HIDDENLAYER PROFESSIONAL SERVICES

PROTECT YOUR ADVANTAGE

> “Require(s) that developers of the most powerful AI systems share their safety test results and other critical information with the U.S. government.”
>
> Biden Executive Order on Standards for AI Safety and Security

> “Common security concerns relate to adversarial examples, data poisoning, and the exfiltration of models, training data, or other intellectual property through AI system endpoints. AI systems that can maintain confidentiality, integrity, and availability through protection mechanisms that prevent unauthorized access and use may be said to be secure.”
>
> NIST AI Risk Management Framework

> “If models meet certain criteria they will have to conduct model evaluations, assess and mitigate systemic risks, conduct adversarial testing, report to the Commission on serious incidents, ensure cybersecurity, and report on their energy efficiency.”
>
> EU AI Act

HiddenLayer’s Professional Services are a multi-faceted services engagement that utilizes our deep domain expertise in cybersecurity, artificial intelligence, and threat research. We couple our industry experts with our patent-pending product offerings to provide our customers with:

- Real-time security assessment of your critical AI models,
- Prioritized recommendations of remediations to ensure your models remain protected against various types of cyber attacks, and
- Improved readiness in the event of a future attack.

WHY NOW?

Artificial Intelligence models present a large, growing, attractive attack surface for adversaries. The risks and costs of an attack on an organization’s AI models are at an all-time high, and the repercussions include damage to the brand, loss of IP, and monetary loss. In addition, escalating AI regulations, including the European AI Act and Biden’s Executive AI Order, have increased the need for organizations to understand their current compliance. HiddenLayer’s offerings complement the organization’s internal corporate efforts to secure AI with the required expertise from data scientists, reverse engineers, threat intelligence, and adversarial engineers.

THE PROFESSIONAL SERVICES TEAM

Our experts have over 50 years of combined experience in cyber, adversarial AI, threat intel, and reverse engineering. Equally important is our team’s front-line experience in responding to significant AI and ML model breaches and researching current vulnerabilities on MLOps platforms before they harm the general public.

OUR APPROACH

AI RISK ASSESSMENT

A detailed analysis of your ML Operations lifecycle and an in-depth review of your most critical AI models to determine the risk your AI investments currently pose to the organization. Findings are mapped to industry best practices such as NIST, MITRE ATLAS, and OWASP to provide actionable guidance for reducing organizational risk.

ADVERSARIAL ML TRAINING

Two-day training to provide data science and security teams with an understanding of Adversarial Machine Learning TTPs and the most effective countermeasures to protect against them. Determine appropriate next steps and modifications required for internal testing processes to include AI models and an overview of offensive AI tooling, including Adversarial Robustness Toolbox (ART), Counterfit, CleverHans, Augly, Foolbox, and more.

RED TEAM ASSESSMENT

The Adversarial Machine Learning Research (AMLR) Team will leverage the same TTPs used by attackers to assess how well the attacks are currently detected and prevented by your existing people, processes, and controls. The red teaming engagement will focus on the following attack techniques to ascertain each model's robustness to Attack: Reconnaissance, Inference, Bypass, Insider Threat, Prompt Injection, Code Audit, and Model Compromise.

AIDR IMPLEMENTATION SERVICES

Professional implementation and integration of HiddenLayer's AI Detection & Response (AIDR) product into the AI environment. It provides the data science and security teams with the functionality and visibility required to prevent attacks, improve responsiveness, and maximize model effectiveness.

SECURITY FOR AI RETAINER SERVICE

An annual retainer service led by our Adversarial Machine Learning Research (AMLR) Team for full-service MLOps Lifecycle support, including an incident response plan, regular risk assessments, and adversarial ML team training.

KEY BENEFITS

HiddenLayer's professional services, rooted in deep expertise across cybersecurity, artificial intelligence, reverse engineering, and threat research - ensure a comprehensive approach to safeguarding your AI investments. From thorough AI Risk Assessments to Adversarial ML Training to Red Teaming for robust model testing, our services provide pragmatic and actionable insights. These services empower teams to proactively safeguard AI investments, offering resilience, security, and aligned business objectives.

- Ensure the security & integrity of your ML Operations Pipeline
- Gain knowledge from Industry Adversarial Machine Learning Experts
- Visibility into the risks & attacks that threaten your AI Models
- Insight into where an attack on your Machine Learning Operations & Models would most likely occur
- Prepare for AI Detection & Response by ensuring a known good state
- Accelerate the use of LLMs within your organization with the proper safeguards in place

WHY HIDDENLAYER?

HiddenLayer, a Gartner-recognized AI Application Security company, is the only platform provider of security solutions for GenAI, LLMs, and traditional models. With a first-of-its-kind, non-invasive software approach to observing and securing GenAI, HiddenLayer is helping to protect the world's most valuable technologies.

The HiddenLayer Team consists of the world's top experts at the intersection of cybersecurity and artificial intelligence. Our collective expertise derives from previous roles at McAfee, Intel, Hewlett Packard, Dell and Cylance. Over the past decade, this team has helped usher in a new era of AI-powered cybersecurity products.