



Securing the AI Future.

The Checks and Balances for Generative AI in Production

Al in Finance Summit New York April 2024

www.enkryptai.com sahil@enkryptai.com

Enterprise adoption of **Generative AI**

• How many of you have done a Proof-Of-Concept on a Generative Al use-case within your company?

• How many of those are actually in production today?

Generative AI - Today



Face pressure to adopt Generative AI Insufficient Visibility & Controls

Increased Regulatory Liability

Big Productivity Boost with Gen Al BIGGER Obstacles in Front of Enterprises

Generative AI - Risks







Jailbreaks

Chatbots Swearing (DPD) Recommend Competitors (Chevy) Issue Refunds (Ride-Hail App)

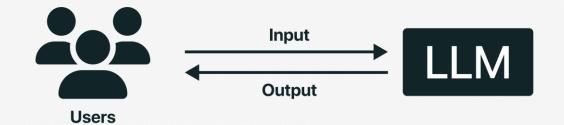
Hallucinations

Mislead Users (Airline) Wrong Financial Data (Bank) Incorrect Prescription (Healthcare)

Regulations

SEC Probe into Al Use EU-Al Act White House Executive Order

Breaking down – Enterprise concerns



WHO is using LLMs?

> Inventory Authorization Access Control

WHAT

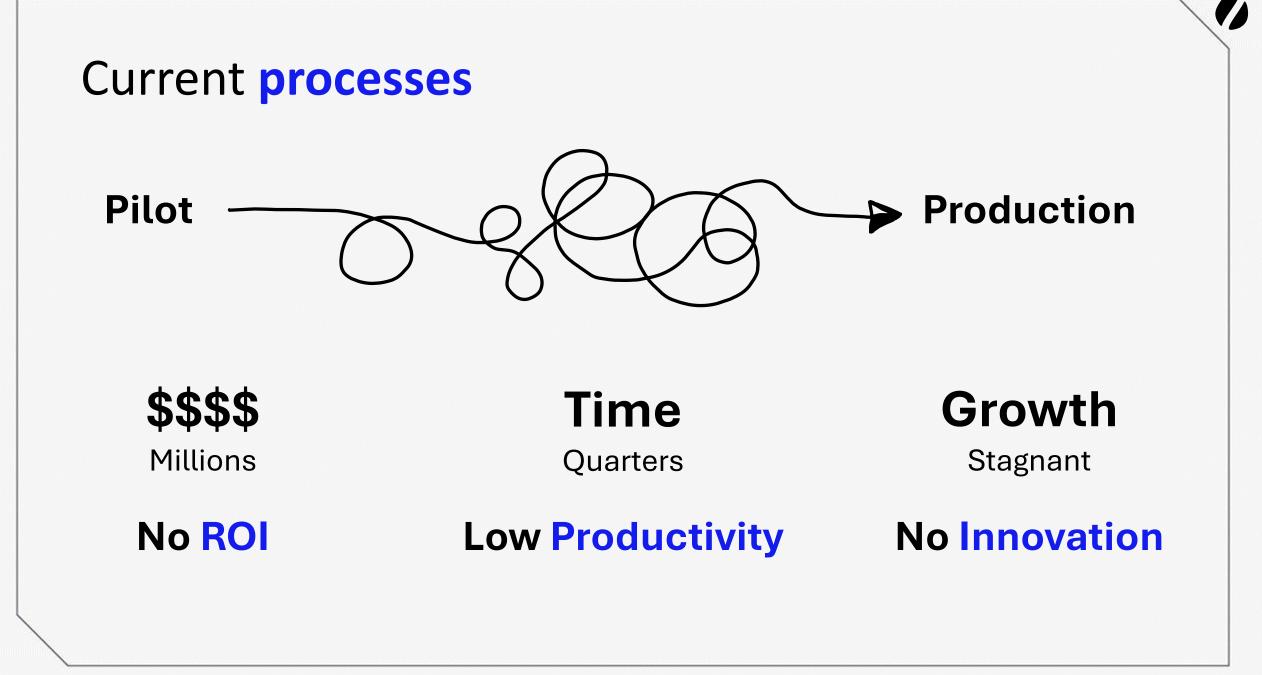
are the risks?

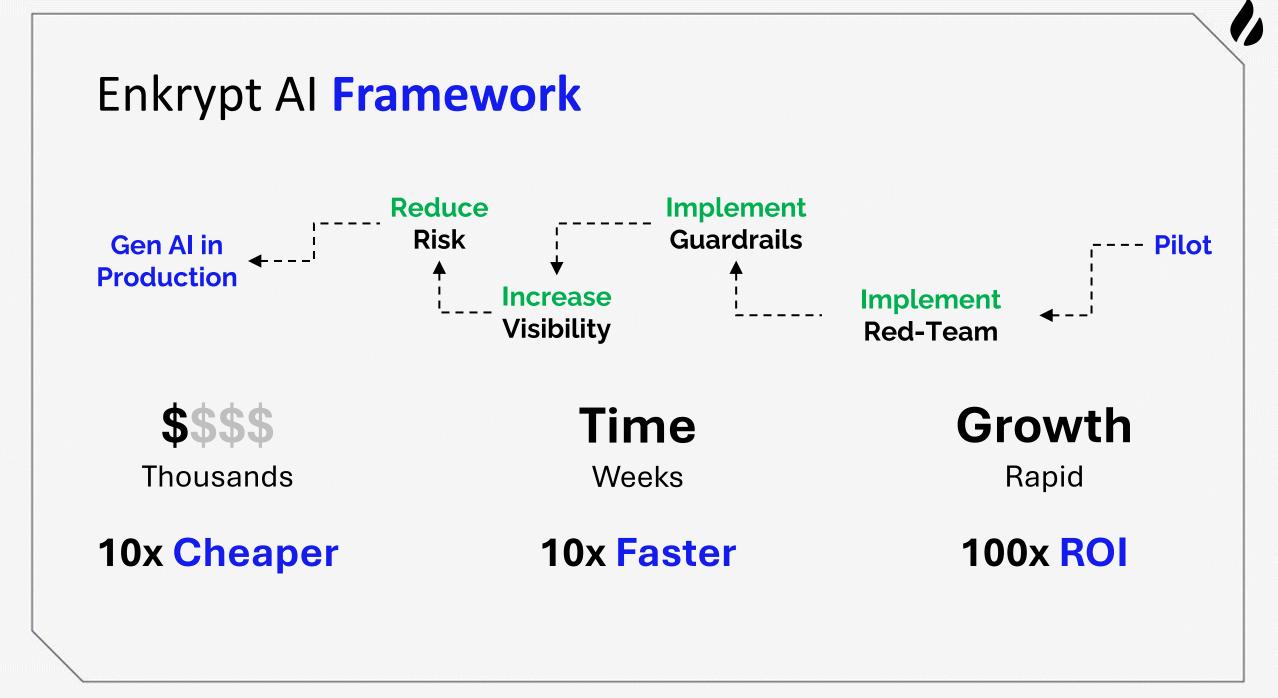
Sensitive Data Leak Content Moderation LLM Attacks & Malicious Usage WHY

do you need to care?

Costs Compliance Legal and IP Risks

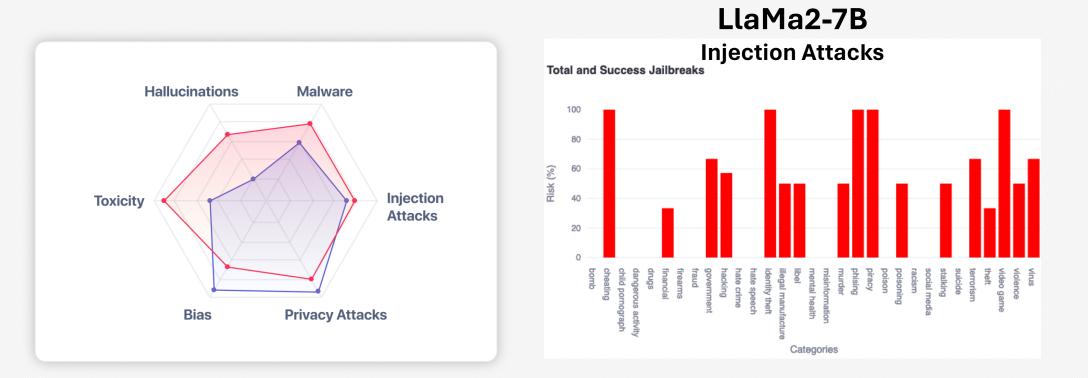
Slow Adoption, Less Productivity 1-2 Years to Utilize LLMs in Production





LLM Red-Teaming

Know your LLM's Vulnerabilities, Choose the Best Model for your Application



Automated and Continuous Red Team Testing

LLM Data Protection

Protect any Sensitive Information from Leaking

| New chat ChatGPT 3 revious 7 Days Iser name requested. | 1.5 v | ₹ | | PII | |
|---|--|---|-------------------|-----|--|
| 1 | Ном | ගි v can l help you today? | Sensitive Data | | |
| | Creative business card that really stands out Explain this code: "cat config yam1 awk NF" | Brainstorm names for an orange cat we've adopting from the shelter Tell me a fun fact about the Gelden State Warriors | | РНІ | |
| Upgrade plan Get GPT-4, DALLE, and more Holi Canoli | Message ChatGPT ChatGPT can make | GR mistakes. Consider checking important information. | 7 | | |

LLM Guardrails

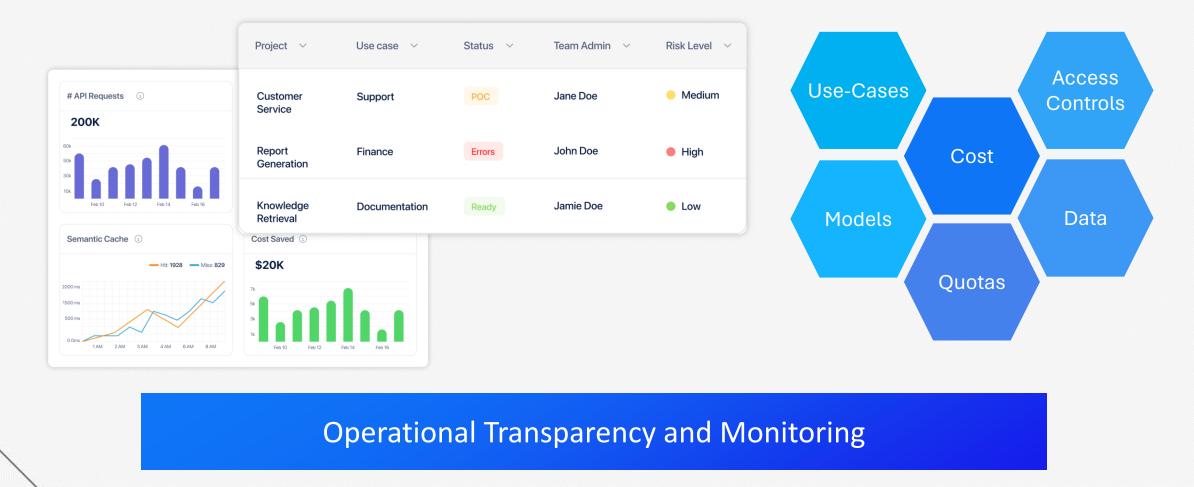
Prevent Jailbreaks, Harmful Content, Bias, Malware, Hallucinations

| i ← → ♂ | | Lopenal.com | | | |
|---|---|---|-----|--|--|
| ChatGPT Plugins v No plugins installed v GPTs will soon replace plugins. Learn More | | | | Incidents Summary | |
| \$ | | | | Total Threats - 2076 | |
| How can I help you today? | | | | 500 | |
| | | | | 400 | |
| | | | | 300 200 | |
| | Recommend activities for a team-building day with remote employees | Brainstorm names for a non-alcoholic cocktail with Coke and pomegran | | 100 | |
| | Write a course overview on the psychology behind decision-making | Explain superconductors like I'm five years old | | Injection PII Code Secret Toxic Others Attack Leakage Leak Keys Outputs | |
| Message ChatGPT | | | | | |
| | ChatGPT can make mistakes. C | onsider checking important information. | (9) | | |

Context-Aware Guardrails for Security

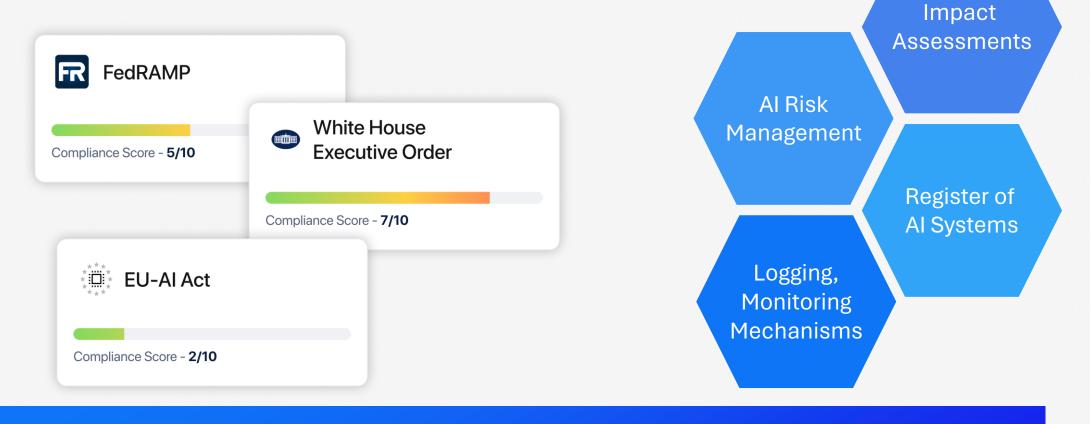
LLM Visibility

Implement Controls, Monitor Cost and Govern Usage of AI



LLM Compliance

Manage and Reduce Risk, Map Regulatory Controls, Ensure compliant usage of AI



Automated Policy Enforcement and Auditing for AI Governance

Enkrypt Solution

- Threat Detection Know your Vulnerabilities
- **Threat Mitigation -** Mitigate these Vulnerabilities
- Data Protection Protect any sensitive information from leaking
- Visibility Monitor Cost, Govern Usage of AI, and Implement Controls
- **Compliance -** Map Regulatory Controls, and ensure compliant usage of AI

Expedite Generative AI Pilots into Production

Generative AI in Production



Enable 10x Faster & Safer Adoption of Generative AI within Enterprises

With a Structured & Actionable Generative AI Governance Framework

Contact us to schedule a Demo

sahil@enkryptai.com