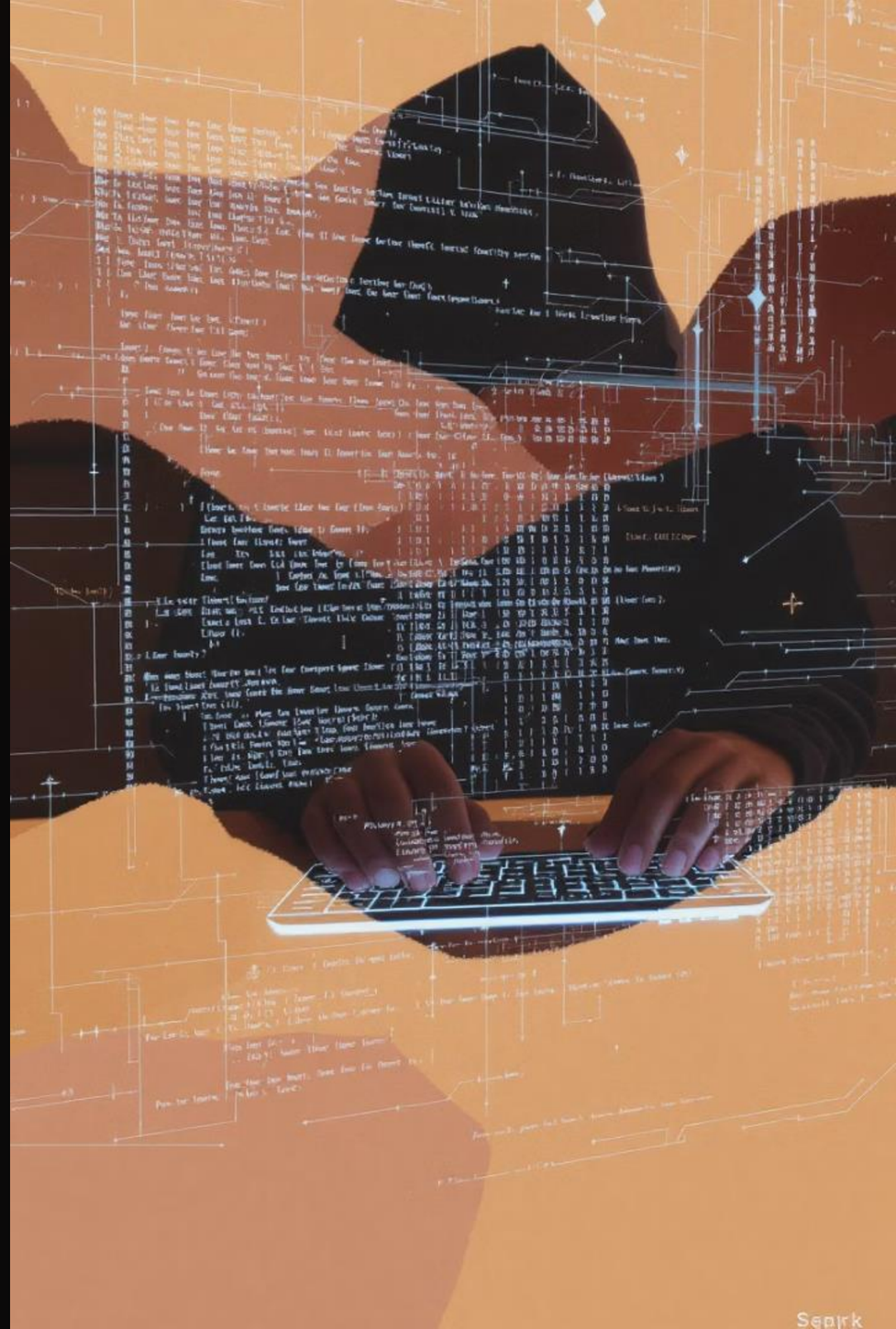
Language Models face prompt injection attacks.

Data leaks possible

Attacks can leak sensitive information.

APIs are targets

Chatbots and APIs are primary entry points.



The Stakes: Why It Matters

Exploiting Trust

Attacks compromise AI model integrity.

Major Risks

Data breaches, reputational damage, regulatory fines.

Industry Impact

41% of firms suffered AI security incidents in 2024.

The Solution: Guardrail Sentinel

Real-time Detection

AI-powered identification of prompt injection.

Continuous Monitoring

Proactive scanning for suspicious input patterns.

Versatile Deployment

Secures APIs, chatbots, and enterprise AI.

Platform Features

Detection Algorithms

Proprietary LLM traffic analysis.

Automated Response

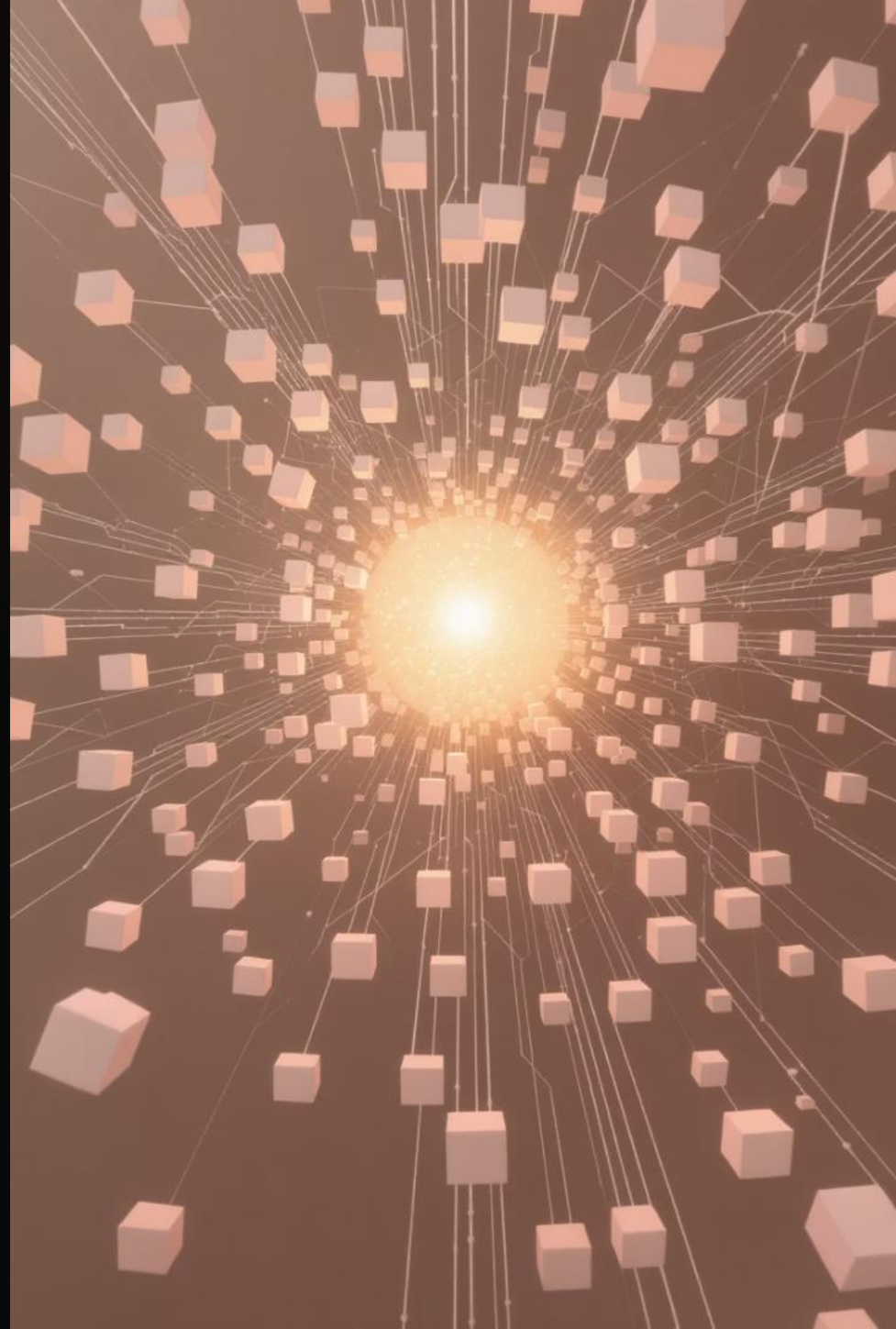
Real-time alerts and workflow automation.

Detailed Dashboard

Risk scoring, incident logs, comprehensive insights.

Seamless Integration

Works with major LLM API gateways.



How It Works: Technical Overview

Input Analysis

Examines prompts for adversarial signatures.

Machine Learning

Flags anomalous queries using advanced ML.

Custom Rules

Supports tailored enterprise risk configuration.

Proven Protection: Case Examples

Financial Chatbot

Stopped credential spoofing in May 2025.

Healthcare API

Detected prompt-leak attempts during testing.

Live PoCs

Zero prompt injection incidents post-deployment in 2025.

Next Steps & Contact

- Deploy Guardrail Sentinel in your AI stack today.
- Start with a complimentary security assessment.
- Contact: mhlongosihle49@gmail.com.