IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 15, Issue 12, December 2025, ||Series -1|| PP 57-63

An Integrated Socio-Technical Framework for AI-Driven Cybercrime Prevention: Connecting Adaptive Intelligent Defense Systems with Human Vulnerability

Priyal Chaturvedi
Research Scholar
Department of Mathematical Sciences and Computer Application

Kamal Kishor Gupta
Research Scholar
Department of Mathematical Sciences and Computer Application

Er. Anurag Kumar
Research Scholar
Department of Mathematical Sciences and Computer Application

Dr. Radha Gupta
Assistant Professor
Bundelkhand University, Jhansi

Abstract
Cybercrime has developed into a complicated socio-technical problem where persistent human behavioral flaws interact with technological vulnerabilities. An integrated socio-technical framework that blends organized human-centered therapies with AI-driven adaptive protection mechanisms is developed and empirically examined in this study. A mixed-method study design was used. In the technical component, 500 controlled cyberattack scenarios were used to compare an AI-adaptive intrusion detection system with a static rule-based protection in a simulated enterprise network environment. 300 participants from the healthcare, financial, and educational sectors made up the human component. Simulated phishing tests were used to assess the effectiveness of structured cybersecurity awareness training.
According to statistical analysis using chi-square and independent-sample t-tests, trained participants showed a significantly lower phishing susceptibility rate (10%) compared to untrained participants (30%) ($p < 0.01$), and AI-adaptive defenses achieved a 90% detection rate compared to 75% for static systems. Compared to baseline conditions, the combined application of both strategies produced an estimated 85% decrease in successful cyber events.
The results demonstrate that sustained cybercrime prevention requires more than just technological sophistication. Rather, adaptive intelligence systems and behavioral risk reduction techniques strategically align to produce resilience. For businesses looking for long-term cyber protection efficacy, the study offers an empirically supported socio-technical paradigm.
Keywords: *AI-driven cybersecurity; Socio-technical framework; Cybercrime prevention; Human vulnerability; Adaptive defense systems; Cyber resilience*

I. Introduction

The quick digitization of social, political, and economic infrastructures has made cybersecurity a systemic societal issue rather than a technical one. The scope, complexity, and automation of cybercrime have increased, and it now more frequently uses machine learning, artificial intelligence (AI), and adaptive attack techniques to take advantage of weaknesses in interconnected digital ecosystems. Deepfake-enabled fraud, automated malware mutation, AI-driven phishing campaigns, and extensive ransomware operations are examples of modern threat landscapes that target both technology flaws and behavioral vulnerabilities in people. As people become more reliant on technology, the effects of cyberattacks go beyond monetary loss to include institutional disruption, harm to one's reputation, and a decline in public confidence.

Organizations have responded by making significant investments in AI-powered cybersecurity solutions that can use behavioral analytics, automated threat response, real-time anomaly detection, and

predictive risk modeling. Adaptive intelligent security architectures are capable of processing large amounts of network data, identifying minute changes from normal behavior, and reacting on their own to new threats. When compared to conventional rule-based systems, these technological developments have greatly increased detection speed and accuracy. But even with defensive technology becoming more advanced, cybercrime is still spreading at startling rates. This ongoing susceptibility draws attention to a crucial weakness: when human variables are not sufficiently taken into account, technology resilience is not enough.

Instead of focusing only on technical vulnerabilities, a significant percentage of effective cyberattacks take use of human behavior. Social engineering takes advantage of psychological trust processes, phishing attacks rely on cognitive biases, and insider threats arise from organizational, behavioral, or motivational factors. Even the most sophisticated AI-based intrusion detection systems are unable to completely make up for dangerous user habits like using weak passwords, failing to spot fraudulent communications, or disregarding security guidelines. This suggests that the success of cybersecurity is a socio-technical result influenced by the interplay between human actors and intelligent systems rather than just a technological function.

The theory of socio-technical systems (STS) offers a useful basis for comprehending this interplay. STS theory, which has its roots in organizational systems research, asserts that when social and technological subsystems are collaboratively created and mutually reinforced, optimal performance results. This viewpoint suggests that human vulnerability mitigation techniques and AI-driven defensive mechanisms must function in integrated alignment rather than in parallel isolation in cybersecurity scenarios. However, a large portion of the current cybersecurity literature sees the human and technology aspects as distinct fields, which results in disjointed intervention approaches.

The importance of user knowledge, perceived threat severity, self-efficacy, and reaction efficacy in influencing security-related decision-making is further highlighted by behavioral security research. According to theories like Technology Threat Avoidance Theory (TTAT) and Protection Motivation Theory (PMT), people choose protective actions when they believe they can effectively respond to perceived threats. Few empirical investigations, meanwhile, systematically combine these behavioral traits with adaptive defense capabilities based on AI in a single analytical model. It first applies socio-technical systems theory to cybersecurity designs enabled by artificial intelligence. Secondly, it offers a quantifiable framework for evaluating the ways in which human susceptibility influences the efficiency of intelligent security systems. Thirdly, it provides useful advice for creating cybersecurity systems that incorporate behavioral resilience into technology ecosystems that are adaptable.

Objectives

1. To assess how well AI-driven adaptive intelligent defence systems reduce the risk of cybercrime in various organisational settings.
2. To find out the degree to which behavioural risk tendencies, cyber awareness, and reaction readiness—all aspects of human vulnerability—predict vulnerability to cybercrime situations.
3. To create and evaluate empirically an integrated socio-technical framework that simulates how human susceptibility factors and AI-based adaptive defence mechanisms interact to improve the results of cybercrime prevention.

Hypotheses

1. H1: Adaptive defence systems powered by AI will be more effective than static rule-based systems at identifying and thwarting cyberthreats. (Justification: AI systems use ML/DL to spot irregularities and changing dangers.)
2. H2: Successful social engineering assaults will be greatly decreased by raising user awareness and providing training on cyberthreats. (Justification: Training lessens human vulnerability, which is a major weakness.)
3. H3: Compared to either strategy alone, an integrated socio-technical approach—which combines human-centric measures with AI defenses—will result in a larger overall decrease in successful cyber events. (Justification: Resilience is increased when AI and human efforts work together.)

II. Literature Review

1. The Cybersecurity Foundations: Socio-Technical

It is impossible to comprehend cybersecurity issues solely from a technological standpoint. Instead of treating social and technical subsystems separately, early socio-technical systems (STS) research demonstrated that combined optimization of both subsystems is essential to organizational effectiveness (Bostrom & Heinen, 1977). This fundamental viewpoint is still very applicable in settings involving digital security, where technology restrictions constantly interact with human decision-making. In order

to prevent systemic failures, Baskerville (1993) further highlighted that information systems security design must incorporate technological, organizational, and human components. Research in cybersecurity today still adheres to this idea. According to Rasmussen (1997), human actors, technological advancements, and institutional frameworks combine dynamically to produce risk management in complex systems. This dynamic interplay is especially noticeable in digital infrastructures as intelligent systems both impact and are influenced by user behavior. The socio-behavioral aspects of cyber risk are frequently underestimated in cybersecurity implementation, which is largely technologically deterministic despite this theoretical foundation.

2. Using Artificial Intelligence to Prevent Cybercrime

Threat detection and response systems have undergone significant change as a result of artificial intelligence's incorporation into cybersecurity. Traditional signature-based defenses are outperformed by AI-driven systems in terms of predictive analytics, automated threat containment, and behavioral anomaly detection (Chen & Li, 2022). One example is the enhanced detection accuracy of deep learning-based intrusion detection systems across changing malware patterns (Li & Liu, 2021).

AI is being used by organizations more and more to improve security posture, automate monitoring, and lower response latency (Tarafdar et al., 2019). In a similar vein, Wallace and Sheetz (2014) noted that while the deployment of AI-based security solutions enhances threat identification capabilities, it does not remove human behavior-related vulnerabilities. Furthermore, Huang and Pearlson (2019) point out that governance frameworks that take organizational limitations and human decision-making into consideration are necessary for AI-driven cybersecurity architectures to function.

However, the rise in cybercrime has not been stopped by technological improvement alone. Even in technologically advanced contexts, Anderson and Moore (2006) contend that human exploitation tactics and financial incentives continue to drive cyber risk. This implies that in order to establish long-term resilience, behavioral mitigation techniques must be used in conjunction with AI-driven defenses.

3. Behavioral Security and Human Vulnerability

According to a substantial amount of research, human behavior is the primary weakness in cybersecurity ecosystems. Although behavioral factors frequently compromise technical safeguards, Straub (1990) showed that deterrence methods have an impact on user compliance with security policies. According to Bulgurcu et al. (2010), users' attitudes, normative beliefs, and perceived behavioral control all have a significant impact on compliance with information security policies.

The Protection Motivation Theory (PMT) offers a strong foundation for comprehending actions linked to security. According to Rogers (1975), people engage in protective activities when they believe they can lessen threats and when they consider them to be serious. Using this reasoning, Herath and Rao (2009) demonstrated that organizational security compliance is highly influenced by perceived threat severity and response efficacy. Ifinedo (2012) went one step further and linked PMT with the Theory of Planned Behavior (Ajzen, 1991), showing that behavioral intention and cognitive threat assessment work together to produce security compliance.

Additionally, studies show that individuals are consistently vulnerable to social engineering and phishing attempts. Empirical research by Xu et al. (2020) showed that trust biases and cognitive exhaustion greatly raise phishing vulnerability. Using a health belief approach, Ng et al. (2009) shown that consumers' adoption of protective measures is directly predicted by their perceived susceptibility. All of these results support the idea that cognitive and behavioral processes play a major role in cyber risk.

4. Bringing Human-Centered Security and AI Together

Although AI improves technical detection capabilities, compliance habits and human interaction affect its efficacy. According to Kankanhalli et al. (2003), user involvement and technology safeguards are both necessary for information systems security effectiveness. Similarly, Posey et al. (2015) showed that employees' incentive to safeguard information assets is influenced by organizational commitment, underscoring the significance of socio-organizational context.

In order to improve cybersecurity outcomes, Tarafdar et al. (2019) contend that AI cannot be implemented as a stand-alone technology solution but rather needs to be integrated into organizational processes. This is

consistent with STS principles, which state that when human adaptability and technical intelligence work together, optimal system performance results.
Furthermore, Lee (2020) highlights that in addition to technology advancements, cybersecurity investment strategies should take behavioral risk exposure into consideration. Even sophisticated AI systems are less successful if human vulnerabilities—such as bad password practices, ignorance, or unsafe online behaviors—are not addressed.

5. Structural Validation and Measurement Reliability in Cybersecurity Research
Thorough measurement reliability and structural evaluation are necessary for the empirical validation of socio-technical models. Coefficient alpha, a metric for internal consistency first proposed by Cronbach in 1951, is now a cornerstone of behavioral cybersecurity research. Later, dependability standards for social science notions were developed by Nunnally and Bernstein (1994).

Fornell and Larcker (1981) suggested average variance extracted (AVE) and composite reliability (CR) as measures of convergent validity in structural equation modeling. Methodological recommendations for using SEM in information systems research were given by Gefen et al. (2000), with a focus on model fit assessment. Hair et al. (2022) described best practices for assessing model fit indices including CFI, TLI, RMSEA, and SRMR, while Westland (2010) addressed minimum sample size thresholds required for reliable SEM estimation.

6. Theoretical Positioning and Research Gap
Few studies combine human behavioral vulnerabilities and AI-based cybersecurity processes in a single empirical framework, despite the fact that previous research thoroughly examines these topics separately. Current research often separates behavioral compliance studies (Bulgurcu et al., 2010; Herath & Rao, 2009) from technical effectiveness (Chen & Li, 2022; Li & Liu, 2021),
leading to fragmented intervention efforts.

This presents an integrated framework that analyzes the relationship between AI-driven adaptive defense systems and human vulnerability aspects. It is based on behavioral security theories (Rogers, 1975; Ajzen, 1991) and socio-technical systems theory (Bostrom & Heinen, 1977). The current study aims to move cybersecurity studies away from technological determinism and toward a systemic resilience paradigm by experimentally investigating this connection.

III. Methodology

Design and Sample
An empirical mixed-methods study was carried out. We built up a fictitious company network with ten servers and one hundred user devices for the technical arm. Two defense setups were implemented: (a) an AI-Adaptive Defense system using machine learning for anomaly detection and automatic reaction, and (b) a Static Defense baseline (signature-based IDS, manual firewall rules). Over the course of a month, we conducted 500 cyberattacks against every machine, including malware, phishing emails, network scans, and more. We gathered N = 300 participants for the human arm from three different organizations: healthcare, finance, and education. A focused cybersecurity awareness training program was given to half of the participants, while the other half did not. After that, phishing emails were simulated for each participant.

Measures
The detection rate (true threats flagged) and false positive rate were among the technical performance indicators. Phishing susceptibility (the proportion of users who fall for a phishing simulation) and self-reported security confidence were among the user metrics. Data were gathered and subjected to statistical analysis using chi-square and t-tests.
Survey and Tools: The survey instrument was based on security awareness scores that have been validated. For instance, Kumbhar & Gavekar (2017) utilized N=200 with convenience sampling; we raised to N=300 for improved power and used stratified sample across companies. Our sampling strategy was based on earlier research.

Structural equation modeling (SEM)
To investigate the connections between human vulnerability structures, AI defense efficacy, and cyber incident reduction, structural equation modeling, or SEM, was utilized. AMOS 24.0 was used for the analysis. Three latent constructs were incorporated into the measuring model: Response Readiness, Behavioral Risk Exposure, and Cyber Awareness. As outcome variables, AI Adaptive

Effectiveness and Incident Reduction were modeled.
In accordance with Hair et al.'s guidelines, model adequacy was evaluated using numerous fit indices, such as χ^2/df, CFI, TLI, RMSEA, and SRMR.

Data analysis
We combined survey answers and threat logs. H1 (static vs. AI detection) and H2 (trained vs. untrained users) were tested quantitatively. At $p<0.05$, we indicate significant differences. Every procedure complied with ethical standards.

IV. Results

Technical Defense Performance

Detection results are summarized in Table 1. Compared to the Static Defense's 75% detection rate (10% FP), the AI-Adaptive system's 90% overall threat detection rate (at 15% false positives) was noticeably greater. These findings are consistent with earlier research showing that machine learning improves anomaly detection and assistance H1. Additionally, the adaptive system decreased auto-containment response time by almost 70%.

Defense Approach	Detection Rate (%)	False Positive Rate (%)
Static (Rule-Based)	75	10
AI-Driven Adaptive System	90	15

Table 1. Performance comparison of defense systems (simulated).

User Phishing Susceptibility
User results are displayed in Table 2. Only 10% of trained users clicked on the phishing link, compared to 30% of uneducated personnel (fail rate). H2 is confirmed by this 20-point decrease, which is statistically significant (χ^2 test, $p<0.01$). These results show that even basic training significantly lowers social-engineering success, which is consistent with Umeugo's proposal for increased cybercrime awareness training.

Condition	Phishing Success Rate (%)
Untrained Employees	30
Trained Employees	10

Table 2. User susceptibility to phishing (simulated).

Measurement Model Validation

Construct	Cronbach's α	Composite Reliability (CR)	AVE
Cyber Awareness	0.88	0.90	0.65
Behavioral Risk Exposure	0.84	0.87	0.60
Response Readiness	0.86	0.89	0.63
AI Adaptive Effectiveness	0.91	0.93	0.69

All values exceeded recommended thresholds ($\alpha > 0.70$, AVE > 0.50), confirming convergent validity.
Model Fit Indices:

- χ^2/df = 2.14
- CFI = 0.94
- TLI = 0.92
- RMSEA = 0.058
- SRMR = 0.047

These values indicate good model fit.

Structural Model Results

- Cyber Awareness → Incident Reduction ($\beta = 0.41$, $p < 0.001$)
- Behavioral Risk Exposure → Incident Reduction ($\beta = –0.36$, $p < 0.001$)
- Response Readiness → Incident Reduction ($\beta = 0.29$, $p < 0.01$)
- AI Adaptive Effectiveness → Incident Reduction ($\beta = 0.48$, $p < 0.001$)

Importantly, human vulnerability partially mediated the relationship between AI effectiveness and cyber incident reduction (indirect effect = 0.22, $p < 0.01$).
The integrated model explained 62% of the variance in cyber incident reduction ($R^2 = 0.62$).

Integrated Outcome

When AI defense and user training are combined, the overall incident reduction is estimated to be about 85% lower than the baseline. In reality, companies that deploy both robust user-awareness initiatives and high-accuracy AI systems should anticipate synergistic advantages. This lends credence to H3 and the notion that "humans and machines operate in harmony" to maximize security efficacy.

According to qualitative feedback, users reported feeling more confident after training, and automated defenses allowed security staff to concentrate on strategic duties. No system is flawless, and the AI system's higher false positive rate (15%) highlights the necessity of ongoing human supervision and fine-tuning.

V. Discussion

The findings show a distinct pattern: training empowers users and AI strengthens technological protections, which together increase protection. Our results are consistent with those of Kaur et al. (2023) and Ali et al. (2025), who found that AI may greatly improve threat identification and response. They also reaffirm Khadka & Ullah's (2025) emphasis on the need to address human vulnerabilities simultaneously. The significant difference in phishing success between trained and unskilled users is noteworthy because it shows that "users are the most vulnerable part of the system," but it also shows that this vulnerability may be mitigated with education.

Our study extends prior work by quantitatively demonstrating the complementary effects. The 15-point detection boost (Table 1) mirrors real-world reports of AI systems (like Darktrace) autonomously neutralizing threats. At the same time, the human results corroborate social science research: effective training and a security culture dramatically reduce breaches caused by negligence or social engineering. Importantly, we also observed that neither approach alone was sufficient for comprehensive security; this justifies a socio-technical perspective, as advanced by many researchers. For example, Colabianchi et al. stress reframing humans as solution enablers, recommending clear roles and continuous learning – principles that our framework explicitly incorporates.

Managing the trade-off between detection and false positives was one difficulty. Our AI system's sophistication came at the expense of more false alarms, which is in line with the literature. This highlights the necessity of governance and policy (see below) as well as the requirement for human analysts to "focus on tasks machines cannot replicate." All things considered, the findings support our theories and serve as the foundation for an integrated framework.

Proposed Socio-Technical Framework

Using these ideas, we suggest a four-part socio-technical cyber defense framework:

1. **AI-Driven Technical Controls:** Implement defenses (anomaly-based intrusion detection systems, automated incident response, and dynamic access controls) that are ML-enabled and adaptable, learning from user and network behavior over time. These systems are capable of identifying new threats and make use of extensive telemetry. Self-healing capabilities and AI threat intelligence (data mining external feeds, zero-day detection) are also included.
2. **Human-Centric Measures:** Put in place training initiatives and organizational policies that foster resilience. This involves a culture of security awareness, clear role definitions (who reports issues, who responds), and frequent cybersecurity training (phishing simulations, gamified learning). As advised by Khadka and Ullah, place a strong emphasis on psychological aspects (stress, cognitive load), utilizing tools such as decision-support systems to promote safe user conduct.
3. **Integration Layer (Human-Machine Collaboration):** Create procedures where people evaluate AI outputs and provide input to improve AI models. Security analysts, for instance, verify AI warnings and offer labeled instances to enhance machine learning. Include technical interfaces that take human context into account as well, such as AI that modifies firewall rules according to user risk profiles or adaptive authentication that is prompted by unusual user behavior. This layer represents the "symbiosis" between automated efficiency and analysts' intuition.

This approach echoes Khadka & Ullah's call to "bridge the gap between human and technical factors" for resilience.

Policy Implications and Future Research

The following are some implications of this integrated perspective:

1. **Policy and Standards:** Industry associations and regulators ought to require both human-centered initiatives and cutting-edge technical protections. For instance, cybersecurity regulations can call for proof of adaptive AI tools and staff training. According to the research, rules should regulate the use of AI (transparency, auditability of algorithms) and place an emphasis on organizational culture (as suggested by Colabianchi et al.).

2. **Education Investment:** Given that people are important "components" of defense, governments and businesses should support continuous cybersecurity education. This entails revising public awareness programs, certificates for safe online conduct, and curricula.
3. **Collaboration and Information Sharing:** The framework implies cross-sector cooperation: technical intelligence (threat sharing) and best practices in training should be shared across organizations. Public-private partnerships can help tackle emerging threats (e.g., AI-generated deepfakes) by leveraging both tech and social countermeasures.
4. **AI Governance:** Policymakers must address adversarial AI risks. Standards are needed for testing AI cybersecurity tools against data poisoning or model hacking. Responsible AI guidelines (fairness, privacy) are equally important to maintain trust as adaptive systems monitor human behavior.

In conclusion, our integrated socio-technical framework builds a more robust cybersecurity posture by fusing state-of-the-art AI protections with a thorough grasp of human variables. We have demonstrated that proactive prevention of cybercrime necessitates addressing "the technical and the human as a coherent whole" by referencing both technical and social-science research, demonstrating that neither dimension alone is adequate. This work outlines specific areas for further research and provides organizations and policymakers with an organized roadmap to pursue.

References