AVERTIUM

AI APPLICATION PENETRATION TESTING SERVICE

APPLICATION-LAYER TESTING ALIGNED TO OWASP AI TESTING GUIDE

As organizations race to deploy AI-powered applications like chatbots, RAG search, and agent interaction, they're facing unprecedented challenges:

- Uncertainty that AI will behave safely
- Fear of leaking sensitive data
- Limited visibility into AI-specific attack paths
- Mounting pressure to protect customer trust

These AI features introduce new ways to manipulate logic, data, and outputs beyond traditional web and API risks, making AI-specific penetration testing essential.

Avertium's AI Application Penetration Testing service tests your AI capabilities and surrounding application layer as one connected attack surface, validating real-world abuse paths and **delivering proof-of-concept findings with prioritized, actionable remediation guidance.**

WHAT WE TEST

Avertium aligns our methodology to the **OWASP AI Testing Guide** and maps AI application risks to the OWASP Top 10 for Large Language Model (LLM) Applications to provide a framework grounded in real attack patterns and widely accepted security guidance:

- → **Prompt injection & jailbreak resistance:** Direct and indirect prompt injection
- → **Sensitive information disclosure:** PII, internal data, system prompts, credentials, proprietary content
- → **RAG/retrieval & embedding security:** Unauthorized retrieval, poisoning/manipulation, cross-context leakage
- → **Unsafe/misleading outputs:** Hallucinations, policy violations, toxic content, misinformation risk
- → **Agentic behavior and misuse:** Excessive permissions, unauthorized actions, privilege escalation paths
- → **Output handling safeguards:** Preventing AI outputs from being blindly executed, published, or trusted

REPORT DELIVERABLES

You receive an executive-ready, engineer-actionable report that translates AI risk into evidence-based findings and a practical remediation roadmap.

1. **Mapped view of the tested AI attack surface and data flows**
2. **Prioritized proof-of-concept findings with business impact**

AI APPLICATION PEN TEST PROCESS

Avertium's analysts operate as an extension of your team, collaboratively scoping, coordinating your project and communicating along the way to ensure we deliver your desired outcomes:

PHASE ONE
Planning & preparation

Align on goals and success criteria, set clear testing rules and scope, and confirm timing and access.

PHASE TWO
Penetration test

Map the AI's interaction surfaces and integrations, then conduct targeted testing to detect risk.

PHASE THREE
Reporting & closeout

Document and convey findings with clear severity and impact explained, provide prioritized remediation guidance, and securely dispose of artifacts.

AI Application Penetration Testing is conducted within an 85-hour timebox, making the engagement predictable, governable, and outcome-focused without sacrificing depth.

WHY AVERTIUM

AI risk lurks at the intersection of the AI feature, application, data sources, and any connected tools. Avertium approaches your environment the way real attackers do, by testing both web application and AI components as a unified attack surface, performing manual validation, and mapping findings to best-practice guidance.

The result is proof-of-concept validated risks, clear remediation guidance, and enterprise-ready documentation that supports governance and auditability.

- See the full, connected attack path across roles, web, APIs, and AI interaction surfaces instead of siloed results.
- Prioritize remediation based on validated impact, with actionable guidance your teams can implement.
- Support auditability and program integration with enterprise-grade reporting, defined data-handling standards, and clear roles/responsibilities.

ABOUT AVERTIUM

Avertium **is an AI security and compliance leader,** delivering comprehensive solution to mid-market and enterprise customers. Our unique "Assess, Design, Protect" approach addresses and improves security strategy, reduces attack surface risk, strengthens compliance, and provides continuous threat protection.

Avertium maximizes customer security investments and enables customers to focus on growth, innovation, and business outcomes, while assuring that their security infrastructure is resilient and adaptive to evolving threats. **That's why customers trust Avertium to deliver better security, improved compliance, and greater ROI.**

© Avertium, LLC. All rights reserved.