Cyber AI Suite (CAIS)

Secure Your AI Innovations, Stay Ahead of AI Threats

What is the Cyber AI Suite (CAIS)?

As AI security concerns shift from theoretical to tangible, the threat landscape evolves rapidly. Corporate data is increasingly at risk of being ingested by third-party models unnoticed. AI-powered applications with internal access introduce new attack vectors, creating a blind spot where innovation outpaces governance.

HolistiCyber's Cyber AI Suite (CAIS) ensures the security of your AI innovations from the ground up. The CAIS service begins with a deep Architecture Review of your RAG pipelines and Vector Databases to identify structural risks. We then apply rigorous Threat Modelling to map potential logic flaws before our specialized AI Red Team actively stress-tests your defenses against adversarial attacks like prompt injection and jailbreaking based on MITRE ATLAS™ and OWASP Top 10 for LLM. Finally, we deliver a robust Governance framework, ensuring your AI systems remain compliant with global standards such as NIST AI RMF and ISO 42001

How does CAIS work for me?

1. **Architecture & RAG Audit**
   An AI model is only as secure as the data it retrieves. Our Architecture Review deep-dives into the 'brain' of your application—specifically the Retrieval-Augmented Generation (RAG) pipeline, Agents deployment, and Database. We verify that tenant data remains isolated (RBAC/RLS), that ingestion pipelines sanitize hidden threats in uploaded files, and that your embedding strategy doesn't inadvertently expose sensitive IP or PII.

2. **AI Penetration Test**
   Standard penetration tests often miss the unique, probabilistic risks inherent to Generative AI. Our Red Team moves beyond traditional vulnerabilities to simulate sophisticated, model-specific attacks. We execute a full adversarial kill chain—launching Prompt Injections to override system instructions, employing Jailbreaking techniques to bypass safety guardrails, and testing for Excessive Agency to ensure your autonomous agents cannot be tricked into performing unauthorized actions. We don't just identify theoretical risks; we prove the impact. We focus on attacks that are AI-oriented using TTPs from MITRE ATLAS™ and OWASP Top 10 for LLM

3. **Security Controls Assessment (Agentic/Embedded AI)**
   Utilizing a proprietary AI Security Framework that maps your specific architecture against 47+ critical controls defined by MITRE ATLAS™, NIST AI RMF, and ISO 42001. Unlike static checklists, our model is dynamic: we ingest your system's unique profile—analyzing factors like RAG usage, data sensitivity, and agent autonomy—to generate a Weighted Maturity Score for each domain like Reconnaissance, Exfiltration, Agent Permissions and Resource Exhaustion. This provides you with a quantifiable, board-ready metric that proves your compliance posture and prioritizes remediation where it matters most.

WHAT SHOULD YOU EXPECT?

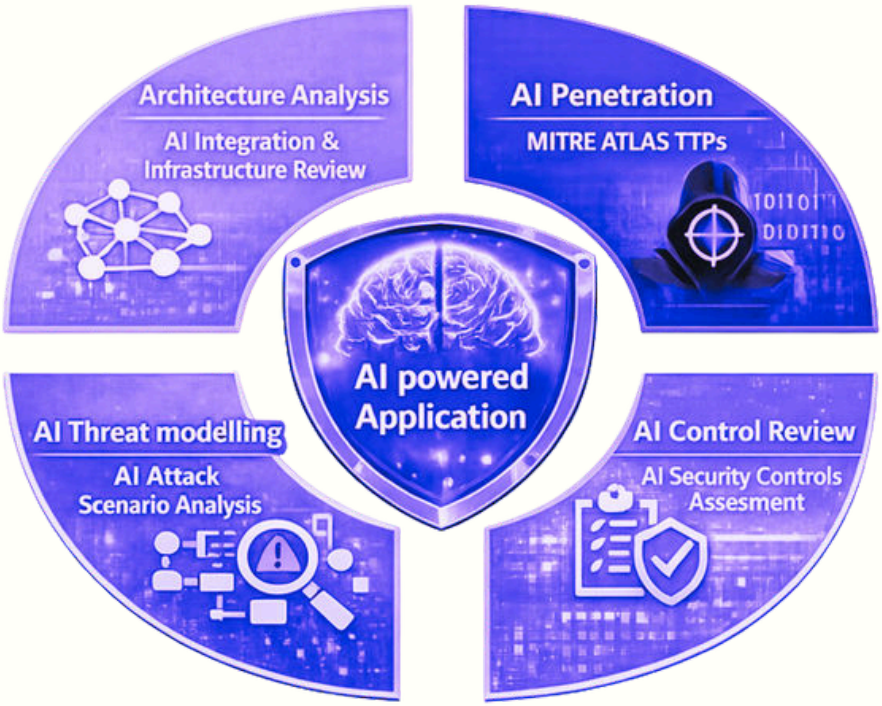
A clear, crafted plan - productive, concise, and prioritized - to quickly identify, assess and provide recommendations to address changing AI threats to meet current and future needs.

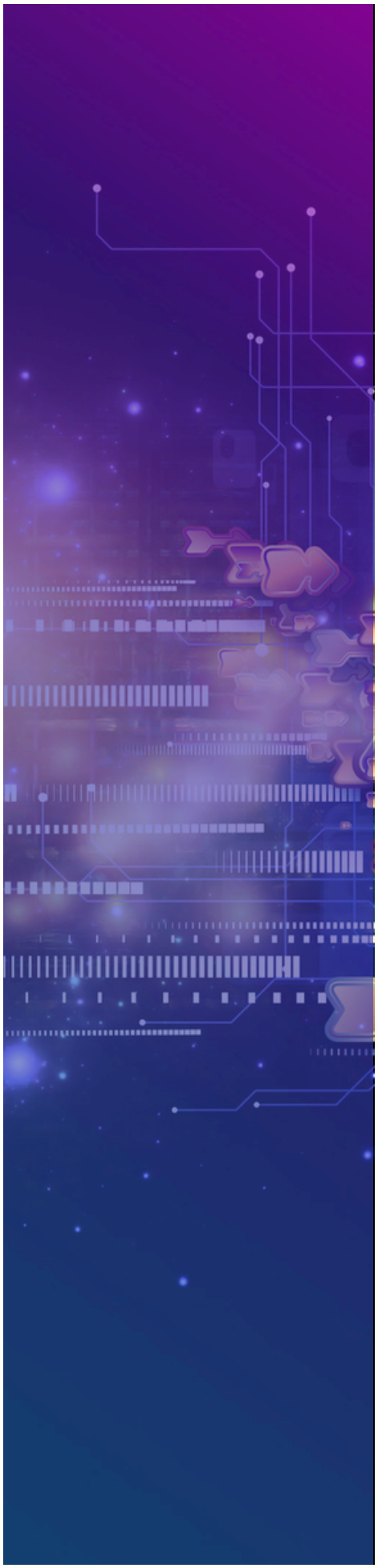
Our experts serve as an extension of your security and engineering teams to mitigate evolving AI threats, readiness levels, and defense mechanisms. Actionable and coherent action plan and recommendations adapted to unique company business needs, priorities, and risk appetite.

Clear prioritization of security threats that need to be handled immediately vs. items that do not pose an immediate risk.

Straight-forward remediation and mitigation tactics and workflows as well as ongoing remediation support where needed.

4. **Threat Modelling**
   Security flaws in Generative AI hide in the logic, not just the code. Our Threat Modelling service focuses on the attack vectors unique to AI. We analyze the full data journey—from Data Ingestion and RAG Retrieval strategies to Model Inference and Output Guardrails. By identifying logic gaps where an attacker could poison your knowledge base, bypass safety filters, or manipulate retrieval context. We identify new threat actors, vectors and the inherent changes on solutions as a result of the adaption of AI technologies.





And the Value...

- **Baseline and visibility:** Executive Scorecard: Risk rating (Critical/High/Medium/Low)
- **Findings**: Detailed steps mapped to MITRE ATLAS
- **Compliance**: Findings cross-referenced with NIST AI RMF 1.0, MITRE ATLAS, and ISO 42001
- **Remediation Guidance**: Actionable deliverables ranging from architectural redesigns and governance policy updates for your security leadership to specific technical controls for your engineering teams

Why HolistiCyber?

1. **Nation-state grade expertise** - our staff of white-hat security experts is composed of former military and government offensive practitioners who can examine the attack surface from the vantage point of the attacker and not only from the company's vantage point. This includes a solid grasp of the sophisticated tooling available to today's attackers and access to those attack tools.
2. **Holistic approach** - compliance, remediation, and mitigation solutions are tied to each company's unique business objectives and workflows. Security should complement productivity and growth and avoid hindering them.