PRAETORIAN GUARD | THE AI-FIRST OFFENSIVE SECURITY PLATFORM

Run offense the way the attacker does. Continuously.

Your annual pen test is a photograph of a moving target. AI has handed adversaries speed and scale a once-a-year test can't see — and can't stop. Praetorian Guard turns that same advantage on your environment: a fleet of autonomous AI agents that attack continuously, prove what's genuinely exploitable, and hand your team the short list that actually matters. Machine speed, elite human judgment, zero noise.

Offense-Proven Security.

REAL RISK

We don't hand you a CVE dump. Guard chains real attacks across your entire enterprise — perimeter, applications and LLMs, cloud, identity, code, and people — and every finding is proven exploitable and validated by an elite operator before it reaches you. You fix the paths an attacker would actually walk, with the evidence in hand.

ZERO NOISE

Human validation is the product. Guard collapses the tens of thousands of "criticals" your tools generate down to the handful that are real — a signal-to-noise ratio on the order of **35,000 to 1**. Your team stops triaging and starts remediating.

ONE PLATFORM

Attack surface management, continuous penetration testing, adversarial exposure validation, vulnerability management, and cyber threat intelligence — unified, with your annual pen test bundled in. Guard meets you where you are: it ingests Wiz, Tenable, Qualys, Censys and more, or runs every capability itself.

Meet Marcus — your conduit into the platform.

Marcus is the orchestrator agent you talk to. Ask a question and he reads across your entire risk graph; give him an objective and he becomes an operator, dispatching specialized sub-agents across every attack surface at once. Marcus runs at the level of control you choose:

- IN **Human in the loop** — read-only. Query 15 years of offensive expertise and your live attack graph. Prompt-injection resistant, constrained in code.
- ON **Human on the loop** — Marcus executes on your behalf, launching live attacks and asking approval before each escalation.
- OFF **Human off the loop** — meet Hannibal.

Hannibal — fully autonomous offense.

Give Hannibal a mandate and a scope, and it runs the engagement itself, with no human in the loop. Point it at your whole infrastructure or a single newly shipped product; it authenticates (SSO and MFA included), hunts, and reports exploitable risk on whatever cadence you schedule. The default mandate is simple — is it reachable, is it exploitable? — or codify your own: PCI, FDA pre-market, whatever "in scope" means to you. Continuous, autonomous penetration testing that never clocks out.

How Guard works — the operator's method, automated.

1 Map the terrain.

Attack surface management is an input, not an outcome. Guard enumerates what you own the way an attacker would — zero-knowledge, outside-in (Pius: EDGAR filings, certificate scraping, zone transfers) — and fingerprints the technology behind it (Nerva reads source to fingerprint exact versions). Then it exposes the gaps between your systems of record: why is Cloudflare seeing an asset your scanner never touched?

2 Attack every surface.

Enterprise risk lives across six surfaces — perimeter, applications and LLMs, social engineering, code, cloud, and internal networks — and Guard's specialized agents cover all of them: Brutus (credentials), Augustus (LLM jailbreaks), Aurelian (cloud), Diana (web and API), Trajan (CI/CD and supply chain), Titus (secrets), Constantine (zero-day). Each agent is bounded by design — a constrained action space enforced in code — so autonomous never means out of scope.

3 Prove it, then prioritize by impact.

Finding is easy now; proof is hard. Guard verifies every finding is a true positive, assembles it into a kill chain, and weighs it against what a bad day actually looks like for your business. Not all criticals are equal — Guard delivers attacker-verified prioritization tied to materiality, not theory.

4 Close the loop.

Offense only matters if defense improves. Guard drives findings from detected to demonstrated, cuts tickets bidirectionally into Jira, and autonomously retests the moment your team says it's fixed. Constantine ships patches as pull requests — unit tests included. And when you connect CrowdStrike, Cortex, or Defender, Guard detonates a kill chain and shows you exactly which steps your controls caught and which they missed — mapped to NIST CSF and ISO 27001 as an evidence-based scorecard for the CISO and the board, benchmarked in reality.

Zero-day hunting, on demand.

When a target is hard, Guard goes hunting. Constantine reads your codebase, threat-models it, and shards the work across agents so it reasons instead of hallucinates. It doesn't just find the flaw — it writes the exploit, verifies it fires, and can open the patch PR. Its little brother Maxentius operationalizes N-days: the moment a CVE hits CISA KEV, Guard already knows which of your assets are exposed.

Built by the humans who find the zero-days.

Guard is engineered by the operators behind MITRE ATT&CK and CWE contributions, a standing responsible zero-day disclosure program, the Microsoft Bug Bounty Hall of Fame, and a National Counterintelligence Award. The AI scales their craft; it never replaces their judgment. Many of the agents — Pius, Brutus, Augustus, Nerva and more — are open-sourced.

You'd never run your EDR two weeks a year. Why run offensive security that way?

See Guard against your own environment. Start with a free, two-week proof of value on your live attack surface — enumerate assets, surface real vulnerabilities, and watch an autonomous fleet find the breach path before someone else does.

praetorian.com/guard