seceon
SECURITY FOR EONS

Seceon aiTRiSM360

AI Activity Monitoring & Protection

Endpoint visibility for generative AI usage on Windows, part of the Seceon EDR / aiXDR stack

At a Glance

- Windows endpoint agent for generative AI activity across browser tabs, extensions and desktop AI apps, extensions, and desktop AI apps
- Detects file uploads, clipboard paste, AI sessions, network access, and prompt security risks
- AI-only scope filters out non-AI destinations (Teams, Drive-only uploads, collaboration tools)
- Telemetry flows to Seceon aiXDR as AI Security events
- Privacy-first: file names and metadata only; prompt hash and tags only, never plaintext payloads, never plaintext payloads
- Flags shadow AI, including unsanctioned chat, code, and local AI tools such as DeepSeek, Ollama, and LM Studio.

Product Overview

aiTRiSM 360 observes activity on the host and sends structured telemetry to the Seceon platform for detection, reporting, and policy workflows. It gives security and IT teams visibility into how users interact with generative AI services such as ChatGPT, Copilot, Perplexity, Claude, Cursor, and dozens of others.

Highlights

BROWSER, EXTENSION AND DESKTOP AI COVERAGE

Sees AI use in browser tabs, extension side panels and background traffic, and desktop AI apps such as Cursor, the ChatGPT app and Ollama.

PROMPT SECURITY DETECTION.

Flags prompt injection, jailbreak, system prompt extraction, data exfiltration, coercion, obfuscation, and authority impersonation.

PRIVACY-FIRST TELEMETRY.

Collects metadata, a prompt hash and security tags only. Prompt text, clipboard contents and file contents are never captured by default.

Native aiXDR Events.

Ten AI Security event types, each mapped to MITRE ATT&CK, flow straight into Seceon aiXDR for detection, reporting and SOAR response.

What Problem It Solves

Organizations need to know when employees:

- Upload files or paste sensitive data into AI services
- Use AI through browser tabs, browser extensions, or desktop agents
- Maintain long-running AI sessions that may involve data exfiltration or policy violations

ADMP answers who, on which machine, used which AI service, how (upload, paste, session, network), and what type of content was involved without capturing AI prompt text or clipboard/file payloads by default.

Detection approach: OS-level signals (file dialogs, keyboard hooks, window titles, UI Automation) plus optional TCP flow observation to known AI API endpoints. ADMP does not perform full packet capture or TLS decryption.

Core Capabilities

Capability	Component	Trigger
File upload to AI	BrowserUploadWatcher	OS file picker (#32770) detected while an AI tab or application is active
Drag-and-drop upload	DragDropUploadWatcher	Files dropped onto an AI browser surface (enabled by default)
AI session tracking	AiSessionWatcher	AI tab or desktop agent opens, progresses, or ends a session
Clipboard paste to AI	AiPasteWatcher	Ctrl+V or Shift+Insert detected on an AI input surface
Prompt injection / jailbreak detection	AiSubmitWatcher + PromptSecurityInspector	Prompt submission via Enter key or suspicious paste activity (enabled by default)
Network-only AI usage detection	AiNetworkWatcher	TCP connections to known AI API hosts, including extensions and background tabs
Browser extension awareness	BrowserAiContextResolver	Detection of extension pages, side panels, or installed AI browser extensions

How It Works

File Upload Detection

- Windows Open File dialog when the user attaches a file in an AI web app
- ADMP hooks the dialog creation and identifies the owning browser or desktop AI process

- Resolves destinations from the address bar, title, or extension context
- Only AI-tagged destinations are recorded; collaboration apps are filtered out
- Ctrl+V / Shift+Insert on AI surfaces records format, size, and security tags

Session, Paste, Network, and Prompt Security

- Session: polls visible AI tabs/apps; emits start, progress, and end events
- Paste: Ctrl+V / Shift+Insert on AI surfaces format, size, and security tags; never plaintext
- Prompt security: Enter submit and suspicious paste scanned for injection, jailbreak, extraction, exfiltration, coercion, obfuscation, and authority impersonation (on by default)
- Network: matches TCP destinations to the forward-DNS cache of AI API hostnames



Browser vs Extension vs Desktop App

Access Mode	How ADMP Knows	Event Signal
Normal browser tab	URL bar or window title analysis	browser_extension omitted
Extension side panel or popup	Title hint detection or chrome-extension:// URL identification	browser_extension: true
Extension background traffic	AiNetworkWatcher combined with AI API hostname detection	browser_extension: true
Desktop AI application (e.g., Cursor, ChatGPT app)	Process name recognition and canonical hostname mapping	browser: null



Supported AI Destinations (Examples)

Category	Services
Chat / Search AI	ChatGPT, Gemini, Claude, Perplexity, Copilot, Bing Chat, You.com, Poe, Phind, DeepSeek
Code AI	GitHub Copilot, Cursor, Windsurf, Codeium, Tabnine, Antigravity
Creative / Other	Midjourney, Stability AI, Runway, Jasper, Copy.ai, Writesonic, Notion AI
Local / Desktop	Ollama, LM Studio, Msty

Out of scope for AI ingestion: cloud storage-only uploads (Drive, OneDrive), code hosts, and collaboration tools (Teams, Slack, Zoom).

Platform Events (aiXDR)

Event ID	Event Type	MITRE ATT&CK	When
1286032	File Upload to AI	T1567.002	Triggered when a file name is present in the event
1286033	Network Access via AI	T1071.001	Triggered during AI session or network activity without file interaction
1286034	Clipboard Paste to AI	T1115	Triggered during the paste phase or when the Clipboard subtag is present, without security tags
1286039	Prompt Injection Attempt	T1059	Triggered when content_tags includes prompt_injection
1286040	AI Jailbreak Attempt	T1059	Triggered when content_tags includes jailbreak
1286041	AI System Prompt Extraction Attempt	T1552	Triggered when content_tags includes system_extraction
1286042	AI Data Exfiltration Prompt	T1213	Triggered when content_tags includes data_exfiltration
1286043	AI Prompt Coercion Attempt	T1059	Triggered when content_tags includes coercion
1286044	AI Obfuscated Prompt Attempt	T1027	Triggered when content_tags includes obfuscation
1286045	AI Authority Impersonation Prompt	T1656	Triggered when content_tags includes authority_impersonation

When multiple content_tags match, platform routing priority is: jailbreak → data_exfiltration → system_extraction → authority_impersonation → coercion → obfuscation → prompt_injection. All tags remain in additionalinfo.content_tags.

All ADMP events: category AI Security, generator admp, origin Host.

Key additionalinfo Fields

Field	Meaning
domain	AI website hostname or service domain
app_name	Friendly service name (for example, Perplexity AI)
browser	Browser used for access such as Chrome, Edge, or Firefox; null for desktop AI applications
browser_extension	true when access occurred through a browser extension
act	Activity phase, such as alert, telemetry, session_start, session_progress, session_end, paste, or submit
content_tags	Security classification tags including prompt_injection, jailbreak, system_extraction, data_exfiltration, coercion, obfuscation, and authority_impersonation
prompt_char_count	Length of the inspected prompt in characters; plaintext content is not stored
prompt_hash	SHA-256 hash of the prompt text; plaintext content is not stored
fname / fpath	Uploaded file name and file path, available only for upload events
file_type	Type of uploaded file, such as image, document, source_code, and others
out_bytes	Estimated number of bytes transferred or clipboard content size
duration	Total AI session duration in seconds

Example Telemetry (Agent JSON)


```
{
 "ts": "2026-06-05T14:22:11.612Z",
 "phase": "alert",
 "tag": "AI",
 "website": "www.perplexity.ai",
 "target": "Perplexity AI",
 "browser": "Chrome",
 "file": "report.pdf",
 "file_type": "document",
 "bytes_out": 4321789,
 "browser_extension": true
 }
```



Seceon aiTRiSM360 continuously monitors AI usage across users, applications, and endpoints to identify risks, enforce governance policies, and protect sensitive data using patented AI/ML-driven analytics.

Privacy and Data Minimization

Collected	Not Collected (by Default)
File name and file path	File contents
Clipboard format and size	Clipboard text or images
Content classification labels (e.g., PII, credentials)	Matched sensitive strings
Network 5-tuple and byte counts	TLS payloads or prompts
Prompt metadata (length and hash, enabled by default)	Prompt plaintext
Prompt security tags (prompt injection and jailbreak heuristics)	Prompt plaintext

About Seceon

Seceon simplifies cybersecurity for MSPs, MSSPs, and IT teams by reducing risk and complexity. Our AI and ML-powered aiSIEM and aiXDR platform provides comprehensive threat detection and response. By integrating data from logs, identity management, networks, endpoints, clouds, and applications, we deliver real-time, accurate alerts and automated remediation. Trusted by over 850 partners, Seceon supports more than 9,800 clients with efficient, high-margin security services and continuous compliance.

Learn more about Seceon

Schedule a Demo www.seceon.com/contact-us/



© 2026 Seceon Inc. All rights reserved. aiSIEM and aiXDR are trademarks and registered trademarks of Seceon Inc. in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners.