SPECTEROPS
Creators of BloodHound

DATASHEET

AI Red Teaming

See Your Environment the Way an AI-Enabled Adversary Does

AI changes what attackers can accomplish, how quickly they can develop new capabilities, and how efficiently they can move from access to impact. Enterprise AI adoption also introduces new models, agents, identities, APIs, data sources, and trust relationships.

SpecterOps AI Red Teaming evaluates the AI systems your organization deploys while testing your broader environment against AI-enabled adversary techniques. Our practitioners work at the leading edge of AI-enabled offensive security, applying original adversarial research and frontier model access to determine what an AI-enabled attacker could exploit, reach, and affect across your enterprise.

Security Testing Beyond the Model

AI systems connect to identities, applications, APIs, data, cloud infrastructure, development pipelines, and production systems. These relationships determine the potential impact of a compromise.

SpecterOps evaluates the full AI attack surface, including:

- Models, applications, and agentic workflows
- Human, non-human, and agentic identities
- Permissions, credentials, and trust relationships
- APIs, data sources, and integration points
- Development pipelines and supporting infrastructure
- Detection, investigation, and response capabilities

Each assessment connects technical findings to realistic adversary objectives and business risk, showing what an attacker could accomplish by exploiting a weakness.

From Initial Access to Business Impact

The central concern is blast radius: what a compromised AI system can reach and how quickly an AI-enabled adversary can move through the environment.

SpecterOps applies attack path analysis to identify how excessive permissions, exposed credentials, vulnerable integrations, and inherited trust can allow an attacker to pivot from initial access to critical resources.

AI increases the consequences of weak access controls. A system with excessive privileges can act across every application, data source, and resource its identity can reach.

Coverage Across the AI Lifecycle

SpecterOps evaluates AI systems from design through production. Engagements can include:

Threat Modeling

Identify likely adversaries, abuse cases, trust boundaries, and business-impact scenarios.

Model and Application Assessment

Test model behavior, inference interfaces, prompt-driven attacks, application vulnerabilities, and misuse scenarios.

Identity and Infrastructure Testing

Assess the identities, permissions, APIs, pipelines, data stores, and infrastructure supporting AI applications.

Production Red Team Operations

Simulate AI-enabled threat actors against live systems to test whether controls, telemetry, processes, and personnel can detect and contain an attack.

Each engagement concludes with a written report, collaborative outbrief, and prioritized recommendations.

The SpecterOps Difference

Proven Adversary Expertise

Our work includes BloodHound, Certified Pre-Owned, SCCM research, GhostWorks, and extensive red team operations. We use AI across our engagements to accelerate vulnerability discovery, zero-day research, and attack path analysis, informed by the same adversary expertise behind our industry-recognized research and tools. These insights inform every assessment.

Frontier AI Access

As red team of record for multiple frontier AI companies, a participant in OpenAI's Daybreak program, and one of the few organizations able to deliver services using Daybreak Red Access, SpecterOps operates at the leading edge of AI-enabled adversary simulation.

Attack Path Methodology

We look beyond isolated vulnerabilities to show how adversaries combine weaknesses, identities, permissions, and trust relationships to reach critical objectives.

Realistic Testing

Our operations test technology, processes, and people under realistic conditions. Collaborative debriefs help teams strengthen detection, response, and institutional knowledge.

Getting Started

SpecterOps can tailor an engagement to your AI systems, enterprise environment, risk priorities, and operational objectives. AI Red Teaming is part of the **SpecterOps AI Adversary Operations Platform**, which also includes:

Building Internal AI Adversary Expertise

AI Red Team Enablement gives security teams the infrastructure, frontier-model access, and hands-on guidance needed to understand and defend against AI-enabled tradecraft.

Evaluating Frontier Models at Enterprise Scale

AI Cyber Ranges provide realistic environments for measuring advanced models across multi-step attack paths and interconnected technologies.

See what an AI-enabled adversary could accomplish in your environment.

specterops.io/contact ›

SPECTEROPS ©2026 Specter Ops, Inc.