The Role of Artificial Intelligence in SOC Operations: Adoption, Perception, and Workforce Impact

Daniel Boughton, and Iain Reid[1]

[1] School of Criminology and Criminal Justice, University of Portsmouth, Portsmouth, United Kingdom

Abstract

This paper examines the incorporation of artificial intelligence (AI) in cybersecurity operations, with a specific focus on applied machine learning within Security Operations Centres (SOCs). A mixed-methods approach was used, surveying 58 cybersecurity professionals from sectors including banking, healthcare, and technology, to investigate adoption, perceived advantages and drawbacks, and the implications for the human workforce. Quantitative results indicate extensive use of machine learning for alert triage and behavioural analytics, while qualitative findings underscore conditional trust, deficiencies in training, and the evolving dynamics of analyst roles. Thematic analysis identified fundamental categories such as AI as copilot, explainability and trust, and ethical risk. These findings indicate that although machine learning improves efficiency and alleviates cognitive burden, its adoption relies on transparent governance, continuous human monitoring, and professional development. This study enhances academic and industrial discussions by prioritising practitioner perspectives and clarifying the socio-technical considerations of machine-learning adoption in SOCs.

Keywords

Artificial intelligence; machine learning; Security Operations Centres; SOC; human factors; workforce1

1. Introduction

The cybersecurity landscape is evolving quickly, as organisations increasingly adopt automation to match the complexity of threat actors and the scale of digital infrastructure [1, 2]. Security Operations Centres (SOCs) function as essential environments tasked with real-time threat detection, investigation, and mitigation. SOC analysts often experience overwhelming alert volumes, false positives, and the cognitive strain associated with continuous operations [3, 4]. These difficulties have accelerated the integration of *artificial intelligence* (AI) as a means to improve operational efficiency and mitigate human fatigue.

The use of AI in cybersecurity includes machine learning, natural language processing, and behavioural analytics systems that can detect abnormalities, prioritise warnings, and facilitate incident response [2, 5]. Despite the extensive technical literature on algorithmic performance and threat detection efficacy, there exists an important gap in research that critically evaluates the impact of AI on SOC operations relating to workforce dynamics and human-machine collaboration [6, 7].

This study aims to ascertain whether the integration of AI into SOC processes has reduced challenges including alert fatigue and manual triage, or if it has created new complexities affecting trust, skill deficiencies, and explainability. Organisations must strategically evaluate if AI enables a reduction in human personnel or involves retraining and role transformation. Researchers such as Taddeo and Floridi (2018) [8] contend that AI ought to augment—not supplant—human decisionmaking, a stance corroborated by experts in contemporary literature and industry evaluations [4, 9].

[*] *STPIS25: The 11th International Workshop on Socio-Technical Perspectives in IS, September 17–18, 2025, Skopje, North Macedonia*

[1*] Corresponding author.

daniel.boughton@myport.ac.uk (D. Boughton); iain.reid@port.ac.uk (I. Reid)

0009-0003-1574-8572 (D. Boughton); 0000-0003-4072-7557 (I. Reid)

This study reports survey results from 58 cybersecurity experts across several areas, including finance, healthcare, and technology. Participants supplied both quantitative data and qualitative insights, presenting a mixed-methods perspective on the degree of AI adoption, trust, and its perceived transformative or disruptive impact. Principal themes explore partial integration, perceived advantages of automation, apprehension regarding algorithmic bias, and ambiguity concerning long-term employment implications.

The structure of this paper is as follows: Section 2 analyses pertinent material regarding AI in SOCs and organisational transformation. Section 3 defines the research method. Section 4 outlines the findings. Section 5 examines issues within the context of broader cybersecurity and technical advancements. Section 6 concludes the paper and explores potential study areas.

In this paper, the term *artificial intelligence* is used in a focused sense to denote the application of machine-learning techniques within security operations centres (for example, anomaly detection and alert triage). It does not extend to generative models such as large language models, which pose different opportunities and limitations and fall outside the scope of this study [23].

2. Background and literature review

The use of AI in cybersecurity has been growing rapidly in recent years, driven by the necessity to navigate increasingly intricate threat landscapes. Multiple studies have evidenced the efficacy of AI-driven tools—especially machine learning (ML) algorithms—in identifying malicious behaviour, detecting anomalies, and enhancing response times [2, 5, 10, 21]. These algorithms can analyse extensive datasets at speeds that far surpass human capacities, making them ideal for tasks such as log analysis, intrusion detection, and threat scoring [1].

SOCs, serving as the frontline of enterprise defence, have increasingly embraced AI-driven solutions to enhance their operations. SOC analysts are often overwhelmed by substantial quantities of low-quality alerts, a significant portion of which yield false positives. This issue, known as alert fatigue, is identified as a primary factor in analyst burnout and errors [3]. AI offers a promising solution to mitigate these pressures by filtering alerts, prioritising those posing the highest risk, and automating initial triage processes [2, 11]. As per Swimlane (2024) [4], SOCs are increasingly implementing AI co-pilots that assist, rather than replace, human analysts by focusing on actionable threats.

A significant portion of the academic literature has been focused on algorithmic efficacy, frequently within the framework of benchmark datasets or controlled testing environments [6, 12]. Nonetheless, there is an absence of studies examining the practical implications of deploying AI in active SOC environments. Research addressing implementation primarily emphasises technological factors, with limited attention given to human and organisational impacts, including alterations in work roles, trust in automation, and ethical accountability [7]. Mittelstadt et al. (2016) [13] contend that although AI might enhance operational effectiveness, it may also pose new problems if openness and accountability are not included into its development and execution.

Human–AI collaboration is a frequent topic in academic journals, with many researchers proposing that AI should enhance rather than replace human decision-making [8]. In this way, analysts utilise AI-generated outputs as decision-support tools yet retain accountability for the final conclusion. Even so, confidence in AI remains conditional. Gunning and Aha (2019) [6] emphasise that the absence of explainability across many AI systems, especially those utilising deep learning, undermines user confidence and may result in underutilisation.

In operational environments, professionals frequently depend on experience and context to inform their decisions. The incorporation of AI necessitates both technological validation and adherence with institutional rules, privacy constraints, and legal frameworks [14]. Organisations implementing AI in SOCs must confront many challenges, which range from ethical considerations to practical questions regarding the separation of human and machine responsibilities [13].

Recent industry surveys highlight a contradictory narrative: while AI is widely embraced, training and organisational capabilities are insufficient compared to deployment. Darktrace (2022) [9] indicates that many analysts using AI tools claim inadequate or nonexistent training, resulting in difficulties with interpretation and accountability in decision-making. The skills gap lowers operational performance and heightens the risk of dependence on tools that users are unable to fully understand.

Literature indicates that the integration of AI is expected to transform the nature of cybersecurity employment throughout time. Rather than eliminating jobs, AI is expected to shift human analysts' focus towards higher-order cognitive tasks, including the examination of complex threats, data contextualisation, and managing of automated systems [4, 8]. This transition demands an update of recruitment, training, and professional development strategies within the security industry.

In summary, though the technological effectiveness of AI in cybersecurity is widely recognised, its organisational, ethical, and human implications remain inadequately understood. A notable gap exists in research regarding practitioners' experiences and assessments of AI in their daily operations, especially within high-pressure settings such as SOCs. This study bridges the gap through exploring the perceptions of cybersecurity professionals, concentrating on trust, usability, and the future of human roles in AI-enhanced security operations.

Alongside enhancing detection capabilities, AI technologies have been suggested as a method for strengthening threat intelligence aggregation. AI systems can discern previously unrecognised patterns or correlations that human analysts might miss by utilising data from various internal and external sources [5, 10, 15]. These methods allow AI to uncover latent relationships or subtle anomalies across diverse datasets that are not readily visible through manual analysis. This functionality is particularly advantageous in advanced persistent threat (APT) situations, where adept adversaries seek to avoid detection by assimilating into routine activities. The incorporation of AI into threat intelligence platforms has demonstrated enhancements in detection latency and expedited reaction times [16].

The organisational context in which AI is implemented is equally crucial. Jalali and Kaiser (2018) [3] assert that institutional culture, leadership endorsement, and change management techniques substantially influence the reception and utilisation of AI solutions. Organisations with a more adaptable or agile security posture may adopt AI more effectively, whereas rigid hierarchical organisations may face pushback, particularly from employees apprehensive about automation and job security. These organisational issues must be evaluated concurrently with technical design and implementation.

The ethical and legal ramifications of AI in cybersecurity have garnered heightened scrutiny. AI-driven surveillance systems can contest conventional standards on data privacy and consent [14]. This has elicited demands for enhanced openness and explainability, not merely as a means of fostering confidence, but as a legal obligation under regulations such as GDPR. Mittelstadt et al. (2016) [13] advocate for the integration of ethical issues into the AI tool development process, assuring alignment with societal values and institutional responsibilities.

Moreover, interdisciplinary research indicates that effective AI integration may be enhanced by cross-functional teams consisting data scientists, cybersecurity experts, legal consultants, and ethicists. This collaborative framework can foresee conflicts and ensure that technological decisions correspond with operational and regulatory requirements [8, 14]. Consequently, the successful implementation of AI in SOCs is not merely a technology issue but a socio-technical one that necessitates synchronised governance, communication, and a collective comprehension across several domains.

New policy frameworks, including the European Union's Artificial Intelligence Act and the NIST AI Risk Management Framework, are beginning to shape organisational governance of AI use. These frameworks emphasise the significance of human oversight, robustness, and transparency— principles that align closely with the issues arising from this study [17, 18]. Their focus on 'high-risk' systems emphasises the need for explainability in cybersecurity, particularly since real-time decisions affect critical infrastructure.

Research in related fields, such as finance and healthcare, shows similar trends. In healthcare, the implementation of AI has enhanced diagnostic efficiency; however, issues of explainability and clinician trust remain as limitations. These findings provide significant parallels to the SOC context, where decisions must be rapid, reasonable, and made under pressure [19].

While much of the foundational literature predates the rapid adoption of generative AI technologies after 2022, these works remain central to understanding machine learning's operational role in SOCs. More recent literature has turned strongly toward generative AI and LLM-based models and their vulnerabilities [24], but this layer is mostly beyond the operational scope of SOC ML systems addressed in this paper.

3. Methodology

This study employed a convergent mixed-methods design to investigate the use and perception of AI within SOCs. The goal was to understand the current adoption of AI in cybersecurity processes, the perceptions of professionals in the field, and whether its implementation is regarded as enhancing or eroding human duties. A summary of the core research aims includes evaluating current AI capabilities within SOCs, assessing practitioner trust in these tools, and understanding whether AI is perceived as a complement to or replacement for human judgement. This methodological approach enabled the triangulation of quantitative trends with rich qualitative insights, enhancing the study's ability to address its complex research questions.

Respondents were asked about the role of "AI" in SOC operations without a technical definition being imposed. While this provides authentic practitioner perspectives, it introduces interpretive variability—some may conflate machine learning with broader notions of AI [23].

A cross-sectional survey was chosen as the principal method for data collection. The survey aimed to collect both quantitative and qualitative data through multiple-choice questions, Likert-scale responses, and open-ended prompts. This approach allowed the identification of statistical trends and in-depth respondent perspectives, illustrating the intricate and complex nature of the issue under investigation [1].

The survey comprised 40 questions, logically organised into sections addressing demographics, organisational context, current use of AI tools, perceived advantages and disadvantages, trust in AI systems, ethical considerations, and expectations of future changes to SOC roles. Adaptive logic was utilised to present specific questions conditionally, dependent upon prior responses; for instance, enquiries on problems with implementation were presented solely to respondents who acknowledged AI utilisation within the company they work for. This rationale reduced fatigue and ensured the relevance of questions to all participants.

Participant recruitment was carried out using professional networking platforms, primarily via LinkedIn. The survey link was disseminated through cybersecurity-oriented discussion forums and direct engagement within the researcher's professional network. This opportunity sampling approach, while pragmatic, presents some inherent limitations around generalisability. A total of 61 replies were gathered; of these, 58 were considered legitimate for analysis after excluding three due to insufficient consent. All eligible participants verified their active engagement in cybersecurity-related positions.

Participants spanned many different industries, including banking, healthcare, technology, and government. The majority of participants had more than five years of experience in cybersecurity, holding roles such as analysts, engineers, CISOs, and policy consultants. The size of organisations varied significantly, with a significant number of individuals employed by large companies (exceeding 1,000 people), while others worked in small-to-medium enterprises or start-ups. This diversity offered broad understanding across several operational contexts.

The University of Portsmouth provided ethical approval for the study. Informed consent was obtained before the start of the survey, and participants were made aware of their right to withdraw at any point before submission. The survey was entirely anonymous—no personal data was gathered, and IP addresses were not recorded. Data was securely stored in accordance with University and GDPR regulations [14].

Quantitative data were analysed using descriptive statistical techniques. Frequencies and percentages were calculated for closed-response questions to identify patterns in adoption, training, and attitudes towards AI. Charts and tables were employed to illustrate essential findings. A thematic analysis methodology was used for the qualitative data, implementing an inductive coding approach. Narrative answers were analysed for common concepts and grouped into themes such as 'trust and transparency', 'alert fatigue', 'AI as co-pilot', and 'training and role evolution'.

This technique enabled the researcher to summarise the narrative content into significant findings consistent with the study's aims [20].

Thematic analysis followed Braun and Clarke’s (2006) [20] six-phase framework, comprising:

- Familiarisation with the data
- Generation of initial codes
- Searching for themes
- Reviewing themes
- Defining and naming themes
- Producing the report

This systematic process enabled the identification of emergent issues that were not directly captured by the quantitative measures, enriching the overall interpretation.

This study had several constraints. Utilising convenience sampling through professional networks has potential selection bias and limits the generalisability of the findings. In addition, the cross-sectional design captures observations at one specific moment in time, which may not reflect evolving perceptions as AI adoption continues. Nevertheless, the results provide significant practitioner focused insights and establish a basis for subsequent, more specific research.

4. Results

The study responses offer an in-depth assessment of the perception and utilisation of AI within the cybersecurity sector, especially in regard to SOC environments. The study identified recurring themes in both quantitative trends and qualitative narratives. Participants had diverse experiences with AI technologies; however, the majority expressed cautious optimism about their potential to enhance threat detection, prioritise alerts, and reduce workload.

Among the 58 valid responses, 68% reported that their organisation currently uses at least one AI-driven technology in cybersecurity operations. Typical uses included behavioural analytics, log analysis, natural language processing (NLP) for report generation and enrichment, and anomaly detection. Among those not already using AI, 64.5% indicated that their organisation intends to begin using such technologies in the future.

Participants were asked about particular AI applications. The primary applications included:

- Detection and triage of threats (82%)
- Analysis of network traffic (71%)
- Behavioural pattern identification (64%)
- Phishing detection and response (49%)
- Automated ticketing and response processes (37%)

Qualitative responses corroborated these findings, with numerous individuals expressing how AI helps with handling alert volume and prioritising the most critical incidents. One participant remarked, "*AI has assisted in reducing the noise. We no longer spend hours on false positives—we focus on what matters.*" Another noted, "*We used to get swamped, now AI helps weed out the nonissues before we even log in.*" The issue of efficiency and reduced cognitive load was prevalent across the dataset.

A significant observation was the effect of AI on alert fatigue. 84% of participants indicated the success of AI in filtering or reducing irrelevant alerts. Several participants provided clear endorsements of the effect AI has had, including: "*My team has stopped burning out every week—AI filters out 80% of the junk now.*" and "*It's the only thing keeping us from drowning in alerts.*" Automation was widely described as having enhanced daily work rhythms and reduced the emotional strain associated with operating in a high-pressure SOC environment.

Notwithstanding the practical advantages, confidence in AI systems remains uncertain. Only 19% of participants indicated a high degree of trust in AI recommendations. 34% expressed partial trust, frequently referencing apprehensions over transparency or 'black-box' results. One respondent stated, "*We use AI to flag issues, but a human still needs to verify. It's useful, but we don't let it run unattended.*" Another noted, "*I trust it for correlation, not for action. It's good at patterns, bad at context.*" This viewpoint corresponds with the broad academic research cautioning against excessive dependence on opaque systems [6,7].

When it comes to role evolution, numerous participants characterised AI as facilitating a transformation in duties. Activities including log review, initial triage, and fundamental correlation were characterised as progressively automated. This enabled analysts to concentrate more on investigative and decision-making activities. One participant commented, "*We spend less time grinding through tickets and more time hunting threats now.*" Another expressed, "*I've shifted from doing routine work to actually analysing complex behaviour.*" A prevalent expression in story answers was "*AI as co-pilot*"—a notion highlighting enhancement rather than substitution. This reflects recent discoveries that AI can enhance, rather than eradicate, the function of human analysts [4, 8].

Training and skill deficiencies were also significant. Fifty-seven percent of respondents indicated that they had not had formal training in AI, despite their organisation implementing such capabilities. This inconsistency was identified as a significant obstacle to successful adoption, with participants advocating for more accessible education specifically designed for cybersecurity professionals. It was observed, "*The instruments exist, yet individuals lack the knowledge to utilise them effectively. That is where errors occur.*" Another participant noted, "*I've got access to powerful tools, but no idea how they work—training has not caught up.*"

Nearly half of the participants expressed ethical concerns. Concerns included the methods of data collection and processing, as well as overarching questions on accountability for AI-generated choices. Many of the participants expressed concerns over privacy and the possibility of AI systems producing false positives as a result of biased training data [13]. Several quotes illustrated this unease:

"*We have no way to audit what data it learned from—so how can we be sure it's fair?*" and "*It flagged a VIP for phishing based on bad training inputs—it caused chaos.*"

Thematic examination of open-ended replies yielded five key recurring themes:

> 1. Alert fatigue: AI mitigates alert frequency and cognitive strain. *"We went from 500 daily alerts to 50—most of which now matter."*
> 2. Trust and transparency: Significant emphasis on explanation and human supervision. *"If I can't see why it flagged something, I don't act on it."*
> 3. Workflow transformation: Transition from technical triage to strategic analysis. *"I spend less time being a log-monkey and more time being an analyst."*
> 4. Training deficiencies: Insufficient organisational support for skill enhancement. *"Tools are great—but we're flying blind on how to use them properly."*
> 5. Ethical and data-related issues: Concerns regarding fairness, accountability, and privacy. *"I worry more about the AI's bias than the hacker's."*

While most respondents regarded AI as beneficial, several emphasised the necessity of careful adoption, ongoing human supervision, and investment in training and ethical governance. These observations reinforce the notion that AI's function in SOCs should be perceived not as a substitute technology, but as an integral component of a broader transition towards enhanced decision-making and flexible operational frameworks.

When asked about the impact of AI on their work, 62% of respondents indicated a significant alteration in their regular routines. This encompassed less time allocated to initial triage and greater involvement in strategic investigation activities. Participants from larger businesses were more inclined to indicate extensive AI integration, whereas individuals in smaller enterprises reported restricted deployment, frequently attributing this to budgetary or skill limitations. A respondent remarked, "*We have some AI capability, but without a dedicated team to maintain and tune the models, it's underutilised.*"

Some professionals expressed unease about the speed of AI-generated answers, voicing concerns that automation would implement containment measures bereft of adequate human context. About 29% of respondents indicated that their organisation restricts AI's autonomous actions to low-impact situations, maintaining human oversight over important response decisions. This indicates an increasing interest in human-centred AI deployment models that reconcile responsiveness with responsibility [6].

A consistent feature in qualitative replies was the variation in trust levels based on the type of AI application. AI technologies utilised for log correlation and enrichment received better trust ratings compared to those producing event severity scores or risk evaluations. Participants stressed the need for transparency and the provision of contextual evidence to substantiate AI choices. One participant remarked, "*If the system provides a recommendation without an explanation, I do not act on it.*" This substantiates findings in the literature that explainability is not solely a technical attribute but an essential condition for operational confidence [13].

The survey revealed that organisations with advanced AI integration exhibited a more robust culture of collaboration between analysts and AI developers. These teams frequently functioned within a feedback loop, utilising human observations to enhance the system's models and settings. This reciprocal relationship enhanced both trust and performance. These techniques are akin to Development, Security, and Operations (DevSecOps) models, wherein the continuous incorporation of feedback into automation pipelines is regarded as a best practice [7]. While some recent literature proposes the evolution toward AI/Machine Learning Security Operations (AI/MLSecOps) as an emerging discipline, this research found no explicit implementation under that specific term.

Variations in responses between industries were also seen. Government and critical infrastructure professionals indicated that the approval processes for AI deployment have become more stringent, frequently including the involvement of legal or compliance teams. The respondents were more inclined to identify regulatory and reputational problems as obstacles to automation. At the same time, private sector representatives emphasised the importance of cost-efficiency and productivity enhancements, with one analyst remarking, "*Automation is our sole avenue to scalability, but we must guarantee it does not inflict more harm than good.*" These disparities show the impact of institutional priorities and limitations on AI implementation.

5. Discussion

This study's findings reinforce key themes in the broader cybersecurity literature while offering novel, practitioner-focused insights into the impact of AI integration in SOCs. Although previous research has established the technical effectiveness of AI in threat detection and automation [2, 5], the study contributes to the growing literature exploring the practical consequences for human roles, organisational workflows, and trust dynamics [6, 7].

The findings suggest that AI is regarded favourably regarding efficiency and efficacy. This corresponds with studies indicating that AI solutions might mitigate warning fatigue and allow analysts to prioritise significant situations [3, 9]. The conditional character of faith in AI, as indicated by most respondents, reflects wider apprehensions over the opacity of algorithmic systems and the potential for misclassification [13]. In practice, numerous organisations appear to have embraced a human-in-the-loop methodology, wherein AI technologies function as decision-support instruments rather than independent agents. This hybrid model posits that the most efficacious cybersecurity measures integrate automation with human discernment [8].

This study highlighted the role of explainability in cultivating trust. Participants expressed reluctance to depend on AI recommendations unless they understood the decision-making process. The problem has been extensively examined in *artificial intelligence* circles, especially with black-box models and the constraints of deep learning systems [6]. Within the context of SOCs, where fast and defensible decisions are crucial, the capacity to scrutinise AI reasoning exceeds mere technicalities, emerging as a practical and ethical obligation.

A notable finding relates to the evolution of analyst roles. Although some respondents voiced concern about automation possibly resulting in redundancy, the majority characterised it as a transition in responsibilities rather than a loss of employment. Analysts dedicate less time to low-value, repetitive tasks and expanding their focus on investigation, contextual analysis, and incident coordination. This is consistent with studies in cybersecurity and related areas, suggesting that AI is more likely to enhance rather than replace human knowledge [4, 8].

The lack of formal AI training in many organisations constitutes an important cause for concern. According to Darktrace (2022) [9] and supported by study participants, the absence of established training hinders adoption and heightens the potential of misuse or excessive dependence on AI systems. Organisations must engage in training programs that combine both technical operations and the development of critical thinking around AI limitations and ethical use. This mirrors broader industry conversations around the increasing necessity for AI literacy among cybersecurity experts [7].

Ethical considerations continue to be a significant problem. Participants expressed concerns regarding data privacy, the equity of AI-based flagging systems, and the absence of explicit accountability in instances of failure. These apprehensions align with ethical critiques in the literature, asserting that inadequately controlled AI systems can replicate or exacerbate existing prejudices [13]. In high-stakes contexts like SOCs, where AI outputs influence real-time choices, the consequences of inaccuracy are significantly elevated. Ethical governance systems must consequently adapt with AI implementation to guarantee appropriate utilisation.

These findings also indicate a wider trend within the IT sector, where AI integration is instigating changes in professional roles, regulatory policies, and organisational culture. In sectors such as healthcare and finance, which are marked by significant risk and regulatory scrutiny, AI has shown the capacity to assist experts without replacing them, conditional upon the establishment of suitable governance and oversight frameworks [12]. The cybersecurity sector seems to be pursuing a comparable path.

This study supports the argument that AI may serve as a strong facilitator in cybersecurity, especially within SOC environments. Despite this, its success hinges not only on the ability of technology but also on the manner of its implementation, governance, and understanding by the human professionals who depend on it. Subsequent research ought to investigate these dynamics more thoroughly, including employing longitudinal studies or ethnographic methodologies to better understand the impact of AI on daily operations over time.

A notable aspect emerging from this study is the need to distinguish between technical trust and operational trust. Although an AI tool may exhibit technical precision in identifying abnormalities, confidence is not completely established unless operators comprehend and endorse its judgements within the operational framework. This notion, occasionally termed 'situated trust', has been examined in associated fields such as autonomous vehicles and healthcare AI, and is now gaining prominence in cybersecurity [12]. Professionals in SOCs need not only operational systems but also systems whose outputs they can substantiate during internal evaluations or post-incident assessments.

Furthermore, the analysis indicates that the effective incorporation of AI into SOCs depends not only on tools but also necessitates cultural and structural modifications. Participants from businesses with integrated security and development teams shown a higher likelihood of expressing satisfaction with their AI solutions. These settings facilitate the interdisciplinary collaboration essential for refining and evolving AI outputs based on analyst comments [7]. In contrast, isolated environments exhibited less trust and less successful execution.

The influence of leadership in moulding perceptions of AI emerged as another significant insight. Organisations in which top leaders actively advocated for AI adoption, allocated resources for training, and positioned automation as an auxiliary tool rather than a substitute, saw diminished apprehensions around job displacement. This corresponds with change management research that underscores the importance of clear communication and psychological safety in the adoption of technological transition [14].

In addition to these organisational and ethical considerations, it is important to reflect on the technical limitations of applying machine learning in SOC environments. These systems can be resource-intensive to train, fine-tune, and integrate effectively, and they may remain susceptible to bias or data drift over time. Moreover, ML-based tools are vulnerable to adversarial attacks and model manipulation [22], requiring additional safeguards beyond traditional engineering controls. Such limitations highlight the need for ongoing monitoring and adaptation when deploying AI in operational settings.

These findings endorse a socio-technical framework for cybersecurity innovation. AI tools are inadequate for achieving transformational value without institutional commitment, ethical safeguards, and continuous education. Policymakers, technology providers, and security executives must collaborate to guarantee that AI implementations enhance performance while preserving professional integrity, user trust, and accountability.

6. Conclusion

The purpose of this study was to investigate the integration of AI into SOC settings, with an emphasis on how practitioners view the technology's benefits, drawbacks, and effects on their jobs. Based on 58 cybersecurity experts' survey answers, it provides a practitioner-led viewpoint that enhances the body of research that has mostly concentrated on technological efficacy.

While recent discourse has centred heavily on generative AI models such as large language models, this study has deliberately focused on applied machine learning in SOC environments. These systems differ fundamentally from generative AI in both purpose and limitation, and separating the two is important for a balanced understanding [23, 24].

The results show that there is broad interest in using AI-driven automation to solve important operational issues including information overload, ineffective triage, and alert fatigue. Many businesses are already seeing real benefits from AI, especially in the areas of incident prioritisation, anomaly detection, and false positive filtering. Rather than eliminating jobs, participants uniformly defined AI as a productivity boost. This supports growing opinions that AI may best complement human decision-making in cybersecurity by enhancing it rather than taking the place of analysts [8].

However, the report also identifies significant obstacles to successful integration. The three most important ones are explainability, skills preparedness, and trust. In particular, when the reasoning behind judgements is unclear, many respondents expressed low or conditional trust in AI outputs. This is in line with more general worries expressed in the literature about the ethical ramifications of opaque decision-making systems and the "black box" aspect of some AI models [6, 13].

Additionally, the results indicate a substantial training gap. More than half of the participants had not received any formal training on how to utilise or interpret AI systems, despite the growing reliance on AI. There is a real risk associated with this mismatch between the application of technology and human preparation. Businesses that ignore this gap risk undermining the return on their AI investments and losing out on an opportunity to strengthen the autonomy of their security teams. Another important realisation has to do with ethical accountability and governance. As AI systems become more extensively used, organisations need to think about not only what these tools can do, but also how they can do it—and how they fit in with institutional values, privacy laws, and fairness standards. According to Razavi et al. (2023) [7], problems including prejudice, false positives, and data misuse highlight the necessity of strong monitoring procedures, ongoing validation, and inclusive policy creation.

At the same time, it is important to acknowledge technical limitations alongside practitioner perceptions. Machine learning models require significant resources to train and integrate, may be vulnerable to adversarial manipulation, and are prone to bias or data drift over time [22]. These issues underline the need for continuous monitoring, model validation, and safeguards to ensure resilience in operational settings.

By presenting the lived reality of AI deployment in SOCs and offering empirical evidence of both enthusiasm and caution, this study adds to the scholarly and professional conversation. It supports the idea that effective AI integration is as much a technical as it is a human and organisational challenge. Organisations need to invest in ethical governance, change management techniques, and staff development in addition to implementing cutting-edge solutions.

Of course, there are restrictions. Obtaining the sample through professional networking sites may have resulted in an over-representation of professionals who are more online. Additionally, the results were recorded at a single moment in time and are restricted to self-reported impressions. However, it provides a useful starting point for future research, particularly in examining how perceptions change as AI technologies become more integrated into security systems.

This study concludes that AI is simplifying workflows, moving analyst roles towards higher-value tasks, and changing the SOC environment in quantifiable ways. However, how it is applied, regulated, and comprehended has a big impact on its efficacy and morality. Although AI will not take the role of cybersecurity experts, those who can work with AI, comprehend its limitations, and challenge its results will be the most successful in the field going forward.

Lastly, this work suggests more cross-disciplinary studies on the governance mechanisms of AI in SOCs. Working together, cybersecurity, law, ethics, and policy may create well-balanced frameworks that protect public trust and professional integrity while fostering innovation. Security executives will face the challenge of staying both technically grounded and ethically flexible as AI develops. While future debates are likely to focus increasingly on generative AI [24], this paper contributes by clarifying the distinct role of machine learning in SOC operations and the socio-technical considerations that accompany its adoption.

Declaration on Generative AI

The author(s) have not employed any generative AI tools.

References

[1] Ahmad, A., Maynard, S. B., & Park, S. (2021). Information security strategies: Towards an organisational multi-strategy perspective. Journal of Strategic Information Systems, 30(1), 101658. https://doi.org/10.1016/j.jsis.2021.101658

[2] Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials, 18*(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

[3] Jalali, M. S., & Kaiser, J. P. (2018). Cybersecurity in hospitals: A systematic, organisational perspective. *Journal of Medical Internet Research, 20*(5), e10059. https://doi.org/10.2196/10059

[4] Swimlane. (2024). Will AI take over cybersecurity jobs? https://swimlane.com/blog/will-ai-take-over-cybersecurity-jobs/

[5] Xin, Y., Kong, L., Liu, Z., Chen, Y., Li, Y., Zhu, H., … & Wang, C. (2018). Machine learning and deep learning methods for cybersecurity. *IEEE Access, 6*, 35365–35381. https://doi.org/10.1109/ACCESS.2018.2836950

[6] Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine, 40*(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850

[7] Razavi, A., Caines, A., & Patel, T. (2023). Human factors in AI-enabled cybersecurity. *ACM Transactions on Cybersecurity, 1*(2), 45–67. https://doi.org/10.1145/3579286

[8] Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science, 361*(6404), 751–752. https://doi.org/10.1126/science.aat5991

[9] Darktrace. (2022). AI and the future of SOCs: Survey results. https://www.darktrace.com/

[10] Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In *2010 IEEE Symposium on Security and Privacy* (pp. 305–316). IEEE. https://doi.org/10.1109/SP.2010.25

[11] Palo Alto Networks. (2023). The role of AI in security automation. https://www.paloaltonetworks.com/

[12] Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., … & Amodei, D. (2018).

The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. *arXiv preprint* arXiv:1802.07228. https://arxiv.org/abs/1802.07228

[13] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21. https://doi.org/10.1177/2053951716679679

[14] Wright, D., & Kreissl, R. (Eds.). (2014). *Surveillance in Europe.* Routledge.

[15] Calders, T., & Žliobaitė, I. (2013). Why unbiased computational processes can lead to discriminative decision procedures. In C. Custers, T. Calders, B. Schermer, & T. Zarsky (Eds.), *Discrimination and privacy in the information society* (pp. 43–57). Springer. https://doi.org/10.1007/978-3-642-30487-3_3

[16] IBM Security. (2023). Transforming SOCs with AI-powered threat detection. https://www.ibm.com/security

[17] European Commission. (2021). Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=CELEX%3A52021PC0206

[18] NIST. (2023). AI Risk Management Framework (AI RMF 1.0). National Institute of Standards and Technology. https://www.nist.gov/itl/ai-risk-management-framework

[19] Tjoa, S., & Tjoa, A. M. (2016). The role of ICT in the resilience of critical infrastructures: Challenges and research directions. In T. Comes, F. Fiedrich, S. Fortier, J. Geldermann, & T. Müller (Eds.), *ISCRAM 2016 Conference Proceedings – 13th International Conference on Information Systems for Crisis Response and Management.*

[20] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

[21] Achuthan, K., et al. (2024). Advancing cybersecurity and privacy with artificial intelligence: A comprehensive review. *Frontiers in Artificial Intelligence, 7*, 11656524. https://doi.org/10.3389/frai.2024.11656524

[22] Musser, G., et al. (2023). Adversarial machine learning and cybersecurity: Risks, challenges, and legal implications. *arXiv preprint* arXiv:2305.14553. https://arxiv.org/abs/2305.14553

[23] Nott, C. (2025). Organizational adaptation to generative AI in cybersecurity: A systematic review. *arXiv preprint* arXiv:2506.12060. https://arxiv.org/abs/2506.12060

[24] Ferrag, M. A., et al. (2024). Generative AI in cybersecurity: A comprehensive review of LLM applications and vulnerabilities. *arXiv preprint* arXiv:2405.12750. https://arxiv.org/abs/2405.12750