INTELLIGENCE INSIGHTS
Expert Insights on National Security Developments

August 2025

Recommendations when Using AI in Insider Risk Management

Presented by
INSA'S INSIDER THREAT SUBCOMMITTEE

EXECUTIVE SUMMARY

Insider threats remain a persistent and evolving concern for the Defense Industrial Base (DIB).[1] As organizations' external defenses have improved, adversaries are increasingly shifting focus inward, exploiting human vulnerabilities and trusted access. According to the *Cybersecurity Insiders 2024 Insider Threat Report*[2], 83% of organizations have experienced at least one insider attack in the past year.

Insider Risk Management (IRM)[3] refers to the policies and practices designed to detect, assess, and mitigate threats from individuals with legitimate access. These threats may result from deliberate malicious actions, accidental mistakes, or stolen credentials.

Artificial Intelligence (AI) is influencing how practitioners approach IRM.[4] Where traditional methods relied heavily on static rules and manual log reviews, AI introduces capabilities for analyzing large datasets, detecting subtle behavioral anomalies, and enabling faster responses. However, these advances come with their own challenges: concerns over bias, privacy implications, operational complexity, and the need for rigorous governance.

This Intelligence Insights outlines key considerations for organizations examining the use of AI in insider risk programs, highlighting potential benefits and trade-offs, drawing on practical observations from use cases, and underscoring the importance of careful, transparent deployment.

INTRODUCTION: UNDERSTANDING INSIDER RISK

Insider risk[5] refers to the potential harm that individuals within your organization can pose, either intentionally or unintentionally. Examples include a disgruntled employee stealing data, a careless new hire clicking on a phishing link, or even a well-meaning executive who reuses passwords across systems.

Historically, insider risk is managed through policies, training, and a perimeter-focused mindset. But today, the threats are subtle. A malicious insider may appear to be a trusted colleague; a negligent insider may unintentionally enable exploitation. Static tools, such as simple keyword alerts or role-based access alone, are increasingly insufficient in addressing this complexity. They may generate false positives, miss context, and most dangerously fail to adapt. For this reason, organizations are exploring AI to augment traditional approaches with technologies that can detect subtle signals, adapt to changing behaviors, and operate at scale.

THE ROLE OF AI IN INSIDER RISK MANAGEMENT

AI has the potential to transform how risk management[6] practices are implemented across industries. The use of advanced algorithms and data analytics, among other techniques, will allow public and private industry stakeholders to engage in more active solutions to address critical threats.

AI excels at identifying patterns within large volumes of data. In the context of insider risk, AI can analyze behavioral signals, such as login habits, file access trends, and communication indicators to develop baseline profiles. Deviations from these baselines may then be flagged for further review.

Potential applications of AI in insider risk management include:

DETECTION
Identifying unusual access patterns or shifts in user behavior that may otherwise go unnoticed.

RESPONSE
Enabling faster reactions through automated alerts or access restrictions when potential threats are detected.

INVESTIGATION
Supporting analysis by providing contextual information, correlating disparate data points, and presenting potential risk narratives.

Techniques such as natural language processing can also examine internal communications for early signs of disengagement or intent to cause harm. Predictive analytics may assist in anticipating behaviors that precede insider incidents.

CONSIDERATIONS & CHALLENGES

While AI offers the promise of more adaptive and precise insider risk detection, its adoption requires careful evaluation of several key considerations:

ACCURACY & FALSE POSITIVES: AI can reduce noise compared to static rules, but like any automated detection approach, it can still misinterpret behavior.[7] Poorly designed or trained models may generate false positives, creating reputational or legal risks if actions are taken without sufficient human validation.[8]

BIAS & OBJECTIVITY: AI systems reflect the data and assumptions on which they are built. Without careful design and oversight, models may reinforce existing biases or produce inconsistent results.

PRIVACY & CULTURE: Continuous behavioral monitoring raises legitimate concerns about employee privacy and trust. Overly intrusive practices can erode workplace morale if not balanced with clear policies, transparency, and proportional safeguards.[9]

ADVERSARIAL EVASION: Sophisticated insiders may attempt to manipulate detection mechanisms or disguise activities to evade monitoring.

OPERATIONAL READINESS: Effective use of AI requires clean, well-integrated data sources and staff who can interpret outputs accurately. AI should complement human judgment, not replace it.

Successful deployment also requires a strong foundation: clean, structured data, system integration, and skilled analysts who understand how to interpret AI findings. Without these elements, even the most advanced tools may fail to deliver meaningful results.

USE CASES & THE ROAD AHEAD

Various industries are employing AI tools to complement their IRM programs.

FINANCIAL SERVICES

AI was used to detect rogue trading long before it became headline news. Algorithms have picked up on shifts in trading frequency and flagged unusual access to sensitive data.[10]

GOVERNMENT

AI is being used to detect insider espionage by matching behavioral patterns with psychological profiling.[11]

PRIVATE SECTOR

Sentiment analysis powered by AI flags emotional disengagement, often a precursor to turnover or sabotage. Disengaged employees don't all follow the same trajectory. AI can help distinguish between different disengagement archetypes, such as:

- **Angry employees:** Frustrated, emotionally reactive individuals who may lash out or leak data in protest.[12]
- **Terminated employees:** Often motivated by financial need or retaliation; they pose high risk immediately following exit.[13]
- **Unwitting employees:** Those unaware they have been compromised or targeted, often manipulated into becoming insider threats.

Each category carries different risk profiles, timelines, and behavioral indicators. AI can assist by flagging early warning signs, tone changes in communication, access pattern shifts, productivity dips, so organizations can tailor their responses. For angry employees, that might mean proactive engagement or HR intervention. For terminated staff, it may involve adjusting access offboarding timelines and increasing post-departure monitoring. For unwitting actors, it could mean tailored security awareness training and closer review of email or download activity. Every disengaged employee should not be treated as a threat. However, every disengagement pattern should be noticed, contextualized, and evaluated through a risk lens.

RECOMMENDATIONS FOR IMPLEMENTATION

The following recommendations are offered to help organizations determine how they may want to use AI as part of their IRM strategy:

PROVIDE HUMAN OVERSIGHT

AI-generated insights should be treated as decision support, not definitive judgments. Human analysts remain essential for interpreting context and determining appropriate actions.[14]

ESTABLISH COLLABORATIVE GOVERNANCE

Successful AI-enhanced IRM programs need functional governance (across IT, HR, Compliance, etc) to ensure coordinated detection, risk scoring and response.[15]

MONITOR DISENGAGEMENT

Organizations may wish to examine how to responsibly identify patterns of disengagement while respecting privacy and avoiding undue assumptions about intent.

MAINTAIN & VALIDATE MODELS

AI models require ongoing tuning to remain effective as threats and behaviors evolve. Regular testing and validation can help maintain relevance and accuracy.

PRIORITIZE ETHICAL USE & TRANSPARENCY

Clear policies on monitoring practices, data handling, and employee communications are essential to maintain trust and comply with legal and ethical standards.

Ultimately, AI should be seen as multiplier amplifying your team's ability to detect, interpret, and respond to insider risk. The organizations that will thrive are those that combine data, people, and ethical intelligence-deliberately, thoughtfully, and decisively.

CONCLUSION

AI presents new opportunities for enhancing insider risk management by complementing traditional security controls with deeper behavioral insights and scalable monitoring. However, its value depends on responsible, transparent, and well-governed implementation.

Organizations within the Defense Industrial Base should weigh the practical benefits of AI tools alongside their operational, ethical, and cultural impacts, recognizing that AI is one component within a broader strategy that must continue to prioritize human judgment, strong governance, and a culture of trust and accountability.

REFERENCES

[1] CrowdStrike. 2025. "Insider Threat." *Cybersecurity 101.* June 12, 2025. https://www.crowdstrike.com/en-us/cybersecurity-101/identity-protection/insider-threat/.

[2] https://www.cybersecurity-insiders.com/2024-insider-threat-report/

[3] (SB)CISA. 2023. "Defining Insider Threats." Cybersecurity and Infrastructure Security Agency CISA. CISA. 2023. https://www.cisa.gov/topics/physical-security/insider-threat-mitigation/defining-insider-threats.

[4] (SB)Center for the Development of Security Excellence. 2025. Review of ARTIFICIAL INTELLIGENCE and the INSIDER THREAT. Edited by Defense Counterintelligence and Security Agency. Center for Development of Security Excellence. https://www.cdse.edu/Portals/124/Documents/jobaids/insider/Artifical-Intelligence-and-the-Insider-Threat.pdf.

[5] (SB) New Jersey Cybersecurity & Communications Integration Cell. 2024. "Insider Threat." Nj.gov. State of New Jersey. 2024. https://www.cyber.nj.gov/guidance-and-best-practices/resources-for-businesses-government/insider-threats.

[6] (SB)Badman, Annie. 2024. "AI Risk Management." Ibm.com. June 20, 2024. https://www.ibm.com/think/insights/ai-risk-management.

[7] CERT Insider Threat Mitigation Guide (2023) https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=665858

[8] NIST AI Risk Management Framework n.d. https://www.nist.gov/itl/ai-risk-management-framework

[9] How to Monitor Your Employees While Respecting Their Privacy n.d. https://hbr.org/2022/01/how-to-monitor-your-employees-while-respecting-their-privacy

[10] Jones, Gareth. 2019. "Banks Use AI to Catch Rogue Traders Before the Act." Financial Times, March 29, 2019. https://www.ft.com/content/be7a5584-2ee7-11e9-80d2-7b637a9e1ba1.

[11] Government: AI in Insider Espionage Detection
The U.S. **Defense Counterintelligence and Security Agency (DCSA)** uses automated monitoring programs that assess behavioral anomalies among cleared personnel to detect potential insider espionage dcsa.mil.

[12] DIA Employee Accused of Trying to Share Classified Info Over Trump Anger 2025

[13] Pryimenko, Liudmyla. 2024. "7 Real-Life Data Breaches Caused by Insider Threats | Syteca." Syteca. February 28, 2024. https://www.syteca.com/en/blog/real-life-examples-insider-threat-caused-breaches.

[14] Can AI Reduce Insider Threats in IAM? A Forward-Looking Approach to Identity Security https://www.avatier.com/blog/can-ai-reduce-insider-threats-iam.

[15] Uniting Forces: Cross-Functional Approaches to Insider Threat Prevention. https://www.corporatecomplianceinsights.com/uniting-forces-cross-functional-approaches-insider-threat-prevention/?

ACKNOWLEDGEMENTS

INSA expresses its appreciation to the INSA members and staff who contributed their time, expertise, and resources to this paper.

INSA MEMBERS

Michael Crouse, *Chair, Insider Threat Subcommittee*

Theresa Campobasso
Vice Chair, Insider Threat Subcommittee

Christopher Hadnagy,
CEO & Chief Human Hacker, Social-Engineer, LLC

INSA STAFF

Suzanne Wilson Heckenberg, *President*

John Doyon, *Executive Vice President*

Bishop Garrison, *Vice President for Policy*

Peggy O'Connor, *Director of Communications and Policy*

Stephanie Bulega-Nasuna, *INSA Policy Intern*

Jacob Hertzinger, *INSA Policy Intern*

ABOUT INSA

The Intelligence and National Security Alliance (INSA) is a nonpartisan, nonprofit membership organization dedicated to advancing collaborative, public-private approaches to intelligence and national security priorities. Through the application of industry expertise, leading-edge academic research, and commercial best practices, INSA seeks to make the Intelligence Community more effective and efficient. Our 175+ member organizations and 4,000+ individual and associate members include senior executives and intelligence experts in the public, private and academic sectors.

ABOUT INSA'S INSIDER THREAT SUBCOMMITTEE

INSA's Insider Threat Subcommittee researches, discusses, analyzes, and assesses counterintelligence and insider threat issues that affect government agencies, cleared contractors, and other public and private sector organizations. The Subcommittee works to enhance the effectiveness, efficiency, and security of government agencies and their industry partners, as well as to foster more effective and secure partnerships between the public, private and academic sectors.