HADRIAN

Hadrian Nova vs Frontier AI Models

THE QUESTION

Frontier AI models like Anthropic's Claude Mythos Preview have demonstrated unprecedented capabilities in vulnerability discovery and exploitation. Thousands of zero-days found across major operating systems and browsers. Working exploits generated with minimal human intervention. The industry is paying attention.

The natural question: if a frontier model can do all that, what role does an agentic pentesting platform play? The answer comes down to the difference between a model and a system.

WHAT FRONTIER MODELS DO WELL

Frontier AI models are raw intelligence. They can reason about code, hypothesize about vulnerabilities, chain techniques together, and write working exploits. Under ideal conditions, with whitebox access, full source code, and a researcher guiding the process, the best models are the most capable vulnerability discovery engines ever built. Each generation gets meaningfully better than the last.

But a model is not a pentest. It doesn't know your attack surface. It doesn't manage scope, credentials, or rate limiting. It doesn't triage findings, track remediation, or produce compliance-ready reports. It doesn't integrate with your ticketing systems. And critically, it doesn't validate its own output. Independent analysis has consistently shown that even the most capable models hallucinate plausible vulnerabilities in patched, correct code. Raw model output requires skilled humans to separate signal from noise.

WHAT NOVA IS

Nova is a platform built for production security testing. It wraps AI capability in everything needed to deliver a real penetration test:

- **Scope management**

  Define the target, exclude sensitive areas, configure rate limiting, provide application context and credentials for multiple roles.

- **Orchestration**

  A hierarchy of specialized agents coordinated by an Orchestrator that allocates effort dynamically, concentrating on the highest-impact vectors and deprioritizing dead ends.

- **Attack surface context**

  Every test starts with what Hadrian already knows about your external exposure: asset inventory, technology stack, configuration history. Agents go in informed, not blind.

- **Human validation**

  Hadrian's offensive security specialists steer the investigation and validate every finding. Zero false positives.

- **Compliance-ready output**

  A full penetration test report with executive summary, methodology, risk ratings, reproduction steps, and remediation guidance. Satisfies SOC 2, ISO 27001, NIS2, and DORA.

- **Remediation workflow**

  Quick assessment and reporting on the organization's security posture and specific risks.

- **Comprehensive integrations**

  Findings flow into Jira, ServiceNow, and your security stack. Retesting is built in at no extra cost.

THE REAL COMPARISON

DIMENSION	FRONTIER AI MODEL	HADRIAN NOVA
WHAT IT IS	General-purpose AI model with security capabilities	Purpose-built agentic pentesting platform
SCOPE MANAGEMENT	None. Requires external scaffolding and expertise	Built in. Define target, exclusions, rate limits, credentials
ATTACK SURFACE AWARENESS	None. Starts from scratch each time	Seeded with Hadrian's continuous external exposure intelligence
FALSE POSITIVE HANDLING	Model output requires manual validation. Hallucination risk well-documented	Zero false positives. Every finding validated by human offensive security team
OUTPUT	Raw findings. No structured report	Full compliance-ready penetration test report
REMEDIATION INTEGRATION	None	Jira, ServiceNow, API, Hadrian Atlas
COMPLIANCE MAPPING	None	SOC 2, ISO 27001, NIS2, DORA
EXPERTISE REQUIRED TO OPERATE	High. Needs scaffolding, guidance, and validation by security professionals	Low. Configure scope and launch. Results in hours

BETTER MODELS MAKE BETTER PLATFORMS

This is not a competition between AI models and pentest platforms. As frontier models improve, platforms like Nova get smarter. Nova's architecture is designed to incorporate advancing AI capabilities while maintaining the orchestration, validation, and compliance infrastructure that turns raw intelligence into trustworthy, actionable results.

The question isn't whether frontier AI can find vulnerabilities. It can. The question is whether your organization has the system around it to turn that capability into a defended infrastructure: scoped, validated, reported, tracked, and remediated.

ABOUT HADRIAN

Hadrian is an external exposure management provider that pioneered the AI attacker's perspective approach. Its agentic engine offers frictionless always-on discovery, validation, and mobilization of organization's most critical cyber risks. Trained by elite hackers with top offensive knowledge, Hadrian adapts the organization's unique environment to continuously probe, discover and validate the risks that attackers can actually exploit.

TO LEARN MORE: HADRIAN.IO

REQUEST A DEMO