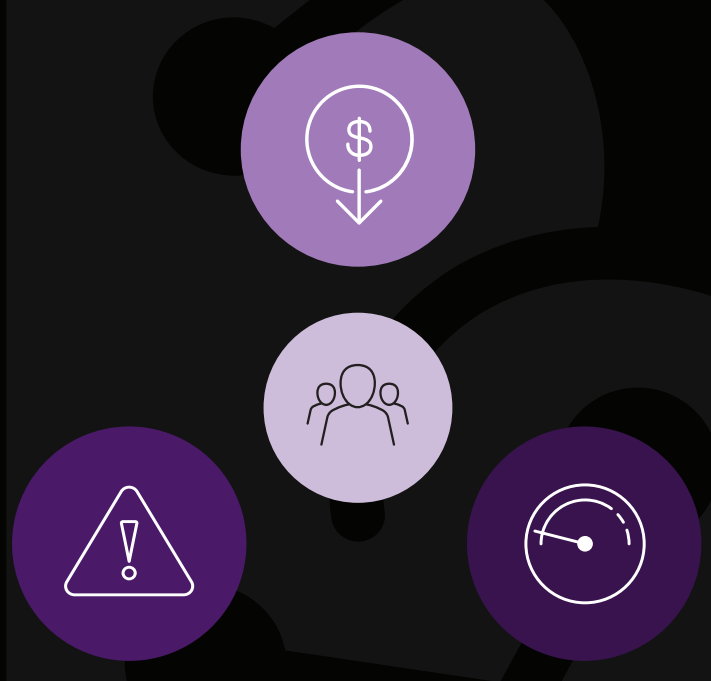
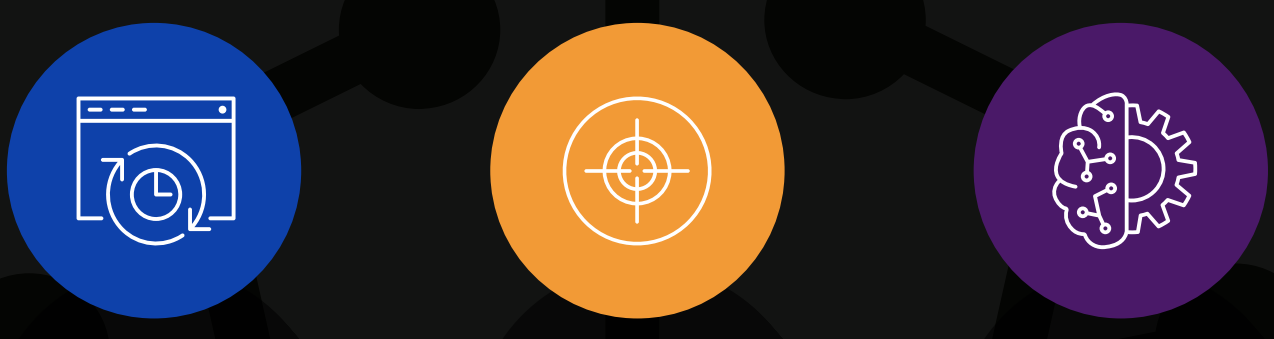
How to Mitigate 5 Top AI Application Risks

Business leaders deploying AI have identified the cost of compute, security risks, and performance as their top concerns.[1] These challenges are interconnected, as security vulnerabilities in AI models and applications can dramatically impact both costs and performance.



F5® AI Gateway addresses these concerns with real-time traffic inspection, automated threat detection, and intelligent resource management to secure and optimize AI apps.



Real-time traffic inspection

Automated threat detection

Intelligent resource management

Use AI Gateway alone or with other F5 application delivery and security solutions to mitigate these key large language model (LLM) risks.

01

Prompt injection

User inputs, whether malicious or accidental, can result in the AI model being manipulated to reveal information, grant access, or change its behavior.

AI Gateway protects against prompt injection attacks by examining both inputs and outputs in real time to detect potential manipulation attempts, block known malicious prompts, and validate input formatting. Through customizable security policies, organizations can also implement their own prompt validation rules and restrictions.

02

Sensitive information disclosure

An LLM's data is at risk of exposure—not just training data but also personally identifiable information (PII) submitted by users.

Reduce the risk of a leak by sanitizing data. AI Gateway scans prompts for sensitive data patterns, including PII, financial information, and confidential business data, preventing this information from reaching AI models. It also analyzes AI responses to detect and redact any sensitive information that might be included in model outputs.

03

Improper output handling

Insufficient validation of LLM outputs creates vulnerabilities that can be exploited through cross-site scripting (XSS) attacks, privilege escalation, or remote code execution.

As part of a zero trust strategy, AI Gateway acts as a security checkpoint to validate LLM outputs before responses reach applications or users. It inspects inbound and outbound traffic for potentially harmful content like embedded scripts, unauthorized commands, or escalation attempts.

04

Misinformation

A key cause of LLM misinformation is hallucination—fabricating content. It also occurs when users place excessive trust in LLM-generated content.

By analyzing AI app outputs, AI Gateway helps prevent hallucinated responses or other misinformation from reaching users. Organizations can add custom validation rules to align with specific business requirements and accuracy standards for different types of AI interactions.

05

Unbounded consumption

When an AI app allows users to conduct excessive and uncontrolled inferences, it can result in denial of service (DoS), slowdowns, or high compute costs.

Rate limiting and load balancing as part of intelligent traffic management in AI Gateway prevent unbounded consumption. Real-time monitoring through the native OpenTelemetry integration identifies potential performance issues, while semantic caching further reduces compute costs and AI token usage.

F5 is ready to help you secure and optimize your AI applications.

To learn more, visit f5.com.

Sources

1. F5, State of AI Application Strategy Report, Jun 2024

DC March 18, 2025 11:06 AM | JOB-CODE-123456789