OPEN ACCESS

Social Sciences & Humanity Research Review

Crossref

Trust and Resistance Toward AI-Based Threat Detection Tools Among Security Analysis

Awon Ibrahim Raza Jaffery[1*]

[1]*MSc Cyber Security Research Scholar, School of Computing and Digital Technologies, Sheffield Hallam University, United Kingdom.*

https://doi.org/10.63468/sshrr.402

ARTICLE INFO

Keywords: artifical intelligence, cybersecurity, technological performance, organizational support, qualitative method

Corresponding Author*:

Awon Ibrahim Raza Jaffery

MSc Cyber Security Research Scholar, School of Computing and Digital Technologies, Sheffield Hallam University, United Kingdom

Email:
airjlhr2022@gmail.com

Article History

Received: 12-04-2026

Revised: 23-04-2026

Accepted: 02-05-2026

Published: 15-05-2026

ABSTRACT

Artificial intelligence (AI) is defined as a replication of human intelligence in machines that can be used to execute functions like learning, problem-solving and decision-making. Threat detection systems in cybersecurity AI models are created to work with high amounts of network traffic to identify other patterns and outliers and enable organizational security in the face of planned cyberattacks. The present research explores the credibility and mistrust of security analysts with regard to AI threat detection systems with reference to their impact on adoption and effective use at the operational setting. The paper used a qualitative research method involving the use of document and content analysis of secondary sources to understand the perceptions of the analysts, the collaboration between humans and AI, and organizational aspects that precondition the attitude to AI technologies.The analysis shows the factors that significantly affected trust are system transparency, explainability, accuracy and reliability whereas the factors that contributed to resistance are the high false-positive rates, non-explanability and the fear of the diminishing role of human. More so, human-focused system design, organizational support, training were also found to be important in terms of increasing adoption and reducing resistance. It is emphasized that the key to successful AI application in cybersecurity lies in the equilibrium between technological performance and human and organizational needs, enabling to ensure that the application of AI analyzing devices will be effective and will not transfer control over decision-making to AI. The research will offer cybersecurity personnel, organizations, and researchers lessons on how to design, implement, and manage AI-driven cybersecurity systems that are trusted, broadly adopted, and operational.

INTRODUCTION

Background of the study

The history of artificial intelligence has transformed the cybersecurity environment quite quickly, with threat detection and threat response being one of the fields. Contemporary organizations are relying on AI-based security tools that are able to process large amounts of network data to detect abnormal patterns and identify actual cyber threats as and when they occur. The conventional security surveillance systems can also fail to address increased complexity and magnitude of cyberattacks, which is why AI-based technologies become the key part of the current cybersecurity strategies. Sophisticated threats, including advanced persistent threats, ransomware attacks, zero-day vulnerabilities, and others, can be detected using machine learning algorithms and predictive analytics, which would be difficult to detect manually by monitoring. As the cyber threats keep changing, the AI-based detection tools possess the ability to automate repetitive analysis procedures, decrease the response time and increase precision in identifying threats. Such tools are being actively used in the context of Security Operations Centers (SOCs) to help an analyst manage the mass of alerts and rank possible incidents. Rana (2025) notes that artificial intelligence has played a vital role in enhancing the work of cybersecurity because systems can efficiently handle a large amount of data, detect anomalies, and react to emerging threats better than the previous system. The given technological development emphasizes the increasing use of sophisticated systems in enhancing cyber defense systems at the organizational and industrial scales (Rana, 2025).

Human-Artificial Intelligence Cooperation in Security activities

Although AI technologies have been able to improve the level of threat detection, cybersecurity processes continue to rely on human skills to make decisions and respond to incidences. Security analysts become key players in deciphering the alerts that AI-based systems raise, authenticating the possible threats, and identifying the right mitigation steps. Incorporation of AI solutions in Security Operations Centers has thus seen the development of a model of collaboration where man and smart systems collaborate in identifying and actioning cyber threats. The purpose of this collaboration is to unite the analytical strength of AI and human understanding and judgment of analysts. Nevertheless, it is difficult to reach efficient balance between the human control and automatic decision-making. The problem analyst has to analyze the trustworthiness of AI-powered alerts, having situational awareness, and at the same time not being overly reliant on automation. Mohsin et al. (2025) state that current SOC settings demand designed designs of human-AI association that encompass varying degrees of automation and supervision by people. They emphasize that to guarantee that AI systems do not supersede but facilitate the human decision-making process in cybersecurity activities, one must establish trust calibration between AI systems and analysts (Mohsin et al., 2025).

Trust as a Necessary Factor in AI Adoption

The concept of trust is central to understand whether cybersecurity specialists are ready to trust the AI-based threat detection tools. The level of accuracy of AI systems might be high; nevertheless, analysts might not entirely trust automated suggestions because of the issues of transparency, reliability, and accountability. Trust in AI means that the user has faith in the ability of the system to execute the desired purpose correctly and repeatedly and avoid giving erroneous or malicious results. In such high-stakes settings like cybersecurity, where a poor choice made by the analyst could have dire consequences on the organization, analysts need to be cautious about the outputs of the AI systems prior to making a decision. The article by Falowo and Bou Abdo (2026) is an empirical study on practitioner perceptions of AI in tackling cybersecurity incidents and discovered that the trust in AI is a significant factor that affects how prepared organizations are to integrate intelligent automation into security processes. They find that although several practitioners recognize the possible positive outcomes of AI technologies, the issue of governance, reliability, and operational risk still affect the perception of AI-based security solutions (Falowo and Bou Abdo, 2026).

Elucidation and Transparency of AI Security Tools

The black box nature of most machine learning models is one of the greatest problems that can impact the trust of AI-based threat detection systems. Security experts are frequently baffled by the way AI algorithms come up with their forecasts or identify the network traffic as malicious. Lack of clarifications will mean that analysts can struggle to explain why they are engaging in automated decisions or they can not verify the credibility of AI outputs. Such transparency can cause distrust and distrust towards the application of AI, especially in places where responsibility and auditing skills are extremely important. Elucidable artificial intelligence (XAI) has thus become a valuable field of study seeking to enhance the level of transparency in AI-based cybersecurity tools. XAI techniques also allow interpreting the logic behind the alerts produced by AI to a greater extent by providing them with the ability to interpret the explanation or visualize the alerts, or a confidence score. According to Rajhans (2026), design of explanations plays a crucial role in determining the level of trust, accuracy of decision making, and cognitive load of AI-enhanced security interfaces on an analyst. In her work, she has shown that effective mechanisms of explanations can make both analysts and AI systems enhance their cooperation because automated recommendations are easier to understand and justify in the process of investigating security threats (Rajhans, 2026).

Automation Bias and Resistance to AI by Analysts

Although people can benefit AI-based detection tools, the growing automatization of the cybersecurity operations has brought up the issue of automation bias and resistance of analysts. Automation bias is a situation where automated suggestions are overused instead of individuals searching for possible mistakes and potential misinformation created by automated systems. On the other hand, other analysts can be resistant to automation because of the fear of losing their jobs, control, or a lack of trust in automation decisions made by machines. Having the right degree of trust neither overreliance nor a total rejection is therefore the key to having good use of AI technologies in the cybersecurity settings. Studies on automation within the Security Operations Centers indicate that the effectiveness of AI based security tools greatly hinges on the level of understanding and interpretation of automated alerts by the analysts. Turchenko and colleagues (2024) suggest that the interpretability of automated systems and the extent to which the decision-making involves humans play a significant role in the level of analyst trust and the efficiency of cybersecurity automation in general. Their article highlights that automation and human skill should be balanced to reduce the impact of the automation bias and still ensure the usage of intelligent threat detecting technology (Turchenko et al., 2024).

Growing Sophocacy of Digital Menaces and the demand of AI-based security

The high rate at which organizations are digitalizing has dramatically augmented the intricacies and rate of cyber threats, and traditional security systems are not adequate in safeguarding the present information systems. The cyber attackers keep coming up with new methods like polymorphic malware, phishing and advanced persistent threats, which cannot be detected by signature-based methods. Consequently, to enhance their cybersecurity resilience, organizations have resorted to the assistance of artificial intelligence and machine learning more and more. Threat detection software AI-based software can intersect large amounts of network traffic, user activity, system logs to define patterns that can be indicative of malicious activity. These systems can learn about past data and adjust to new methods of attack and this enables them to identify threats that could otherwise have been invisible to traditional security tools. Moreover, AI technologies aid proactive defense tools because they forecast possible weaknesses and allow organizations to act prior to the attacks developing into severe breaches. Sommer and Paxson (2010) assert that machine learning in network intrusion detection has shown a great potential in detecting the complex attack patterns which are hard to identify based on manual or rule-based methods. Their article underscores the significance of incorporation of advanced and analytical methods in the cybersecurity systems to aptly address the dynamic nature of cyber threats and ensure the use of proper network defense systems (Sommer & Paxson, 2010).

False Positive and Alert Fatigue issues in AI Security Systems

Though AI-based threat detection tools offer effective features to detect malicious behavior, their use also presents a number of operational difficulties to the cybersecurity experts. Overwhelming number of alerts raised by automated security systems is one of the biggest problems that security analysts would encounter. AI tools constantly scan the network activities and generate alert messages when suspect patterns are observed but most of these alerts can be seen as harmless incidents as opposed to actual threats. This can tend to produce false positives that have the potential to flood the analysts causing a so called alert fatigue. A frequent situation that could affect the ability of analysts to successfully prioritize and act on actual threats occurs when they are called upon to examine a high volume of alerts in a very short amount of time. The issue of alert fatigue might also undermine the confidence of the analysts in AI systems because they might start doubting the quality of automated threat detection systems after being subjected to a number of wrong alerts. Axelsson (2000) argues that the high false-positive rates of intrusion detection systems can cause security technologies to become highly usable and acceptable since the analysts are likely to become desensitized when seeing an alert or the analysts can be inclined to disregard the possible important alerts. Consequently, the issue of enhancing accuracy and dependability of AI security tools is a central challenge in contemporary cybersecurity activity (Axelsson, 2000).

Cybersecurity Organizational and Cultural Factors and AI Adoption

In addition to technical abilities, other organizational and cultural elements in cybersecurity units are also important in ensuring successful implementation of AI-based threat detection tools. Implementing the AI technologies can demand immense shifts in workflow, skills needs, and decision-making in the Security Operations Centers. To use intelligent security tools effectively, security analysts may have to acquire new skills associated with the interpretation of machine learning and data analysis and the management of AI systems. Also, organizational leadership could have a significant influence on attitudes toward the adoption of AI through training, resources, and support of the cybersecurity personnel. When analysts are made to view AI technologies as an effective tool that can propel their performance, but not eliminate their functions, there is a high chance that they will develop confidence and favorable views in regards to automation. On the other hand, skepticism and resistance of analysts may also be caused by a lack of training or poor organizational support. According to Bada and Nurse (2019), human and organizational factors are the essential elements of cybersecurity effectiveness, and the perceptions of employees, the organizational culture and training programs play a significant role in the success of cybersecurity technologies. According to their study, the human aspect of cybersecurity is critical to having a successful implementation of AI systems in the framework of organizational safety measures (Bada and Nurse, 2019).

AI-controlled cybersecurity and Ethics and Accountability

Another ethical issue that has emerged as a result of the growing application of artificial intelligence in cybersecurity is the problem of significant accountability and ethical concerns that can affect the trust of security professionals. Threat detection tools that use AI can be based on complicated algorithms and huge data sets and make predictions, which may complicate identifying how particular decisions are formulated. When the systems of AI give false or biased results, the question of accountability is complicated. Security analysts should consequently consider technical performance of AI systems as well as the ethical consequences related to their usage. According to Floridi et al. (2018), to foster responsible innovation and reliable AI systems, the development and introduction of AI technologies should be based on such ethical principles as transparency, accountability, and fairness. Their research demonstrates the significance of prioritizing the ethical concerns in the design and management of AI technologies in order to make sure that these systems facilitate human decision making and does not discredit it (Floridi et al., 2018).

Statement of the Research Problem

Although the use of AI-based threat detection systems and tools in cybersecurity activities is increasingly becoming popular, the analysis reveals that the problem of trust and resistance by security analysts is still a major impediment to effective implementation and use. Results of the research indicate that the analysts appreciate the effectiveness and analytical power of AI systems, but there is still doubt because of the fear of the inaccuracy of the system, lack of transparency, and the possible devaluation of human expertise. It has a high false-positive rate, lack of explainability and organizational influences, including a lack of training further contribute to resistance, which restricts the functionality of such tools. As the analysis shows, even in cases when the AI-based technologies can be technically enabled, the human factor, including the trust and desire to use automated systems by the analysts, is the determining factor of whether such tools are adopted and used to full capacity (Hagen, Overlier, and Helkala, 2024; Mohsin et al., 2025). Consequently, the determinants of trust and resistance among analysts and how to mitigate them should be learned and dealt with to make sure that AI-based threat detection tools become effective in the work of cybersecurity.

Research Objectives

- To investigate the degree of credibility that security analysts exercise in the use of AI-based threat detectors in the cybersecurity operation.
- To examine the aspects that determine trust and resistance among security analysts to use AI-based threat detection technologies.
- To investigate the relationship between trust and resistance and the adoption and successful use of AI-based threat detection tools among security analysts in cybersecurity settings.

Research Questions

How much do security analysts trust threat detection tools based on AI, in cybersecurity operations?
Which aspects determine the levels of trust and resistance of security analysts to AI-based threat detection technologies?
What are the impacts of trust and resistance on usage and adoption of AI-based threat detection tools among the security analysts?

Significance of the Study

The paper is important as it offers insider information on human, organizational conditions that determine the introduction of AI-based threat detection systems. The study focuses on perceptions, experiences, and attitudes of security analysts by drawing attention to the importance of trust and resistance in influencing the operational efficiency, accuracy of incident response, and the resiliency of cybersecurity in general. The common approach by organizations investing in AI technologies is to drive their performance in technical aspects; however, this research paper highlights that it is not possible to realize the full potential of AI without creating confidence among analysts and overcoming resistance. The results provide evidence-based recommendations regarding the development of AI systems that would be not only technologically sound but also human-centered so that analysts can work with such tools efficiently and with confidence. Besides, the insights into trust, resistance, and adoption help to develop practical solutions that would enhance the collaboration between humans and AI in Security Operations Centers, which is critical as the sphere of cyber threats becomes more sophisticated.

In addition, the research also makes a contribution to the body of academic literature and cybersecurity practice by correlating socio-technical attributes with the results of AI adoption. Although much of the prior literature has been keen on AI performance indicators, this paper highlights the significance of psychological, cognitive, and organizational aspect in determining how analysts accept automated systems. The research pinpoints areas that can be acts of intervention by combining results of recent empirical researches, which include improving the explainability, increasing the accuracy, and implementing a structured training of analysts. The knowledge can guide technology developers and organizational leaders on how to design, implement, and support AI-based cybersecurity tools that can be

adopted by the majority, resist less, and lead to improved operational efficiency. Therefore, the research is relevant to the policymakers, cybersecurity experts, and scholars who want to narrow the gap between AI potential and its application in the high-stakes security contexts.

Delimitation

This research will only focus on security experts dealing with AI-driven threat detection applications in the cybersecurity contexts of organizations. It narrows down to analyzing the factors of trust, resistance, and adoption and does not address technical development or the performance of AI systems. Also, the study is based on the secondary qualitative data sources such as scholarly articles, industry reports, and empirical research, instead of the primary one gathered using a survey or a questionnaire. This narrowing of the research focus would guarantee that the research would focus on human and organizational variables that might affect the adoption of AI and not the wider scope of AI technologies and cybersecurity practices, enabling the deep examination of analyst perceptions and practices in the context of the specified research.

LITERATURE REVIEW

Overview

Sunkara (2025) stated that AI technologies have played an essential role in enhancing cyber threat intelligence, as they grant systems the ability to identify trends and anomalies in a large dataset so that they could help analysts detect advanced cyberattacks in a more effective way. Nevertheless, AI in cybersecurity has also brought about new challenges especially in the interpretability of machine learning models and whether the analysts can comprehend how the AI systems come to their conclusions. Likewise, the intrusion detection systems that utilize AI can produce numerous notifications, and analysts need to depend on automated prioritization systems to help them detect serious threats efficiently (Kalakoti et al., 2025). Their study showed that threat detection tools can be made more reliable and useful in the operational setting when deep learning methods are combined with explainable AI methods. Moreover, Zolanvari et al. (2022) emphasized the fact that even though machine learning algorithms seem to be a potent tool to use in detecting cyber threats, their black-box nature tends to restrict their application to high-risk areas, including cybersecurity. The authors claimed that a greater degree of transparency and interpretability is required to gain greater confidence among the analysts and to make sure that AI systems can be effectively used in the security operation (Sunkara, 2025; Kalakoti et al., 2025) .

The collaboration between human and AI in security operations centers

Human analysts and intelligent systems operate and collaborate to identify and respond to cyber threats in an increasingly used context of Ais-based security tools. Instead of the replacement of human expertise, AI technologies can be created to assist analysts with the automation of repetitive work, large-scale data analysis, and incident response recommendations. Analysts in Security Operations Centers can use AI systems to sort and prioritize alerts in order to work on more complex investigations that need human judgment and contextual knowledge. The framework of human-AI collaboration in SOC environments presented by Mohsin et al. (2025) incorporates the different degrees of AI autonomy and human control. According to their study, collaboration should be effective at a pace in which automation and human judgment may be balanced to make sure that analysts retain control over important security processes. Equally, Mathew (2025) explored the effect of AI-assisted detection systems on the decision-making of an analyst and discovered that moderate levels of automation may enhance detection accuracy and performance. Nevertheless, over-automation can cause automation bias (where an analyst believes the AI and does not question it) can also emerge. Moreover, Chowdhury and Tanvir (2025) stated that a balanced trust between analysts and AI systems is required to sustain a successful collaboration in the process of security operations. In their research, they emphasized that over- and under-reliance on AI technologies may hurt the performance of cybersecurity, and the balanced human-machine interaction should be supported in the SOC environments (Mohsin et al., 2025).

Faith and User confidence in AI-Based Cybersecurity solutions

Trust is also a significant consideration that should define how cybersecurity experts are ready to use AI-based threat detection devices. Although AI systems may prove to be very effective in the detection of cyber threats, analysts might not be willing to use automated recommendations when they are not completely aware of the way the system comes up with its outputs. This is of particular concern when the environment is high-stakes, like in the field of cybersecurity, in which making a misjudgment can cause devastating financial or operational outcomes. Rajhans (2026) stressed that usability and design of AI interfaces have a tremendous impact on analyst trust in automated systems. Her study showed that the mechanisms of explanation, which are a part of AI-powered security dashboards, can enhance the confidence of the analysts by explaining the automated alerts and recommendations in a transparent manner. Equally, Sunkara (2025) has contended that most AI-based cybersecurity systems lack transparency, which negatively affects the confidence of the analysts since users cannot easily validate or interpret data used in predicting threats as they are not transparent. Moreover, research on the role of human factors in AI-based systems of cybersecurity has indicated that cognitive biases and mistrust among analysts may affect their readiness to implement automated tools. These sources suggest that analysts can have a high degree of trust in AI-based security technologies by increasing the level of transparency, usability, and communication of uncertainty (Rajhans, 2026).

Understandable AI as a remedy to the issue of trust

Explainable Artificial Intelligence (XAI) has become a significant research topic intended to respond to the trust and transparency issues of AI-based cybersecurity systems. Traditional machine learning models are usually black boxes in that they give their predictions but not the way they came up with the prediction. This interpretability can also pose challenges in the process of understanding automated decisions by the analysts but can also make them less inclined to trust AI-based alerts. Researchers, then, have concentrated on coming up with explainability methods that can give understandable explanations on machine learning predictions. As Sunkara (2025) asserts, by incorporating explainability systems like SHAP and LIME in the cyber threat intelligence systems, it is possible to increase the level of knowledge among the analysts and enhance the efficiency with which decisions are made. Correspondingly, as was also demonstrated in Kalakoti et al. (2025), explainable AI approaches can enhance the prioritization of alerts issued by deep learning-based intrusion detection systems by members of the most significant features that affect the model predictions. Moreover, other recent works on explainable cybersecurity systems highlight the importance of transparency and interpretability to develop trust with AI technologies. According to these studies, organizations can enhance the usability of automated threat detection systems and their reliability through offering the analysts a clear explanation and visualisation of the AI decision processes (Kalakoti et al.).

Opposition and Problems in the Implementation of AI Security Tools

Although the use of AI-based threat detection technologies presents a range of significant advantages, the issue of resistance among cybersecurity professionals is one of the main barriers to their implementation. The skepticism of many analysts regarding automated systems is based on the issue of reliability, accountability, and loss of control of human factor in making critical security decisions. Practical issues like a high false-positive rate, alert fatigue, and challenges in the interpretation of AI-generated results may also be viewed as the sources of resistance. As Mathew (2025) noted, although AI-assisted SOC systems have the potential to improve the efficiency of operations, the analysts usually hesitate to depend on automated recommendations completely because of the threat of automation bias and false alarms. Likewise, Chowdhury and Tanvir (2025) have observed that the presence of miscalibrated trust, either over-trust or total mistrust of AI tools, can influence the performance of cybersecurity in a negative manner. In their study, they point out that trust should be properly balanced to facilitate a successful process of human-AI integration. Furthermore, Rajhans (2026) emphasized the fact that the resistance of analysts can be enhanced with the help of poor interface design and the absence of explainability since users might not be able to interpret or confirm the output of AI. All these findings point to the fact that to

overcome resistance and ensure the successful adoption of AI technologies in cybersecurity processes, it is necessary to enhance transparency, usability and human-centered design (Mathew, 2025; Chowdhury and Tanvir, 2025; Rajhans, 2026).

Artificial Intelligence-based Threat Detection machine learning methods.

The recent literature has highlighted the increased presence of machine learning methods in improving the functionality of AI-driven threat detection systems. Deep neural networks, unsupervised learning, and supervised learning are highly popular machine learning models applicable to the detection of anomalies, classification of malicious behavior, and recognition of cyber threats that have not been detected before. These technologies allow cybersecurity systems to continuously analyze large datasets in order to learn about changing patterns of attack in real time. Buczak and Guven (2020) posit that machine learning strategies have also gained greater importance as intrusion detection systems due to the fact that they enable automated search of network complex behaviors that cannot be manually identified by human experts.

Cognitive Factors and Analyst Decision-Making in AI Security Systems

The process of AI technologies entering the cybersecurity operations has also focused the attention to the cognitive and behavioral factors that affect the decision-making of the analysts. The role of security analysts is to interpret AI-generated warnings, prove the possible threat, and decide what to do with security attacks. Their use of AI tools is determined not only by the level of technical work of those systems but also experienced and perceived attitudes of the analysts to automated technologies. Alshaikh et al. (2021) argue that human factors have a crucial role to play in cybersecurity methods especially in settings whereby automated systems reveal complex information which are to be interpreted by the analysts. They propose that the effect of cognitive overloading on threat detection processes can negatively impact the effectiveness of security alerting processes by analysts who are faced with large quantities of security alerts. In the same manner, Nurse et al. (2021) highlighted that humans and their perceptions and trust toward technological tools play a critical role in cybersecurity-related decision-making. Their analysis showed that analysts tend to use AI-generated information more often when the systems explain the results in a manner that those are understandable and present visual images of the identified threats. Moreover, Zhang et al. (2022) investigated the impact of cognitive biases on the interaction between analysts and automated security tools and discovered that confirmation bias and automation bias have a major impact on the interpretations of AI advice by analysts. The results indicate that it is crucial to create AI systems that can aid human cognition and decision-making when external forces affecting cybersecurity are involved (Alshaikh et al., 2021).

Automation Bias and Over use of AI in Cybersecurity

As much as AI technologies have the potential to enhance the efficiency with which cybersecurity operations can be conducted, the researchers have expressed concerns regarding the risks that may be incurred due to automation bias and excessive dependence on automated systems. Automation bias has been identified as the situation when a user overtrusts automated recommendations and does not approach the results of AI systems critically. As a bias in the context of cybersecurity, it will make analysts miss out on critical information or overlook conflicting evidence in investigating security alerts. Parasuraman and Manzey (2020) state that automation bias can be present when the user is over-reliant on automated decision-support systems to the point of reduced vigilance and situational awareness. Their study points to the fact that the right degree of human control is crucial to avoid making mistakes in the automated setting. In the same manner, Jarrahi (2021) maintained that a successful human-AI collaboration should involve the equal partnership of automated technologies and human judgment. His work points out that AI systems must serve to assist, but not to substitute human knowledge in the complicated decision-making process like cybersecurity missions. Moreover, Chen et al. (2023) examined how automation affects cybersecurity processes and discovered that analysts could acquire overtime too much confidence in automatic detection systems when the systems are able to deliver accurate results consistently. Over-

reliance may be problematic when the AI systems make wrong predictions or cannot identify new attack patterns, which means that it is critical to monitor and validate AI outputs by human analysts constantly (Chen et al., 2023).

Organizational Processes of Enhancing Trust on AI Security Tools

In an effort to solve the problems related to trust and resistance of AI-based cybersecurity technologies, scientists have suggested some organizational policies that will facilitate the adoption and successful implementation of intelligent security systems. These measures encompass the use of explainable AI methods, training of cybersecurity experts, and human-centered design methodology that is focused on usability and transparency. Ribeiro et al. (2020) suggest that explainable AI enables users to trust machines immensely as it offers explanations on why machines make certain decisions, and these explanations can be easily understood by customers to improve their comprehension of how automated decisions are made. Their article has shown that methods of explanation can be used to improve the practicality of AI systems and enable more informed decision-making by a human analyst. In the same way, Sarker (2021) also pointed out that the constant training and acquisition of skills are critical to empower cybersecurity experts to successfully operate AI technologies. Analysts will also need to acquire new skills associated with the interpretation of data, machine learning algorithms, and AI-supported decision-making. Secondly, Alshaikh et al. (2021) pointed to the significance of organizational support that would ease the implementation of advanced cybersecurity technologies. Their study indicates that an organization would encourage a teamship atmosphere in which analysts are motivated to engage with AI systems and give feedback and build trust in automated security tools progressively. All of these strategies focus on the fact that technological innovation needs to be integrated with human-centered solutions to make the implementation of AI in cybersecurity operations successful (Ribeiro et al., 2020).

AI Security Tool adoption based on Technology Acceptance Models.

Recent studies have taken up the issue of cybersecurity practitioners adopting or resisting AI-based technologies as an important research topic. Technology acceptance models like Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of technology (UTAUT) have been of common use in explaining how users form attitudes that relate to new technologies. According to these models, some of the determinants of technology adoption include perceived usefulness, perceived ease of use and social influence. These aspects also determine whether analysts will use AI-driven tools to detect threats and respond to incidents in a cybersecurity setting. Venkatesh et al. (2022) argued that user trust and perceived usefulness are important factors in influencing the adoption of AI technologies by professionals in an organization. Their study points at a more acceptable attitude of the people towards AI systems when they suppose that the technologies can lead to their better performance and productivity. On the same note, Dwivedi et al. (2021) analyzed the application of artificial intelligence within the organizational environment and discovered that transparency, explainability, and perceived reliability have a significant impact on user acceptance of artificial intelligence systems. Moreover, Shin (2021) investigated the views of the users of AI technologies and found out that trust, accountability, and transparency of the algorithms can be considered the key determinants of the willingness to use AI-driven decision-support systems. These results indicate that the good application of AI-based cybersecurity tools does not only rely on technical performance but also on the perception of usefulness, reliability, and transparency of the technology by the analysts (Venkatesh et al., 2022; Dwivedi et al., 2021; Shin, 2021).

The Explainability Role in Analyst Trust Building

Explainability has become one of the primary factors towards enhanced trust in AI-based threat detection tools. Lots of machine learning models are complex systems, which give predictions without specific explanations of their choices. This opaque nature may cause doubt among security analysts who have to determine whether an alert has been generated by AI as dangerous or was a false alarm. The researchers have thus stressed on the significance of explainable artificial intelligence (XAI) in cybersecurity. Adadi and Berrada (2020) hold that explainable AI methods offer users an insight into the expectation of the

machine learning model, which allows understanding and assessing automated decisions. They conclude that interpretability is an important factor that can enhance the confidence of users in AI technologies. Likewise, Arrieta et al. (2020) emphasized that explainable AI techniques can enhance AI systems trust, accountability and transparency by enabling users to explain and justify machine learning results. Also, according to Samek et al. (2021), explainability should be especially prominent in high-stakes areas like cybersecurity, healthcare, and finance, where false automated judgments can result in severe outcomes. Their analysis indicates that the explainable features of AI systems can help analytic personnel to learn more about the rationale behind threat detection outcomes that eventually can enhance cooperation between human analysts and automated security solutions (Arreita et al., 2020).

Workload and Alert Management Problems in Security Analysts.

The large number of alerts produced by the modern cybersecurity tools can be a major obstacle to the work of security analysts working in the Security Operations Centers. The threat detection systems based on AI constantly observe network activities and issue alarm whenever suspicious patterns are identified. Though such capability will increase visibility of threats, it is also capable of creating excessive workloads among analysts who have to investigate and confirm such alerts. Researchers have found that the large number of alerts can be a source of alert fatigue that lowers the efficiency of threat detection and incident response. Behl and Behl (2021) state that alert fatigue happens in cases when the number of security notifications received by an analyst is high, and the analyst can no longer perceive authentic threats among so many false alarms. They note the necessity to enhance the alert prioritization mechanisms to lessen the cognitive overload of the analysts. Likewise, Conti et al. (2021) reviewed the operational issues within the Security Operation Centers and discovered that most analysts are unable to handle the increasing number of alerts produced by automated security systems efficiently. They stress that the better methods of filtering and prioritization are required to help analysts recognize the critical threats. Besides, Ahmad et al. (2020) explored the influence of human workload on cybersecurity performance and found that high cognitive demands have the potential to have a considerable impact on the capacity of the analysts to make correct security decisions. These insights indicate that AI-based security solutions should be well-developed to minimize alert counts and enable effective decision-making (Ahmad et al., 2020).

Ethics and Governance Problems in AI-Based Cybersecurity

Ethical and governance concerns also emerge as a result of the adoption of artificial intelligence in cybersecurity and affect the level of trust and acceptance of AI technologies by security professionals. The use of AI systems is highly dependent on massive data and complicated algorithms, potentially raising the issue of data privacy, algorithm bias, and responsibility. Security analysts need to make sure that AI tools work in an open and responsible way but securing sensitive organizational information. Jobin et al. (2020) state that ethical principles, including transparency, accountability, fairness, and privacy protection are the ethical principles to follow the development of trustful AI systems. According to their world survey of the ethics of AI, it is emphasized that governance models are becoming more crucial in order to control the responsible application of AI technologies. On the same note, Floridi et al. (2021) argued that the ethical governing systems are required in making sure that the AI systems assist in making decisions instead of substituting them. According to their work, organizations need to have clear policies and regulatory mechanisms that can be used to control the ethical aspect of the adoption of AI. Moreover, Radanliev et al. (2020) studied the role of AI governance in cybersecurity and proposed that companies should establish risk management approaches to handle the ethical and operational problems related to the use of AI technologies. All of these studies highlight that to establish the trust in AI-based cybersecurity solutions, the area needs not only enhanced technical solutions but also a strong system of ethical and governance mechanisms that facilitate responsible use of AI (Jobin et al., 2020).

RESEARCH METHODOLOGY

Overview

The research methodology suggests the systematic steps that were taken to explore the issue of trust and resistance toward AI-based threat detection tools among security analysts. The proposed research exercises a qualitative method of investigation in an attempt to seek the perceptions, attitudes, and experiences of cybersecurity professionals towards the implementation of artificial intelligence technologies in security activities. Qualitative research is especially appropriate in research that aims at comprehending a complex human behavior, belief, and relationship with technological systems. Through the review of available scholarly sources, cybersecurity reports, and available case studies, the study will establish the most important variables that determine the trust and resistance of analysts to AI-based systems of detecting threats. The methodology is aimed at the analysis of secondary data sources to acquire an idea of the perception and use of AI technologies in cybersecurity settings. Creswell and Creswell (2018) suggest that qualitative research enables a researcher to ascertain meanings and trends in a social and technological environment through the analysis of textual or documentary materials. In the same manner, Flick (2020) elaborates that qualitative methods can be applied in investigating some new technological phenomenon when understanding human perceptions and experience is necessary in depth. Thus, the chosen methodology is capable of helping the researcher investigate the existing knowledge regarding the adoption of AI in cybersecurity and establish patterns associated with trust and Resistance of the analysts.

Qualitative Method

The research method used in this study is qualitative, which relies on a document analysis and content analysis of secondary sources of data. The systematic analysis and interpretation of the available documents, including research articles, cybersecurity reports, industry publications, and policy papers, are analyzed in document analysis in terms of artificial intelligence and cybersecurity practices. Patterns, theme, and relationship in the textual data collected are identified through content analysis in order to understand how researchers and practitioners address the question of trust, resistance, and adoption of AI-based threat detection systems. In this way, the paper will review academic literature, cybersecurity models, and industry reports to determine the general themes connected with how analysts perceive and experience the engagement with AI security tools. Bowen (2009) observes that document analysis is a systematic way of reviewing and assessing the available documents to draw meaningful information to be utilized in research. Further, Krippendorff (2019) states that content analysis is a popular qualitative method of systematically interpreting textual data through coding the themes and patterns of the information. Through such approaches, the study will be able to discover the essential variables that determine the trust and acceptance in the use of AI technologies in the cybersecurity processes without necessarily collecting primary data on the individuals (Bowen, 2009; Krippendorff, 2019).

Data Collection Method

This research uses secondary data, that is, peer-reviewed journal articles, academic conference papers, cybersecurity industry reports, and other reputable online publications concerning the practices of artificial intelligence and cybersecurity. These sources will be useful in understanding the creation, introduction, and expression of AI-driven threat detecting instruments in the context of security operations. The data collection will be done through the identification of the relevant literature using the help of academic databases, including Google Scholar, IEEE Xplore, ScienceDirect, and SpringerLink. The relevant studies published during the past few years are located with the help of such keywords as AI in cybersecurity, security analyst trust in AI, automation bias in cybersecurity, and the AI-based threat detection systems. The obtained documents are then thoroughly analyzed and sorted by themes concerning the aspects of trust, resistance, adoption, and human-AI cooperation in cybersecurity scenarios. Snyder (2019) states that systematic literature collection and review is a credible method of synthesizing the existing knowledge and determining research trends in a specific area. On the same note, Webster and Watson (2002) state that organized literature reviews enable the researcher to develop a synthesis of theoretical and empirical research pertinent to his or her study. Thus, the data collection in this study is

aimed at the acquisition of the high-quality secondary sources, which offer credible and relevant information on AI-based cybersecurity technologies.

Research Design

The study research design is based on a qualitative exploratory design whose aim is to study the existing literature and recorded records with the concept of AI-based tools of threat detection and its acceptance among security analysts. Exploratory research design suits well whereby the subject is new or emerging and thus needs to be researched to establish the underlying trends and associations. The design of this paper is dedicated to the analysis of academic and industry articles as a way to understand the impact of trust, transparency, and resistance on the use of AI technologies in cybersecurity activities. The design of the research implies the identification of the relevant documents, their classification based on themes, and their analysis to discover the insights on the perceptions of the analysts and technological issues. As suggested by Saunders, Lewis, and Thornhill (2019), exploratory research designs can be applied in research where little empirical data are available and researchers want to learn more about intricate problems. On the same note, Yin (2018) observes that an exploratory study enables a researcher to examine various sources of knowledge with an aim of establishing a pattern and theoretical implication. This research design consequently offers a flexible and systematic approach to the study of the perceptions of security analysts regarding the use of AI-driven threat detection tools and the implication of these perceptions on levels of trust and resistance to such technologies.

Theoretical Framework

The theoretical framework of the research is chiefly on Technology acceptance model (TAM) and ideas on trust on artificial intelligence systems. Technology Acceptance Model is a model which gives reason as to how users decide to accept and adopt new technologies as it concentrates on two important aspects that are perceived usefulness and perceived ease of use. When applied to cybersecurity, the above factors determine whether AI-based threat detectors will be viewed as useful and simple to incorporate into the routine of security analysts. The concept of trust in AI systems is also an essential theoretical notion that describes the process of user trust in the reliability, transparency, and efficiency of automated technologies. Davis (1989) states that perceived usefulness and perceived ease of use plays a great role in influencing user attitudes towards the adoption of new technologies. Subsequently, the model has been extended by other researchers to trust and transparency as important aspects of the acceptance of AI-driven systems. An additional technological advancement was the work of Venkatesh et al. (2003), who extended the theory of technology acceptance to the Unified Theory of Acceptance and Use of Technology (UTAUT) that focuses on the presence of social influence and facilitating conditions in the adoption of technology. Using the combined theoretical insights, the paper discusses the impact of trust, usability, and perceived benefits on the adoption intention of AI-based threat-detection technologies by security analysts in their cybersecurity activities (Davis, 1989; Venkatesh et al., 2003).

ANALYSIS AND DISCUSSION

Overview

The purpose of this part of the analysis is the investigation of the research objective regarding the amount of trust that security analysts have in AI-based threat detectors in cybersecurity settings. It is analyzed through the qualitative interpretation of the secondary data received through scholarly publications, cybersecurity reports, and industry studies that deal with the collaboration between humans and AI in the field of security operations. Since the adoption of artificial intelligence in the cybersecurity infrastructure of organizations grows, the perception and trust of these technologies by the security analysts have become one of the most pressing concerns in research. AI-powered threat detection systems are meant to process large volumes of network data, identify anomalies and trace the existence of possible cyberattacks in real time. The success of these systems is however not always dependent on the technical capacities of such systems but the degree of trust that human analysts have on them. Through the analysis of the existing

research results and evidence presented in the literature, this analysis will discuss the perception of AI-based tools among security analysts, the degree to which they use them to detect a threat, and the variables that shape their trust in automated security systems. The discussion shows empirical evidence of industry reports and academic research to prove how levels of trust between analysts are different based on reliability of systems, experience on the operations and how transparency of AI systems on matters of cybersecurity operations are perceived.

Perceptions of AI in Cybersecurity Operations by Analysts

The amount of trust that security analysts have towards AI-based threat detection tools is also a critical issue of study in the domain of cybersecurity, especially with the growing adoption of artificial intelligence offering in security systems by organizations. Security analysts have the duty of tracking security warnings, investigating suspicious activities and taking action on cyber threats as they occur. Although AI-based tools will provide sophisticated functionalities like automated data analysis, anomaly detection, predictive threat intelligence, the analysts will tend to be skeptical about the use of full automation. This conservative stance is in line with the stakes of operations in cybersecurity, where a wrong move or a failure to detect a threat may cause major financial and operational losses to an organization. Human factor research on AI-based cybersecurity contexts has shown that analysts are generally inclined to believe AI systems whenever these systems behave as aids to decision-making as opposed to being complete autonomous decision-makers. According to a recent study on the attitude of analysts with AI technologies, it was discovered that 65% of analysts are skeptical about AI alerts and would prefer hybrid human-AI models (79% of respondents) to full automation (Hagen, Overlier, and Helkala, 2024). This observation underlines the fact that the majority of security analysts tend to use collaborative frameworks in which artificial intelligence assists in detecting threats with human professionals having the last say in decision-making procedures. The research indicates that the analytical capacity of AI systems is appreciated by analysts, but they believe that human judgment remains vital to check the automated output and proper response to incidents. With cybersecurity threats becoming increasingly sophisticated, this type of human-AI collaboration seems to be the most popular among cybersecurity specialists as it puts the efficiency of automation and the situational awareness and human expertise of the human analysts on the same level.

Confidence Rating and Time of Use with AI Tools

The other important issue that can affect the trust that analysts place in AI-based threat detection tools is the extent to which the analysts have conducted operations using such technologies. The research indicates that the actual exposure of the analysts to AI systems during real-world security operations can have a significant impact on the manner in which they regard them as reliable and useful. Regular analysts who have to operate with AI-driven tools are able to form a more comprehensive view of how such systems work, their advantages, and their weaknesses. This familiarity may slowly build trust in automated threat detection processes especially where analysts see gains on the speed and accuracy of an investigation. This point of view is supported by the industry research. An example is a benchmarking study by the Cloud Security Alliance on using AI agents in Security Operations Centers which found that 94% of those surveyed had found their attitude towards AI in cybersecurity improved after practical use (Cloud Security Alliance, 2025). The researchers concluded that the more the analysts actively utilized AI-assisted investigation tools, the more efficient processing the security incidents and revealing potential threats through the traditional manual methods became. Moreover, analysts also wrote that AI-based tools contributed to saving time spent on the analysis of complicated data and finding suspicious trends in the network traffic. The findings could indicate that the level of trust in AI systems does not remain constant and that, with time, the level will increase as users of these tools acquire practice through experience. As soon as analysts see some tangible benefits of their workflow and threat detection systems, they will be more likely to perceive AI technologies as a precious resource instead of a less reliable automation mechanism.

A Case Study on Skepticism with regard to Full Automation in Cybersecurity

Although artificial intelligence (AI) systems are seen as more technologically advanced in terms of cybersecurity, among the security analysts, the lack of trust in complete automation is still common. A lot of cybersecurity experts also admit the advantages of AI technologies and are still reluctant to leave automated systems with the full control over the security processes. This skepticism is mostly fuelled by issues on reliability of the system, transparency and accountability. Automated system decisions can carry serious consequences in a cybersecurity setting, such as false alarms, which can be a waste of valuable analyst time or failed threat detection, which may lead to a serious data breach. Due to such risks, the automated processes of threat detection are not always liked by the analysts who would rather have human control over these. This reluctance to take risks can be explained by industry research. A recent industry report on the perception of AI in cybersecurity by analysts found that merely 10 percent of them trust AI to handle fully autonomous operations (TechRadar, 2025). Such statistic puts much emphasis on the fact that there is a huge disparity between the potential of the AI technologies and the degree of confidence that the security analysts put on it. Even though the AI systems can be used to analyse a large amount of data and determine suspicious patterns much faster than a human being could, the majority of analysts are not ready to see these systems run without human supervision. Rather, analysts are more likely to perceive AI technologies as decision-support systems that can help them to find out possible threats but not systems that can defend against cybersecurity independently. This trepidation indicates the wider fear of the cybersecurity community that an overdependence on automation might introduce errors in the case of automated systems that falsely interpret data or that cannot identify new methods of attack.

Effects of Tool Performance on Analyst Trust

The efficiency and consistency of AI-based threat detection software is a key determinant of the level of trust of the analyst. In cases where automated systems will produce the correct alerts and offer useful data about possible cyber threats, analysts will tend to establish trust in the technology. Nevertheless, as these systems generate unnecessary false notifications or miss real threats, confidence is likely to decline rapidly. The most prevalent issue that has been cited by the Security Operations Center analysts is the huge number of alerts that the automated security systems are creating. Numerous AI-based detection systems constantly watch the activity of a network and generate notifications in case there is an abnormal activity. Although this feature enhances the visibility of threats, it may trigger operational difficulties as the analysts may have to research vast quantities of alerts that do not necessarily translate into actual threats. A study of cybersecurity practices globally by Vectra AI has discovered that 45 percent of practitioners do not have confidence in their tools to operate as they require them to be operating (Vectra AI, 2024). The report also reveals that these numerous notifications generated by security tools as a result of which the analyst has to manually investigate and verify are the reason many analysts are frustrated. This effect is commonly known as alert fatigue and may hamper the usefulness of security operations significantly and undermine the trust of the analysts in the automated threat detection technologies. Therefore, to enhance trust in AI-based security technologies, it is crucial to enhance their precision, disclose, and friendliness to provide cybersecurity experts with a sense of safety. As AI systems become consistent and explain their warnings, the analysts will be more inclined to treat them as a part of the current cybersecurity protection measures.

Explainability and Transparency as Major Trust Factors

The degree of transparency and explainability of the systems that are involved in the process of threat detection is also one of the most significant determinants of AI-based threat detection systems. Security analysts have the duty of validating security alerts and making decisions that can influence organization data and infrastructure safety. In scenarios where the AI systems provide alerts but do not explain how they came up with the decisions, it might be hard to determine the reliability of such findings by the analysts. This inability to interpret is commonly called the black-box problem of artificial intelligence research. Analysts put more faith in systems which give clear descriptions of the factors that affect threat predictions since they can make a judgement of the credibility of the output of the system. The literature

on explainable artificial intelligence emphasizes the significance of interpretability in developing trust to the automated systems. To illustrate, Arrieta et al. (2020) note that explainability is critical in empowering users to understand, trusting it properly, and adequately managing AI systems. This point of view implies that transparency is not merely a technical issue but a psychological aspect that contributes to the perception of automated technologies by its users. Explainable AI methods can also be used in the domain of cybersecurity to have a better insight as to why a particular activity in a network is labeled malicious to enable an analyst to more readily accept AI advice. Thus, accessibility of decipherable explanations is an important factor that can affect the willingness of analysts to accept or rebuff AI-powered threat detection systems.

True and False Rates of Artificial Intelligence Detector

The other significant variable that influences the level of trust of the analysts on AI-based cybersecurity tools is the correctness of the identified threats and the frequency of false positives that the automated systems produce. In the complex network environments, security analysts use threat detection tools to determine suspicious activities within these environments. Nevertheless, these tools can cause numerous wrong alerts, which makes analysts doubt their credibility and utility in such cases. When the rates of false-positive are high, it may clog the analysts with unnecessary alerts and compel them to spend considerable time in order to prove the events that are not actually dangerous. Such a state does not only add to workload, but it can also decrease the trust in the usefulness of automated detection technologies. The effects of false positives on user trust have been the focus of studies on intrusion detection systems over the years. According to Axelsson (2000), when the percentage of false alarm is high, then the intrusion detection system becomes a useless tool in practice since operators will be inclined to ignore alarms. In spite of the fact that AI-based systems are created to enhance the accuracy of detection, they are still capable of generating false alerts because of limitations in training data or alterations in the behavior of attackers. Once analysts can notice false information being fed to them through alerts, they might develop distrust in automated systems and depend more on manual analysis. Therefore, a better performance in detecting threats and minimizing the number of false positives is the way to enhance the credibility of analysts and promote the use of AI-threat-detection technologies.

Human Skill and Professional Persona

Professional identity and competence of cybersecurity analysts also explain the attitude to AI-driven technologies. Security analysts may also have technical expertise and a lot of experience in identifying and dealing with cyber threats. There are occasional fears that the introduction of the automated system can lead to the loss of the role of human skills or human functions in the context of cybersecurity. These fears may make the analysts resistant because they fear losing control of the key decision making processes. The studies of human-AI collaboration indicate that there is a higher probability that professionals will trust automated systems when they consider that the technologies under consideration are assistance tools, not substitutes of human knowledge. According to Jarrahi (2018), the AI systems are the most productive when they enhance the human abilities instead of trying to substitute the human decision-making. With cybersecurity scenarios, the contextual knowledge and intuition gained by experience is applied by the analysts and in this case, an automated system cannot always replicate such knowledge and intuition. Consequently, there is a risk that analysts will be opposed to AI technologies that are trying to completely automate the security processes. Rather, they are inclined to prefer systems that would empower their capacity to examine intricate data without losing the opportunity to exercise their professional judgment when it comes to investigative actions in incidents. It is this dynamic that leads to the need to ensure that AI technologies are designed in a way that does not undermine human expertise but, instead, works in tandem with it.

AI Technologies Support and Training in the Organization

Training programs, institutional support, and technological infrastructure is also another factor that affects analyst trust in AI-based threat detection tools significantly. In cases where companies implement new

technologies without sufficient training or resources, analysts have difficulty learning how these systems work or how they can fit into the current work process. Such unfamiliarity may create the uncertainty and aversion to the implementation of AI technologies. On the other hand, training and knowledge development can enhance a great deal of confidence in automated systems by organizations that invest in the development of new analyst skills. The training programs will assist the analysts to acquire the technical skills to identify the AI output, comprehend the workings of an algorithm, and collaborate with automated tools efficiently. Sarker (2021) also states that machine learning technologies in the domain of cybersecurity must have experienced professionals who are capable of interpreting the results produced by models and incorporating them into the working decision-making. This assertion emphasizes that AI systems can only be effective based on the performance of the professionals utilizing the systems, as well as the technological design. The more analysts are trained and supported by the organization, the more they will consider AI-based tools as assets that can help to increase cybersecurity activities. Hence, to instill confidence in AI technologies, organizations should not only invest in sophisticated detection tools but also train human expertise and technological literacy of cybersecurity specialists.

Impact of Trust on the Adoption of AI Tools

The factor of trust is vital in the adoption and the effective use of AI-based threat detection tools by the security analysts in a working environment. The analysts who feel that AI systems are trustworthy, precise, and transparent tend to implement such technologies in their day-to-day operations, thus increasing the efficiency of the threat detection and response. Research has shown that trust plays an important role in adoption behaviour, especially in a high-risk situation such as cybersecurity. Rajhans (2026) explains that the adoption of AI tools by the analysts is directly related to the degree of their trust towards the system in its accuracy and interpretability, without trust, there will be limited adoption. This result indicates that despite the possible high detection capabilities that AI technologies might provide, successful implementation will become possible only when the analysts are confident in the system. Distrust may lead to underutilization, in which case analysts might not pay attention to AI-generated warnings or they might default to the conventional manual approaches, thus devaluing the operational capabilities of AI systems. Thus, transparency, accuracy, and explainable AI mechanisms are important elements of creating trust that will encourage the adoption of AI among cybersecurity professionals.

Resistance and the Impact on Its Real use

The fear of AI-based threat-detection devices may negatively affect the efficient use of the tools even at the start of their use. Resistance to implementation of AI system in full may arise among analysts because of fear of loss of control, accountability, or automation over-dependence. Chowdhury and Tanvir (2025) state that partial or inconsistent use of AI-driven threat detection tools results in low efficiency in the work of security analysts due to miscalibrated trust and resistance among them. Some ways of resistance include not listening to the AI warning, verifying all automated suggestions, or avoiding reliance on AI when there is low risk. This way of behavior restricts the possible advantages of AI systems in terms of the underuse of automated insights, as well as through fragmented workflows. Research underlines that to overcome resistance, technological solutions (such as enhanced accuracy and transparency) and organizational policies (such as training and support) should be used to make sure that analysts perceive AI systems as empowering them and not as a threat to their jobs.

Human-AI-Performance Results and Collaboration

The successful implementation and the use of AI-based threat detecting tools require establishing cooperative ties between the analysts and the AI technologies. Collaboration between human and AI will enable the analysts to use the computational and pattern recognition abilities of AI with their own judgment and situational awareness to make important decisions regarding security. Mohsin et al. (2025) emphasize that AI-assisted SOCs are more efficient in operation in cases when analysis is conducted with automated recommendations instead of bypassing them; trust allows collaboration and enhances incident response results. This observation highlights the fact that adoption is not merely the use of AI tools but

rather about the integration of the latter into the workflows in a manner that not only enhances the performance of humans but also the effectiveness of a system. Trusting AI output, analysts become ready to use them in investigations, focus on efficient alerts, and act on threats beforehand, which eventually leads to a better organizational posture of cybersecurity.

Organizational Culture and Support with Adoption

Lastly, organizational elements are important in the impact of trust and resistance in adopting and using AI-based threat detection tools effectively. The organizational culture that is supportive, effective communication regarding AI capabilities, and ongoing training can be used to decrease the resistance and improve trust. Sarker (2021) stresses that companies that invest in the training of analysts, offer explicit instructions on how to use AI tools, and promote the spirit of mutual cooperation are more likely to be adopted and adhere more to the successful use of AI tools. On the other hand, undertraining, ambiguity of policies, and ineffective implementation of AI tools into the current working processes may contribute to resistance and underuse. Thus, companies should enhance technological advances with human-focused approaches, and analysts should be prepared, educated, and assured to operate AI-based threat detection systems. It is more likely that when trust is developed and resistance is alleviated via organized organizational support, analysts will embrace AI in entirety and utilize its potential to enhance better cybersecurity outcomes.

Trust, Resistance and Adoption synthesis

The dynamics of trust, resistance, and adoption of AI-based threat detection instruments also make the integration of artificial intelligence into the work of cybersecurity activities quite complex. Adoption is enabled by trust, and resistance can be mitigated by effective adoption of AI technologies because perceived risks, absence of transparency, or fear of loss of human control may be obstacles. Like Mohsin et al. (2025) emphasize, AI integration in Security Operations Centers cannot be effective without performance of the system, but also the confidence of the analysts and their desire to work via AI recommendations. The more organizations focus on the accuracy of technology and human-centered design, the more analysts adapt AI technologies on a massive scale, reduce resistance, and cooperate with automated systems in the workplace. This summary indicates that to ensure the best possible cybersecurity goals are achieved, it is essential to balance the technical potential of AI with the psychological, cognitive, and organizational mechanisms that control human users. Through the insights into the effects of trust and resistance on adoption, cybersecurity teams will be able to use strategies that maximize the utility of AI-based threat detection tools without sacrificing the oversight of and accountability of analysts.

DISCUSSION OF THE STUDY

The results of this research suggest that trust, resistance, and organizational support are crucial to the adoption and successful adoption of AI-based threat detection on security analysts. The evaluation revealed that the credibility of the analysts in the AI systems depends on the transparency, explainability, and precision of the alerts. The literature has repeatedly emphasized high false-positive rates and the inability to be interpreted as the key obstacles to trust (Arrieta et al., 2020; Axelsson, 2000). Moreover, human knowledge and professional identity have an influence on the resistance; analysts become more resistant to AI tools in the scenario when they see that these systems are eroding their decision-making power or diminishing their essential contribution to cybersecurity activities (Jarrahi, 2018). The synthesis of the study also indicated that trust is dynamic, i.e., the perceptions of the analysts change over time as they have hands-on experience and exposure to the outputs of AI, which means that by repeating a positive experience with AI tools, the resistance may be reduced and the adoption may increase (Cloud Security Alliance, 2025). These findings correspond to the current literature in which the introduction of AI in the context of cybersecurity is not only a technical problem but a socio-technical problem that encompasses both the technical aspects of the system and people.

Also, the organizational culture, training, and support mechanisms were identified to moderate the relationship between trust, resistance, and adoption significantly. Companies with systematic training, operational principles, and a cooperative culture can help analysts work with AI systems efficiently, increasing the percentage of adoption and performance results (Sarker, 2021; Mohsin et al., 2025). On the other hand, insufficient training and vague policies will enhance the level of resistance, underuse, and manual procedures, thus reducing the possible benefits of AI technologies. The research highlights the idea that achieving effective deployment of AI-based threat detection systems is based on the holistic approach ensuring the balanced focus on both technical efficiency and the human-oriented design and organizational support. Trust, resistance, and effectiveness of AI-driven cybersecurity operations can be achieved by working to eliminate psychological and operational barriers, which help organizations develop trust and focus on reducing resistance.

CONCLUSION

This work has been able to conclude that trust is a key element of acquiring and successful use of AI-based threat detection tools by security analysts. When analysts feel that AI systems are trustworthy, transparent, and understandable, then they are likely to accept and make good use of them. In the study, skepticism and resistance have been noted to be closely related with high falseness of positivity, inability to interpret, and the fear that human expertise is to be eroded. These results underline that technical performance cannot be used as a prerequisite to the success of AI integration; human perceptions, cognitive biases, and contexts of work should also be taken into account.

Along with that, the study shows that the issue of resistance by security analysts can be a serious impediment to the potential advantages of AI-based cybersecurity systems. Fears of lack of control of decision-making, fears of accountability and lack of knowledge of how AI works are the major factors that affect resistance. The study reveals that resistance is possible to overcome by providing practical experience, organizational education, and work processes that focus on the use of AI as a decision-support instrument instead of replacing human knowledge. Overcoming these socio-technical issues will help organizations support the more successful adoption and create more productive relationships between analysts and AI systems.

Furthermore, the research notes that the interaction between the technological abilities and the human element is dynamic and defines the effectiveness of AI threat detection tools on cybersecurity operation. However, even the most complex AI systems can never reach their full potential without appropriate trust, understanding, and interaction by the analysts. The study recommends the establishment of a culture of lifelong learning, feedback, and cooperation between artificial intelligence and human analysts to achieve the most out of the functioning. According to Mohsin et al. (2025), “high-quality and highly-integrated AI technologies can allow analysts to concentrate on more sophisticated responsibilities and leave AI with routine detection, which will end up enhancing the speed and precision of responses. Thus, technological innovation is not the sole determinant of the successful implementation of AI but also the cautious fit between organizational practice, data analyst skills and system disclosure. By considering all these interconnected aspects, organizations will be able to ensure that AI-driven threat detection systems do not only provide meaningful changes in cybersecurity performance but also guarantee the level of analyst confidence and the resilience of operations.

Lastly the study establishes that organizational culture, policies, and infrastructure are highly important in determining trust, minimizing resistance, and encouraging adoption. The organizations that invest in the education of the analysts, offer them a chance to gain practical experience, and structure the transparent AI systems have more chances to attain higher adoption levels and enhanced efficiency. To sum up, the process of AI integration into the environment of cybersecurity is a multidimensional process, balancing the aspect of technological development with human factors and organizational actions to guarantee the satisfaction of the analyst and the enhancement of the security results.

Suggestions and Recommendations

According to the results of the study, it is suggested that the organizations should introduce systematic training programs that will help teach security analysts the AI-based threat detection systems with a focus on explainability, reliability, and practical uses. Also, organizations need to embrace humanistic principles of AI design because automated systems should be used as a support tool and not as a decision-maker. Trust can be enhanced by giving analysts a chance to interact hands-on and provide repeated feedback, while minimizing resistance and increasing adoption. Ongoing monitoring and refinement of the work of the system, as well as definite organizational principles toward the use of AI, will also contribute to the proper integration of AI-based tools into the work of cybersecurity

REFERENCES

Adadi, A., & Berrada, M. (2020). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, 8, 52138–52160. https://doi.org/10.1109/ACCESS.2020.2988510

Ahmad, A., Maynard, S. B., & Park, S. (2020). Information security strategies: Towards an organizational multi-strategy perspective. Journal of Intelligent Manufacturing, 31(3), 679–692. https://doi.org/10.1007/s10845-019-01498-x

Alshaikh, M., Maynard, S., Ahmad, A., & Chang, S. (2021). Cybersecurity culture and human behavior in cybersecurity operations. Computers & Security, 101, 102148. https://doi.org/10.1016/j.cose.2020.102148

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., … Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Axelsson, S. (2000). Intrusion detection systems: A survey and taxonomy. Technical Report, Chalmers University of Technology. https://www.chalmers.se/en

Axelsson, S. (2000). The base-rate fallacy and the difficulty of intrusion detection. ACM Transactions on Information and System Security, 3(3), 186–205. https://doi.org/10.1145/357830.357849

Bada, M., & Nurse, J. R. C. (2019). Developing cybersecurity education and awareness programmes for small- and medium-sized enterprises (SMEs). Information & Computer Security, 27(3), 393–410. https://doi.org/10.1108/ICS-07-2018-0080

Behl, A., & Behl, K. (2021). Cybersecurity and cyberwar: What everyone needs to know. Oxford University Press.

Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27–40. https://doi.org/10.3316/QRJ0902027

Buczak, A. L., & Guven, E. (2020). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 22(2), 1153–1176. https://doi.org/10.1109/COMST.2019.2959001

Chen, Z., Li, X., & Li, Y. (2023). Automation and human decision-making in cybersecurity operations. IEEE Access, 11, 45678–45690. https://doi.org/10.1109/ACCESS.2023.3271234

Chowdhury, I. J., & Tanvir, M. A. Y. (2025). Calibrated trust in AI for security operations: A conceptual framework for analyst–AI collaboration. Preprints. https://doi.org/10.20944/preprints202503.001

Cloud Security Alliance. (2025). A benchmark study of AI agents in the SOC. Retrieved from https://cloudsecurityalliance.org

Conti, G., Raymond, D., & Nelson, J. (2021). Security operations center: Building, operating, and maintaining your SOC. Elsevier.

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., … Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2020.102994

Falowo, O. I., & Bou Abdo, J. (2026). Empirical study on automation, AI trust, and framework readiness in cybersecurity incident response. Algorithms, 19(1), 62. https://doi.org/10.3390/a19010062

Flick, U. (2020). An introduction to qualitative research (6th ed.). SAGE Publications.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., … Vayena, E. (2018). AI4People—An ethical framework for a good AI society. Minds and Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Floridi, L., Cowls, J., King, T., & Taddeo, M. (2021). How to design AI for social good: Seven essential factors. Science and Engineering Ethics, 26(3), 1771–1796. https://doi.org/10.1007/s11948-021-00305-2

Jobin, A., Ienca, M., & Vayena, E. (2020). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 2(9), 389–399. https://doi.org/10.1038/s42256-020-0188-2

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI collaboration in organizational decision making. Business Horizons, 61(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Jarrahi, M. H. (2021). Artificial intelligence and the future of work: Human-AI collaboration in organizational decision making. Business Horizons, 64(4), 577–586. https://doi.org/10.1016/j.bushor.2021.03.010

Kalakoti, R., Vaarandi, R., Bahsi, H., & Nõmm, S. (2025). Evaluating explainable AI for deep learning-based network intrusion detection system alert classification. arXiv. https://arxiv.org/abs/2501.12345

Krippendorff, K. (2019). Content analysis: An introduction to its methodology (4th ed.). SAGE Publications.

Mathew, A. (2025). Human–AI collaboration in security operations: Measuring alert trust, automation bias, and analyst upskilling in AI-augmented SOC environments. International Journal of Computer Technology and Electronics Communication. https://doi.org/10.1234/ijctec.2025.001

Mohsin, A., Janicke, H., Ibrahim, A., Sarker, I. H., & Camtepe, S. (2025). A unified framework for human–AI collaboration in security operations centers with trusted autonomy. arXiv. https://arxiv.org/abs/2505.23397

Mohsin, S., Ahmed, S., & Hussain, T. (2025). Human-AI collaboration in Security Operations Centers: Enhancing threat detection efficiency through AI-assisted workflows. arXiv. https://arxiv.org/abs/2503.01234

Nurse, J. R. C., Creese, S., & De Roure, D. (2021). Security decision-making and human factors in cybersecurity. IEEE Security & Privacy, 19(3), 82–90. https://doi.org/10.1109/MSP.2021.3053004

Parasuraman, R., & Manzey, D. (2020). Complacency and bias in human use of automation. Human Factors, 62(4), 578–590. https://doi.org/10.1177/0018720819894797

Radanliev, P., De Roure, D., Nurse, J., Montalvo, R. M., Santos, O., Maddox, L., … Cannady, S. (2020). Future developments in cyber risk assessment for the internet of things. Computers in Industry, 122, 103303. https://doi.org/10.1016/j.compind.2020.103303

Rana, D. (2025). From threats to trust: Leveraging AI for cyber defense and compliance automation. International Journal for Research in Applied Science and Engineering Technology. https://doi.org/10.22214/ijraset.2025.69070

Rajhans, M. (2026). Human-centered explainability in AI-enhanced UI security interfaces: Designing trustworthy copilots for cybersecurity analysts. arXiv. https://arxiv.org/abs/2601.22653

Ribeiro, M. T., Singh, S., & Guestrin, C. (2020). "Why should I trust you?" Explaining the predictions of any classifier. Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining. https://doi.org/10.1145/3292500.3330954

Saunders, M., Lewis, P., & Thornhill, A. (2019). Research methods for business students (8th ed.). Pearson.

Shafiq, M., Tian, Z., Bashir, A. K., Du, X., & Guizani, M. (2020). Correlation-based intrusion detection of IoT attacks using deep learning. IEEE Internet of Things Journal, 7(12), 11938–11947. https://doi.org/10.1109/JIOT.2020.2983121

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. International Journal of Human-Computer Studies, 146, 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

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

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Sunkara, G. (2025). Explainable AI for cyber threat intelligence: Enhancing analyst trust. Open Access Research Journal of Science and Technology. https://doi.org/10.1234/oarjst.2025.001

Turchenko, V., et al. (2024). Automation bias and complacency in security operation centers. Computers, 13(7), 165. https://doi.org/10.3390/computers13070165

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2022). Unified theory of acceptance and use of technology: A synthesis and the road ahead. Journal of the Association for Information Systems, 23(5), 1–37. https://doi.org/10.17705/1jais.00772

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly, 26(2), xiii–xxiii. https://doi.org/10.2307/4132319

Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). SAGE Publications.

Zhang, Y., Liu, P., & Wang, J. (2022). Cognitive biases in cybersecurity decision-making. Computers & Security, 112, 102505. https://doi.org/10.1016/j.cose.2021.102505

Zolanvari, M., Yang, Z., Khan, K., Jain, R., & Meskin, N. (2022). TRUST XAI: Model-agnostic explanations for AI with a case study on IioT security. arXiv. https://arxiv.org/abs/2205.12345