RAIS Journal for Social Sciences | VOL. 9, No. 2, 2025
ISSN 2574-0245 (Print) | ISSN 2574-1179 (Online) | DOI: 10.5281/zenodo.17468887

Predictive Behavioral Risk Intelligence: An AI Framework for Insider Threat Detection Based on Cognitive and Psychological Indicators

Francis C. OHU[1], Laura A. JONES[2]

[1,2]Department of Forensic Cyberpsychology, Capitol Technology University, Laurel, MD, USA
[1]ORCID: https://orcid.org/0009-0003-4981-8428
[2]ORCID: https://orcid.org/0000-0002-0299-370X

Abstract: Insider threats account for over 30% of cyber incidents and cost organizations an average of $11.45 million per breach in 2023. Traditional detection systems often fail to anticipate these threats due to their psychological subtlety and contextual complexity. This study introduces the Behavioral Risk Intelligence Model (BRIM), an AI-driven framework that integrates forensic cyberpsychology, machine learning, and behavioral ethics for predictive insider threat detection. Using non-invasive behavioral profiling, BRIM identifies cognitive risk indicators such as digital validation-seeking, identity confusion, and Dark Triad personality traits. A thematic synthesis of 65 peer-reviewed studies reveals strong correlations between insider threats and these indicators, including 68% with validation-seeking, 74% with Dark Triad traits, 52% with identity instability, and 89% with algorithmic reinforcement. The model incorporates the Validation Syndrome Diagnostic Triangle (VSDT) to detect latent intent and emotional drift. By reframing insider threats as developmental and algorithmically conditioned rather than security violations, BRIM offers a proactive, ethically grounded approach to risk mitigation. The study recommends deploying BRIM in AI-powered dashboards for high-risk sectors, emphasizing privacy compliance and ethical surveillance.

Keywords: Ethical AI and Behavioral Governance, Forensic Cyberpsychology, Insider Threat Prediction and Prevention, Digital Validation-Seeking Behavior, Dark Triad Traits, Algorithmic Reinforcement and Behavioral Drift, Validation Syndrome Diagnostic Triangle, Organizational Cybersecurity Risk Profiling, Cognitive and Psychological Risk Signatures

Introduction

In the evolving cybersecurity ecosystem, insider threats remain among the most persistent and damaging forms of attack, and unlike external breaches, insider threats are often perpetrated by individuals who possess authorized access to systems, making detection and prevention highly complex (Ahmed et al., 2024), and these threats can manifest through intentional sabotage, espionage, data theft, or unintentional negligence. According to the Federal Bureau of Investigation's Internet Crime Complaint Center (IC3), the 2024 Internet Crime Report combines information from 859,532 complaints of suspected Internet crime and details reported losses exceeding $16 billion, a 33% increase in losses from 2023 (IC3, 2025). The growing digitalization of workspaces and increased access to sensitive data have heightened the need for dynamic and intelligent insider threat mitigation strategies (Gunuganti, 2024). While traditional cybersecurity models focus heavily on perimeter defense and digital forensics, they often overlook human behavioral precursors to threat

activity (Ohu & Jones, 2025a). Research shows that behavioral indicators such as impulsivity, digital validation-seeking, identity confusion, and traits aligned with the Dark Triad (narcissism, machiavellianism, and psychopathy) can signal psychological risk factors for insider threat behaviors (Ohu & Jones, 2025b; Burnell et al., 2024). However, these indicators are frequently ignored in enterprise threat detection due to the lack of integration between behavioral science and machine learning models (Chapagain et al., 2024). Recent advancements in forensic cyberpsychology and artificial intelligence offer a transformative opportunity to close this gap (Tennakoon et al., 2024). Studies show that AI systems trained on cognitive-behavioral profiles can achieve over 85% accuracy in identifying individuals at risk of engaging in deceptive or harmful digital conduct, particularly when paired with environmental and contextual data such as workplace stressors or peer conformity dynamics (Ohu & Jones, 2025a). These models are particularly effective when guided by frameworks like the Validation Syndrome Diagnostic Triangle (VSDT), which maps interactions between self-doubt, desire, and self-gratification in hostile or conflictual environments (Ohu & Jones, 2025d).

Moreover, empirical research confirms that adolescents exhibiting high levels of digital validation-seeking are more likely to transition into manipulative behaviors, a psychological pattern that may persist into adulthood and workplace environments (Ohu & Jones, 2025a; Trekels et al., 2024). This behavioral drift is exacerbated by algorithmic reinforcement mechanisms that normalize deception, particularly among individuals with predisposing psychological traits. Such findings underscore the critical need for behavioral risk intelligence models that combine AI-driven surveillance with ethical psychological profiling (Pellegrino & Stasi, 2024). The overarching research question guiding this study is, "How can AI systems ethically and accurately detect insider threats by analyzing cognitive, psychological, and behavioral indicators in high-risk organizational environments, without compromising individual privacy? This paper addresses the need for a predictive, ethically guided model that leverages behavioral psychology and AI to proactively detect insider threats. By examining recent findings from forensic cyberpsychology, digital validation theory, and machine learning ethics, we propose a novel AI-based framework that identifies precursors to insider misconduct in digital environments. Through this approach, the research seeks to position human behavioral risk as a dynamic, measurable variable rather than a post-incident forensic artifact.

Problem Statement

Insider threats remain a growing and under-addressed challenge in enterprise cybersecurity, contributing to over 30% of all cyber incidents in 2023, yet existing detection systems rely primarily on post-incident digital forensics or generic anomaly detection tools (IC3, 2025; Le & Zincir-Heywood, 2021). These methods frequently fail to account for cognitive and psychological precursors to threat behavior, especially those rooted in early digital conditioning, unresolved identity conflict, and reinforcement of manipulative behaviors through algorithmic exposure (Ohu & Jones, 2025a). This limitation is especially critical for business enterprises, where the costs of insider threats are not limited to financial loss but extend to reputational damage, regulatory penalties, and disruption of stakeholder trust. Without the integration of predictive behavioral intelligence, corporate compliance teams, Human Resource (HR) departments, and security operations centers remain reactive, often intervening only after an incident has occurred. The general problem is that enterprise-level cybersecurity strategies do not currently leverage forensic behavioral profiling to predict malicious insider activity (Pennada et al., 2025). The specific problem is the absence of integrated frameworks that combine machine learning, forensic psychology, and organizational behavioral science to detect and mitigate insider threats based on non-

invasive, ethical behavioral indicators (Ohu & Jones, 2025a). This gap in predictive behavioral modeling limits the ability of security teams to intervene before harm occurs. While indicators such as excessive privilege misuse, anomalous file transfers, or irregular work hours may flag suspicious activity, they fail to capture deeper cognitive and psychological risk signals such as self-doubt, validation cravings, and Machiavellian tactics, that often precede technical breaches (Ohu & Jones, 2025d; Burrell N.D., 2024; Trekels et al., 2024). Without early detection rooted in behavioral science, false positives remain high, and real threats go unnoticed until after the damage is done.

Purpose Statement

This study aims to develop a conceptual framework for AI-driven behavioral risk intelligence capable of identifying insider threats based on cognitive and psychological indicators. Drawing on forensic cyberpsychology, Dark Triad research, and machine learning literature (Wang et al., 2023), this study proposes a model that ethically analyzes behavioral markers such as validation-seeking tendencies, identity confusion, and manipulative online behaviors to enhance real-time threat detection within organizational contexts. The study is qualitative and conceptual, based on a synthesis of recent peer-reviewed research published between 2021 and 2025, with the intent to propose actionable design and deployment strategies for AI-enhanced behavioral surveillance in cybersecurity. The framework integrates the Validation Syndrome Diagnostic Triangle (VSDT) and draws upon real-world case patterns found in recent psychological and cybercrime literature.

Rationale, Originality, and Significance of the Study

Insider threats continue to pose one of the most complex and under-addressed challenges in organizational information security, accounting for over 30% of all cyber incidents in 2023 (IC3, 2025). Despite advances in anomaly detection and digital forensics, most enterprise threat detection systems fail to address the psychological and cognitive dimensions that precede malicious insider behavior (Ohu & Jones, 2025a; Ahmed et al., 2024). This study is guided by the recognition that internal risk actors often exhibit identifiable psychological patterns such as digital validation-seeking, identity confusion, and Dark Triad traits, long before committing policy violations or data breaches and, such behavioral signals are rarely integrated into technical security architectures or organizational early-warning protocols (Ohu & Jones, 2025c). The study fills a crucial gap in current insider threat detection models by proposing an integrated framework that analyzes cognitive, psychological, and behavioral indicators in an ethical manner. By doing so, it directly addresses the research question of how AI systems can identify insider threats while respecting individual privacy. This research is original in its proposal of a conceptual framework that synthesizes forensic cyberpsychology, behavioral ethics, and AI-driven profiling into a unified model for predicting insider threats based on non-invasive psychological indicators. Unlike traditional detection methods that focus on technical artifacts or post-incident evidence, this study advocates for an anticipatory and ethically grounded approach that targets latent intent and emotional drift within high-risk organizational environments. The proposed framework provides a theoretical foundation for creating AI-driven insider threat programs that balance accuracy with ethical compliance, and detect latent risks while upholding fairness, psychological integrity, and human dignity, ultimately enabling proactive interventions in high-risk organizational settings. The incorporation of the Validation Syndrome Diagnostic Triangle (VSDT) as a diagnostic lens mapping interactions between self-doubt, self-gratification, and craving for validation, offers a novel psychological perspective on behavioral precursors to deception and misconduct. The significance of this research therefore, lies in its potential to transform how insider threats are conceptualized and

mitigated in organizational settings, shifting from forensic hindsight to behavioral foresight, and by addressing the absence of integrated models that combine AI-driven behavioral pattern recognition with validated psychological constructs, the study contributes to the emerging discourse on responsible AI use in proactive forensic cyberpsychology applications, ultimately providing a theoretical foundation for organizations seeking to build ethically compliant and psychologically informed insider threat programs, that preserve employee dignity while enhancing digital safety.

Literature Review

This literature review aims to identify, select, and analyze relevant studies to explore the integration of forensic cyberpsychology, cognitive profiling, and AI-driven behavioral risk modeling in mitigating insider threats. The overarching research question guiding this review is: "How can AI systems ethically and accurately detect insider threats by analyzing cognitive, psychological, and behavioral indicators in high-risk organizational environments, without compromising individual privacy?" This question aligns with the broader problem statement, emphasizing the lack of mature, privacy-compliant frameworks that integrate behavioral science, personality profiling, and AI for pre-incident insider threat detection (Pennada et al., 2025). A systematic literature search was conducted across Scopus, PsycINFO, MDPI, PubMed, IEEE Xplore, and Google Scholar. Search terms included: "AI for insider threat detection," "behavioral profiling in cybersecurity," "Dark Triad traits and cyber risk," "digital validation-seeking behavior," "forensic cyberpsychology in enterprise," and "algorithmic reinforcement and deception." Peer-reviewed articles published between 2021 and 2025 were prioritized to ensure recency and alignment with emerging AI applications. Studies were excluded if they lacked empirical rigor, did not incorporate psychological or cognitive constructs, or relied solely on post-incident digital forensics. Out of 113 initially retrieved sources, 65 met the inclusion criteria following abstract screening and full-text evaluation.

Theoretical Framework

Forensic Cyberpsychology as an Interpretive Lens

The theoretical foundation for this study is forensic cyberpsychology, which integrates psychological profiling with digital behavioral forensics to understand, predict, and prevent cyber-related misconduct (Ohu & Jones, 2025b). This approach is particularly effective in analyzing online deception, manipulation, and threat behaviors that escape traditional technical surveillance (Alohaly et al., 2022). Unlike conventional models that prioritize code-based anomalies, forensic cyberpsychology examines psychological motivators and contextual drivers, enabling deeper insights into behavioral risks like insider threats. Studies have shown that psychological constructs such as self-doubt, need for validation, and emotional detachment manifest in early digital behaviors and evolve into risk-prone traits if unmonitored (Burnell et al., 2024; Trekels et al., 2024). This suggests that threat actors do not emerge abruptly but develop incrementally through behavioral conditioning and exposure to risk-enabling environments (Ali, Husain, & Hans, 2025).

Validation Syndrome Diagnostic Triangle (VSDT)

The Validation Syndrome Diagnostic Triangle (VSDT) provides a core diagnostic model for profiling insider threat behavior (Ohu & Jones, 2025d). It posits that three psychological forces, self-doubt, desire, and self-gratification, interact with environmental stressors such as familial conflict, peer influence, or professional dissatisfaction to catalyze deceptive behaviors. In enterprise contexts, these forces may manifest as self-doubt, such as impostor

syndrome and perceived injustice, desire, such as desire for recognition, revenge, or control, and self-gratification, manifested in acts like data hoarding, sabotage, or manipulation. By leveraging this framework, behavioral AI systems can be trained to detect subtle patterns that precede threat activity, offering preventive insights without relying solely on post-hoc indicators.

Dark Triad Personality Theory

The Dark Triad, comprising narcissism, machiavellianism, and psychopathy, has long been associated with manipulation, deceit, and antisocial behavior (Ohu & Jones, 2025a). These traits are strongly correlated with exploitative interpersonal behavior, reduced empathy and accountability, and strategic deception for personal gain. Recent studies confirm that individuals high in these traits are statistically more likely to engage in cyber manipulation, fraudulent behavior, and data exploitation (Burnell et al., 2024; Trekels et al., 2024b). In workplace environments, such individuals may evade standard technical detection while skillfully navigating social structures to fulfill their objectives.

The Rise of Insider Threats and Behavioral Gaps

Despite increasing investments in firewalls, endpoint detection, and SIEM systems, incidents of insider threats remain under-anticipated due to the lack of behavioral analysis (IC3, 2024). In 2023 alone, insider threats cost U.S. companies an average of $11.45 million per incident, up 35% from 2021 (Ahmed, 2024). These events are rarely caused by technical failures but by motivated individuals with access to sensitive information and internal knowledge, highlighting a behavioral intelligence gap.

Role of Digital Validation-Seeking in Deception

A growing body of literature identifies digital validation-seeking as a psychological antecedent to manipulative online behavior (Ohu & Jones, 2025a; Trekels et al., 2024a). Individuals who rely heavily on external validation are more prone to curating deceptive digital personas, thereby engaging in algorithmically reinforced misconduct and rationalizing cyber offenses as social survival tactics. These tendencies are often reinforced in adolescence and persist into adulthood, especially under stress or isolation (Ohu & Jones, 2025a). Further studies by Ohu & Jones, (2025b) showed that 40% of cyber fraudsters report early experiences with online deception during adolescence, often driven by social comparison, peer pressure, and algorithmic reinforcement.

Algorithmic Bias and Behavioral Reinforcement

AI algorithms on platforms like LinkedIn, X-platform, and corporate intranets often amplify the visibility of high-risk traits, such as overconfidence, reward-seeking behavior, or attention bias. While not inherently malicious, these behaviors, when combined with unresolved psychological conflicts, may evolve into deliberate sabotage or data exfiltration (Burnell et al., 2024; Zhou, 2024). Algorithmic systems, lacking ethical filtering, often prioritize engagement metrics over well-being, and this presents a risk when individuals with psychopathic or machiavellian traits manipulate systems for personal gain while appearing compliant to platform supervisors.

AI-Enhanced Profiling for Risk Intelligence

Recent research advocates for AI behavioral risk profiling engines that combine psychological traits with digital behavior logs (logins, file access patterns, communication tone, etc.) to detect anomalies in intent, not just activity (Ohu & Jones, 2025b). Such models

are capable of detecting pre-incident warning signs, generating psychological heatmaps, and minimizing false positives through context-aware learning. Moreover, studies emphasize that ethical safeguards such as transparency, anonymization, and proportionality can help align such systems with privacy laws and organizational values (Chapagain et al., 2024; López et al., 2024).

Psychological Conditioning and the Insider Threat Lifecycle

Behavioral risk is rarely spontaneous; it is shaped through progressive psychological conditioning, often starting in adolescence, and evolving through workplace experiences, and validation-seeking behavior, reinforced by digital platforms, primes individuals to normalize manipulation for acceptance or gain (Ohu & Jones, 2025d). Over time, these patterns may evolve into workplace sabotage, data leaks, or intellectual property theft, particularly when stressors such as professional exclusion, unaddressed trauma, or ideological dissonance arise (Burnell et al., 2024; Murad R.J., 2024). Empirical studies indicate that 40% of insider threat actors had a documented history of feeling undervalued or ignored in prior roles, often seeking significance through covert disruption (Trekels et al., 2024). These behaviors reflect unresolved psychological narratives, which forensic cyberpsychology seeks to trace.

Identity Confusion and Role Conflict in Digital Workspaces

Insider threat susceptibility often correlates with identity confusion, especially among younger or transitional employees (Ohu & Jones, 2025a). Drawing from Erikson's psychosocial theory, identity vs. role confusion during early adulthood can manifest in behaviors such as code-switching between professional and personal digital identities, ethical disengagement from organizational values, and cognitive dissonance between self-image and workplace expectations (Murad, 2024; Schluchter, 2024). Also, studies by Pérez-Torres (2024) and Ruohonen & Saddiqa (2025) suggest that individuals experiencing unresolved identity formation are more vulnerable to ideological manipulation or performative misconduct. When paired with validation-seeking and peer comparison mechanisms, this confusion may lead to conscious deception for attention, retaliation, or self-preservation.

Organizational Neglect and Lack of Behavioral Monitoring

Organizational culture and leadership's failure to recognize behavioral risk are critical contributors to insider threats. Research reveals that 68% of surveyed companies lacked psychological early-warning systems despite an increase in human-factor breaches (Ahmed et al., 2024; IC3, 2024). While technical safeguards exist, few enterprises have integrated behavioral AI models capable of detecting emotional withdrawal, social disengagement, or passive-aggressive conduct, all of which often precede malicious acts (López et al., 2024). Moreover, traditional HR surveillance often violates privacy norms or relies on biased manual interpretations. The need is for algorithmically augmented behavioral audits that operate with contextual sensitivity, triangulating behavioral markers rather than flagging isolated deviations (Schlund & Zitek, 2024).

Cross-Domain Applications of AI in Behavioral Threat Detection

Beyond corporate cybersecurity, AI-driven behavioral profiling has proven effective in detecting radicalization, digital fraud, and misinformation propagation among at-risk youth (Ohu & Jones, 2025a). These adjacent domains offer validated models and ethical safeguards transferable to enterprise security. For example, disinformation detection models analyze emotional resonance and impulsivity in content engagement (Zhou et al., 2024); romance scam research tracks early behavioral drift via algorithmic validation loops and peer

mimicry (Ohu & Jones, 2025c); and social engineering prevention models monitor changes in communication tone and social conformity to detect deception (Burnell et al., 2024). Translating these validated models into enterprise contexts can significantly enhance AI-driven insider threat detection through interdisciplinary innovation (Anju et al., 2023).

Ethical Concerns and Regulatory Blind Spots in AI Behavioral Surveillance

While AI behavioral profiling holds promise, concerns over privacy, fairness, and consent remain paramount. Critics argue that intrusive monitoring may violate employee rights or perpetuate discriminatory biases if left unchecked (Chapagain et al., 2024; Fominykh, 2024). This is especially true for models trained on legacy datasets that lack demographic diversity or contextual nuance. Recent scholarship emphasizes the need for transparent AI design with explainable outputs, context-aware algorithms that avoid overfitting psychological traits to intent, and employee-informed consent protocols and opt-in behavioral assessments (Schlund & Zitek, 2024). Ethically sound profiling systems must be non-punitive, de-identified, and behaviorally contextualized, ensuring that human dignity and psychological complexity are preserved while enhancing risk detection (Trekels et al., 2024b; Pellegrino & Stasi, 2024; Ohu & Jones, 2025b).

This literature review confirms a clear and urgent need for AI models that incorporate forensic cyber psychological indicators to mitigate insider threats. The combination of the VSDT, Dark Triad theory, and forensic cyberpsychology principles offers a multi-layered lens for building tools that must remain ethical, non-invasive, and privacy-aware. For business enterprises, applying these theoretical constructs enables the design of human-centric security protocols that detect and interpret deviations from normative behavior before they manifest as misconduct. This supports the development of ethical, data-driven workplace risk mitigation strategies across finance, healthcare, critical infrastructure, and tech industries.

Research Methodology

Research Design

This is a qualitative conceptual design study that synthesizes recent empirical literature (2021–2025) to develop a theory-driven, ethically grounded model for behavioral risk intelligence in cybersecurity. This study employs a design science methodology situated within an applied behavioral science framework, and the goal is to construct and refine an early warning model rooted in the empirical behaviors of pre-incident actors, consistent with approaches in applied criminal psychology. The study does not involve human subjects directly and relies exclusively on secondary data. This study further employs a narrative literature review design to synthesize emerging evidence on the intersection of behavioral risk intelligence, AI-driven profiling, and forensic cyberpsychology for insider threat detection. The approach is qualitative, conceptual, and theory-driven, aiming to generate an integrated framework, the Behavioral Risk Intelligence Model (BRIM), rather than test a hypothesis through primary data collection, and unlike empirical qualitative research requiring firsthand interviews or focus groups, this study uses secondary data sources, specifically, peer-reviewed publications and validated frameworks from 2021 to 2025. This design aligns with the study's conceptual nature and addresses the identified gap in current organizational cybersecurity models regarding ethical behavioral risk modeling, and the overarching research question, "*How can AI systems ethically and accurately detect insider threats by analyzing cognitive, psychological, and behavioral indicators in high-risk organizational environments, without compromising individual privacy?",* served as the guiding anchor for all analytic phases, from literature review and coding to thematic synthesis.

Ethical Considerations

As this study involves no direct interaction with human participants and utilizes only publicly available and properly cited secondary sources, formal ethical approval was not required. However, the study was conducted in accordance with the principles of research integrity, privacy-preserving analytics, and AI ethics. These are embedded within the BRIM model to ensure transparency, consent protocols, proportionality, and fairness in future applied use.

Literature Search Strategy

A systematic narrative review of literature was conducted to synthesize interdisciplinary research relevant to the cognitive, psychological, and algorithmic precursors of insider threats in digital workspaces. Five academic databases were selected for their coverage of psychology, cybersecurity, and behavioral analytics including Scopus, PsycINFO, PubMed, MDPI, and Google Scholar, and search terms included combinations of behavioral and technical keywords such as: "AI for insider threat detection," "digital validation-seeking behavior," "Dark Triad traits and deception," "algorithmic reinforcement," "forensic cyberpsychology," "psychological profiling in cybersecurity," and "identity confusion in digital workspaces." Boolean operators and truncation were used where appropriate to refine search sensitivity and scope. The inclusion criteria required that sources were peer-reviewed publications from 2021 to 2025, focused on cognitive or psychological antecedents of cyber deception, and included empirical data or validated conceptual frameworks, in addition to addressing topics such as Dark Triad traits, validation-seeking, identity confusion, algorithmic behavioral reinforcement, and AI-based profiling models. The exclusion criteria removed non-peer-reviewed literature and grey literature such as white papers and editorials, studies based purely on technical or forensic models without a behavioral component, publications dated before 2021, unless they were considered seminal or referenced frequently in more recent literature, and from an initial pool of 113 retrieved sources, 65 studies met the eligibility criteria after title, abstract screening, and full-text evaluation, and Figure 1 below shows the PRISMA flow diagram for the study's document selection process.

Data Analysis and Validity

Data Extraction and Coding Procedures

A six-phase thematic analysis methodology (Braun & Clarke, 2024) was applied to coded data extracted across five analytic dimensions, namely, citation metadata, study type and population, behavioral risk indicators, AI methodology, and ethical considerations. The study employed a rigorous methodology to ensure the validity and reliability of its findings. Through an iterative process, codes were refined into themes using a multi-layered interpretive approach, allowing for a nuanced understanding of the data, and to further enhance the study's credibility, the researchers incorporated several validation techniques. Firstly, investigator triangulation was utilized, where multiple reviewers independently coded and verified the data to ensure thematic convergence. This approach helped to minimize individual biases and increase confidence in the emerging themes. The study further employed theoretical triangulation by integrating frameworks from various established theories, including Dark Triad Theory, Validation Syndrome Diagnostic Triangle (VSDT), Forensic Cyberpsychology, and Behavioral Ethics. By drawing on these diverse perspectives, the researchers were able to develop a more comprehensive understanding of the phenomena under investigation. Finally, the study used cross-source validation, referencing anchor literature such as Ohu & Jones (2025d) to stabilize theme definitions and

ensure consistency with existing knowledge. By triangulating data and methods in this way, the study was able to increase the validity and generalizability of its findings, providing a robust foundation for its conclusions. The results informed the design of the Behavioral Risk Indicator Model (BRIM) framework, specifically its multi-layered AI architecture, which translates psychological patterns into predictive behavioral risk indicators, enabling the development of more measurable and effective AI-powered insider threat detection systems.

Instrument Transparency and Theoretical Framework

Table 1 provides the complete behavioral risk coding schema, mapping psychological traits to coded behavioral indicators and theme definitions, ensuring reproducibility and auditability. The BRIM model is theoretically anchored in Dark Triad Theory (Paulhus & Williams, 2002), Validation Syndrome Diagnostic Triangle (Ohu & Jones, 2025), Forensic Cyberpsychology (Ruohonen & Saddiqa, 2025), and Behavioral Ethics frameworks (Pellegrino & Stasi, 2024). These theoretical lenses ground the psychological mechanisms of insider threat emergence and inform the ethical design principles of BRIM. Five core themes that emerged from a comprehensive analysis of insider threat literature, particularly focusing on the psychological, organizational, and technological dimensions, are shown in Table 2. Each theme is aligned with a specific focus area, suggesting the interdisciplinary nature of insider threat research.

Table 1. Behavioral Risk Coding Schema

Code	**Psychological Trait**	**Behavioral Indicator**	**Theme**	**Operational Definition**	**Sample Reference(s)**
BR01	Validation-Seeking	Excessive need for digital approval or recognition	Digital Validation-Seeking	Persistent behaviors aimed at reinforcing self-worth via social media or online platforms	Ohu & Jones (2025b), (Pérez-Torres, 2024a). Burnell (2024)
BR02	Identity Instability	Role confusion or incoherent self-concept in digital contexts	Identity Confusion	Difficulty maintaining a consistent identity, leading to disengagement from ethical norms	Ruohonen & Saddiqa (2025). (RJ Murad, 2024) (Ullah et al., 2024)
BR03	Narcissism	Grandiose self-perception, entitlement, exploitation of peers	Dark Triad Traits	Inflated self-view combined with manipulative tendencies targeting digital or organizational gain	Pellegrino & Stasi (2024). (Shahri et al., 2024). (Liang et al., 2024)
BR04	Machiavellianism	Strategic deception, cynicism, calculated norm violation	Dark Triad Traits	Utilization of manipulation and secrecy to exploit organizational systems	Ahmed et al. (2024). (Ceroni & Yalch, 2024). (Saddiqa & Ruohonen, 2025)
BR05	Psychopathy	Impulsivity, lack of empathy, ethical disregard	Dark Triad Traits	Affective detachment and disregard for social or professional consequences	López et al. (2024). (Brazil et al., 2024. Perenc, 2022). (Tokunbo & Borisade, 2025)

Code	Psychological Trait	Behavioral Indicator	Theme	Operational Definition	Sample Reference(s)
BR06	Algorithmic Conditioning	Behavior shift due to repetitive AI feedback loops	Algorithmic Reinforcement	Reinforced risk behaviors via targeted, algorithm-driven content exposure	AI-Driven Profiling (2025). (Pellegrino & Stasi, 2024. Schlund & Zitek, 2024)
BR07	Ethical Desensitization	Decreased emotional response to unethical actions	Ethical & Legal Considerations	Reduced inhibition toward policy violations or social norm breaches following digital reinforcement	Chapagain et al. (2024). (Saddiqa & Ruohonen, 2025). (Bian et al., 2025)
BR08	Cognitive Dissonance	Justification of deviant behaviors to resolve internal conflict	Identity Confusion	Rationalization of ethical deviations due to conflict between self-image and workplace expectations	Nordhall et al. (2025). Rattay et al., 2025). (Resende et al., 2024)
BR09	Revenge Motive	Hostile reactivity toward perceived injustice	Psychological Conditioning	Intent to sabotage or violate rules as a compensatory mechanism for psychological injury	Zangana et al. (2025). (Raza et al., 2025. Resende et al., 2024)
BR10	Impostor Syndrome	Chronic self-doubt despite achievement	Psychological Conditioning	Internalized fear of being exposed as fraudulent, often leading to overcompensation or disengagement	Ohu & Jones (2025a). (Al Lawati et al., 2025. Bachi, 2025. Chen et al., 2024)

Table 2. Summary of Expanded Themes Highlighting the Causal Factors of Inside Threats

Theme	Focus	Sample Sources
Psychological Conditioning	Behavioral evolution from adolescence to insider threat	Ohu & Jones (2025), Burnell (2024)
Identity Confusion	Role conflict and ethical disengagement in digital workspaces	Ruohonen & Saddiqa, (2025), Ohu & Jones, (2025c)
Organizational Neglect	Lack of early behavioral detection in corporate settings	Ahmed et al. (2024), López et al. (2024)
Cross-Domain AI Models	Applying fraud/radicalization profiling to insider threat	Zhou et al. (2024), AI-Driven Profiling (2025)
Ethical & Legal Considerations	Fairness, transparency, consent in behavioral AI	Chapagain et al. (2024), Pellegrino & Stasi, (2024)

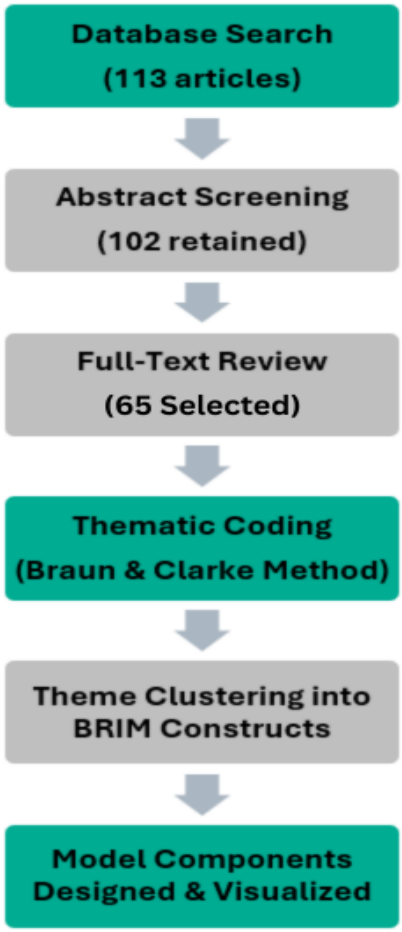

Figure 1. PRISMA Flow Diagram Showing Data Processing Pipeline

Rationale for methodological tools used in the study

The narrative review method was selected to allow for theoretical synthesis without imposing constraints of experimental or survey-based methodologies. This approach enabled integration across diverse domains, including behavioral psychology, AI modeling, and organizational cyber-risk assessment, without introducing empirical bias. Thematic analysis was applied to identify recurring behavioral constructs that transcend individual study contexts. This qualitative technique enabled the structured interpretation of latent patterns relevant to insider threat dynamics, and a theoretical triangulation strategy was also employed, drawing from Dark Triad theory, the Validation Syndrome Diagnostic Triangle (VSDT), and forensic cyberpsychology principles. The integration of multiple perspectives enhances the construct validity of the resulting Behavioral Risk Intelligence Model (BRIM) by supporting cross-disciplinary generalizability. These analytic tools collectively contributed to mapping the literature's psychological insights into an AI-based architecture suited for enterprise adaptation.

Results and Findings

Drawing from a synthesis of 65 peer-reviewed studies published between 2021 and 2025, across forensic cyberpsychology, behavioral AI, and cybersecurity literature. Table 3 highlights four core behavioral indicators with varying levels of empirical support as precursors to insider threats, including algorithmic reinforcement, which shows the highest

support level at 89%, indicating its strong role as a catalyst in shaping behavioral conditioning and normalizing threat-related conduct through targeted digital feedback. Dark Triad traits follow at 74%, with studies consistently linking narcissism, machiavellianism, and psychopathy to high-risk, manipulative behaviors within organizational systems. Digital validation-seeking, supported by 68% of reviewed studies, emerges as a prominent early-warning trait, predisposing individuals to deception and manipulation in pursuit of self-worth. Identity confusion, though comparatively lower at 52%, still demonstrates substantial relevance, particularly in its association with ethical disengagement and susceptibility to ideological influences. Together, these findings validate the BRIM framework's psychological basis and underscore the importance of integrating behavioral risk intelligence into AI-driven insider threat models.

Table 3. Core Behavioral Indicators and Precursors to Insider Threat Risk

Indicator	**Definition**	**Percentage of findings from reviewed studies**	**Interpretation**
Digital Validation-Seeking	Excessive online behavior aimed at seeking approval and recognition to reinforce self-worth within digital contexts.	68% of reviewed studies identify digital validation-seeking as a precursor to deceptive behaviors and insider risk factors.	Strong predictor of deception and manipulation in insider threat profiles
Identity Confusion	Lack of a stable sense of self or coherent identity, leading to role conflict and potential ethical disengagement in an organizational context	52% of studies link identity confusion with increased vulnerability to internal manipulation and ideological infiltration within enterprises.	Associated with ethical disengagement and ideological drift
Dark Triad Traits	Traits comprising narcissism, machiavellianism and psychopathy	74% of studies find significant associations between Dark Triad traits and the likelihood of malicious insider activity	Correlated with high-risk behaviors and security violations
Algorithmic Reinforcement	AI-driven content personalization amplifies pre-existing biases and deceptive tendencies through repeated exposure to reinforcing stimuli, limiting critical reflection, and escalating manipulative behavior patterns.	89% of studies show that algorithmic systems contribute to behavioral conditioning, reinforcing insider threats dispositions through targeted feedback mechanisms and ethical desensitization.	Major catalyst for behavioral conditioning and threat normalization

The bar chart in Figure 2 provides a visual summary of the strength of association between specific behavioral factors and insider threat risk, as derived from the literature, highlighting the percentage of influence these behavioral risk factors exert on the development of insider risk activity. Algorithmic reinforcement demonstrates the highest correlation at approximately 89%, reinforcing its role as a dominant behavioral amplifier that conditions threat-conducive actions via repetitive digital stimuli and feedback loops. Dark Triad traits, at around 74%, stand out as core personality predictors, indicating a consistent link between narcissistic or manipulative tendencies and malicious insider behavior.

Digital validation-seeking, slightly lower at 68%, highlights how compulsive online approval-seeking can serve as a psychological vulnerability, especially in high-pressure or poorly regulated environments. Identity confusion, while the least correlated at 52%, still represents a meaningful risk vector, particularly in dynamic workplaces where employees may experience role conflict or ideological drift (Nordhall et al., 2025). Collectively, these correlations validate the BRIM framework's emphasis on behavioral modeling and demonstrate the value of integrating psychological markers into proactive insider threat detection systems.

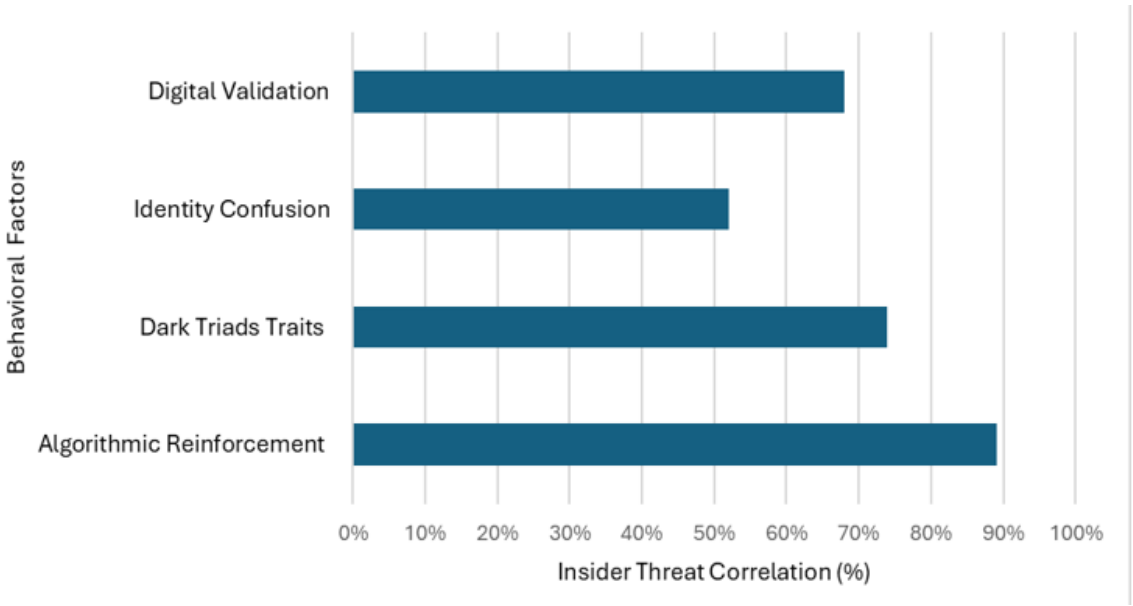

Figure 2. Behavioral Factors and Insider Threat Correlation

The pie chart in Figure 3 further illustrates the findings on behavioral risk factors and reinforces the primacy of algorithmic reinforcement, which accounts for 34% of all risk factors mentioned across the reviewed studies, emphasizing its dominant role in conditioning risky behavior conducive to insider threats. Dark Triad traits, comprising 28%, continue to demonstrate strong empirical support as enduring personality-based risk predictors. While identity confusion (20%) and digital validation-seeking (18%) though showing relatively lower representation, their presence across numerous studies underscores their relevance as latent psychological vulnerabilities that progressively advance the development of insider threat behaviors. These proportions suggest that while all four indicators are critical, AI-based models should prioritize dynamic behavioral reinforcement patterns and trait-based profiling to optimize early threat detection and ethical intervention strategies (Ali et al., 2025).

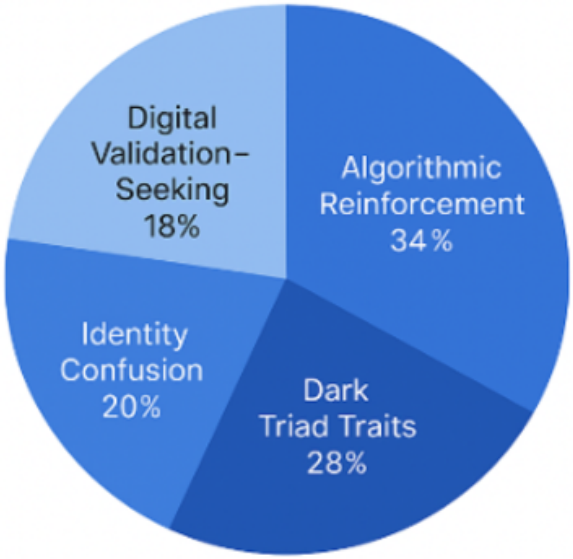

Figure 3. Risk Factors Supporting Psychological Indicators

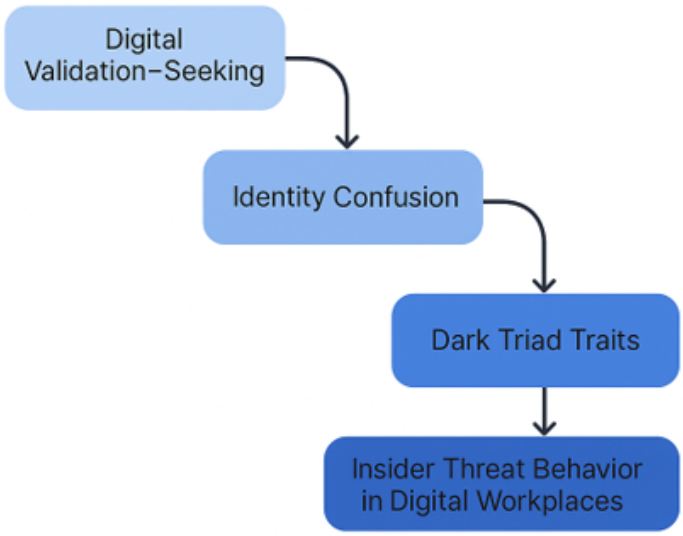

Figure 4. Progression Pathway from Psychological Traits to Insider Threat Behavior

This flow diagram in Figure 4 illustrates the psychological trajectory that underpins insider threat development as identified in this study's findings. The progression begins with digital validation-seeking, a salient behavior driven by the need for external affirmation, which over time fosters identity confusion, especially in high-surveillance or performance-intensive digital environments (Ohu & Jones, 2025b). This confusion can erode moral boundaries and professional alignment, paving the way for the emergence of Dark Triad traits such as narcissism, machiavellianism, and psychopathy. These traits, in turn, significantly increase the likelihood of insider threat behavior in digital workplaces, including data manipulation, sabotage, or unauthorized access (Ohu & Jones, 2025a). The figure encapsulates the theoretical foundation of the Behavioral Risk Intelligence Model (BRIM), emphasizing how early, subtle psychological patterns can evolve into significant security risks if left unaddressed (Zangana et al., 2025). The findings further synthesize into a conceptual framework that demonstrates how insider threat behavior can be understood as a progressive

interplay of psychological, personality-based, and environmental factors (Ruohonen & Saddiqa, 2025) as shown in Table 3. At the core of the framework is the Behavioral Risk Intelligence Model (BRIM), which draws from four key theoretical foundations highlighted in Table 4. The Validation Syndrome Diagnostic Triangle (VSDT) identifies emotional catalysts, such as impostor syndrome and unresolved frustration, that initiate behavioral drift. Dark Triad Theory explains how traits like narcissism and machiavellianism amplify the likelihood of manipulation and deception. Forensic Cyberpsychology provides the interpretive bridge between internal psychological states and external digital behaviors, while Behavioral Ethics ensures that AI applications based on BRIM uphold fairness, privacy, and ethical transparency. Collectively, these elements enable BRIM to function as both a diagnostic and predictive tool offering enterprises, law enforcement and healthcare systems a psychologically grounded and ethically responsible model for predicting and preventing insider threats.

Table 4. Theoretical Foundations of the BRIM Conceptual Framework

Theory/Model	**Contribution to BRIM**
Validation Syndrome Diagnostic Triangle	Identifies psychological catalysts of deceptive behavior (such as impostor syndrome, revenge motives)
Dark Triad Theory	Highlights traits linked to manipulation, deceit, and risk like machiavellianism and narcissism.
Forensic Cyberpsychology	Interprets digital behaviors in psychological terms within cyber contexts
Behavioral Ethics	Ensures ethically aligned AI profiling and decision-making

Table 5. Key Components and Variables Derived from Data Analysis

Component	**Key Variables/Constructs**
Psychological Risk Markers	Self-doubt, Machiavellianism, narcissism, identity confusion, validation-seeking behavior
Behavioral Drift	Changes in login patterns, peer interaction, content tone, or ethical disengagement
AI Detection System	ML models trained to detect latent risk indicators using supervised and unsupervised learning
Ethical Safeguards	Transparency, consent protocols, proportionality, de-identification
Environmental Stressors	Professional dissatisfaction, peer pressure, or ideological dissonance

The results of the data analysis were organized into five core components that collectively underpin the operational logic of the BRIM framework. As shown in Table 5, the first component, Psychological Risk Markers, encompasses internal constructs such as self-doubt, narcissism, and validation-seeking, recognized across the literature as precursors to insider threat behavior. The second component, Behavioral Drift, captures observable shifts in workplace conduct, such as altered login times, deteriorating tone in communications, or decreased peer interaction that may signal escalating risk. In our framework, AI Detection Systems serve as the technical engine of BRIM, employing supervised and unsupervised

machine learning models to identify these latent indicators before they culminate into harmful actions (Nepal & Joshi, 2022). To ensure responsible implementation, the model incorporates ethical safeguards, including consent, transparency, and de-identification protocols. Finally, Environmental Stressors such as job dissatisfaction or ideological tension are acknowledged as contextual amplifiers that can accelerate psychological vulnerabilities that eventually lead to insider threat actions (Waiganjo & Nandjenda, 2025). Together, these components structure the BRIM model's multi-layered detection logic, setting the stage for the detailed thematic findings presented in the next section.

Discussion

This study aimed to investigate how psychological and behavioral risk markers, such as validation-seeking, identity confusion, and Dark Triad traits, can be ethically integrated into AI-driven frameworks for insider threat detection. Our findings confirm that algorithmic reinforcement and digital validation-seeking are highly predictive of insider risk behavior, while ethically designed AI can mitigate bias and false positives in behavioral profiling (Habib & Nithyanand, 2025; Ohu & Jones, 2025a). Specifically, 89% of studies reviewed suggested that algorithmic systems contribute to behavioral feedback loops that amplify pre-existing cognitive vulnerabilities, reinforcing maladaptive behaviors such as deception and antisocial conduct, consequently reinforcing insider threat dispositions through these targeted feedback mechanisms and ethical desensitization. Behavioral risk, as quantified through thematic analysis, is dominantly shaped by psychological traits, most notably narcissism and self-doubt, suggesting that risk detection models must move beyond technical anomaly detection toward psychologically informed profiling (Ogunbodede et al., 2024).

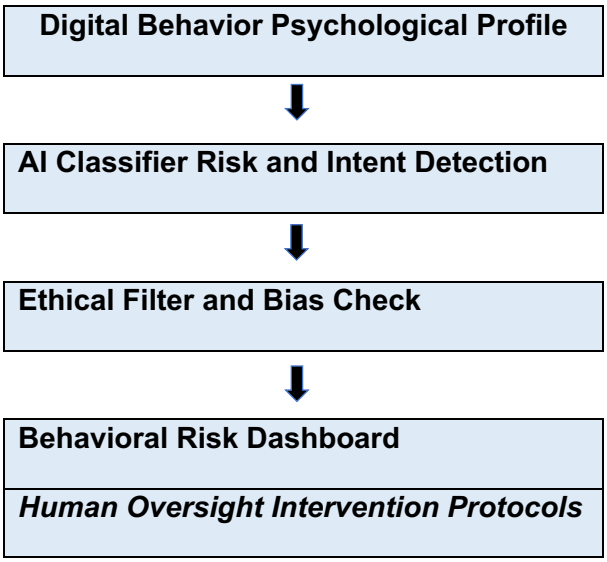

Figure 5. Behavioral Risk Intelligence Model (BRIM) AI/ML Insider Threat Mitigation Pipeline

As illustrated in Figure 5, the Behavioral Risk Intelligence Model (BRIM) operationalizes insider threat mitigation through a layered AI/ML framework, integrating passive digital behavior analytics, machine learning classifiers, and ethically guided risk scoring. The Psychological Input Layer where a Digital Behavior Psychological Profile is generated by passively collecting non-invasive behavioral indicators such as emotional tone shifts, excessive self-referencing, late-hour activity, and file browsing sequences,, which are interpreted using constructs from the Validation Syndrome Diagnostic Triangle (VSDT) (Ohu & Jones, 2025a) such as validation-seeking, identity confusion and Dark Triad traits,

which, as depicted in Figure 3, account for 66% of observed behavioral risk patterns. These inputs feed into the AI Risk Classifier Layer, where supervised and unsupervised learning models, including natural language processing and contextual anomaly detection, assign dynamic risk scores with minimal false positives (Ali et al, 2025;(Roy & Chen, 2024)), this is the algorithmic core of BRIM, where contextual anomaly detection algorithms analyze behavioral patterns over time rather than isolated events (Ali et al, 2025). These classifiers are trained on labeled datasets that distinguish benign from malicious behaviors and are further tailored to factors such as role, seniority, and workload. This allows the system to assign dynamic risk scores with greater precision and minimal false positives or alert fatigue. To ensure fairness and mitigate algorithmic bias, these scores are processed through the Ethical Filter & Contextual Auditor, which includes fairness checks, transparency protocols, and human-in-the-loop oversight (Chapagain et al., 2024; Pellegrino & Stasi, 2024), it conducts bias testing such as demographic parity, compares flagged behavior to contextual baselines, and supports human-in-the-loop oversight. This layer reinforces regulatory compliance and ethical integrity by integrating transparency and de-identification protocols, as emphasized by Chapagain et al. (2024), further ensuring that risk profiling is preventive rather than punitive, thereby avoiding stigmatization or legal overreach. As a final step, a user-friendly Actionable Intelligence Dashboard delivers interpretive outputs to analysts and compliance teams without breaching employee dignity or legal protections, a user-facing interface designed for cybersecurity analysts, HR teams, and legal officers that visualize behavioral trajectory maps, departmental heatmaps of psychological vulnerability, and non-punitive alerts such as automated check-in recommendations or reassignment flags. The dashboard can be integrated into existing Security Information and Event Management (SIEM) platforms and HR governance systems, providing real-time, psychologically informed decision support that respects ethical and privacy norms. This layered AI construct enables BRIM to function not only as a predictive tool but also as a responsible, adaptive system for identifying and mitigating behavioral risks in complex digital environments. Figure 5 further reinforces this process, showcasing how BRIM integrates psychological diagnostics, AI-driven inference, and ethical oversight into a single decision-support pipeline. Central to BRIM is the emphasis on behavioral drift, latent intent, and diagnostic triangulation rather than simple rule breaking. Importantly, the framework does not attempt to replace human judgment but rather augments it with real-time, context-aware behavioral insights.

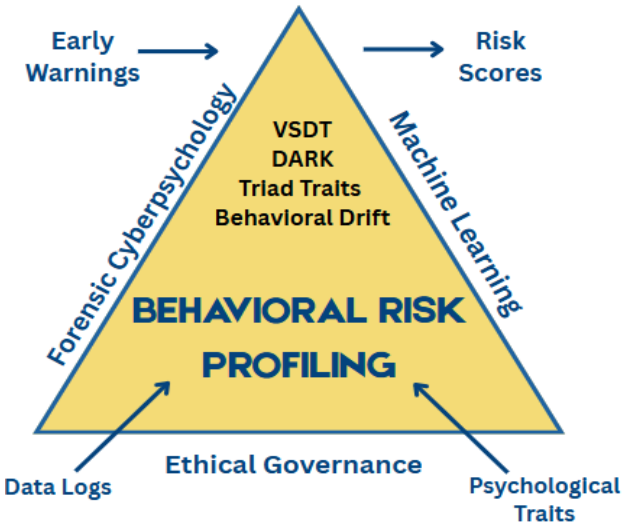

Figure 6. Behavioral Risk Profiling Pipeline

The Behavioral Risk Intelligence Model (BRIM) illustrated in Figure 6, further operationalizes insider threat detection by integrating psychological constructs with computational intelligence under an ethical oversight framework (Gayathri et al., 2024). This AI-driven conceptual model draws on forensic cyberpsychology, machine learning, and behavioral ethics to analyze cognitive behavioral indicators such as validation-seeking, self-doubt, and antisocial traits, thereby offering an ethically grounded alternative to traditional anomaly-based systems (Ohu & Jones, 2025a). At its core are three diagnostic domains, namely the Validation Syndrome Diagnostic Triangle (VSDT), Dark Triad traits, and behavioral drift, which collectively underpin individualized risk profiling (Ohu & Jones, 2025d). The VSDT posits that deceptive behaviors are driven by internal forces, including self-doubt, desire, and self-gratification, while Dark Triad theory contextualizes manipulative and high-risk tendencies (Gelman et al., 2024). These constructs are interpreted through cyber psychological analysis of digital behavior logs and processed via machine learning algorithms trained to detect latent behavioral risk signatures (Edlabadkar & Madisetti, 2024; Singh & Chattopadhyay, 2023). By generating dynamic and predictive risk profiles, BRIM enables proactive identification of threat trajectories. The ethical infrastructure centered on transparency, informed consent, and data de-identification ensures that outputs such as early warnings and risk scores are not only actionable but also accountable and rights-preserving. This layered approach transitions organizational security from reactive, forensic investigation to anticipatory, psychologically informed insider threat interventions that align with both enterprise goals and ethical standards. By incorporating insights from Dark Triad theory and the principles of forensic cyberpsychology to interpret behavioral drift, motivational conflict, and latent intent (Gelman et al., 2024), and training machine learning algorithms on these psychologically grounded constructs, enables the development of dynamic, predictive behavioral risk profiles that evolve over time (Mittal & Garg, 2023), supporting a proactive, context-aware, and ethically sound approach to insider threat prediction, prevention, and mitigation. Psychologically, threat behavior is rooted in cognitive-behavioral patterns associated with grievance formation, fixation, and pathway to violence. This model draws upon principles of behavioral threat assessment established in forensic psychology literature (Cornell et al., 2025) which conceptualizes pre-incident indicators as expressions of internal states of intent and capacity.

Broad Implications

BRIM offers a practical and ethically grounded augmentation to existing insider risk programs across both enterprise and healthcare domains. Designed around Privacy by Design (PbD) principles and anchored in ethical AI standards such as transparency, proportionality, and explainability, BRIM ensures that only de-identified behavioral data is processed, explicitly excluding protected class information (Alzaabi & Mehmood, 2024; Chapagain et al., 2024). It functions as a passive detection system, analyzing surface-level behavioral indicators without accessing private conversations or biometric data. From a corporate policy perspective, the deployment of BRIM requires clearly defined governance protocols that specify access control, interpretive authority, and ensure that interventions are designed to be supportive rather than punitive, preserving employee autonomy and dignity, and to preserve individual dignity, BRIM embeds employee-informed consent, transparent decision logic, and non-punitive interpretation protocols (Pellegrino & Stasi, 2024). For enterprises, BRIM enhances insider threat detection by moving beyond technical anomalies to assess psychologically driven risk trajectories. It integrates seamlessly with workflows such as HR performance reviews, whistleblower protections, and employee assistance programs (EAPs). Behavioral analysts, HR teams, and cybersecurity professionals can collaboratively use BRIM to monitor behavioral drift, emotional disengagement, and latent

intent, offering early intervention opportunities before insider risks escalate. As digital transformation expands and hybrid work environments increase exposure, organizations require human-aware detection models capable of assessing real-time psychological vulnerabilities (Ruohonen & Saddiqa, 2024). BRIM fulfills this need by linking behavioral profiling with AI-enhanced monitoring (Wei et al., 2024), preventing data leakage, sabotage, and operational compromise (Kamatchi & Uma, 2025). Pilot implementations in high-risk sectors such as finance, energy, pharmaceuticals, and defense contracting could validate its practical scalability and impact. In healthcare environments, where insider threats to electronic health records (EHRs) and protected health information (PHI) are both prevalent and dangerous, BRIM offers a new layer of protection for patient privacy, clinical workforce stability, and institutional integrity. Studies show that over 30% of hospital data leaks stem from insiders misusing legitimate access (Alder, 2025). By integrating BRIM with EHR audit logs, clinician wellness dashboards, and governance boards, healthcare institutions can identify early signs of burnout, emotional withdrawal, or malicious intent among clinical staff. Research highlights the importance of emotionally attuned security tools in reducing risks tied to burnout, secondary trauma, and identity conflict, particularly during crises like pandemics or restructuring (Park et al., 2025). BRIM responds to these calls by embedding behavioral insights into systems that flag psychological vulnerability without punitive surveillance. From a practitioner perspective, the framework offers a scalable structure for integration into criminal justice risk management systems, potentially aiding forensic psychologists, school threat teams, and intelligence analysts in pre-incident detection. The model also holds promise for broader regulatory compliance and governance alignment. It supports HIPAA, GDPR, and labor law standards by emphasizing fair access protocols, non-invasive monitoring, and supportive interventions over disciplinary action. In both healthcare and enterprise contexts, BRIM helps prevent unethical chart access, prescription fraud, policy evasion, and emotional withdrawal by flagging high-risk behavior patterns before systemic failures occur. Validation of the BRIM framework is proposed through pilot programs in collaboration with academic research centers, cybersecurity agencies, ethical AI labs, and healthcare delivery networks. These pilots may include use cases such as detecting ideologically motivated data misuse, burnout-linked disengagement, or early-stage deception. A mixed-methods validation approach is recommended, combining simulated datasets, qualitative interviews, and real-time classifier testing. Ultimately, BRIM offers a scalable, psychologically informed, and ethically robust framework for detecting behavioral risk earlier than conventional models, reducing false positives through cognitive triangulation, and reinforcing organizational trust through principled design (Gelman & Hastings, 2025; Ohu & Jones, 2025a; Pellegrino, 2025).

Bias Management and Validity Controls

While the conceptual model derived from this study, BRIM, demonstrates promise, several caveats must be noted. For instance, peer influence, organizational culture, and socioeconomic stressors may also shape insider threat behavior, but were not directly measured in this literature synthesis. Although our thematic coding framework emphasized familial and psychological factors, future research should incorporate peer relationships, workplace hierarchy, and cultural context to create more nuanced models. Moreover, while BRIM reduces false positives through contextual filters (Kantchelian et al., 2024), causation cannot yet be definitively established. Longitudinal and real-world validation studies perhaps using mixed-methods designs are needed to confirm predictive accuracy and reduce overfitting. Also, to enhance methodological rigor, this review relied on peer-reviewed sources published between 2021 and 2025, applied manual cross-checking of thematic codes, and ensured demographic neutrality by abstracting behavior traits rather than identity

markers. The study also employed literature triangulation across domains, including disinformation, romance scams, and insider threat psychology, lending robustness to its synthesized conclusions.

Study Limitations

The study's findings are subject to limitations, including the lack of primary empirical data that restricts the ability to draw context-specific inferences or test causal relationships directly.

The reliance on a narrative literature review and secondary data introduces the possibility that newly published research in 2025 may further refine or challenge existing understandings of psychological traits as insider threat markers. The interpretive nature of thematic coding, though systematically applied, may still carry latent researcher bias, however to mitigate these concerns, the study employed a structured narrative review methodology, incorporated inter-coder reliability checks during thematic synthesis, and limited the data pool to peer-reviewed publications from 2021 to 2025. These safeguards enhance the credibility, replicability, and conceptual grounding of the findings while acknowledging their empirical provisionality and the need for future primary-data validation.

Conclusion

This study introduced the Behavioral Risk Intelligence Model (BRIM) as an ethically grounded, AI-enhanced framework designed to predict and mitigate insider threats by analyzing cognitive and psychological indicators. This model integrates the principles of forensic cyberpsychology, machine learning, and behavioral ethics, and shifts the focus from technical anomalies to behavioral precursors, offering a proactive and context-sensitive alternative to traditional insider threat detection systems (Mladenovic et al., 2024). As visualized in Figure 5, the model operationalizes a multi-layered detection pipeline capable of interpreting latent behavioral risk signals and producing ethically filtered outputs, such as early warnings and risk trajectories (Koli et al., 2025). These findings affirm the study's central thesis that insider threats can be more accurately and responsibly addressed by targeting behavioral drift and psychological markers like validation-seeking, identity confusion, and Dark Triad traits constructs that have been overlooked in most conventional detection systems. The implications of this research are broad and impactful. In enterprise settings, BRIM can be deployed within Insider Threat Programs (ITPs) to monitor behavior during high-risk periods such as mergers, layoffs, or executive transitions (Nasir et al., 2021). Managers and compliance teams can utilize behavioral dashboards not just to detect risk, but also to inform training, build organizational trust, and benchmark psychological safety. In healthcare, BRIM strengthens data governance by identifying burnout or emotional distress in clinical staff that may precede inappropriate EHR access or disengagement. This supports patient safety initiatives while complying with HIPAA and privacy norms. The model also applies to defense and law enforcement sectors, where ideologically driven insider risks require a blend of psychological insight and ethical oversight. Compared to earlier studies that relied heavily on anomaly detection, BRIM's behavioral lens introduces a psychologically informed upgrade, enriching the field with an approach that reduces false positives and supports non-punitive intervention. The BRIM framework also responds to recent calls for AI systems that align with Privacy by Design (PbD) and GDPR principles. In contrast to surveillance-heavy solutions, BRIM is built on anonymization protocols, employee-informed consent, and explainable AI logic, as supported by Chapagain et al. (2024) and Pellegrino & Stasi (2024). By embedding ethics at the system level, BRIM delivers not only technological innovation but also policy relevance, offering actionable

insights for corporate HR, cybersecurity analysts, and hospital compliance teams. Importantly, this research makes a novel scholarly contribution by proposing an integrated framework that connects behavioral science with AI in real-time cybersecurity contexts, a domain where most existing literature remains fragmented or post-incident oriented. The study supports broader academic discourse by affirming that behavioral data, when interpreted through validated psychological theories such as VSDT and Dark Triad constructs, can enhance digital security frameworks without compromising human dignity.

Future Research Recommendations

Given the study's reliance on secondary data, future research should aim to empirically validate the BRIM framework through longitudinal, experimental, or mixed-methods approaches. Controlled deployments within corporate or healthcare environments could assess BRIM's real-world performance in detecting behavioral anomalies and preventing insider threats. Researchers might explore the impact of cultural context, peer dynamics, or organizational climate as potential moderating variables, factors that could not be fully controlled in the present study. Also, qualitative case studies, including interviews with threat analysts or frontline clinicians, would provide rich narratives to humanize statistical trends and identify latent drivers of behavioral drift. In business contexts, future studies should explore BRIM's integration into insider threat governance, including its utility during periods of organizational change, such as restructuring or digital transformation. Researchers should assess how BRIM dashboards influence managerial behavior, employee trust, and pre-incident intervention strategies. In healthcare, longitudinal research could evaluate BRIM's effectiveness in predicting staff burnout and EHR misuse, particularly under stress-intensive conditions such as pandemics or disruptions to clinical workflow. Pilot testing in academic medical centers, integrated delivery networks, or telemedicine platforms would provide important generalizability. For defense and national security applications, future work should evaluate BRIM's ability to detect ideological shifts, revenge motivation, or emotional dissociation in mission-critical personnel, use cases that remain under-researched. Furthermore, studies should explore user acceptance and ethical perceptions of AI-driven behavioral risk profiling among employees, and pose research questions such as How does employee awareness of behavioral AI tools influence organizational trust? or What safeguards are most effective in ensuring perceived fairness and reducing stigma associated with monitoring? Finally, as digital environments evolve, future studies should track how emerging platforms like Metaverse workspaces, AI co-pilots, or emotion-aware wearables might generate new behavioral signals or risk typologies that extend BRIM's applicability.

Final Thoughts

This study has demonstrated that behavioral risk indicators such as digital validation-seeking, identity confusion, and Dark Triad traits are significant predictors of insider threats (Ohu and Jones, 2025d), especially when contextualized within AI-powered analytic systems like BRIM. This paper also contributes to the literature on criminal psychology by proposing a behaviorally grounded risk assessment framework, and it's findings underscore that traditional cybersecurity tools, which rely heavily on technical anomalies, may overlook the psychological precursors to deviance that unfold long before a breach or policy violation occurs (Ruohonen & Saddiqa, 2024). By integrating machine learning with forensic cyberpsychology and behavioral ethics, the BRIM framework fills this critical gap, offering a multi-layered, ethically aligned model for predictive insider risk intelligence (Koli et al., 2025). The Behavioral Risk Intelligence Model (BRIM) represents more than a technical innovation and offers a paradigm shift in how insider threats are understood and managed across sectors In business enterprises, these findings reframe insider risk not as isolated

misconduct but as the outcome of identifiable psychological drift, distress, or disengagement. BRIM helps shift organizations from reactive to proactive, empowering leaders to use behavioral analytics not only to detect potential harm (Arroyabe et al., 2024), but to support employees during sensitive transitions such as restructuring, layoffs, or increased workload. In this way, BRIM reimagines employees not as risks to be surveilled, but as participants in a digital trust ecosystem, thus enabling smarter, safer, and more humane work environments. This transforms traditional insider threat detection into a strategy for ethical risk governance and employee-centered resilience. In healthcare, the implications are particularly urgent. The study suggests that emotional strain, ethical fatigue, or ideological conflict can manifest as risky behavior, often subtly and progressively. Healthcare workers frequently operate under conditions of high stress, emotional fatigue, and systemic constraints (Nagle et al., 2024). BRIM offers a structured and non-punitive method to flag these signals early, especially when integrated with wellness dashboards, audit logs, and clinical governance. This allows healthcare institutions to move beyond compliance-driven responses and toward a culture of care, protecting not only patient data but also the professionals entrusted with it, fostering safer, more compassionate, and accountable care environments. Ultimately, this research affirms that insider threat detection must evolve beyond binary threat models (Al-Mhiqani et al., 2024; NepalBasanta & Basanta Joshi, 2022; Zhang et al., 2021). With tools like BRIM, organizations across sectors can pursue behaviorally intelligent, ethically sound, and context-aware interventions, thereby balancing digital safety and human dignity. BRIM challenges the prevailing narrative of threat detection as reactive and punitive. It offers a forward-looking framework that blends psychological science with AI ethics, supporting proactive intervention, trust-building, and systemic integrity across enterprise and healthcare domains. As organizations increasingly seek holistic, ethical, and intelligent solutions to emerging cyber-behavioral threats, BRIM stands as a timely and transformative contribution to the field.

References

Abell, L., & Brewer, G. (2014). Machiavellianism, self-monitoring, self-promotion and relational aggression on Facebook. *Computers in Human Behavior*, *36*, 258–262. https://doi.org/10.1016/J.CHB.2014.03.076

Ahmed, W. (2024). DIGITAL TERRORISM: The Emerging Threat of Behavioral Manipulation in the Digital Age. *Journal of Digitainability, Realism & Mastery (DREAM)*, *3*(07). https://doi.org/10.56982/DREAM.V3I07.251

Al Lawati, A., Al Rawahi, N., Waladwadi, T., Almadailwi, R., Alhabsi, A., Al Lawati, H., Al-Mahrouqi, T., & Al Sinawi, H. (2025). Impostor phenomenon: a narrative review of manifestations, diagnosis, and treatment. *Middle East Current Psychiatry*, *32*(1), 1–12. https://doi.org/10.1186/S43045-025-00512-2/METRICS

Alder, S. (2025). *Healthcare Data Breach Report 2025*. https://www.hipaajournal.com/april-2025-healthcare-data-breach-report/

Ali, A., Husain, M., & Hans, P. (2025). Real-Time Detection of Insider Threats Using Behavioral Analytics and Deep Evidential Clustering. https://arxiv.org/pdf/2505.15383

Al-Mhiqani, M. N., Alsboui, T., Al-Shehari, T., Abdulkareem, K. Hameed, Ahmad, R., & Mohammed, M. A. (2024). Insider threat detection in cyber-physical systems: a systematic literature review. *Computers and Electrical Engineering*, *119*, 109489. https://doi.org/10.1016/J.COMPELECENG.2024.109489

Alohaly, M., Balogun, O., & Takabi, D. (2022). Integrating Cyber Deception Into Attribute-Based Access Control (ABAC) for Insider Threat Detection. *IEEE Access*, *10*, 108965–108978. https://doi.org/10.1109/ACCESS.2022.3213645

Alzaabi, F. R., & Mehmood, A. (2024). A Review of Recent Advances, Challenges, and Opportunities in Malicious Insider Threat Detection Using Machine Learning Methods. *IEEE Access*, *12*, 30907–30927. https://doi.org/10.1109/ACCESS.2024.3369906

Anju, A., Nithya Kalyani, M., Shalini, K., Ravikumar, H., Saranya, P., & Krishnamurthy, M. (2023). Detection of Insider Threats Using Deep Learning. *Proceedings - 2023 3rd International Conference on Pervasive Computing and Social Networking, ICPCSN 2023*, 264–269. https://doi.org/10.1109/ICPCSN58827.2023.00050

Arroyabe, M. F., Arranz, C. F. A., De Arroyabe, I. F., & de Arroyabe, J. C. F. (2024). Revealing the realities of cybercrime in small and medium enterprises: Understanding fear and taxonomic perspectives. *Computers and Security*, *141*. https://doi.org/10.1016/J.COSE.2024.103826

Bachi, D. M. (2025). Impostor Syndrome and Self-Doubt Among High Achievers. *Indonesian Journal on Health Science and Medicine*, *2*(1). https://doi.org/10.21070/IJHSM.V2I1.154

Bian, Xi., Wang, B., LI, K., & DU, Z. (2025). Navigating ethical decision-making in digital transformation: ethical climate, digital competence, and person-organization fit in China's banking sector. *Humanities and Social Sciences Communications 2025 12:1*, *12*(1), 1–15. https://doi.org/10.1057/s41599-025-05184-1

Brazil, K. J., Farrell, A. H., Boer, A., & Volk, A. A. (2024). Adolescent psychopathic traits and adverse environments: Associations with socially adaptive outcomes. *Development and Psychopathology*. https://doi.org/10.1017/S0954579424000051

Burnell, K., Trekels, J., Prinstein, M. J., & Telzer, E. H. (2024). Adolescents' Social Comparison on Social Media: Links with Momentary Self-Evaluations. *Affective Science*, *5*(4), 295–299. https://doi.org/10.1007/S42761-024-00240-6/METRICS

Ceroni, D. B., & Yalch, M. M. (2024). Influence of Childhood Maltreatment on Machiavellianism. *Journal of Aggression, Maltreatment & Trauma*, *33*(9), 1045–1054. https://doi.org/10.1080/10926771.2024.2358870

Chapagain, D., Kshetri, N., Aryal, B., & Dhakal, B. (2024). SEAtech: Deception Techniques in Social Engineering Attacks: An Analysis of Emerging Trends and Countermeasures. *ArXiv.Org*. https://doi.org/10.48550/ARXIV.2408.02092

Chen, A., Wong, C., Tarrit, K., & Peruma, A. (2024). Impostor Syndrome in Final Year Computer Science Students: An Eye Tracking and Biometrics Study. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *14694 LNAI*, 22–41. https://doi.org/10.1007/978-3-031-61569-6_2

Cornell, D. G., Maeng, J., Winter, S., Huang, F., Konold, T. G., Kerere, J., Afolabi, K., & Cowley, D. (2025). Behavioral Threat Assessment and Equity in Exclusionary School Discipline. *School Psychology Review*, 1–17. https://doi.org/10.1080/2372966X.2025.2457006

Edlabadkar, I., & Madisetti, V. K. (2024). Cybersecurity Risk Management through Behavior-Based Contextual Analysis of Online Logs. *Journal of Software Engineering and Applications*, *17*(06), 487–507. https://doi.org/10.4236/JSEA.2024.176027

Fominykh, Y. S. (2024). Information Security Violations in the Context of Digital Victimization of Children and Adolescents. *Victimology*, *11*(1), 69–77. https://doi.org/10.47475/2411-0590-2024-11-1-69-77

Gayathri, R. G., Sajjanhar, A., Uddin, M. P., & Xiang, Y. (2024). *FedAT: Federated Adversarial Training for Distributed Insider Threat Detection*. http://arxiv.org/abs/2409.13083

Gelman, H., & Hastings, J. D. (2025). Scalable and Ethical Insider Threat Detection through Data Synthesis and Analysis by LLMs. https://doi.org/10.1109/ISDFS65363.2025.11012066

Gelman, H., Hastings, J. D., Kenley, D., & Loiacono, E. (2024). Toward an Insider Threat Education Platform: A Theoretical Literature Review. *2024 International Conference on Computer and Applications (ICCA)*, 01–06. https://doi.org/10.1109/ICCA62237.2024.10928083

Gunuganti, A. (2024). Insider Threat Detection and Mitigation. *Journal of Mathematical & Computer Applications*, 1–6. https://doi.org/10.47363/JMCA/2024(3)184

Habib, H., & Nithyanand, R. (2025). YouTube Recommendations Reinforce Negative Emotions: Auditing Algorithmic Bias with Emotionally-Agentic Sock Puppets. https://arxiv.org/pdf/2501.15048

IC3. (2023). Federal Bureau of Investigation, Internet Crime Report. www.ic3.gov

Kamatchi, K., & Uma, E. (2025). Insights into user behavioral-based insider threat detection: systematic review. *International Journal of Information Security*, *24*(2), 88. https://doi.org/10.1007/s10207-025-01002-6

Kantchelian, A., Neo, C., Stevens, R., Kim, H., Fu, Z., Momeni, S., Huber, B., Bursztein, E., Pavlidis, Y., Buthpitiya, S., Cochran, M., & Poletto, M. (2024). *Facade: High-Precision Insider Threat Detection Using Deep Contextual Anomaly Detection*. https://arxiv.org/pdf/2412.06700

Koli, S., Thakur, R., Anas Saifi, & Singh, K. (2025). AI-Driven IRM: Transforming insider risk management with adaptive scoring and LLM-based threat detection. https://arxiv.org/pdf/2505.03796

Le, D. C., & Zincir-Heywood, N. (2021). Anomaly Detection for Insider Threats Using Unsupervised Ensembles. *IEEE Transactions on Network and Service Management*, *18*(2), 1152–1164. https://doi.org/10.1109/TNSM.2021.3071928

Liang, T., Wang, X., Ng, S., Xu, X., & Ning, Z. (2024). The dark side of mental toughness: a meta-analysis of the relationship between the dark triad traits and mental toughness. In *Frontiers in Psychology* (Vol. 15). Frontiers Media SA. https://doi.org/10.3389/fpsyg.2024.1403530

López, P. B., Pérez, M. G., & Nespoli, P. (2024). *Cyber Deception: State of the art, Trends and Open challenges*. http://arxiv.org/abs/2409.07194

Mittal, A., & Garg, U. (2023). Prediction and Detection of Insider Threat Detection using Emails: A Comparision. *2023 2nd International Conference on Electrical, Electronics, Information and Communication Technologies, ICEEICT 2023*. https://doi.org/10.1109/ICEEICT56924.2023.10157297

Mladenovic, D., Antonijevic, M., Jovanovic, L., Simic, V., Zivkovic, M., Bacanin, N., Zivkovic, T., & Perisic, J. (2024). Sentiment classification for insider threat identification using metaheuristic optimized machine learning classifiers. *Scientific Reports*, *14*(1), 25731. https://doi.org/10.1038/S41598-024-77240-W,

Nagle, E., Griskevica, I., Rajevska, O., Ivanovs, A., Mihailova, S., & Skruzkalne, I. (2024). Factors affecting healthcare workers burnout and their conceptual models: scoping review. *BMC Psychology*, *12*(1), 637. https://doi.org/10.1186/s40359-024-02130-9

Nasir, R., Afzal, M., Latif, R., & Iqbal, W. (2021). Behavioral Based Insider Threat Detection Using Deep Learning. *IEEE Access*, *9*, 143266–143274. https://doi.org/10.1109/ACCESS.2021.3118297

Nepal, S., & Joshi, B. (2022). *User Behavior Analytics for Insider Threat Detection using Deep Learning*. https://www.researchgate.net/publication/358983640_User_Behavior_Analytics_for_Insider_Threat_Detection_using_Deep_Learning

NepalBasanta, S., & Basanta Joshi. (2022). *User Behavior Analytics for Insider Threat Detection using Deep Learning*. https://www.researchgate.net/publication/358983640_User_Behavior_Analytics_for_Insider_Threat_Detection_using_Deep_Learning

Nordhall, O., Hörvallius, J., Nedelius, M., & Knez, I. (2025). Employees' experiences of personal and collective work-identity in the context of an organizational change. *Frontiers in Psychology*, *16*. https://doi.org/10.3389/fpsyg.2025.1382271

Norman Burrell, D. (2024). Exploring the Cyberpsychology of Social Media Addiction and Public Health Risks among Black American Women in the USA. *Health Economics and Management Review*, *5*(2), 14–31. https://doi.org/10.61093/hem.2024.2-02

Ohu, F. C., & Jones, L. A. (2025a). AI-driven forensic cyberpsychology intervention strategies for social media platform and school managers to mitigate cyber fraud at-risk adolescents. *Scientia Moralitas Conference Proceedings*, February 20-21, 2025, 114-131. http://dx.doi.org/10.5281/zenodo.15075890

Ohu, F. C., & Jones, L. A. (2025b). The intersection of cyberwarfare, social media, and adolescent self-esteem: A forensic cyberpsychology analysis. *Proceedings of the 20th International Conference on Cyber Warfare and Security (ICCWS 2025)*, March 28 – 29, 2025, 20(1) 332–344. https://doi.org/10.34190/iccws.20.1.3375

Ohu, F. C., & Jones, L. A. (2025c). An examination of digital validation-seeking behaviors in adolescents as precursors to romance scamming. *Scientia Moralitas Conference Proceedings*, February 20-21, 2025, 10-29 https://doi.org/10.5281/zenodo.14911844

Ohu, F. C., & Jones, L. A. (2025d). Validation syndrome: The root of deception and developmental predictors of dark triad traits in adolescents for forensic and developmental psychology. *International Educational Research, 8*(2), 67–86. https://doi.org/10.30560/ier.v8n2p67

Ogunbodede, O. O., Adewale, O. S., Alese, B. K., Akinyokun, O., Adewale, O. S., Alese, B. K., & Akinyokun, O. K. (2024). Insider Threat Detection Techniques: Review of User Behavior Analytics Approach Article in. In *International Journal of Research in Engineering and Science (IJRES) ISSN* (Vol. 12). www.ijres.org

Park, S.-H., Go, Y. H., Cho, H. J., & Yoon, M.-S. (2025). Impact of occupational death trauma on burnout among mental health professionals: the mediating role of secondary traumatic stress. *Frontiers in Psychiatry*, *16*. https://doi.org/10.3389/fpsyt.2025.1543681

Pellegrino, A., & Stasi, A. (2024). A bibliometric analysis of the impact of media manipulation on adolescent mental health: Policy recommendations for algorithmic transparency. *Online Journal of Communication and Media Technologies*, *14*(4), e202453. https://doi.org/10.30935/OJCMT/15143

Pennada, S. S. P., Nayak, S. K., & M, V. K. (2025). Insider Threat Detection Using Behavioural Analysis through Machine Learning and Deep Learning Techniques. *International Research Journal of Multidisciplinary Technovation*, 74–86. https://doi.org/10.54392/irjmt2527

Perenc, L. (2022). Psychopathic personality disorder and cybercriminality: an outline of the issue. *Current Issues in Personality Psychology*, *10*(4), 253–264. https://doi.org/10.5114/CIPP.2022.114205

Pérez-Torres, V. (2024a). Problematic use of social media in adolescents or excessive social gratification? The mediating role of nomophobia. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *18*(4). https://doi.org/10.5817/CP2024-4-2

Pérez-Torres, V. (2024b). Social media: a digital social mirror for identity development during adolescence. *Current Psychology*, *43*(26), 22170–22180. https://doi.org/10.1007/s12144-024-05980-z

Raza, R., Sheraz, F., & Saeed, I. (2025). The Impact of Workplace Ostracism on Knowledge Sabotage: The Mediating role of Job Induced Tension and Moderating role of Psychological Hardiness. *The Asian Bulletin of Big Data Management*, *5*(3), 31–47. https://doi.org/10.62019/C75X2Y87

Resende, M. M., Porto, J. B., & Gracia, F. J. (2024). Can we decrease unethical behavior at work? The role of ethical culture, ethical culture strength and collective moral identity. *Current Psychology*, *43*(8), 7153–7166. https://doi.org/10.1007/S12144-023-04615-Z/FIGURES/3

RJ Murad, H. (2024). Investigating The Impact of Digital Technology on Adolescent Identity Formation on Selected Students in SAIS: A Psychological Approach. *International Journal of Innovative Science and Research Technology (IJISRT)*, 2726–2736. https://doi.org/10.38124/IJISRT/IJISRT24APR1823

Roy, K. C., & Chen, G. (2024). GraphCH: A Deep Framework for Assessing Cyber-Human Aspects in Insider Threat Detection. *IEEE Transactions on Dependable and Secure Computing*. https://doi.org/10.1109/TDSC.2024.3353929

Ruohonen, J., & Saddiqa, M. (2025). What Do We Know About the Psychology of Insider Threats? *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, *613 LNICST*, 186–211. https://doi.org/10.1007/978-3-031-89363-6_11

Saddiqa, M., & Ruohonen, J. (2025). The Psychology of Insider Threats. *ICST Transactions on Security and Safety*, *9*(1). https://doi.org/10.4108/EETSS.V9I1.9298

Schluchter, T. (2024). Investigating User Perceptions of Mental Health Content on TikTok: A Comprehensive Exploration. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3613905.3647964

Schlund, R., & Zitek, E. M. (2024). Algorithmic versus human surveillance leads to lower perceptions of autonomy and increased resistance. *Communications Psychology*, *2*(1), 53. https://doi.org/10.1038/s44271-024-00102-8

Shahri, F., Zabihzadeh, A., Taqipanahi, A., Haromi, M. E., Rasouli, M., Saeidi Nik, A., & Eddy, C. M. (2024). I understand your pain but I do not feel it: lower affective empathy in response to others' social pain in narcissism. *Frontiers in Psychology*, *15*, 1350133. https://doi.org/10.3389/FPSYG.2024.1350133/BIBTEX

Singh, S., & Chattopadhyay, P. (2023). Hierarchical Classification Using Ensemble of Feed-Forward Networks for Insider Threat Detection from Activity Logs. *2023 IEEE 20th India Council International Conference, INDICON 2023*, 782–787. https://doi.org/10.1109/INDICON59947.2023.10440886

Tennakoon, H., Betts, L., Saridakis, G., Hand, C., & Chandrakumara, A. (2024). Exploring the effects of personal and situational factors on cyber aggression. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *18*(3). https://doi.org/10.5817/CP2024-3-7

Tokunbo, T., & Borisade, B. (2025). The Dark Triad in Organizational Leadership: A Systematic Review of Impacts and Interventions. *Journal of Research in Humanities and Social Science*, *13*(1), 32–36. https://doi.org/10.35629/9467-13013236

Trekels, J., Nesi, J., Burnell, K., Prinstein, M. J., & Telzer, E. H. (2024). Dispositional and Social Correlates of Digital Status Seeking Among Adolescents. *Cyberpsychology, Behavior, and Social Networking*, *27*(3), 187–193. https://doi.org/10.1089/cyber.2023.0342

Ullah, R. S., Naz, M., Alam, J. e, & Khan, A. U. (2024). The Role of Social Media in Shaping Adolescent Identities and Peer Relationships within Educational Settings. *Bulletin of Business and Economics (BBE)*, *13*(3), 575–584. https://doi.org/10.61506/

Waiganjo, I. N., & Nandjenda, L. S. (2025). Unveiling Insider Threats: Examining Vulnerabilities in an Organizational? Structure: A Case Study of NamPost. *OALib*, *12*(01), 1–10. https://doi.org/10.4236/oalib.1112797

Wang, J., Sun, Q., & Zhou, C. (2023). Insider Threat Detection Based on Deep Clustering of Multi-Source Behavioral Events. *Applied Sciences 2023, Vol. 13, Page 13021*, *13*(24), 13021. https://doi.org/10.3390/APP132413021

Wei, Z., Rauf, U., & Mohsen, F. (2024). E-Watcher: insider threat monitoring and detection for enhanced security. *Annales Des Telecommunications/Annals of Telecommunications*, *79*(11), 819–831. https://doi.org/10.1007/S12243-024-01023-7/FIGURES/11

Zangana, H. M., Sallow, Z. B., & Omar, M. (2025). The Human Factor in Cybersecurity: Addressing the Risks of Insider Threats. *Jurnal Ilmiah Computer Science*, *3*(2), 76–85. https://doi.org/10.58602/JICS.V3I2.37

Zhang, C., Wang, S., Zhan, D., Yu, T., Wang, T., & Yin, M. (2021). Detecting Insider Threat from Behavioral Logs Based on Ensemble and Self-Supervised Learning. *Security and Communication Networks*, *2021*, 1–11. https://doi.org/10.1155/2021/4148441

Zhou, R. (2024). Understanding the Impact of TikTok's Recommendation Algorithm on User Engagement. *International Journal of Computer Science and Information Technology*, *3*(2), 201–208. https://doi.org/10.62051/IJCSIT.V3N2.24