CROWDSTRIKE

CrowdStrike Products

Data Sheet

Falcon AI Detection and Response

Secure the AI attack surface with AI detection and response (AIDR) from the pioneer in detection and response for endpoint, cloud, and managed services

AI now spans every part of the digital estate, creating a rapidly expanding attack surface that includes a new and uniquely exposed domain: the prompt and agentic interaction layer that traditional security tools were never designed to monitor or control.

Employees often use AI tools without IT's knowledge, creating a shadow AI gap where sensitive data may be exposed. At the same time, engineering teams are developing homegrown AI models, agents, and workflows and securing them via open-source or DIY guardrails that introduce risk and can slow product delivery.

One platform. Unified AI security.

CrowdStrike Falcon® AI Detection and Response (AIDR) redefines AI security with comprehensive protection for both employee adoption of generative AI (GenAI) tools and runtime security for homegrown AI development. Built on the AI-native CrowdStrike Falcon® platform, Falcon AIDR provides unified visibility, real-time threat detection, data protection, access controls, and automated response across endpoints, applications, AI agents, Model Context Protocol (MCP) servers, AI/API gateways, and cloud environments — all managed through a single sensor and console.

By consolidating AI security into a single platform, Falcon AIDR accelerates AI adoption and innovation without compromising security or adding complexity.

Key benefits

- **Unified AI visibility and control:** Gain real visibility into how employees use AI and how AI agents operate. Map relationships between users, prompts, models, agents, and MCP servers with runtime logs for compliance, investigations, and continuous monitoring.
- **Real-time threat detection:** Detect and stop direct and indirect prompt injection, unsafe content, and attempted model manipulation in real time. Leverage global telemetry and CrowdStrike's analysis of over 300,000 adversarial prompts and 180+ prompt-injection techniques for high detection efficacy and low latency.
- **Comprehensive data protection:** Prevent sensitive data from being shared with models, agents, or external AI systems. Automatically identify and block confidential information, including credentials and regulated data, before it's uploaded, shared, or processed.

Securing workforce adoption and development

As employees adopt AI tools for productivity and organizations build agents and infrastructure with AI, Falcon AIDR provides consistent, real-time protection.

Secures Workforce Use of AI

Protect employee adoption of GenAI tools by enabling visibility, providing access controls to enforce AI governance policy, and delivering detection and response capabilities to defend against AI-related breaches and attacks on AI. Falcon AIDR illuminates shadow AI by discovering employee AI tool use across browsers, applications, and cloud services. Organizations can use Falcon AIDR to block sensitive data leaks to third-party large language models (LLMs) and enforce employee AI access policy by applying granular controls based on user, application, model, and context.

Secures AI Development

Protect AI-enabled software and infrastructure developed by organizations, including agentic AI platforms and AI factories, to enable scalable AI innovation across teams and business units. Falcon AIDR stops prompt injection and jailbreak attacks, prevents sensitive data leaks to LLMs and users, ensures AI behavior conforms to organizational policies and ethical guidelines, and generates comprehensive audit trails of AI runtime behavior for compliance, forensics, and continuous improvement.

Key capabilities

Illuminate Shadow AI and Enforce Governance

- **Discover shadow AI:** Gain complete visibility into employee AI tool usage across browsers (e.g., Chrome, Edge, Firefox), applications, AI/API gateways (e.g., LiteLLM, Kong, Apigee, Azure API Management), MCP servers, and cloud environments (e.g., AWS).
- **Map AI relationships:** Visualize connections between users, prompts, models, agents, and MCP servers to understand AI usage patterns and identify risky behaviors.
- **Enforce attribute-based access controls:** Apply granular AI security policies across users, agents, tools, and models based on user, application, model, use case, and custom attributes — without introducing friction to workflows.
- **Capture runtime logs:** Maintain comprehensive audit trails for compliance, forensics, and continuous monitoring of AI interactions.

Detect and Stop AI-Specific Threats

- **Block prompt injection attacks:** Detect and prevent direct and indirect prompt injection, jailbreaks, and model manipulation attempts using CrowdStrike's industry-leading taxonomy of adversarial prompt techniques.
- **Identify malicious entities and indicators of compromise (IOCs):** Detect malicious URLs, domains, IP addresses, and entities within prompts and AI responses in real time.
- **Validate MCP interactions:** Monitor and validate MCP server communications to prevent unauthorized tool execution and agent behavior manipulation.
- **Detect harmful content:** Identify and block toxic content, negative sentiment, self-harm and violence, weapons, criminal conduct, sexual content, and other harmful topics across prompts and AI responses.
- **Enforce content alignment:** Ensure prompts and responses conform to organizational policies, with topic detection for financial advice, legal advice, religion, politics, health coverage, roleplay, and custom topics.

Protect Sensitive Data Across AI Interactions

- **Prevent data leaks to third-party LLMs:** Automatically detect and block confidential information, personally identifiable information (PII), secrets, keys, and regulated data before exposure to AI models.
- **Redact sensitive information:** Apply multiple redaction methods — replacement, masking, partial masking, hash, and format-preserving encryption (FPE) — to protect data while preserving AI workflows.
- **Detect code and competitors:** Identify code in 26 programming languages and mentions of competitors to prevent intellectual property leakage.
- **Create custom entity detectors:** Build custom detectors for organization-specific sensitive data types and enforce tailored protection policies.

Respond and Investigate Faster

- **Enforce real-time policy actions:** Block unsafe interactions, report violations, and transform (redact or encrypt) sensitive data automatically across browsers, applications, gateways, and MCP proxies.
- **Integrate with next-gen SIEM:** Stream Falcon AIDR findings directly to CrowdStrike Falcon® Next-Gen SIEM for unified security operations and cross-domain correlation.
- **Visualize AI activity:** Leverage interactive Sankey diagrams to trace attribute-based relationships, monitor metrics, track policy detections, and view active collectors in real time.
- **Accelerate investigations:** Provide security teams with detailed context around AI events — including user, application, model, prompt content, and detector findings — for faster, evidence-backed response.
- **Track usage and compliance:** Monitor AI usage measured by API calls, payload size, and activity tracking, and maintain alias mapping for human-readable reporting.

Enable Comprehensive Collection and Enforcement

Falcon AIDR provides flexible deployment options to protect AI wherever it runs:

- **Browser extensions:** Chrome, Edge, Firefox, and Atlas for employee AI tool usage
- **Application instrumentation:** Software development kits (SDKs) for Python, Node.js, Go, Java, and C# applications, and supports OTel for application logging
- **AI/API gateway integrations:** LiteLLM, Kong, Apigee, and Azure API Gateway
- **MCP proxy:** Secure MCP interactions via stdio transport
- **Cloud log analysis:** AWS logs (S3)

Schedule an AIDR demo →

About CrowdStrike

CrowdStrike (Nasdaq: CRWD), a global cybersecurity leader, has redefined modern security with the world's most advanced cloud-native platform for protecting critical areas of enterprise risk — endpoints and cloud workloads, identity and data.

Powered by the CrowdStrike Security Cloud and world-class AI, the CrowdStrike Falcon® platform leverages real-time indicators of attack, threat intelligence, evolving adversary tradecraft and enriched telemetry from across the enterprise to deliver hyper-accurate detections, automated protection and remediation, elite threat hunting and prioritized observability of vulnerabilities.

Purpose-built in the cloud with a single lightweight-agent architecture, the Falcon platform delivers rapid and scalable deployment, superior protection and performance, reduced complexity and immediate time-to-value.

CrowdStrike: We stop breaches.

Learn more: https://www.crowdstrike.com/
Start a free trial today: https://www.crowdstrike.com/free-trial-guide/