AiFort

DATASHEET

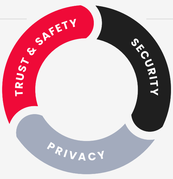
Securing LLM AI Applications with Intelligence-Driven Red Teaming

NEW OPPORTUNITIES AND EMERGING RISKS

The Generative AI market's growth from $12.9 billion to $1.8 trillion by 2030 highlights its broad integration but also brings to light significant risks, including Trust & Safety, Security and Privacy risks that require advanced protective measures for AI systems.

GENERATIVE AI RISKS

Trust & Safety
- Harmful content
- Hallucinations
- Misinformation



Security
- Prompt injection
- Jailbreaks
- Prompt leakage

Privacy
- Data Leakage
- Intellectual property
- PII

AIFORT - SECURE THE FUTURE OF AI APPLICATIONS

AiFort provides organizations with a powerful security solution for MLOps and Generative AI applications, leveraging KELA's intelligence and an advanced AI verdict engine. This ensures comprehensive protection for Large Language Models (LLMs), safeguarding against threats and vulnerabilities in the evolving landscape of AI technology.

Detect Risks at an Early Stage in Development

AiFort's advanced platform allows detection of vulnerabilities from the early stage of the development phase, ensuring a secure development and production of AI models.

Discover Emerging Risks in Production Models

Continuous monitoring capabilities of AiFort spot emerging threats in deployed models, maintaining their integrity against evolving risks.

Validate Models at Scale

With AiFort, the validation of AI models is both swift and thorough, ensuring they meet the highest standards of security and trustworthiness before deployment.

A Customizable Testing Frameworks

AiFort enables tailored testing to adopt emerging AI vendors, across various industries, regulations, and internal policies, moving beyond a one-size-fits-all approach.

AIFORT'S SOLUTIONS

Multifaceted Approach to Enhancing Model Security and Performance

AiFort RED

Automated, Intelligence-Based Red Teaming

AiFort Red simulates adversarial attacks on AI models to test their robustness against real-world threats. This automated AI-based approach ensures that GenAI systems are robust, trustworthy, and responsible. AiFort allows running tests at scale to ensure AI applications are secure against potential misuse.

Automated red teaming tool: AiFort Red employs automated **red teaming** to meticulously scrutinize AI models against a spectrum of LLM threats, ensuring comprehensive vulnerability identification and fortification.

Integration of LLM Threat Landscape: By integrating the LLM threat landscape with your AI attack surface, AiFort provides holistic threat visibility, allowing proactive risk mitigation.

Advanced Benchmarking Capabilities: AiFort enables advanced benchmarking, facilitating comparison of model performance against other versions or vendors, empowering informed decision-making and optimization.

Detailed Scoring Reports: Receive clear score analysis reports to mitigate vulnerabilities and fine-tune your model and dataset.

Customizable testing platform: Create custom tests to meet your specific security standards and regulatory requirements across various industries.

Protect your customers and employees:
- Organizations that develop their LLM models.
- Organizations that integrate LLM to build their in-house applications:
  Customer-facing AI applications • AI applications for **employees'** internal use

AiFort DATA

Kela's data lake offers a unique advantage for training Generative AI models, providing a wealth of real-world cyber threat intelligence.

Unique Access to Past & Current CTI Data: Train Gen AI models on a diverse array of datasets from underground forums, messaging apps, and illicit marketplaces, crucial for the deep and comprehensive development of AI capabilities.

Contextualization of Threats: Provides an understanding of threat actors' tactics to prevent misuse of AI applications.

Detect Impending Attacks and Mitigate Risks: Utilize cybercrime data to refine the model's abilities, and enhance its capability to proactively identify and mitigate risks.

Empowering Gen AI Data models: Stay informed with the criminal terminology and eliminate negative contexts by analyzing cybercrime discussions.

AiFort PROTECT

COMING SOON

Helps organizations to detect and mitigate AI risks with real-time AI prompt filtering.

Discovery: Identify all the AI tools that are being used in your organization.

Detection: Get real-time detection and alerting of adversarial attacks and prompt injections.

Response: The firewall blocks malicious inputs in real-time.

WELCOME TO THE
FUTURE OF AI SECURITY.
WELCOME TO AIFORT.

Ready to secure your AI journey? Choose AiFort by Kela for an unparalleled blend of intelligence, innovation, and integrity in AI security.

Visit our website for detailed information on pricing and plans, and to start securing your AI models today.

Visit **kelacyber.com**

Empowering Diverse Industries: From retail to finance, healthcare to government, AiFort's versatile platform ensures that every sector can deploy AI applications with confidence, safeguarding against financial loss, compliance violations, operational disruptions, and more.

KELA

www.kelacyber.com · marketing@ke-la.com · +972-3-970-2720

Copyright© 2024 KELA