SOLUTION OVERVIEW

# Data leakage detection and prevention for enterprise AI models and applications

Safeguard sensitive data in AI workflows with the F5® Application Delivery and Security Platform to enable real-time detection, policy enforcement, and compliance at scale.

**Key features**

**Scalable AI governance**
Continuously audit and monitor logs and transmissions to meet regulatory compliance requirements.

**Robust security controls**
Enforce role-based access control (RBAC), credential management, and secure authentication with detailed audit logs for oversight.

**Inline data leak protection**
Inspect and mitigate risks in prompts and responses by detecting, blocking, or redacting sensitive data in real time.

**Processor labs program**
Access pre-release processors (AI security guardrails) through the processor labs program.

**Python-based SDK**
Develop and integrate custom inspection, routing, and policy enforcement logic using the processor framework, an extensible SDK for implementing custom AI guardrails.

# The challenge of protecting and detecting data leakage in the AI era

The widespread adoption of AI models across organizations of all sizes has introduced a complex set of challenges in protecting sensitive data as it traverses increasingly distributed and dynamic environments. Unlike deterministic outputs from traditional APIs, AI workflows generate non-deterministic and context-dependent results, such as natural language responses or unstructured data, which makes classifying and protecting sensitive information far more difficult during transit. These workflows involve continuous exchanges between applications, APIs, and models, creating intricate, real-time traffic patterns that change unpredictably based on user interactions, roles, and credentials. Traditional security tools, designed for static systems, lack the capability to inspect these real-time transactions or enforce granular, context-sensitive policies at the point of interaction. As a result, sensitive information can be unknowingly exposed, highlighting critical blind spots and increasing risks of data leakage.

Regulatory compliance further amplifies these challenges. Frameworks such as GDPR, HIPAA, and other data protection mandates require organizations to classify, audit, and tightly control sensitive data throughout its lifecycle, including as it moves through complex AI workflows. However, the dynamic and decentralized nature of AI-driven interactions creates significant visibility gaps. Many organizations lack real-time tools to track which data is being accessed, processed, or shared during AI operations, particularly as these workflows span multiple models, APIs, and third-party integrations. Legacy classification engines and data leakage prevention tools were designed for static, predictable environments and struggle to keep pace with the high-speed, context-dependent, and unpredictable nature of AI systems. This misalignment leaves critical blind spots where sensitive data can be exposed without detection or control—putting organizations at risk of compliance violations and data breaches.

To effectively mitigate these challenges, architects, developers, and security teams require a solution capable of deep inspection, dynamic policy enforcement, and real-time data classification without compromising operational performance or scalability. Failing to address these gaps will lead to escalating risks ranging from internal data leaks to regulatory non-compliance and operational disruptions. As organizations continue to adopt AI at scale, securing these workflows is no longer optional but imperative for long-term resilience and success.

**Key features**

**Audit logging**
Track detailed audit trails of prompts and metadata mutations to ensure compliance and enable debugging.

**Upstream retry and failover**
Maintain service availability by retrying failed requests or switching traffic to backup models automatically.

**Enterprise-grade secrets management**
Integrate smoothly with industry-standard Kubernetes secrets management solutions.

**Configuration validator**
Preempt errors with pre-deployment checks for smoother and more reliable rollouts.

# A new standard for securing AI workflows with inline data protection and governance

To address the pressing challenges of securing AI-driven workflows, F5® AI Gateway introduces robust data protection capabilities to the F5 Application Delivery and Security Platform (ADSP), establishing it as a powerful enforcement point for safeguarding AI models. By leveraging advanced features, the F5 ADSP integrates proprietary data classification technology directly into the data path. This functionality allows organizations deploying enterprise AI applications to inspect every AI prompt and output in real time, enabling the identification and management of sensitive information such as personally identifiable information (PII), protected health information (PHI), and source code during inference. With highly configurable policies, enterprises can automatically redact or block sensitive content—ensuring that data always remains secure and under their control.

AI Gateway addresses critical blind spots in AI workflows by securing interactions at the point of inference. This allows organizations to enforce guardrails for both application-layer and AI-layer communication. Sensitive data protection prevents both inadvertent and malicious exposure of sensitive data. Furthermore, the platform integrates seamlessly with existing workflows by streaming enriched detection results to SIEM and SOAR systems. This empowers security teams to respond rapidly to incidents while demonstrating auditability to meet regulatory requirements.

These capabilities extend beyond basic detection and mitigation, offering essential tools that scale effectively across distributed AI environments. Features such as enhanced prompt logging provide detailed audit trails of interaction history, while RBAC ensures that permissions are strictly enforced across different environments. As a result, organizations can safely evaluate emerging functionality and implement custom workflows.

Additionally, AI Gateway offers enterprise-grade enhancements, including upstream retry and failover integrations, advanced secrets management, a Python-based SDK for customization, and configuration validation tools. These features enable enterprises to securely extend and optimize their AI deployments without compromising visibility, performance, or compliance. By providing these comprehensive security measures, AI Gateway empowers organizations to confidently innovate in the rapidly evolving landscape of AI-driven applications.

## Protecting sensitive data in AI workflows

**Organizations can mitigate critical risks like data leakage and compliance gaps without sacrificing performance.**

The convergence of AI innovation and increasing data protection requirements demands solutions that provide real-time visibility, control, and enforcement. By integrating advanced sensitive data protection functionality into AI Gateway as part of the broader F5 ADSP, organizations gain a unified security architecture designed for modern, AI-driven workflows. This solution empowers enterprises to mitigate critical risks like data leakage and regulatory non-compliance while maintaining the performance, scalability, and flexibility needed to drive AI adoption.

As organizations continue to adopt AI at scale, AI Gateway delivers the tools needed to protect sensitive data, ensure compliance with regulatory frameworks, and support enterprise-level innovation. With inline detection, dynamic governance, and seamless operational integration, the F5 ADSP enables teams to confidently secure their AI-powered applications and infrastructure, positioning them for a future where both AI and data integrity coexist to fuel sustainable growth.

Deploy AI applications anywhere with observability, security, and cost containment for your organization.

**Learn more at f5.com/products/ai-gateway.**


DC 06.2025 | JOB-CODE-OV-1682213214