cyberhaven

The Insider Risk Incident *Response* Playbook

Detect, investigate, and contain insider risk before data leaves your organization

Section 01

Why Insider Risk Occurs

Insider risk is not an anomaly. It is a structural feature of how modern organizations operate. Sensitive data is created, shared, and acted on by employees every day.

The access that enables productivity is the same access that creates exposure.

Security teams cannot eliminate insider risk by restricting access further; that would make work impossible. The goal instead should be to understand why it occurs, identify the conditions that amplify it, and build a program that detects and contains it before damage is done.

39.7%	77%	90%+
of AI interactions involve sensitive data	more likely: office workers to exfiltrate vs. remote workers	false positive reduction with data lineage

Source: Cyberhaven Labs

The Three Root Causes

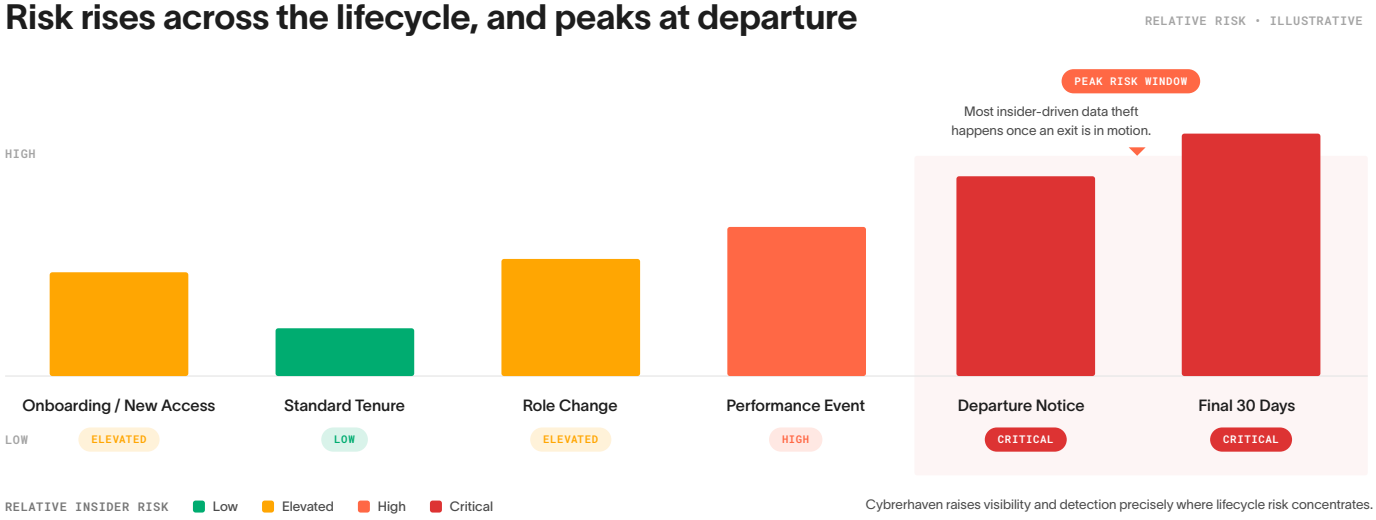
Insider incidents trace back to three underlying dynamics, each of which requires a different response from security teams.

Negligent or Unintentional Behavior	Most insider incidents are not malicious. Employees paste sensitive data into AI tools to work faster, email files to personal accounts to finish work from home, or misconfigure sharing settings without realizing the exposure they create. Intent is not the risk: the data movement is the risk.
Malicious or Financially Motivated Insiders	A smaller but higher-impact category involves employees who deliberately take data for personal gain, to benefit a future employer, or to cause harm. These incidents tend to be premeditated, happen over extended periods, and often involve data that is difficult to classify through traditional content inspection alone.
Departing Employees	The window surrounding an employee's departure concentrates risk. In the 30 days before and after a resignation notice, data movement activity rises sharply. Employees copy project files, customer data, and proprietary work to personal storage before access is revoked. Offboarding without behavioral monitoring creates a gap that traditional tools cannot close.

When Risk Is Highest: The Employee Lifecycle

Insider risk is not uniformly distributed across an employee's tenure. Three lifecycle stages account for the majority of data loss events.

Risk rises across the lifecycle, and peaks at departure



Cybrerhaven raises visibility and detection precisely where lifecycle risk concentrates.

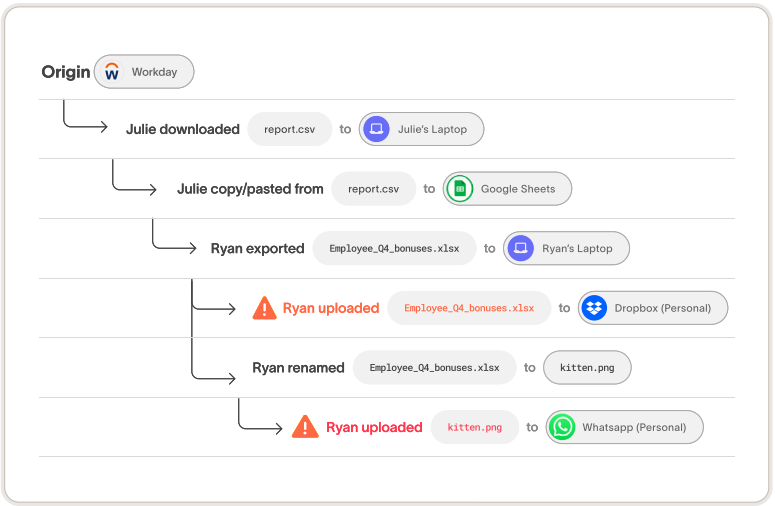
Why Traditional Tools Struggle

Legacy data loss prevention (DLP) tools were built for a world where data lived on defined endpoints and moved through predictable channels. Today, that model has broken down. Data moves through cloud applications, AI platforms, browser sessions, and agentic workflows in ways that file-hash-based or keyword-match approaches cannot track.

Traditional IRM tools face a related problem. **They focus on what users do, not what happens to data.**

A user accessing a file outside normal hours is an anomaly. Whether that file contained regulated customer data, a product roadmap, or routine correspondence makes all the difference in how seriously the incident needs to be treated.

Without content awareness and data lineage, security teams are left investigating alerts without the context to prioritize them. The result is analyst fatigue, delayed response, and incidents that only become visible after data has already left the organization.



Data lineage highlights provenance, user behavior, and data's movements across an enterprise.

Section 02

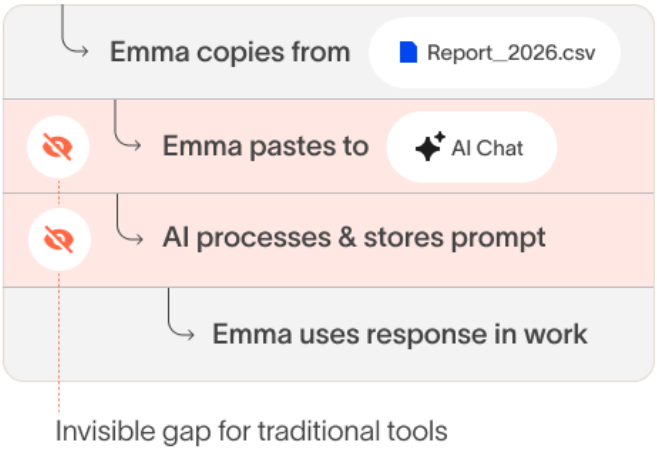
How Agentic AI Changes What Insider Risk Looks Like

Generative AI and agentic workflows have changed the threat model for insider risk in ways that most security programs have not yet adapted to. The shift fundamentally changes where data moves, how it moves, and who is accountable for it.

The Rise of the AI Insider Threat

An AI insider threat occurs when employees use generative or agentic AI tools in ways that expose sensitive corporate data. Unlike traditional insider threats involving intentional sabotage or theft, AI insider threats typically result from legitimate work activity. Employees are trying to work faster, not to cause harm.

The exposure emerges from what happens to data after it is submitted to an AI platform. Depending on the provider, that data may be stored, reviewed by humans, used to train future models, or accessible to other users under certain configurations. Once data has been submitted, organizations may lose all control over it.



Shadow AI: Uncontrolled AI Adoption at Scale

AI adoption has outpaced enterprise governance. Employees discover AI tools through social media or colleague recommendations, start using them without IT approval, and begin processing sensitive data before any policy is in place.

This unstructured, user-initiated data movement remains largely invisible to traditional security tools. Sensitive business data is copied and pasted into platforms that may store prompts, use them to train models, or expose them through third-party data sharing arrangements.

BY THE NUMBERS: AI ADOPTION RISK

Source: Cyberhaven Labs 2026 AI Adoption & Risk Report

300+
GenAI tools are used by organizations with the highest AI adoption rates

40%
of all AI interactions involve sensitive data

Most
employees using AI tools have no awareness of data retention policies

What Changes with Agentic Workflows

Agentic AI raises the stakes further. Traditional AI interactions are user-initiated: an employee decides to paste data into a prompt. Agentic workflows are automated: an AI agent acts on behalf of the user across multiple systems, making decisions and moving data without direct human oversight at each step.

This creates three new insider risk scenarios that did not exist in the pre-AI era:

Scenario	Why It Is Harder to Detect
Agent-initiated data access	Agents access data at machine speed across dozens of systems. Baseline behavioral patterns do not apply.
Prompt injection via document content	Malicious instructions embedded in documents direct AI agents to exfiltrate or alter data without user awareness.
Authorized agent, unauthorized destination	A user grants an agent access to send emails or upload files. The agent routes sensitive data to an unintended external destination.

What Data-Centric Security Looks Like in an AI-Native Environment

Monitoring user behavior is no longer sufficient. When AI agents act as autonomous intermediaries between users and data, the agent's behavior is indistinguishable from the user's at the behavioral layer. Security programs need visibility into what data was involved, not just which account was active.

Data lineage, which tracks the full lifecycle of a file from creation through every copy, rename, transformation, and transmission, provides the thread that connects user intent to actual data outcome. It is the only approach that remains meaningful when the "user" acting on data is an AI agent operating under delegated authority.

Why Behavior-Only IRM Falls Short in an AI Environment

User-activity monitoring cannot distinguish an AI agent's actions from the user's baseline

Behavioral anomaly detection requires historical baselines that do not exist for new AI tooling

Alert volumes from AI interaction monitoring create analyst fatigue without data context

Without content awareness, security teams cannot assess the severity of AI-related incidents

Section 03

Streamlining Insider Risk Investigations with Cyberhaven

Insider risk investigations are time-consuming, legally sensitive, and frequently inconclusive. The challenge is not detecting that something happened. It is determining what data was involved, where it went, whether it was intentional, and what the organization's exposure actually is.

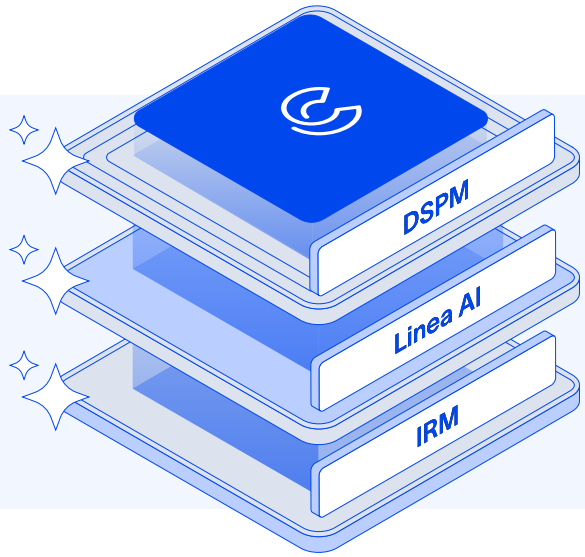
Cyberhaven approaches insider risk through the lens of data, not just user behavior. By combining complete data lineage, content inspection, and real-time enforcement, Cyberhaven gives security teams the context they need to investigate faster and respond with precision.

From Posture to Behavior: Connecting DSPM and IRM

DSPM tells security teams what sensitive data exists and where it is stored. Data lineage tells them what has happened to that data. Insider risk investigations require both.

A posture finding reveals a shared drive accessible to departing employees. Data lineage shows which files were accessed, copied, and sent to a personal cloud account the week after a resignation was submitted.

Without that connection, security teams can describe the risk but not the incident. With it, they can reconstruct the chain of events with enough precision for HR, legal, and executive review.



Data Discovery & Classification
Find and understand sensitive data across all locations and platforms

Data Lineage Tracking
Track how sensitive data moves and transforms across users, apps, and environments

Behavioral Context
Add user behavior, intent signals and lifestyle context to assess risk of data exposure

The Cyberhaven Advantage

Traditional IRM / DLP Approach	Cyberhaven Data-Centric Approach
Monitors user activity and metadata	→ Tracks the full lifecycle of every sensitive file
Loses track of data that is renamed, copied, or transformed	→ Lineage persists through renames, copies, and format changes
Alerts without blocking; blunt lockout as the only enforcement option	→ Real-time, context-aware blocking across all channels
High false positives from behavior-only detection	→ 90%+ false positive reduction using content and lineage context
Cannot see data movement inside AI tools	→ Monitors data flow into and out of GenAI and agentic platforms
Investigation requires manual reconstruction from logs	→ Complete data trail available instantly for investigation

The Insider Risk Incident Response Checklist

Use the following checklist to guide investigations from initial detection through resolution. Cyberhaven capabilities are mapped to each step.

	Action	Cyberhaven Capability
1	**Identify the triggering alert or signal** Review the initial alert: data type accessed, user involved, destination or channel flagged	Linea AI risk scoring and real-time alert console
2	**Assess user context** Check employee status: active resignation, performance event, role change, or off-hours activity	HR lifecycle event integration; user risk timeline
3	**Classify the data involved** Determine if the data is regulated, IP-classified, or critical to business operations	Content inspection with EDM, OCR, and lineage-based classification
4	**Establish data lineage baseline** Identify the origin of the file or data in question: where it was created, who has accessed it	Data Lineage tracks origin, interactions, modifications, and derivatives

	Action	Cyberhaven Capability
5	**Reconstruct the full data trail** Trace every copy, rename, transformation, and transmission of the file or dataset	Complete lineage graph: file origin through every derivative and destination
6	**Map all exfiltration channels** Identify if data moved via email, USB, cloud upload, AirDrop, print, or AI platform	Cross-channel monitoring: apps, cloud, email, USB, print, GenAI tools
7	**Assess scope of exposure** Determine how much data moved, to how many destinations, and whether recipients are external	Lineage-based scope analysis; destination classification
8	**Review AI and agentic activity** Check whether data was submitted to GenAI tools or acted on by an AI agent	AI Security: monitoring of GenAI interactions and agentic data flows
9	**Correlate with DSPM posture findings** Identify whether the incident exploited a known posture gap (overexposed repo, misconfigured permissions)	DSPM integration: posture gap correlation with behavioral data
10	**Apply targeted blocking or restriction** Block further data movement for the user or data in question without disrupting the broader organization	Granular, context-aware real-time enforcement across all channels
11	**Engage HR and Legal** Share anonymized behavioral timeline with HR; determine if legal review is required for IP or compliance incidents	Exportable investigation reports; HR coordination workflow
12	**Preserve evidence chain** Lock down the data trail and export documentation for legal hold or regulatory response	Immutable data lineage record; exportable audit trail
13	**Notify relevant parties** Determine notification obligations based on data type (PII, PHI, financial data) and jurisdiction	Data classification with regulatory tagging (GDPR, HIPAA, PCI)
14	**Remediate posture gaps** Close the access or configuration issues that made the incident possible	DSPM-driven remediation recommendations

Section 04

Building a Program That Scales

Insider risk is not a problem you can solve with a single tool or a one-time audit. It requires a program: defined processes, cross-functional coordination between security and HR, and a platform that provides the data context investigators need without creating unsustainable alert volumes.

Five Principles for a Mature Insider Risk Program

1. Lead with data, not behavior
Behavioral monitoring surfaces anomalies. Data lineage explains them. The investigation starts with the data, traces back to the user, and maps the full path of exposure. Behavior without data context generates noise. Data context without behavioral signals misses patterns.

2. Integrate HR lifecycle events
The highest-risk windows in an employee's tenure are HR events: resignation notice, performance improvement plan, role change. Security teams that receive no advance notice of these transitions are operating blind during the periods that matter most. Define a formal handoff protocol between HR and security before an incident occurs.

3. Build for AI-era threats from the start
Shadow AI adoption is accelerating. Agentic workflows are becoming standard. An insider risk program that does not cover AI interactions, AI agent activity, and GenAI platform data flows has a gap that will grow larger over time. Content and lineage monitoring needs to extend to AI tools by design, not as an afterthought.

4. Reduce false positives before scaling monitoring
High alert volumes are the primary reason insider risk programs stall. Teams cannot investigate everything, so they investigate nothing. Precision is more valuable than recall in most environments. A smaller number of high-confidence alerts, grounded in content and lineage context, is more actionable than a large volume of behavior-only signals.

5. Align response protocols before an incident, not during one
The investigation checklist in this document is most useful when it is socialized, agreed on, and rehearsed before the first alert fires. Security, HR, Legal, and executive leadership should all have defined roles in the response protocol. Incidents that require cross-functional coordination under time pressure are the ones most likely to be mishandled.

INSIDER RISK MANAGEMENT · MATURITY MODEL

From reactive alerts to autonomous response

INCREASING MATURITY →

01 Reactive	02 Aware	03 Proactive	04 Optimized
Alerts only. Teams respond after the fact, with no visibility into where data went.	Behavior is observed. Anomalies surface earlier, but context is still partial.	Risk is understood in context. Lineage and identity make intent legible before loss.	Detection and response run themselves. AI covers the surface humans can't
■ Rule & signature-based alerts ■ Manual triage & unvestigation □ No data lineage	■ Behavioral baselines ■ Some content inspection ■ User activity monitoring	■ End-to-end data lineage ■ HR & identity integration ■ DSPM coverage	■ AI-native detection ■ Agentic coverage ■ Automated response

WHERE CYBERHAVEN OPERATES
Data lineage and AI-native detection move teams into Proactive and Optimized maturity.

See Cyberhaven in Action

Find out how Cyberhaven's Data Lineage helps security teams detect, investigate, and contain insider risk before sensitive data leaves your organization.

Request a demo

About Cyberhaven

Cyberhaven is the AI-powered data security company revolutionizing how companies detect and stop the most critical insider threats to their most important data. Until now, data security products were limited to scanning data content and looking for specific user actions. Our AI technology analyzes billions of workflows to understand every piece of data within an organization, when it's at risk, and takes action to protect it. It's like nothing that's come before and protects data like nothing else. For more information, visit cyberhaven.com.