# Comprehensive GenAI Security with Red and Blue Teaming

**Lasso**
- Threat Protection
- Mitigating Risks
- Policy Enforcing

**splx**
- Vulnerability Scanning
- Security Reporting
- Assessing Compliance

## Introduction

As Generative AI continues to revolutionize the world, organizations face an increasing number of cybersecurity challenges and risks. The rapid adoption of Generative AI technologies brings not only innovative capabilities but also potential vulnerabilities that malicious actors can exploit.

**1.5bn+**
Global users engage with conversational AI applications.

**80%**
Of conversational AI systems will embed GenAI by 2025.

**64%**
Would use Generative AI more if it was more safe/secure.

TIDIO Gartner salesforce

**45%**
Companies don't implement chatbots due to privacy, security and legal concerns.

**42%**
Teams face data privacy and security challenges when implementing chatbots.

**4X**
Annual rise in attacks on chatbots due to GenAI vulnerabilities.

accenture accenture SECURITY

## The Challenges

### Model and Data Poisoning
Attackers continually seek new methods to inject malicious data into the training process or input data to manipulate the model's behavior.

### Brand Reputation Risk
Organizations and employees can fall victim to misinformation or AI-generated malicious content. As a result, businesses face the potential for reputational damage, loss of trust, and legal consequences.

### Prompt Injection
Prompt injection is a critical vulnerability where attackers manipulate the input prompts to alter the chatbot's behavior, potentially exposing sensitive data or executing unauthorized actions. This can lead to severe security breaches and operational disruptions.

### Model Leakage
Model leakage is a critical threat in AI chatbots, exposing the model's architecture, parameters, training data, personal user data, and proprietary company data through sophisticated attacks.

### Hallucination
AI hallucinations refer to instances where a chatbot generates responses that are factually incorrect or nonsensical. These errors can mislead users and undermine the reliability of the AI system.

## The Solution

Addressing these risks requires a robust and comprehensive security strategy that includes both offensive and defensive measures. In this context, the synergy between Lasso Security's blue teaming solutions with SplxAI's red teaming expertise provides a comprehensive security mechanism for organizations leveraging Generative AI.

### Blue Teaming

Lasso Security's solution emphasizes defense and mitigation of existing and emerging GenAI threats. Lasso Security's blue teaming solutions are designed to safeguard Generative AI applications and LLMs from a wide array of cyber threats.

Lasso Security ensures that organizations can securely harness the power of GenAI without compromising their security poster.

1. **Always-on Shadow LLM™**
   Uncover every LLM interaction and facilitate the precise identification of active tools, models, and their users within the organization.
2. **Real time Response and Automated Mitigation**
   Swift alert mechanisms provide timely notifications to both users and security teams for rapid response and protection against real-time threats.
3. **Tailored Policy Enforcement**
   Lasso enables organizations to implement customized security policies that align with their unique requirements and regulatory. No coding, development or data science expertise is required.
4. **Privacy Risk Reduction**
   Prioritizes data protection right from the initial stages of deployments.

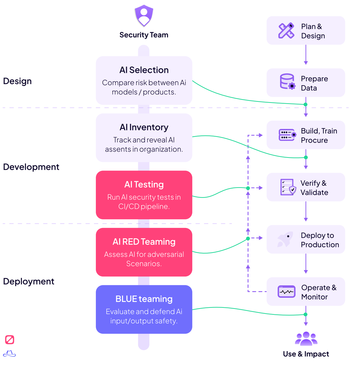
### Red Teaming

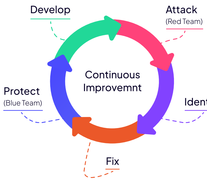
SplxAI specializes in AI red teaming, an offensive security practice that simulates real-world attacks to identify and exploit vulnerabilities within an organization's Generative AI applications.

By mimicking the tactics, techniques, and procedures (TTPs) of adversaries, SplxAI helps organizations understand their exposure to potential threats.

1. **Automated Scans**
   Eliminate the need for months of manual testing and reporting, significantly optimizing time efficiency and ensuring on-demand continuous protection against evolving threats.
2. **Compliance Mapping**
   Assess your applications' conformance to 10+ critical AI security frameworks, including OWASP LLM Top 10, MITRE ATLAS, EU AI Act, GDPR, and NIST AI RMF.
3. **Comprehensive Reporting**
   Detailed reports provide actionable insights into vulnerabilities, including the potential impact and recommended remediation steps.
4. **Continuous Improvement**
   SplxAI's red teaming services are iterative, ensuring that organizations stay ahead of emerging threats and refining their AI security posture.

## How we fit in your AI Development Lifecycle




## A Unified Defense Strategy for The GenAI Era

The combination of Lasso Security's security suite and SplxAI's red teaming create a holistic defense strategy for Generative AI. This better together approach ensures that organizations can not only identify and understand their vulnerabilities but also effectively defend against and mitigate potential threats.

- **Enhanced Security Posture**
  By addressing both sides of the security equation, organizations can achieve a more robust and resilient security posture.
- **Strategic Insights**
  The combination of offensive and defensive expertise provides strategic insights that can inform long-term security planning and investment.
- **Continuous & Automated Risk Management**
  The integrated approach ensures that all potential risks are identified, analyzed, and mitigated, leaving no stone unturned.
- **Compliance benefits**
  The iterative process of red teaming and the defensive measures of blue teaming create continuous improvement cycle, keeping security measures up-to-date with evolving threats.

By leveraging the strengths of both approaches, companies can achieve security and resilience and confidently embrace the potential of GenAI without compromising on safety.

### About Lasso Security

Lasso Security is a pioneering comprehensive Large Language Models (LLMs) cybersecurity solution that is committed to addressing the evolving threats and challenges that businesses are beginning to encounter with Generative AI and LLMs at every touchpoint within their organizations.

### About SplxAI

SplxAI develops AI security software to identify, assess, and mitigate risks associated with artificial intelligence. The company specializes in providing offensive security for enterprises that deploy GenAI applications, helping them identify vulnerabilities and ensure a safe user experience. Founded in 2023, SplxAI is based in the United States.