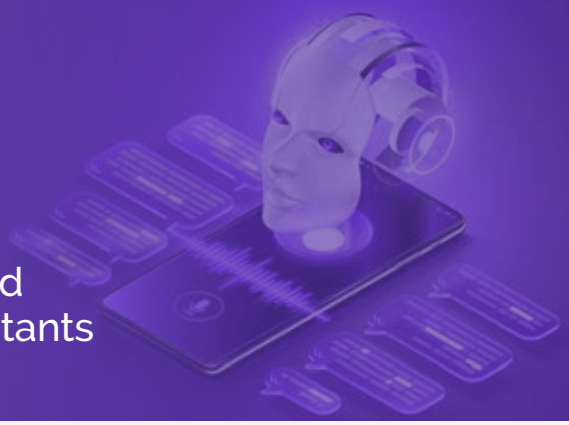
Arkose Labs



Arkose GPT Protection

Comprehensive defense against threats aimed at LLM platforms and generative AI chat assistants

The Critical Need for Arkose GPT Protection

Driving your business forward hinges on delivering innovative features, but the rise of generative AI technologies like large language models (LLMs) and AI chat assistants presents significant challenges. While they offer new opportunities for customer engagement and efficiency, they also introduce risks that can jeopardize your strategic goals. Emerging threats like GPT prompt compromise and LLM platform abuse aren't merely technical issues – they pose direct threats to your business outcomes.

Arkose GPT Protection is engineered to defend your AI assets against these sophisticated attacks, ensuring your innovations continue to drive growth while safeguarding your business's future.

What Is Arkose GPT Protection?

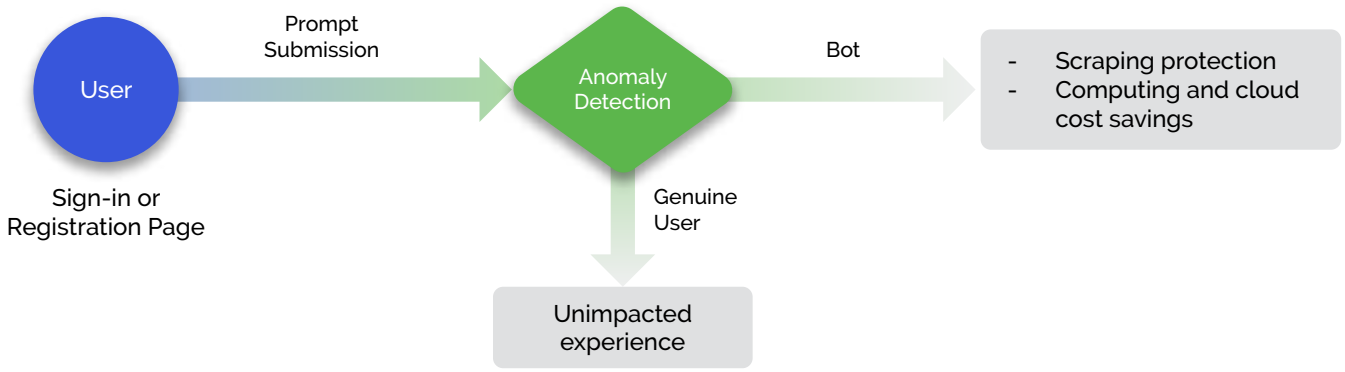
Arkose GPT Protection is a specialized capability of Arkose Bot Manager, tailored to defend generative AI platforms and chat assistants from cyberattack. While Arkose Bot Manager is designed to protect your digital ecosystem from a wide range of automated threats, Arkose GPT Protection hones in on the unique vulnerabilities introduced by LLMs and AI-driven systems. To learn more about Arkose Bot Manager's broader capabilities, visit our product brief.

Key Capabilities

GPT Prompt Compromise Protection

- **The Problem:** GPT prompt compromise is an attack type where bots are able to programmatically submit prompts and scrape the response with an intention to either train their own models, resell similar services or gain access to proprietary, confidential and personal information. This poses severe risks to data security and intellectual property.

- **Our Solution:** Arkose GPT Protection offers a context-aware, AI-resistant challenge configuration feature that shields your chat assistants from being scraped. By detecting and disrupting these scraping attempts, we ensure your data remains secure.

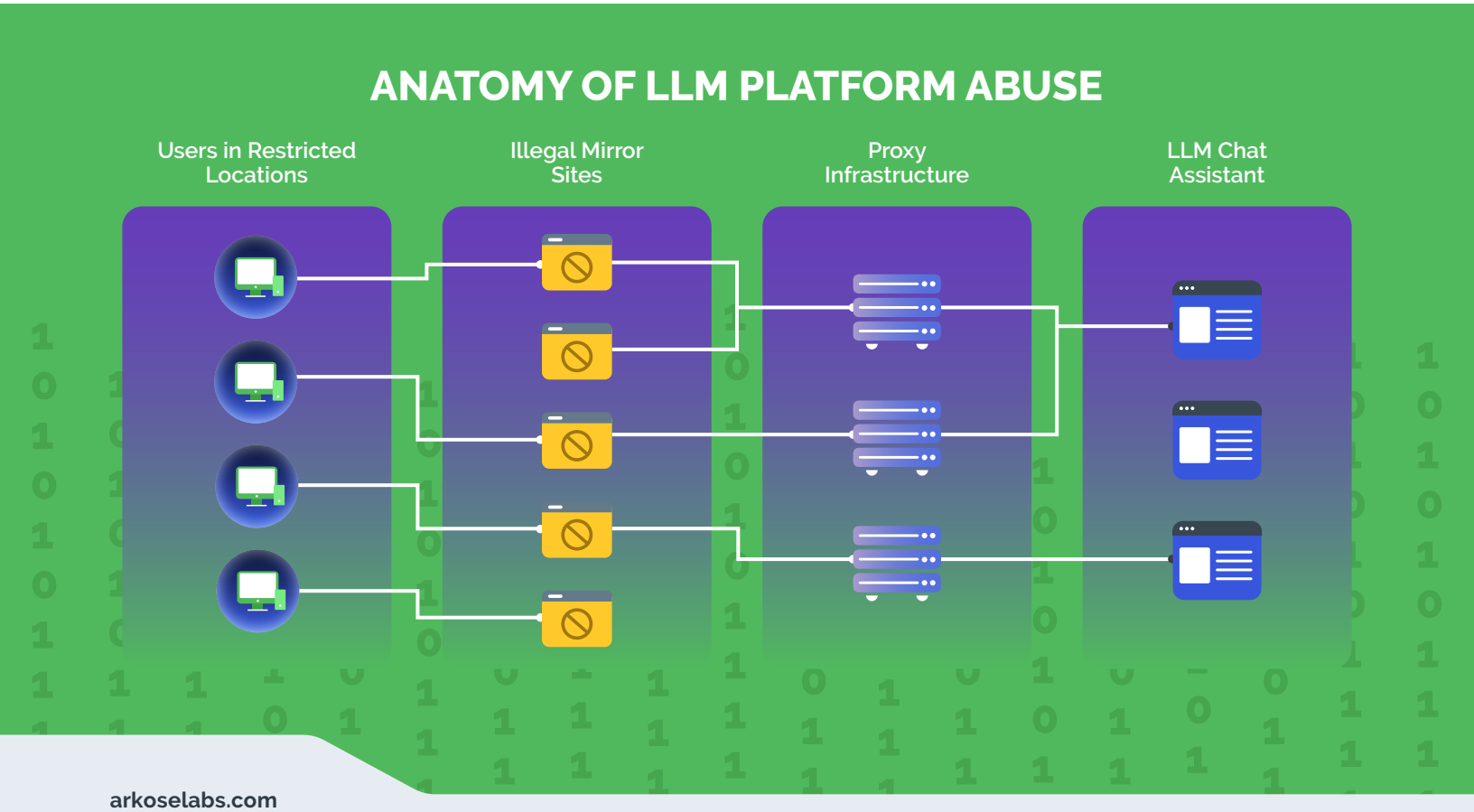


How GPT prompt compromise detection and mitigation works

LLM Platform Protection

- **The Problem:** LLM platform abuse is an attack type where malicious actors utilize reverse proxy infrastructures and mirror services to bypass geo-restrictions and conceal their activities, which may include creating knock-off services that are increasingly used to generate phishing emails, create deep-fake videos and conduct other illicit acts. This makes it difficult for AI platform providers and regulatory authorities to track and mitigate these actions, leading to potential legal and reputational risks.

- **Our Solution:** Arkose GPT Protection defends against illegal reverse proxy infrastructure and mirror services that facilitate LLM platform abuse. Our solution helps maintain the integrity of your AI systems by preventing unauthorized access and use, even in regions where such activities are prohibited.



How Arkose GPT Protection Works

Detection

Arkose GPT Protection leverages the advanced detection capabilities of Arkose Bot Manager, including:

- Traffic pattern anomaly detection
- Browser and device anomaly detection
- Network and geo anomaly detection
- Behavioral anomaly detection
- Workflow anomaly detection
- API instrumentation

Multiple prompts are required for scraping attacks to be fruitful; our system detects and challenges these, thus rendering such attacks ineffective and unsustainable. These stacks often operate across multiple data centers, targeting specific AI assistants. By analyzing hundreds of risk signals, we distinguish between legitimate users and malicious actors, ensuring that only authorized users interact with your AI systems.

For LLM Platform Abuse, Arkose GPT Protection detects and manages traffic from illegal reverse proxy infrastructure-based mirror services, for real-time detection of cyberattacks.

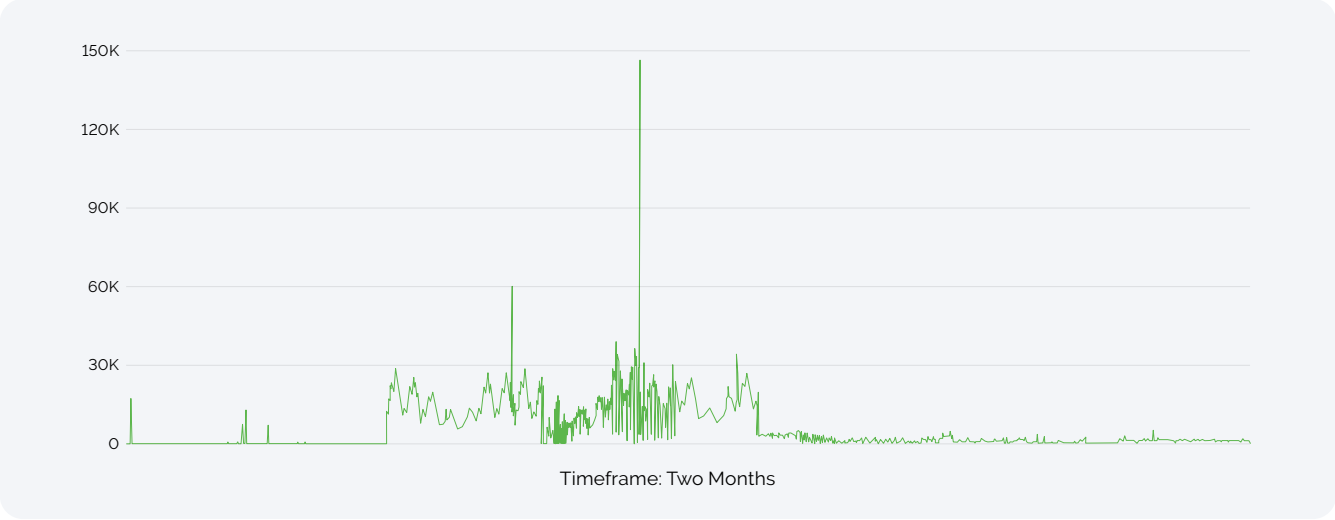
Mitigation

Our dynamic challenge methodologies make it cost-prohibitive at best, and unsolvable at worst, for bots and scripts to operate at scale. In chat use cases, we monitor and can take action, balancing user experience with security. Our approach is designed to interrupt malicious activities, making it increasingly difficult for attackers to profit from their efforts.

For LLM platform abuse, our solution not only stops immediate threats but also provides valuable data for your cybersecurity teams to develop long-term strategies. This proactive approach helps dismantle the networks behind these malicious activities, ensuring lasting protection.

Real-World Example of Arkose GPT Protection in Action

This real-life example shows chat assistant proxy traffic being detected and mitigated.



Reverse proxy activity spike and mitigation

Arkose Labs Proof of Value

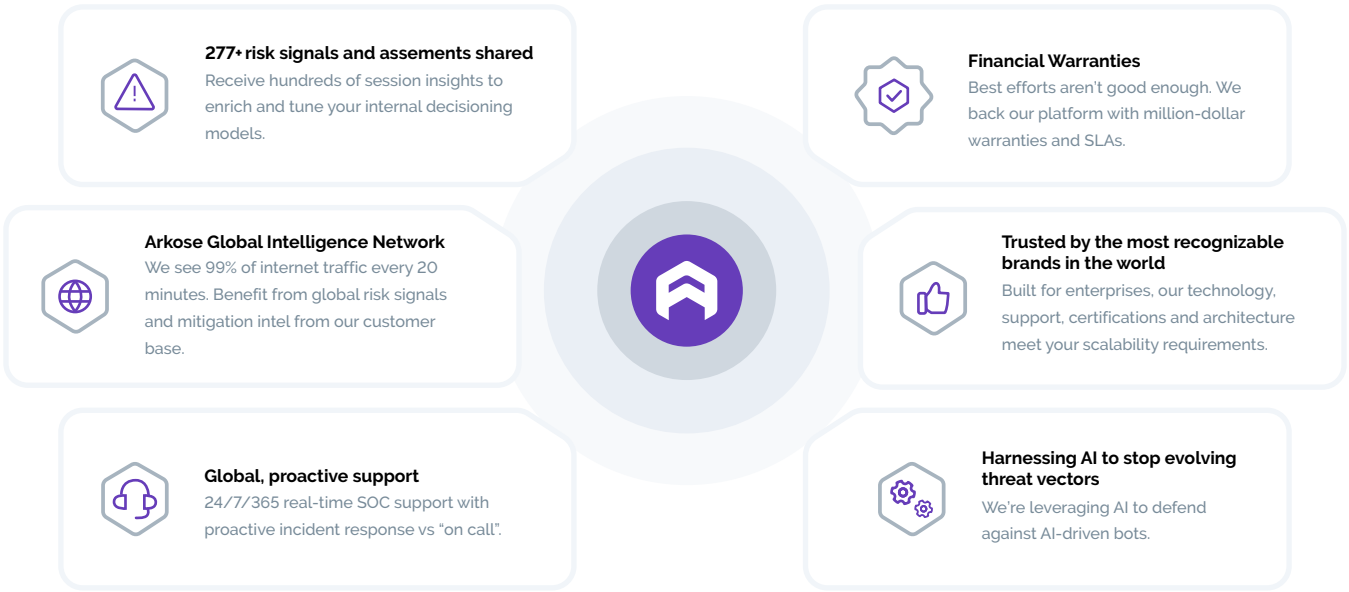
The Arkose Labs proof of value (POV) process offers your business a hands-on opportunity to experience the platform's capabilities. During the POV, Arkose Labs provides expert guidance and consultation tailored to your specific needs, ensuring you can test the platform's effectiveness in real-world scenarios. This process allows your business to define and track its own success metrics, such as fraud reduction, improved user experience or cost savings, giving you a clear view of the value Arkose GPT Protection can deliver.

Seamless Integration: No Engineering Delays

We understand the challenges of securing engineering time, which is why Arkose Labs is designed for effortless integration. If you're using a CDN like Cloudflare or Akamai, our pre-built CDN workers allow you to deploy our solution without any code changes to your application. This low-code implementation ensures you can get up and running quickly, without needing to pull your engineers away from their critical projects.

If you're already working with one of our partners, enabling Arkose Labs is even simpler – our integration is ready to go, requiring minimal effort on your part. This means you can deploy robust protection without the usual delays, keeping your business secure and not waiting months for engineering resources.

The Arkose Labs Advantage



ACTIR and the Arkose Labs SOC: Proactive Defense

Arkose Labs operates as an extension of your team, rapidly countering attacks and providing actionable insights without overburdening your internal resources. The Arkose Cyber Threat Intelligence Research (ACTIR) unit conducts proactive threat hunting, risk intelligence gathering and other counterintelligence methods to provide vital, fresh intelligence.

Meanwhile, the 24/7/365 Security Operations Center (SOC) team focuses on identifying and stopping both sophisticated low-and-slow and large-scale attacks immediately.

The SOC continuously monitors for new threats and collaborates with ACTIR. This feedback loop ensures a seamless collaboration between the SOC and ACTIR, enhancing the overall accuracy, timeliness and effectiveness of your cybersecurity defense.

Arkose Bot Manager in Action

Global AI Company Elevates User Experience

A global AI research and deployment company faced unprecedented cyberattacks, including LLM platform abuse, SMS toll fraud, account takeover and new account fraud. These attacks exhausted its processing capacity, costing tens of millions monthly and preventing genuine users from accessing services while bots ran rampant. Additionally, the company needed to ensure its services weren't accessible in prohibited countries.

The initial focus was on the registration flow, where Arkose Bot Manager rapidly curtailed fake account sign-ups and reduced the impacts of SMS toll fraud. As soon as that flow became more difficult for attackers to target, they doubled their efforts against the company's premier chat prompts via existing accounts. As the cost to conduct attacks on this flow increased, the attacks shifted to other parts of the company's business. The company and Arkose Labs teams then pivoted to safeguarding additional flows, including login, forgot password, a developer portal and profile update.

This approach proved to be so effective against the cybercriminals, the bad actors gave up and shut down their repositories within weeks. Meanwhile, legitimate users experienced no friction whatsoever.

Results with Arkose Labs

- → 2 billion bot attacks detected and mitigated in first 6 months
- → >99% reduction in LLM platform abuse
- → 99% good user throughput

BOOK YOUR DEMO

Arkoselabs.com