COMPANY > BLOG

Independent tests reveal F5 provides strong protection against traditional, emerging, and targeted AI app attacks

F5 ADSP | MARCH 23, 2026

JOHN MADDISON
Chief Marketing Officer | F5

For years, the industry narrative was simple: the future of apps is 100% cloud. But operational, regulatory, cost, and performance needs pushed enterprises toward a more pragmatic approach—one that is hybrid and multicloud. Enterprises now have apps, APIs and workloads split across multiple public clouds, essential on-premises data centers and edge deployments. They are embracing more AI-driven services that are by their very nature distributed.

This landscape has increased the cyberthreat surface exponentially. APIs, for example, are growing in usage and importance to modern apps, but bring substantial security risk with them. As organizations adopt microservices and event-driven designs, API sprawl increases, leaving visibility gaps that attackers can exploit.

Meanwhile, bots have grown highly sophisticated, often mimicking human behavior to commit fraud, scrape data, or overwhelm systems. AI has shattered the skill floor for novice attackers and expanded the ceiling for the most sophisticated threat actors. The words and prompts flowing between users, agents, and models create an entirely new attack surface for social engineering in the form of prompt injection and net-new jailbreak techniques.

Reducing the risk and complexity of today's modern app landscape

Now more than ever, there is a universal need across all organizations for greater consistency, visibility, and security across all environments. This need requires heavily on a platform-centric approach that reduces reliance on numerous, disparate delivery and security point products that lack centralized control and management.

At F5, our mission is simple: help customers deliver extraordinary digital experiences that are fast, available, and—above all—secure. That mission comes to life through our **Application Delivery and Security Platform (ADSP)**, a unified approach designed to simplify and strengthen security across increasingly complex environments. F5 ADSP converges industry-leading capabilities spanning application, API, and AI security, zero trust access, bot defense, performance optimization, and multicloud networking—in a single, extensible platform built for the realities of today's threat landscape.

> **"SecureIQLab determines F5 AI Guardrails achieves a security score of 98.36%."**

SecureIQLab: F5 AI Guardrails Security Efficacy Validation Test

The power of third-party validation

Security requires confidence and resilience. That's why F5 regularly participates in independent third-party testing, to validate the effectiveness and resiliency of our security services under real-world like conditions.

These evaluations expose products to adversarial scenarios that mirror actual attacks—bot campaigns, API abuse attempts, DDoS events, application exploits, and model manipulation scenarios like prompt injection, excessive agency, and more. Independent testing gives customers confidence that our solutions perform not only in controlled lab conditions but also in complex, stress-tested environments that resemble their own hybrid and multicloud production deployments.

We have recently conducted two security evaluations with the independent security testing laboratory SecureIQLab to test our **F5 AI Guardrails** and **F5 Web Application and API Protection (WAAP)** solutions.

F5 AI Guardrails achieves security score of 98.36%

Across 10 tested threat categories, F5 AI Guardrails successfully blocked 19,356 of 19,679 adversarial payloads.

The strongest results came in the categories that represent the highest volume of real-world AI misuse:

- Toxic output protection: 99.72%
- Harmful content and bias: 99.67%

- Prompt injection: 99.33%
- Sensitive data leakage: 99.01%

“Excessive agency,” the category that evaluates whether a solution can enforce operational boundaries and block unauthorized tool and API access in agentic deployments, came in at 98.68%. That result matters in context: in our **SecOps sentiment research,** concerns about excessive agency grew 13x over the past six months, making it one of the fastest-rising threats on security teams' radar.

F5 WAAP + F5 AI Guardrails earn a total security score of 97.09%

SecureIQ Lab ran roughly 40 test cases, encompassing approximately 2,000 attacks and benign payloads, with enterprise traffic running in the background to recreate a typical customer environment. The testing lab determined that F5 provides strong protection against traditional and emerging application attacks, automated bot activity, and application-layer DoS techniques, with a total security score of 97.09%.

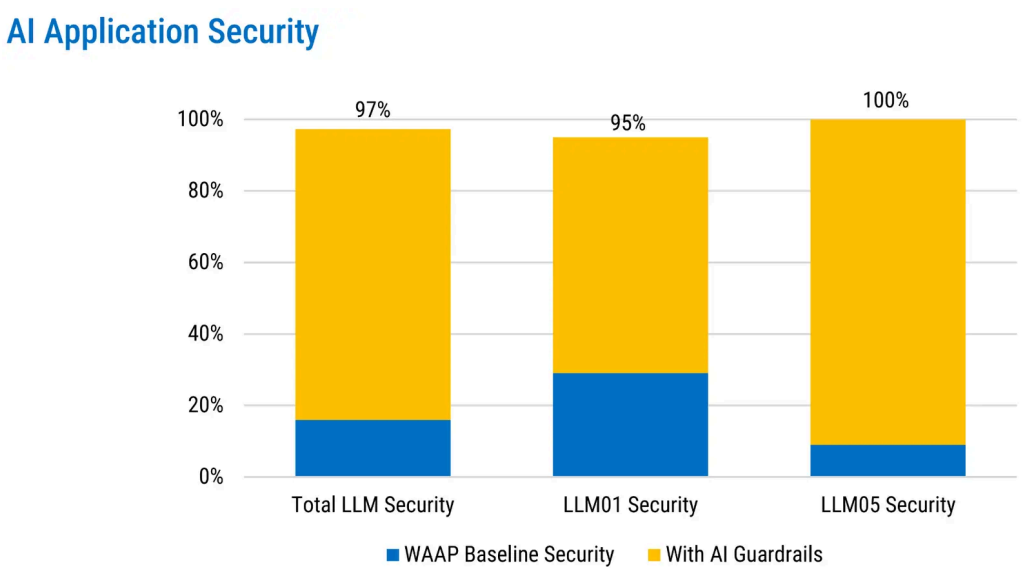
Here are just some of the notable findings from the report. F5 received a perfect score, protecting with 100% accuracy, for the following OWASP WAF and OWASP API Top 10:

- A 02:2025 Security Misconfiguration
- A 03:2025 Software Supply Chain Failures
- A 05:2025 Injection
- A 06:2025 Insecure Design
- A 07:2025 Authentication Failures

- A 10:2025 Mishandling of Exceptional Conditions
- API1:2023 Broken Object Level Authorization
- API2: 2023 Broken Authentication
- API3:2023 Broken Object Property Level Authorization
- API6:2023 Unrestricted Access to Sensitive Business Flows
- API7:2023 Server-Side Request Forgery
- API9:2023 Improper Inventory Management
- API10:2023 Unsecure Consumption of APIs

F5 WAAP also earned a perfect score of 100% for “Bot Attack Mitigation” and “Layer 7 DoS Protection.”

When AI Guardrails was integrated into the security architecture, protection against AI-specific threats increased dramatically, improving LLM security effectiveness from 19% to 97%. Additionally, 100% of LLM05 security attacks (“improper output handling” in OWASP Top 10 for LLM applications) were addressed.



F5 AI Guardrails, added to F5 WAAP, blocks 97% of all LLM attacks—including 95% of prompt injection (LLM01) and 100% of improper output handling (LLM05) attacks.

View the reports

SecureIQLab evaluations help us at F5 to continually refine our capabilities. We learn from these rigorous assessments so we can ensure our solutions remain aligned to real-world customer challenges—not theoretical ones. SecureIQLab's methodology aims to set a global standard for WAAP testing, providing enterprise security leaders with transparent, actionable, and 100% vendor-neutral WAAP data. (Learn more about the SecureIQLab methodology on the organization's website).

SecureIQLab testing demonstrates that F5 WAAP and AI security capabilities provide strong protection against traditional and emerging application attacks—delivering OWASP WAF and API Top 10 protection, advanced application protection from malicious bots and Layer 7 DoS attacks, and protection from attacks targeted at AI applications.

We invite you to read the full reports and reach out to our team to further evaluate how our solutions can help you improve your security posture:

SecureIQLab F5 WAAP 2026 Test Briefing
SecureIQLab: F5 AI Guardrails Security Efficacy Validation Test

Featured Blog Posts

Three things every CISO should know about API security →

F5 completes acquisition of CalypsoAI, introduces F5 AI Guardrails and F5 AI Red Team →

F5's announcement to acquire CalypsoAI builds towards TRiSM framework →

Tags: AI Security, Web Application & API Protection

About the Author

John Maddison
Chief Marketing Officer | F5

John Maddison is F5's Chief Marketing Officer. John has more than 30 years of experience leading product teams and corporate marketing functions in cybersecurity, cloud, and telecommunications. Prior to F5, John spent 12 years at Fortinet, most recently as Chief Marketing Officer and EVP of Product Strategy. He previously held leadership positions at Trend Micro, Vina Communications, and Octel Communications. John holds a B.S. in Engineering degree from Plymouth University in Devon, England.

More blogs by John Maddison →