We Really Need the Help of AI: A Case Study of AI Adoption in Cybersecurity

STEFANI SLAVOVA, TU Delft, Delft, Netherlands and Rapid Circle, Amsterdam, Netherlands
YURY ZHAUNIAROVICH, TU Delft, Delft, Netherlands

We investigate the sociotechnical factors influencing the adoption of AI-based tools in cybersecurity operations within a large international financial organization, using a reflexive thematic analysis grounded in a Sociotechnical Systems (STS) framework. Our qualitative case study involved 15 interviews with security analysts, data scientists, and departmental leaders to explore end-user perspectives, organizational culture, and technical constraints shaping AI adoption. Drawing on established models, we analyze barriers such as mistrust in AI systems, ineffective feedback mechanisms, lack of domain knowledge, and job security concerns. The study reveals a disconnect between the availability of AI tools and their actual use, primarily driven by human-centric resistance and structural inefficiencies rather than technical limitations. These findings emphasize the importance of aligning AI development with analysts' workflows, increasing explainability, and making design processes more collaborative. We propose a targeted suite of interventions – including training, cross-functional mentorship, and enhanced feedback channels – to support the responsible and effective integration of AI. Our research contributes a theory-informed and empirically grounded understanding of AI adoption challenges in cybersecurity, with practical implications for organizations navigating the human-AI interface in corporate environments.

CCS Concepts: • **General and reference** → **Empirical studies**; • **Human-centered computing** → **Field studies**; **User studies**; • **Social and professional topics** → *Management of computing and information systems.*

Additional Key Words and Phrases: AI Adoption, Cybersecurity, Innovation Resistance

1 Introduction

The ongoing digital transformation – characterized by the integration of digital technologies into business processes – has fundamentally reshaped how organizations operate. While this evolution brings numerous advantages, it also introduces significant cybersecurity risks. Cyberattacks, often driven by motives such as financial gain (e.g., a recent Bybit hack resulted in $1.5 billion in losses [20]) or operational disruption (as seen in the attack on Change Healthcare, which left customers struggling for months to recover [44]), exploit weaknesses of computer systems and human beings and represent an escalating threat. As organizations adopt digital solutions, implementing robust cybersecurity measures becomes essential to protect against potential attacks [69, 89]. However, this task is complicated by the evolving nature and growing volume of cyber threats. Attackers continuously devise new methods to infiltrate networks, applications, and data, thus, undermining the confidentiality, integrity, and availability of information [17].

In this high-stakes environment, the integration of Artificial Intelligence (AI) offers a promising direction [10]. By providing advanced analytical capabilities, AI has the potential to transform cybersecurity by automating routine tasks, prioritizing alerts, and speeding up incident response [35, 85, 93]. In recent years, the potential of AI to enhance digital protection mechanisms has been increasingly recognized [22, 59, 68]. Researchers constantly suggest new approaches and tools [32, 56, 91]. On the commercial front, multiple vendors, e.g., Darktrace [29] or

Authors' Contact Information: Stefani Slavova, TU Delft, Delft, Netherlands and Rapid Circle, Amsterdam, Netherlands; e-mail: stefani.slavova@rapidcircle.com; Yury Zhauniarovich, TU Delft, Delft, Netherlands; e-mail: y.zhauniarovich@tudelft.nl.

ACM 1557-6051/2026/2-ART
https://doi.org/10.1145/3799707

Cisco [26], offer AI-empowered solutions that provide security detection and response services. Nonetheless, some organizations opt to develop custom AI-enabled security systems tailored to their specific needs [87, 88].

Despite significant advancements in AI and the availability of AI-powered cybersecurity solutions, *their adoption by cybersecurity teams remains notably slower* than the rate at which threat actors are exploiting them [77]. Researchers have identified several factors contributing to it, including the absence of compatibility and configuration issues with legacy systems [4, 85], the need for training and building resilience among security professionals [35], and the difficulties in ensuring transparency and accountability in AI-driven systems [93]. While these factors clarify why cybersecurity teams resist to deploy AI-enabled systems, they do not explain why end-users refuse to use *available* AI-enabled security solutions or adopt only a small portion of their functionality in their daily activities. This gap underscores the need for further research to understand the obstacles to AI adoption by end-users and how they can be addressed to enhance cybersecurity defenses.

In this study, we aim to help narrow this gap through a case study approach, focusing specifically on a large international organization[1]. Similar to the aforementioned study [77], our preliminary observation[2] also revealed a low adoption rate of AI solutions among end-users within its cybersecurity department, despite the availability of such tools and the organization's in-house capacity to develop and enhance them through its team of data scientists. *This indicates that the issue lies not in the availability of these tools but in the resistance of end-users to adopting or utilizing them.*

To discover what factors contribute to this and how they can be overcome, we ran 15 interviews with key stakeholders, including security analysts, data scientists, and leaders, ensuring a diverse range of perspectives, and analyzed them using a reflexive thematic analysis [27]. The main contributions of this study are the following:

- We provide an empirical evidence that AI tools, despite being readily available, remain underutilized in the cybersecurity department of a large organization. This finding underscores the importance of independently and regularly assessing the utilization of these tools.
- We conducted a case study involving interviews with 15 experts to understand why AI tools remain underused, identifying key barriers such as distrust in AI outputs, lack of model interpretability, concerns over job security, and limited feedback mechanisms.
- We perform a root cause analysis showing how misaligned technical development, the organizational culture and business processes, and analysts' psychological and experiential factors contribute to low this adoption.
- We propose concrete interventions to improve adoption, including simplified AI explanations tailored to analysts' needs, career mentorship to reduce uncertainty, and cross-functional collaboration and education to align AI development with operational realities.

2 Background

2.1 Innovation Adoption Theories

Over the past several decades, the adoption of innovation has been a heavily researched topic [78]. Several models have been developed to predict whether users will adopt an innovation, some of which are considered below.

Technology Acceptance Model (TAM). Proposed by Davis [31], TAM is based on the Theory of Reasoned Action (TRA) [36], a psychological theory that explains the relationship between attitudes, intentions, and behaviors. TAM specifically applies TRA to the context of technology use, identifying the factors that lead to a user's acceptance or rejection of a technology. The theory posits that the use of a technology is primarily

[1] Due to a confidentiality agreement with the collaborating organization, its name cannot be disclosed in this paper.

[2] As part of the exploratory phase of this research, we conducted a short informal poll among employees in the cybersecurity department. The poll asked whether they were using the available AI tools. While not a systematic survey, the responses suggested limited adoption, motivating the present qualitative study.

determined by two factors: *perceived usefulness* and *perceived ease of use*. These factors shape the user's attitude towards using a system, defined as the impact of an individual's positive or negative emotions on performing a specific behavior [92].

TAM plays a significant role in providing theoretical insights into the behaviors related to the use and acceptance of Information and Communication Technology (ICT) [24, 65]. Researchers have applied it to explore the adoption of various technologies, establishing it as a key theory in the field [6]. While being acknowledged as an effective, credible, and highly reliable model [57, 90], TAM has also been openly criticized for being too simple and leaving out important variables [14]. For instance, it does not sufficiently address *subjective norms* (the influence of peers or organizational culture on behavior), nor does it consider *facilitating conditions* such as training or resource availability, or *barriers* such as perceived risks and trust concerns [97]. These omissions have led to the development of several extensions of TAM that attempt to capture social and organizational factors influencing technology adoption.

Theory of Planned Behavior (TPB). TPB [2], developed by Ajzen [2] as an extension of his TRA [36] by incorporating the concept of perceived behavioral control, is a widely recognized psychological model designed to predict and explain human behavior within specific contexts. TPB considers three main factors that influence an individual's behavioral intention: *attitude toward the behavior* (positive or negative evaluation of performing the behavior), *subjective norms* (perceived social pressure to perform or not the behavior), and *perceived behavioral control* (individual's perception of their ability to perform the behavior taking into account both internal and external factors). In its turn, an individual's behavioral intention is a predictor of actual behavior, meaning that stronger intentions lead to a higher likelihood of performing the behavior [37]. However, intentions are not the sole determinant of behavior. TPB does not sufficiently consider the impact of environmental or situational factors that can shape behavior in practice. Examples include supportive organizational policies, the availability of resources like training and technical support, and time or workload pressures – all of which may strongly affect whether intentions are translated into action, particularly in workplace technology adoption.

Technology-Organization-Environment (TOE) Framework. The TOE framework, developed by Tornatzky and Fleischer [98], approaches technology adoption from an organizational standpoint, unlike other technology adoption models, such as TAM or TPB, which focus on the individual perspective. The TOE framework suggests that the decision to adopt a technological innovation and the implementation process are influenced by factors in three distinct contexts within an organization [15]: *technology* (internal and external technologies relevant to the organization, including existing technologies in use and those available in the market), *organization* (characteristics and resources of the organization such as connections among employees, communication practices within the firm, the size of the company, and the availability of surplus resources), and *environment* (setting in which the organization operates, including industry characteristics, market structure, the regulatory environment, and the presence of technology suppliers and competitors).

TOE, however, may oversimplify the complex processes involved in technology adoption by categorizing factors into just three contexts. It primarily focuses on organizational-level factors and does not explicitly account for individual-level variables such as employee attitudes or resistance to change, which may affect adoption decisions [11]. This framework also does not specify particular influencing factors for each context. Hence, the exact factors for a given research inquiry should be established by referencing prior research and theoretical insights, as various types of innovations are influenced by distinct factors affecting their adoption [15]. For example, in the case of AI in cybersecurity, adoption may depend on environmental factors such as regulatory requirements or organizational factors such as the availability of skilled personnel.

Unified Theory of Acceptance and Use of Technology (UTAUT). The Unified Theory of Acceptance and Use of Technology (UTAUT), developed by Venkatesh and colleagues [101], consolidates elements from eight models, including TRA, TAM, and TPB, that previously tried explaining technology acceptance and its use. It identifies four key factors influencing technology acceptance: *performance expectancy* (belief that the technology will improve performance), *effort expectancy* (ease of use), *social influence* (perception of social pressure to use the technology), and *facilitating conditions* (availability of resources and support). Additionally, the model considers four moderating variables, such as *age*, *gender*, *experience*, and *voluntariness of use*, that may change the effects of factors. UTAUT is widely used in technology adoption research for its comprehensive approach to understanding and predicting the acceptance and use of new technologies [101]. However, it has been also criticized for having too many variables that may complicate the prediction of intentions and behaviors [14]. Despite its breadth, UTUAT does not explicitly account for certain variables that are relevant in specific contexts, such as trust and perceived risk (especially salient in cybersecurity) or emotions (e.g., anxiety or enjoyment), and these omissions limit the model's applicability in contexts where such factors shape technology adoption.

Innovation Resistance Theory (IRT). The existing theories on innovation adoption primarily focus on the factors driving innovation adoption, with limited attention given to the barriers. This "pro-innovation bias" arises from the assumption that every innovation is inherently beneficial and will eventually be adopted. This makes it challenging for researchers to identify resistance to innovations and understand the factors behind it [58, 74]. One theory that studies this problem is IRT. Developed by Ram and Sheth [86], it focuses on understanding why individuals or organizations resist adopting innovations. It seeks to explain the barriers that hinder the acceptance and integration of new technologies, products, or practices, even when they offer potential benefits. IRT emphasizes that resistance is not merely opposition to change but can arise from rational concerns or psychological factors. Although several extensions have been proposed to enhance IRT [25, 75, 95], it still remains a primary theory in this field.

IRT identifies two main categories of barriers: *functional* (more practical and related to the perceived usefulness or feasibility of the innovation) and *psychological* (more subjective and linked to emotions, perceptions, or social influences). The former category includes *usage* (concerns about the complexity, difficulty, or compatibility of the innovation with existing systems, workflows or habits), *value* (concerns about whether the innovation offers enough relative advantage or value compared to existing solutions) and *risk* (concerns regarding potential side effects, e.g., financial, functional, or social consequences, of using the innovation) barriers. The latter category includes *tradition* (degree to which an innovation forces adopters to change established traditions or culture) and *image* (concerns about how the adoption might affect one's self-image or social identity, especially if the innovation is associated with negative stereotypes or perceptions). The study [86] also suggests strategies on how to cope with different types of barriers.

2.2 Related Work

The integration of AI into our digital world continues to deepen. While its adoption in cybersecurity initially lagged behind other industries due to the skepticism about its ability to meet the domain's high requirements, this process has accelerated significantly in recent years. Comprehensive overview of AI applications within the cybersecurity sector is provided in the works of Li [59], Chan et al. [22], and Mohamed [68]. However, understanding both the theoretical and practical implications related to this shift remains essential. In this section, we analyze the related literature to highlight existing findings and identify challenges of AI adoption in the cybersecurity domain. To synthesize these studies and clarify how our work differs from prior research, we summarize their methodologies, contributions, and limitations in Table 1.

Adoption of AI in Cybersecurity. Several studies have explored the benefits of AI adoption in cybersecurity. For instance, Sontan and Samuel [93] explores how AI transforms cybersecurity by enhancing threat detection, vulnerability analysis, and incident response. The study reveals AI's effectiveness in automating vulnerability scans, prioritizing threats, and speeding up incident responses, thus reducing human error and strengthening

Table 1. Comparison of Related Work and This Study

Study	Methodology	Contributions	Limitations / How This Study Differs
Sontan & Samuel [93]	Conceptual analysis	Highlights AI's role in threat detection, vulnerability analysis, and incident response.	Lacks empirical evidence; our study provides real-world case study insights.
Familoni [35]	Literature review	Stresses ethical AI use, human expertise, and professional training.	No empirical data; our study explores adoption barriers from end-users.
Gusman [41]	10 semi-structured interviews (US)	Identifies the importance of human decision-makers and training for AI integration.	Small sample size and focus on a specific region.
Al-Dosari et al. [4]	9 semi-structured interviews (Qatar banking)	Identifies AI benefits and challenges in banking cybersecurity.	Small sample size and focus on a specific region.
Gonaygunta [39]	Survey using UTAUT (banking sector)	Examines adoption factors such as performance expectancy and facilitating conditions.	Theoretical limitation; our study investigates adoption factors beyond UTAUT variables.
Radebe et al. [85]	11 semi-structured interviews (South Africa)	Reports benefits like automation, reduced false positives, and higher productivity.	Small sample size and focus on a specific region.
This study	Case study with 15 semi-structured interviews (international organization)	Examines AI adoption in cybersecurity using IRT framework; provides practical insights for overcoming barriers and fostering human-AI collaboration.	Extends prior literature by focusing on end-users, resistance and collaboration; covers a large international organization.

security. Nonetheless, it also points out significant challenges, such as ethical and privacy issues in automated decision-making, potential biases in AI models, and the difficulties in ensuring transparency and accountability in AI-driven systems.

Familoni [35] highlights AI's role in improving threat detection, authentication, and response. However, it is also admitted that AI introduces new risks through AI-driven attacks on machine learning algorithms. Familoni emphasizes the importance of human expertise in the ethical use of AI, advocating for robust training and skill development for cybersecurity professionals. The study also stresses the need for adherence to ethical standards, awareness of cognitive biases, user-focused design, and building resilience and adaptability within cybersecurity teams.

Gusman [41] conducted a qualitative study on the deployment of AI and Machine Learning (ML) in cybersecurity, focusing on their impact on intelligent decision-making. Through 10 semi-structured interviews with cybersecurity professionals in the United States, three key themes emerged: the ongoing importance of human decision-makers due to AI limitations, the increasing use of AI-driven decisions in cybersecurity, and the necessary learning curve and training for effective AI integration. Minor themes included a strong interest in learning about AI and ML systems and the significant adjustments needed for their adoption. Despite the rise of AI and

ML, IT professionals expect to remain the primary decision-makers due to current technological limitations. The study's limitations include its small sample size and potential biases inherent in qualitative research.

Al-Dosari et al. [4] conducted a qualitative study on AI applications and challenges in the cybersecurity of Qatar's banking sector. Through thematic analysis of interviews with nine experts, four main themes emerged: the importance of AI in enhancing cybersecurity, challenges in AI deployment, potential misuse of AI, and vulnerabilities in AI-based tools. The study found that AI is crucial for defending against web-based attacks and fraud but faces obstacles like inefficiencies in in-house development, compatibility issues with legacy systems, and regulatory compliance challenges. Additionally, AI poses risks through adversarial machine learning[3] and AI-powered malware, and inherent vulnerabilities such as data accumulation and privacy risks in chatbots. The study's limitations include its small sample size and potential qualitative research biases.

Gonaygunta [39] explored the factors influencing the adoption of ML algorithms for cyber threat detection in the banking industry using the Unified Theory of Acceptance and Use of Technology (UTAUT) model. The study examines how performance expectancy, effort expectancy, social influence, and facilitating conditions affect IT professionals' intentions to use ML in cybersecurity. Key findings show that performance expectancy and facilitating conditions positively influence the intention to adopt ML, while social influence has a negative impact, and effort expectancy has no significant effect. Despite ML's potential to improve cybersecurity through efficient threat detection and response, its adoption is hindered by challenges such as organizational readiness, system compatibility, and regulatory compliance. The study is limited by its focus on a single geographic region and reliance on self-reported data.

Radebe et al. [85] studied the perceptions of cybersecurity experts on AI-enabled tools in large South African enterprises, using the Expectation Confirmation Model (ECM) and semi-structured interviews with 11 professionals based in South Africa. The findings indicate that experts see substantial benefits in AI tools, such as automation, reduced human intervention, insightful reporting, fewer false positives, decreased risk, and enhanced productivity through quicker data gathering. However, challenges like data privacy concerns, alert fatigue, configuration issues, the risk of AI misuse by cybercriminals, marketing hype, and costs were also identified. Overall, the experts expressed high satisfaction with AI tools and a strong intention to continue using them, underlining AI's crucial role in improving cybersecurity in large organizations. The study's limitations include its small sample size and focus on a specific region, indicating the need for broader research across various countries.

Despite the growing body of research highlighting the potential of AI to transform cybersecurity, a significant gap remains in the literature regarding how organizations transition to developing and using in-house developed AI-based cybersecurity solutions. The papers by Sontan and Samuel [93] and Familoni [35] provide comprehensive reviews of the benefits and challenges associated with AI in cybersecurity, but these studies are largely theoretical and lack real-world insights into the practical implementation and transition processes. The empirical studies conducted by Gusman [41], Al-Dosari et al. [4], Gonaygunta [39], and Radebe et al. [85] deliver findings based on survey and interview data. These studies reveal how AI is perceived in cybersecurity among professionals in the field, identifying key challenges and benefits of its adoption. In this work, we contribute to the latter body of work by conducting a case study on the resistance factors to AI adoption among end-users of a cybersecurity department and sharing practical insights on how these barriers can be overcome.

Resistance to AI Adoption. Given our observation that end-users in cybersecurity departments resist to adopt the AI technologies, although they are available at hand or can be developed in house, the Innovation Resistance Theory (IRT) seems the most relevant theory for our study. This theory has recently gained significant popularity due to the rapid emergence of IT innovations and the resistance to their adoption. Below, we describe the works that apply IRT to AI adoption cases.

[3] *Adversarial machine learning* is a field that focuses on the security vulnerabilities of machine learning models, particularly how they can be deceived or manipulated by carefully crafted inputs.

Yang et al. [103] investigated how the anxiety induced by so-called AI-based smart technologies, e.g., front desk reception robots or autonomous delivery, in particular, anxiety due to AI delegation or AI surveillance, impacts IRT functional barriers, which, in their turn, contribute to the resistance of Unmanned Smart Hotels (USH) technology adoption. By conducting an online survey of 355 participants, they demonstrate that functional barriers and the usage barrier related to the ease of use of the technology are important determinants of consumer resistance to USH. There are also high concerns regarding the risks introduced by these technologies. The anxiety due to delegation and surveilance only increase the perception of these functional barriers.

Jiang et al. [47] proposed a model to investigate how perceived anthropomorphism (similarity to human behavior) and perceived intelligence (degree of smartness) of AI components impact the resistance towards adoption of AI-enabled fintech innovations. Unfortunately, the model has not been tested experimentally; therefore, the results are unknown.

Chia-Lun Lo [62] empirically investigated the attitude towards AI in the job market. The study explores how factors like global (an innate trait showing the willingness of a person to adopt innovations) and technological (innovation acceptance in a given field) innovativeness and technology affinity ("gadget loving") influence attitudes toward AI, with four resistance barriers (usage, value, risk, and tradition) acting as moderators. A survey conducted among recent graduates in Taiwan (230 questionnaires) revealed that higher innovativeness and technology affinity positively influence attitudes toward AI, while innovation resistance dampens this effect, i.e., the higher each of the innovation barriers the lower the attitude toward AI.

A very relevant study was done by Mahmud et al. [63], who investigated algorithm aversion (non-acceptance of the results of algorithmically-made decisions) among bank managers by analyzing the impact of functional and psychological IRT barriers. Survey data from 167 bank managers show that value, tradition, and image barriers significantly contribute to algorithm aversion, while usage and risk barriers are less impactful due to managers' familiarity with technology and risk management.

Mou et al. [71] investigated how the AI "image" can influence the resistance of consumers to its adoption. The authors found that if users see AI as a "substitutor" they perceive it as a higher threat and thus, tend to resist its adoption, while if they see AI as a "facilitator" the perceived threat is lower, as is their resistance. It was also shown that the communication ('servant' rather than 'partner') and language ('literal' rather than 'figurative') styles can moderate the level of the perceived threat, lowering the resistance towards AI adoption.

Human-AI Collaboration in Cybersecurity. AI is often framed either as a potential replacement for human expertise or as a neutral tool to automate repetitive tasks. Such binary views overlook a third and increasingly important perspective – AI as a *collaborator*. In this framing, AI systems and humans form a joint team, referred to as *hybrid intelligence* [3], working in coordination to complete a common task while combining their respective strengths and offsetting each other's limitations [28].

Parasuraman et al. [81] outline a taxonomy of automation ranging from fully manual (level 1) to fully autonomous (level 10), illustrating varying degrees of human control. Later research conceptualizes different configurations of human-AI teaming, including humans supervising AI systems, collaborating as equal partners, or being overseen by AI systems [34]. Related notions such as human-in-the-loop (HITL), human-on-the-loop (HOTL), and human-out-of-the-loop (HOOTL) similarly capture how responsibility and authority are distributed across human–AI systems, emphasizing that collaboration exists along a continuum rather than as a fixed state.

The synergy between humans and AI becomes particularly valuable in decision-making environments marked by uncertainty, complexity, and ambiguity [45]. Cybersecurity exemplifies such a context, where incomplete information, rapidly evolving threats, and multiple plausible interpretations challenge individual cognition. While humans excel at intuitive judgment and contextual reasoning under uncertainty, AI systems contribute by processing vast datasets and generating analytical insights in real time. Their collaboration combines human guidance with machine precision, enhancing decision quality in dynamic threat environments [16].

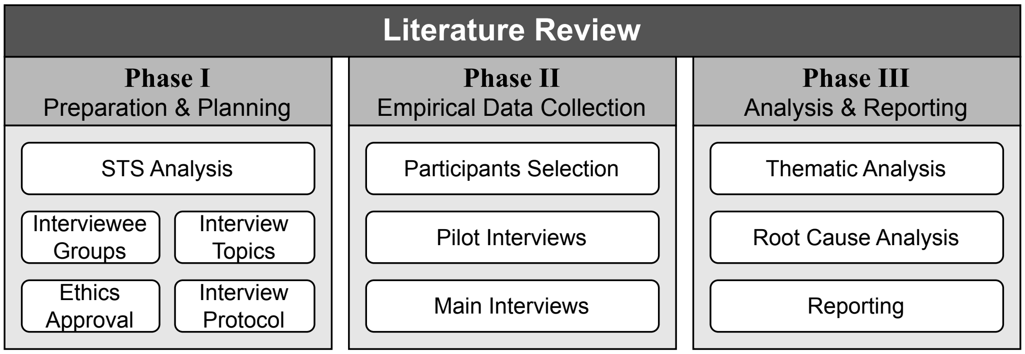

Fig. 1. Overview of our research methodology.

Human–AI teaming has gained growing attention across disciplines, including cybersecurity [38], leading to the development of several frameworks aimed at enhancing collaboration. Osholake et al. [80] proposed an AI augmentation framework that combines machine learning with human expertise to optimize analysts' cognitive load, evaluated through a mixed-methods design. The study integrated quantitative performance metrics with qualitative insights from security operations center (SOC) practitioners across 47 enterprises in the United States, structured in three phases: framework development, pilot implementation, and large-scale evaluation. The participating organizations reported a 52% reduction in false alerts, a 38% improvement in mean time to detection, and a 41% decrease in analyst burnout.

Similarly, Baruwal Chhetri et al. [16] introduced a framework supporting adaptive decision-making in SOCs through seamless transitions between *automated*, *augmented*, and *collaborative* modes of operation. This approach enables automation for routine tasks, AI augmentation to accelerate expert decision-making, and human–AI collaboration to address complex or novel threats. A subsequent study [96] validated the framework using benchmark datasets: on the KDDCup intrusion detection dataset, performance improved from 33.43% under full automation to 35.18% with augmented deferral, and further to 87.04% through collaborative exploration, demonstrating the model's effectiveness in managing uncertainty. Nonetheless, these results were obtained in controlled simulation settings, and real-world deployments may yield different outcomes due to greater variability in expert interactions and contextual factors.

Successful implementation of human–AI collaboration frameworks ultimately depends not only on technical design but also on organizational readiness and cultural adaptation. As demonstrated by Osholake et al. [80], effective deployment requires comprehensive change management strategies addressing both technological and human dimensions of SOC transformation, supported by targeted training and stakeholder engagement. Looking ahead, advancing cybersecurity defense calls for augmenting human intellect through AI systems that are not only collaborative and adaptive but also responsible and explainable, ensuring that automation enhances rather than replaces human judgment in complex, high-stakes decision-making environments [3].

3 Methodology

Figure 1 provides an overview of the methodology applied in this case study, which comprises three main phases: (i) Preparation & Planning, (ii) Empirical Data Collection, and (iii) Analysis & Reporting. Throughout the study, we continuously referred to relevant literature to identify applicable theories, methodologies, and related work. A summary of the insights gathered during our literature review activity is presented in Section 2.

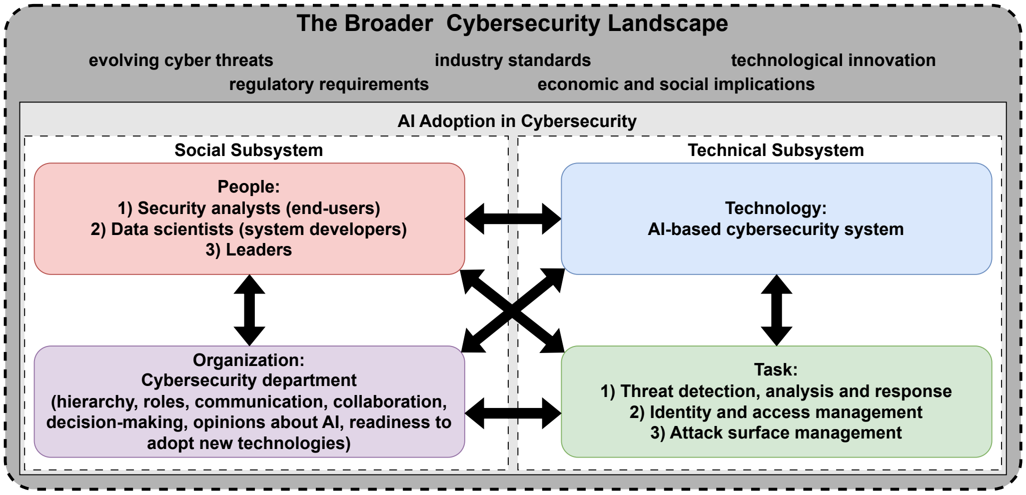

Fig. 2. Sociotechnical System analysis.

3.1 Phase I: Preparation & Planning

The goal of this phase is to acquire relevant knowledge and to prepare the foundation for conducting our study. For this research on how users in organizations adopt AI-based cybersecurity solutions, we have employed a *case study* approach. The challenges and strategies involved in developing and adoption of AI-based solutions for cybersecurity are multifaceted and influenced by many contextual factors. A case study allows for an in-depth exploration of complex phenomena within their real-life context, and this method is particularly suitable for answering *how* and *why* research questions [104].

In this study, we explore the case of a large international financial organization, specifically its cybersecurity department. *The Organization*[4] is a major international financial services provider, offering a wide range of retail and wholesale banking solutions. It operates globally across numerous countries and employs tens of thousands of people. Like any leading financial institution, the Organization faces significant cybersecurity threats and invests substantial resources to enhance its security measures. Its cybersecurity department has several hundred employees working across various countries. The department uses both vendor-provided security solutions, some of which incorporate AI-based services, and custom AI-based tools developed by the team of data scientists in-house.

To better understand the case of our study, we conducted an analysis using the *Sociotechnical System* (STS) theory. The STS theory splits sociotechnical systems into two components – 'social' and 'technical' – and studies the interaction between them [102]. At its heart is the concept that new systems can be optimized and will only function effectively when social and technical elements are combined and considered as interdependent parts of a work system [102]. The theory focuses on the joint optimization of both subsystems [18, 99] and highlights how important humans in the organization are to solve complex issues instead of relying on technical solutions only [79]. According to this theory, we split our system into technical and social components and outline all possible interactions between them. Figure 2 shows the main components of our STS. In the supplementary document[5], we describe these components in detail and characterize identified interactions among them.

[4]To protect the confidentiality and anonymity of the participants and organization involved, the organization's name is not disclosed. Throughout the article, it will be referred to as *the Organization*.
[5]Please refer to the supplementary materials for this publication.

Table 2. Interview Topics per Interviewee Group

Interviewee Group	Topics Covered
Security Analysts	• Current use of AI tools for cybersecurity • Integration into daily workflows • User experience • Participation in AI development projects • Trust in AI technologies for cybersecurity • Concerns and barriers for non-users • Perceptions of AI's impact on job roles and decision-making
Data Scientists	• Technical and organizational challenges in developing AI models • Collaboration and communication with other teams • Ensuring user-friendliness and workflow integration • Gathering and incorporating feedback • Factors influencing AI adoption among end-users • Future AI trends in cybersecurity • Trust and accountability in AI-assisted decision-making
Leaders	• Benefits and challenges of AI in cybersecurity • Ideation and adoption of new AI innovations • Communication and feedback mechanisms • Factors influencing AI adoption rates • Future impact of AI on job roles and responsibilities • Accountability and liability issues related to AI-assisted decision-making

Based on our STS analysis, we identify the groups of employees from the Organization's cybersecurity department and the topics to discuss with them. We identified three groups of stakeholders whose opinions we would like to consider, namely, *Security Analysts*, *Data Scientists*, and *Leaders*. Security analysts are selected because they are the end consumers of different AI tools, therefore, their input about why they do not use particular AI solutions is critical for our research. Data scientists are heavily involved into the development of new approaches and tools to be used by security analysts, hence, their perspective on where they obtain the input from and whether they obtain and incorporate the feedback is important. Finally, leaders formulate the agenda of the cybersecurity department. They are involved in the decision-making on what AI solutions to procure or to develop in-house. Table 2 displays a summary of the topics for each interviewee group identified performing the STS analysis.

As the data collection method, we decided to employ *semi-structured interviews*. Semi-structured interviews rely on a predefined set of questions whose order or phrasing is not set [84]. This interview method was chosen because the research aimed to uncover rich details about the AI implementation process and AI's acceptance in the Organization's cybersecurity department. This includes personal perspectives and details about the organizational dynamics, which might not have been properly captured if, for example, a survey was used. Semi-structured interviews were preferred over structured or unstructured, as this method allows for exploring specific topics in detail and adapting to the participants' responses. This way, non-anticipated insights can be identified too [7], which was one of the goals of our study.

Based on the identified topics, we developed initial interview questions for each group. To determine the sequence of the questions, the guidelines, developed by Annette Lareau [55], were followed. The interviews started with standard questions to make the participants feel at ease, such as asking them to talk about their job responsibilities. More sensitive or speculative questions were asked towards the end of the interview. For more sensitive topics, participants were also asked to reflect on the experiences of their colleagues. This approach is supported by sociological guidelines that emphasize the importance of creating a safe environment for interviewees. Indirect questioning can reduce the perceived personal risk of disclosing sensitive information [23]. This measure reduces social desirability bias, one of the most common sources of bias that affects the validity of research findings [73].

We sought the feedback from subject-matter experts to review the initial version of the interview questions, ensuring their clarity and relevance. Following this feedback, the interview protocol was adapted, and the number of questions was reduced. The final version of the interview protocol is provided as a supplementary document.[6]

All interviewees received detailed information about the study before consenting to participate. A comprehensive data management plan was implemented to ensure the proper handling, storage, and security of the collected data. All interview transcripts were anonymized to protect participants' identities. The data was stored on the Organization's secure servers, accessible only to the principal researcher. The data will be retained for two years after the completion of the project, according to ethical guidelines, and after will be securely disposed. Ethics approval was obtained from the Human Research Ethics Committee (HREC) of the research institution involved in this study.

3.2 Phase II: Preparation and Planning

Research shows that empirical saturation – the point where new interviews do not bring new ideas, indicating that additional data collection is no longer needed – is usually reached within the range of 9-17 interviews [42]. For this study, participants were recruited through *purposive sampling*, with the aim of capturing diverse perspectives on AI integration in cybersecurity. Selection was based on their usage of AI tools, involvement in AI-related projects, and/or cybersecurity expertise. This ensured the inclusion of security analysts, data scientists, and leaders whose roles and experiences provided complementary viewpoints:

- **Security analysts** – *Participants A1-A6* – are directly involved in day-to-day security operations and threat management. Some of them dedicate limited time (e.g., a day per week) to collaborate with data scientists on machine learning projects;
- **Data scientists** – *Participants D1-D6* – are responsible for developing and deploying AI solutions in the cybersecurity department;
- **Leaders** – *Participants L1-L3* – oversee operations in the cybersecurity department.

The interviewees were contacted via direct invitations or internal referrals. Due to the sensitivity of the provided information, participants' job titles, years of experience, and cybersecurity areas of expertise are not disclosed in the report to prevent their identification among their colleagues.

Notably, two-thirds of the respondents are involved in the Security Detection and Response (SDR) area, where AI-integration efforts have been ongoing for the longest time within the Organization's cybersecurity department. This focus provided a rich source of data on the long-term challenges and successes of AI adoption.

The study includes the opinions of four security analysts who either use machine learning-based applications for their tasks or have been involved in machine learning projects as domain experts, as well as two analysts who do not use AI tools in their work. The analysts who use machine learning tools provide direct insights into the practical challenges of AI integration, while those who do not use AI shed light on barriers and resistance to adoption. By capturing both user and non-user experiences, we aimed at a thorough analysis of the sociotechnical factors influencing AI adoption.

The leaders selected for the study hold significant influence within the cybersecurity department. Their responsibilities include making critical decisions about AI adoption and implementation, overseeing operational

[6]Please refer to the supplementary materials for this publication.

activities, and shaping the strategic direction of cybersecurity initiatives. These leaders offer a mix of strategic and operational perspectives, contributing to a more holistic understanding of the factors influencing AI adoption.

We conducted 3 pilot interviews with a representative of each interviewee group to refine the clarity of the questions. The results of these interviews are not included in the results.

Within this study, a total of 15 main interviews were conducted either face-to-face or virtually, based on the participants' location and preference. On average, the interviews lasted 45 minutes. All interviews were recorded with the consent of the participants and transcribed for analysis.

3.3 Phase III: Analysis and Reporting

The interview data were analyzed using the Reflexive Thematic Analysis (RTA) framework developed by Braun and Clarke [27]. The reflexive approach to thematic analysis emphasizes the researcher's active role in creating knowledge [19]. It involves six steps: 1) Familiarization with the data; 2) Generating initial codes; 3) Generating initial themes; 4) Revision and refinement; 5) Development and definition of themes; 6) Reporting. During the first step, the responsible researcher re-read the transcripts a couple of times, ensuring the obtained transcripts were accurate, complete, and free of errors. Additionally, interesting passages and points for analysis were noted at this step. At the second step, each interview was coded systematically using the ATLAS.ti [12] software using an *inductive* (bottom-up) approach. During the initial theme generation step, initial categories were created to organize the codes into groups with similar meanings. Then, at the next step, these categories were refined through continuous engagement with the data. During this process, some codes were merged, while others were disregarded. At the fifth step, the themes were reviewed to ensure coherence and alignment with the overall data set. This involved two levels of review. First, at the coded data level, each theme was examined to determine if the coded data excerpts formed a coherent pattern. Secondly, at the data level, the coded excerpts were re-read to confirm that the themes accurately reflect the meanings present within the data. Finally, we provided clear definitions of the identified themes. In the final step, the focus shifted to producing the report. This involved integrating the themes into a cohesive narrative, providing a detailed account of the findings.

To gain a deeper understanding of challenges and doubts regarding the AI adoption for cybersecurity operations, we conducted a *Cause-and-Effect Analysis*, also known as a *Root Cause Analysis (RCA)*. This is a collective term describing a range of systematic approaches applied to identifying the underlying causes of a problem and determining the actions required to resolve it. RCA distinguishes between:

- **Undesirable Effects:** These are the visible problems that emerge as symptoms of deeper issues within the system.
- **Causal Factors:** Actions, conditions, or events that directly contributed to the problem occurring. These are the intermediate elements that contribute to the undesirable effects, acting as links between the root causes and the observed symptoms.
- **Root Causes:** The most fundamental system, process, or design flaws which, if corrected, would prevent recurrence of the problem (or similar problems). Often, removing a causal factor only fixes one instance, while fixing a root cause prevents many potential future instances.

Thus, RCA aims to investigate the problems and find their underlying causes, ensuring that proposed solutions address the fundamental issues rather than merely treating the symptoms of a problem [9]. In the context of this study, RCA aims to move a level deeper and use the interview findings to identify the logical relationships between identified themes and sub-themes, as well as understand the actual causes of the problems mentioned. It should be noted that these relationships are not empirically proven but rather logical connections or potential associations. Establishing causality would require further statistical studies and empirical validation, which were not possible to do within this work due to the small interview sample.

4 Findings

Table 3 provides an overview of the main benefits, challenges, and interventions emerging from the interviews. It highlights how participants perceive the impact of AI adoption on cybersecurity practices, the barriers that hinder effective integration, and the organizational mechanisms that may foster sustainable human–AI collaboration.

Table 3. Summary of Benefits, Challenges, and Interventions Identified in the Study

Category	Key Insights	Impact
Benefits of AI adoption	• Automation reduces repetitive workload and decision fatigue. • Improves threat detection, prioritization, and anomaly recognition. • Enables faster and more comprehensive threat response through self-learning models.	Analysts reported reduced manual effort and improved focus on complex incidents.
Challenges of AI adoption	• Cultural resistance and legacy mindset. • Ineffective feedback mechanisms. • Insufficient management support and strategy alignment. • Data-related barriers (quality, labeling, fragmentation). • Analysts' limited AI literacy and trust. • Data scientists' limited cybersecurity knowledge. • Job-security concerns.	Difficulties realizing AI projects and lack of trust in ML outputs; feedback loops often ineffective.
Proposed interventions	• **Education & training:** foster AI literacy, collaboration, and innovation culture. • **Reward innovation:** recognize employee contributions and co-creation. • **Improved feedback mechanisms:** enhance openness and psychological safety. • **Strategic communication:** align leadership vision with practice. • **Clear process structure & transparency:** ensure scalability and explainability. • **Career mentorship:** address AI-related career anxiety.	Interventions collectively target sociotechnical barriers and aim to build sustainable human–AI collaboration.

4.1 Benefits of AI Adoption in Cybersecurity Operations

In general, participants were enthusiastic about AI's transformative potential for cybersecurity operations. The use of AI in cybersecurity has the potential to revolutionize the field, bringing considerable efficiency gains and improving threat management. For many security analysts, the daily responsibilities involve repetitive and redundant tasks that can be time-consuming and monotonous. Implementing AI-based solutions can alleviate this burden by automating routine tasks, allowing analysts to redirect their focus toward more critical issues. For instance, analyst A2 shared: *"We have so many events that the model could help with this manual work, and we will have more time to do another task. And these are events that you just scroll down."* This highlights how AI can take over mundane tasks, giving analysts more time for higher-level analysis. Another analyst A3 reflected on their early career experiences: *"When I started [my job], it was really routine, I was scrolling through the data and (. . .) I think, we were all overwhelmed by this routine. (. . .) We are really grateful that the company decided to focus on us and eliminate the manual work."* This opinion highlights the *significant relief AI can provide by removing the repetitive aspects of analysts' jobs.*

This shift reduces the workload and also helps to combat decision fatigue, a common issue in the cybersecurity environment. Analyst A2 further explained *"It [the model] actually reduces sort of your workload. It helps to make it easier and to focus just on the things that require more attention."* Decision fatigue occurs when individuals are required to make too many decisions in a short period, often leading to errors and decreased productivity. By filtering and prioritizing tasks, AI helps mitigate this problem, enhancing overall efficiency.

Beyond automating tasks, AI's most frequently mentioned contribution to cybersecurity is its ability to improve threat detection and management. ML models are used to filter out false alerts and to prioritize potential threats, ensuring that analysts' attention remains focused on the most important events. For instance, data scientist D2 noted: *"There are things that they [analysts] cannot do on their own. For example, they have huge amounts of alerts from something, and they need to prioritize them, or they need to filter out."* This points to AI's capability to handle large volumes of data and streamline the threat management process.

AI's advanced computational capabilities enable it to process vast amounts of data quickly and efficiently, a critical requirement in the cybersecurity domain. Analyst A2 supported this: *"And of course, in the case of so many events, there's a possibility that you won't recognize something (...) so there is a chance that the model would choose the events that you missed."* Data scientist D2 also highlighted AI's anomaly detection ability: *"ML was found to be very useful in finding anomalies."* This underscores the role of AI in ensuring comprehensive threat coverage. The need for speed and comprehensive analysis was further highlighted by analyst A2: *"We really need that because of so many events, we really need the help of AI to find everything, and the second thing is the quicker time."* An AI system's ability to continuously learn from new data and adapt to evolving threats further enhances its effectiveness. This is a major differentiator between using AI systems and traditional rule-based approaches, which prevail in current solutions.

AI's self-learning capabilities ensure that it constantly evolves, adapting to new threats and improving its performance over time. This dynamic nature of AI is a significant advantage in the ever-changing landscape of cybersecurity. Traditional rule-based systems lack this adaptability, making AI a superior choice for modern threat detection.

4.2 Challenges of AI Adoption and Their Root Cause Analysis

Despite the promising potential of AI outlined in the previous subsection, various factors hinder its effective application. To gain a more in-depth understanding of these challenges, we conducted their Root Cause Analysis (RCA). Figure 3 presents the diagram of our findings.

In Figure 3, undesirable effects (or symptoms) are listed in dusty rose boxes with arrows pointing toward them. In our analysis, we identified two primary undesirable effects, or symptoms: 1) the difficulties in realizing AI projects in practice; and 2) the lack of trust in ML models. These effects are influenced by various underlying factors, which are depicted as grey boxes. While we consider these undesirable effects separately, we admit that they can be interconnected. For instance, difficulties in realizing AI projects in practice can contribute to the lack of trust in ML models and vice versa, creating a feedback loop that further hinders successful AI development and implementation. This potential bidirectional relationship highlights the complexity of the challenges faced and underscores the importance of addressing both issues simultaneously to achieve meaningful improvements.

Potential root causes are enumerated in blue boxes. While the term root cause analysis seems to point to a single cause, it is rare, or even impossible, to identify only one reason for a recurring problem [83]. For instance, within this study, we identified seven possible root causes, which we describe in detail below.

Organizational Culture and Mindset. A traditional organizational culture that resists change can hinder the adoption of innovative technologies like AI. Many professionals are accustomed to established methods and processes and are skeptical about new approaches.

Participants indicated that even if models are technically successfully developed and implemented, people struggle to understand exactly how to use the insights generated by them, as they are accustomed to their old ways of working. For instance, data scientist D4 noted the department's reliance on dashboards as a primary solution: *"The other challenge that's specific for analytics is the way that people think about it because the Organization is a world of dashboards. So, for everyone, the solution is always another dashboard and then that's how people think about things."*

This *resistance to change makes it difficult for the Organization to embrace AI solutions, leading to slow progress in AI project implementation.* A couple of participants indicated that some of the reluctance to adopt AI in cybersecurity comes from the culture within the department. This aligns with broader organizational-level variables described in the TOE framework, where *internal culture and historical dependencies on existing technologies*

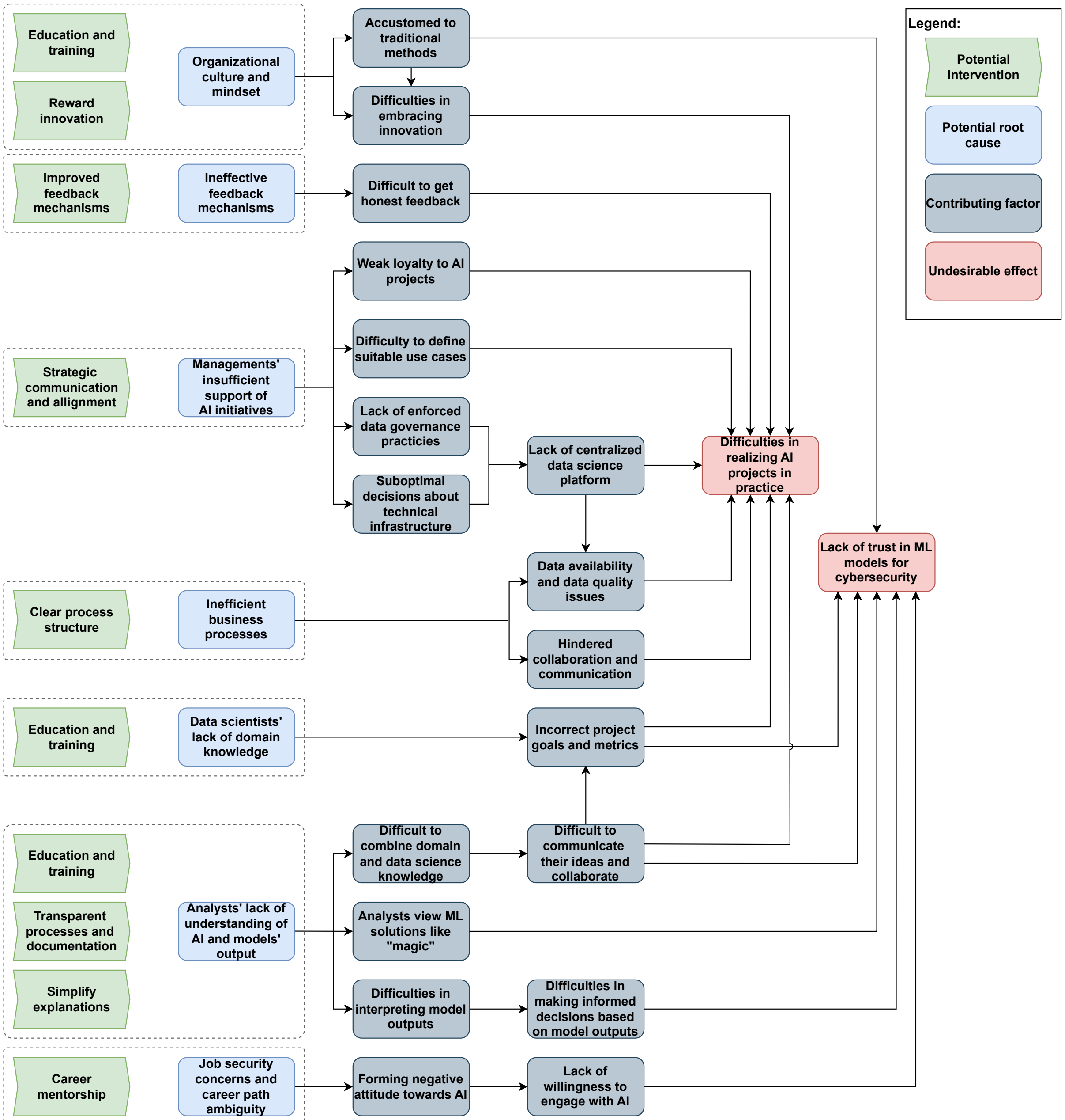

Fig. 3. Root cause analysis and potential interventions.

shape attitudes towards innovation adoption. For cybersecurity practices, the Organization has relied on rule-based solutions for a long time. Data scientist D6 highlighted: *"…another barrier might be legacy because usually, we don't do analytics in cybersecurity."* Hence, the integration of AI is considered very innovative. The highly regulated nature of the department further contributes to its reluctance to adopt new technology. Another participant added that this resistance is stronger among more experienced teams with older employees.

Ineffective Feedback Mechanisms. Ineffective feedback processes can lead to communication gaps between stakeholders. For instance, the data scientists expressed concerns about not receiving honest feedback from analysts regarding the models' usefulness to them and their performance. *"I think they are sometimes afraid of telling us that they are not using all of the models,"* said data scientist D3. Feedback on model performance is regularly collected during recurring meetings that include data scientists, managers, and security analysts. Unexpectedly, this feedback mechanism was found to be often ineffective, with analysts hesitant to provide honest feedback in formal settings. Additionally, feedback tends to come from the most experienced analysts, which means it might not reflect the broader team's perspectives, leading to a partial understanding of the models' usability.

This suggests that current processes for collecting and integrating user feedback are not effectively fostering open communication, indicating a need for adjustments. Unfortunately, *without honest feedback, data scientists might not fully understand the practical needs and challenges faced by analysts, leading to the development of AI models that do not address the most critical issues or meet the actual needs of the end-users.*

Leaders also suffer from a lack of honest feedback that affects their decision-making process. For instance, leader L1 emphasized the importance of having clearly defined success criteria during the development phase and explained: *"I think we should already at this [early] phase have a clear understanding of what we are trying to achieve. (…) Of course, there is always part of experimentation there, but we should also say, OK, at this point we stop, if it does not work, then we will not develop it further."*

Managements' Insufficient Support of AI Initiatives. We also observed weak loyalty to AI projects among management. The lack of understanding of how AI functions and the historical context of cybersecurity practices significantly affects the acceptance of AI in this area. Unlike many technologies, which we use in daily life without fully understanding them, cybersecurity has traditionally relied on rule-based models, which are inherently easier to comprehend. These models operate on clear, predefined rules that are easy to understand, modify, and predict, providing a sense of control and transparency. *The shift to AI-driven approaches introduces a level of complexity and opacity that contrasts with the simplicity of rule-based systems. However, the adoption of AI by malicious actors underscores the necessity for defensive teams to also integrate AI into their operations.* This ensures that defenders can keep pace with attackers who are continually evolving their tactics using advanced technologies. As leader L1 pointed out: *"Simply to be able to keep up with the attackers, we need to start onboarding and using as many new technologies as we can, so that will definitely change the way we operate, how we build the detections, but also how we do the security analysis."*

Identifying effective use cases for AI in enterprise-level cybersecurity is a challenging task. While academic research often provides theoretical solutions, these do not always translate seamlessly into practical applications. Models that generate too many false positives, fail to provide new alerts, or do not operate in real time are often rejected by security teams. Leader L2 noted: *"Almost every paper that we found on the subject that we were looking for was useless. Because when they apply it to a large organization, it's simply much harder to tune it and in many cases, it turns out that you gotta figure out something on your own from scratch."*

Participants indicated that top managers often lack a deep understanding of the infrastructure and data requirements for AI technologies. This knowledge gap can lead to suboptimal decisions regarding technical infrastructure and resource allocation, placing additional pressure on data scientists and hindering the success of AI projects.

The absence of a centralized platform for developing, testing, and deploying AI models adds another layer of difficulty. Data scientist D3 emphasized: *"With rule-based models, I believe that you can do whatever you want because we have those detection systems where you can write something simple, and for machine learning models, you need to invest in the platform."* Therefore, resource allocation and investment in building such a platform, which would be a non-trivial undertaking, should be supported by the management at all levels.

Inefficient Business Processes. *The effectiveness of AI is highly dependent on data availability.* Obtaining data within the Organization is a significant challenge due to data being spread across various departments. *"Every system that you need to get data from is different or [has] different data owners and is a gigantic challenge to be able to get, not even to talk about good quality data,"* data scientist D4 pointed out. This fragmentation also complicates data integration and consistency. AI models – particularly those using supervised-learning approaches – require labeled data to perform effectively. However, in the field of cybersecurity, such labeled data is often scarce. *"One of the biggest challenges for ML in security is the lack of labeled data,"* noted data scientist D2.

Another critical challenge is the quality of the data, which directly affects the performance of AI models. Issues such as missing data, schema imperfections, and bugs can significantly impair model performance. As data scientist D6 observed: *"When people create datasets or the data structure in the back-end of the system, for example, they don't really think about the analytics side. So, they just create the field, and they don't care about the quality, just capturing data."* This underscores the importance of strong data governance practices to ensure high data quality. The absence of well-defined business processes or their inefficiency contributes significantly to these factors.

Unclear business processes can also impede communication and collaboration, leading to tension and eroding trust. For instance, analyst A1 shared a story about the development of a model where analysts were neither actively engaged nor properly informed, and how this discouraged them from using it: *"We only got the information that [the model] was prepared and we didn't get any details about this project. I mean, what is the reason to create it? How can it interact with us? I think the [data science] team has some ideas, but they don't verify with us during the development. I think that, in some way, this model is a very far project from us. (. . .) As analysts are treated only as the side of the project, they don't know about it."*

Half of the participants shared that the lack of time is among the biggest challenges that they face when collaborating on AI projects and their integration into the Organization's cybersecurity department. Data scientists mentioned that scheduling meetings with analysts is challenging because of the analysts' busy schedules. Data scientist D4 explained: *"Everyone is up to their eyeballs in work and priorities (...) so getting their availability is the hardest."* Additionally, some analysts work on three shifts, including evenings and nights, which complicates scheduling further. Analysts who are interested in machine learning and want to participate in these projects also reflected on this issue. Analyst A5 highlighted the challenge of balancing regular tasks with AI project work: *"So far, it was mostly the time issue, in the context of actually getting some time for this project along with regular tasks. (. . .) There wasn't much time available from our side, even though we really wanted to do it."*

Data Scientists' Lack of Domain Knowledge. Participants indicated the complexity of combining highly specialized cybersecurity knowledge with data science. Data scientist D5 described this: *"If you are doing data science for a grocery store, you can most of the time understand what's going on. But if you're doing machine learning applied to logs of web servers, you need all these people who deal with web servers and know exactly what is going on."*

A lack of domain knowledge among data scientists can result in a focus on incorrect goals or the use of inappropriate metrics, often leading to unnecessary or unrecognized efforts. "In some cases, you know, you work on some models for a long period of time, and then you put in the effort to really making this model do some good things. And this is just trashed to the side, you know, because it's generating too many false positives," noted data scientist D1. Knowing when to discontinue the efforts is crucial, especially given the time constraints faced by analysts. Analyst A4 reflected: *"There was sometime when they [the data scientists] produced a lot of additional work for us. (…) Not all the models worked and brought us good-quality results. So, you want to create something great, but after years, you finally quit it because the results are bad and there is no possibility to use machine learning in such a way."*

The difficulty in developing effective AI models directly impacts trust and adoption among cybersecurity professionals. Analysts may be hesitant to embrace AI-based solutions, particularly if previous models have not met expectations. Building trust requires demonstrating the reliability and utility of AI in real-world scenarios. Data scientist D2, reflecting on the challenges faced with an internally developed model, illustrated this point: *"The model simply wasn't necessary and wasn't well-performing. And I think they lost a lot of confidence in us. (…) So I think it's really important to be very cautious now because if we again send something, it needs to be really good so we can have their confidence back. They can know it's not useless, it's worth their time."*

Analysts' Lack of Understanding of AI and Models' Output. Many analysts indicated that they do not understand how machine learning works, causing them to see ML solutions as "magic". This lack of understanding affects both the implementation of AI solutions and the analysts' trust in these technologies. Indeed, when analysts cannot grasp how ML models arrive at their conclusions, it becomes difficult for them to trust these outcomes. Analysts may feel more confident relying on their own judgment rather than on an AI system they do not fully understand. This was nicely described by security analysts A4: *"I think it's confusing me. I don't know what to analyze (…) so I'm not really confident to work with them [models] because I don't know how they work exactly. (…) Maybe I feel more confident with my own analytics to verify all data, and I know cases from my past that we are looking for something, and when I saw all logs in a table on a screen, I saw here is the problem. (…) I'm pretty sure that AI is not going to find it."*

When analysts understand how AI works, they can better interpret unexpected results. They can identify specific reasons for these results instead of assuming that the system is faulty. Data scientist D2 explained: *"… and they [analysts] know how it works, so they're not surprised if they see a peak in the number of results because they can think, more or less, well, this is because of this and that. And they are less angry with us when they see more outcomes and more work for them."*

Job Security Concerns and Career Path Ambiguity. Many employees fear that AI technologies might replace their roles or render their skills obsolete. Most participants expressed a general fear about the future of the security analyst profession. They highlighted that within the Organization's cybersecurity department, there is a specific concern about AI potentially replacing security analysts' jobs. As we mentioned in Section 4.1, AI benefits cybersecurity by reducing false positive alerts and automating routine tasks. While this automation will not eliminate the need for analysts due to the high-stakes nature of the field, there is a fear that the demand for security analysts may decrease in the future. This fear was illustrated by analyst A4 saying: *"The day is going to come when in analytics, some of us will be replaced with AI. That's it, I would say. Because this is the future, unfortunately."* This reflects the "image" and "tradition" barriers outlined in IRT, where perceived threats to role identity or career longevity foster resistance to adoption, even when the technology offers demonstrable benefits.

This statement illustrates that analysts recognize AI's growing capabilities in the field and its potential to outperform humans in certain analytical tasks. The use of the word "unfortunately" suggests that while AI's integration is seen as a progression, it also raises fears about the potential decrease of human roles in the industry. This fear creates resistance and scepticism as people worry about their professional future. Analyst A6 shared: *"They [analysts] are scared about losing their job position. (…) This is the main reason why they don't want this technology in cybersecurity. (…) When I speak with people, this is one of their arguments."*

This aligns with survey results that underscore high anxiety levels about automation [1]. In the context of this study, this concern leads to negative attitudes toward AI, resulting in a lack of willingness to engage with these technologies. This phenomenon can be understood through the lens of Innovation Resistance Theory (see

Section 2.1). In this case, if analysts believe that AI will replace their jobs, it creates psychological discomfort that makes it difficult for them to trust or cooperate with AI technologies.

4.3 Potential Interventions

Within our study, we have also identified potential interventions that, based on the experience of our participants, can mitigate the identified issues. These recommendations are grounded in both participant experiences and the sociotechnical systems framework used throughout our analysis, highlighting the interdependence of social and technical subsystems in successful AI adoption. In this section, we describe these measures.

Education and Training. Our findings suggest that *education and training can address two key issues. First, they can promote a culture of innovation by changing the organizational mindset and creating a more welcoming environment for AI technologies. Second, they can enhance analysts' understanding of AI and data scientists' domain knowledge, effectively bridging the gap between data science and cybersecurity.* Notably, the Cybersecurity Data Science (CSDS) practice is a new discipline that is currently developing from practical experience rather than being a purely theoretical or synthetic program. For its success, cross-training and collaborative teaming are essential components [70]. It is evident that individuals influence the organization by contributing to the collective skills and expertise of the workforce. This collective expertise determines how effectively the organization integrates and leverages new technologies [48]. Empirical evidence supports this view: organizations that implemented comprehensive training programs focused on effective collaboration with AI – combining technical integration, AI-partnership skills, and cognitive load management – reported significantly improved performance outcomes compared to those with minimal training investments [80]. Given this evidence, implementing targeted educational measures becomes essential. Organizations can:

- *Invest in AI fundamentals education*
- *Organize innovative events*
- *Provide continuous learning opportunities*
- *Run serious games*

From the empirical analysis, it became apparent how important it is to invest resources in educating cybersecurity analysts about AI fundamentals because employees often lack a clear understanding of what AI is and how it can support their work that was also observed by Zhang and Lee [105]. The interviews revealed that analysts' increased knowledge of AI has significantly enhanced teamwork dynamics between analysts and data scientists, improving the communication between them. Data scientist D3 illustrated the last point: *"We know how to prepare data for the model, how to prepare features for the model because we know what the feature is and why it is important, and I think for the end-users, who don't have knowledge around the data science, it's very difficult to start thinking which features we could use."* While participants acknowledge that there is still much progress to be made, they agree that foundational training is an essential building block of this journey.

Events such as the organized hackathons were also recognized as positive measures, providing several benefits. First, they encourage creative thinking. Cybersecurity professionals are accustomed to traditional analytical methods, while hackathons push them out of their comfort zones, challenging them to create innovative machine learning use cases. Second, it gives participants the chance to collaborate and gain practical experience. Working on hands-on projects helps them see how machine learning is applied in real-world scenarios, as illustrated by analyst A2: *"The hackathon was one of the main points of my journey to get into the topic. (...) I had the chance to work with other people and see how it [machine learning] really works."* Additionally, they allow participants to experience working in a team on a data science project based on a use case they create. This breaks the routine way of working and shows the benefits of close interdisciplinary team collaborations [72]. Hence, we advise organizations to organize such events, and benefit from their outcomes. Besides hackathons, this can also be achieved through workshops.

While training, hackathons, and workshops can introduce people to the topic, effective learning occurs through continuous practical application. This aligns with the educational theory of constructivism, an action-oriented approach to learning, which states that learners should construct knowledge rather than passively absorb information [13]. Data scientist D2 highlighted that by saying: *"I think in order for a person to learn something, they need to do something (. . .), but people don't have time for that. So, this needs to be something that they do as a part of their job."* Among the interviewed analysts, several expressed their interest in learning about the intersection of machine learning and cybersecurity and building a career in this field. It is important to identify such individuals and provide them with opportunities to expand their knowledge and practice. This can be achieved by limited involvement in AI projects that do not interfere with their regular security tasks. This aligns with the recommendation about career mentorship.

To bridge the gap between cybersecurity and data science, it can be equally important to provide opportunities for data scientists to understand the work of security analysts better too. A way to achieve this is through serious games. Gamification involves utilizing game-like mechanics, visual elements, and strategic game principles to engage individuals, encourage participation, enhance learning, and address challenges [43, 49]. In this context, serious games can simulate real-world cybersecurity scenarios [53, 76], allowing data scientists to understand the basics, experience the challenges and thought processes involved in cybersecurity analysis. This experiential learning method can enhance empathy, improve communication, and foster better collaboration between data scientists and security analysts.

Reward Innovation. *Another intervention, which can stimulate employees to be more adaptive and embrace innovation, is a good rewarding system for innovations.* Kerr and Rifkin [50], emphasize that the design of reward systems plays a crucial role in shaping employee behavior and organizational outcomes. They argue that effective reward systems align with organizational goals and motivate employees to perform desired behaviors. The authors highlight that *there is no one-size-fits-all approach, and if reward systems are implemented wrong, these can lead to expensive and unwanted outcomes.*

Malek et al. [64] study new product development performance and empirically demonstrate that financial rewards may negatively impact intrinsic motivation. In contrast, they find that recognition and social rewards positively influence intrinsic motivation. In our study, we discovered that analysts gain a sense of ownership over the product when they participate in its development process. Their influence on the development increases their satisfaction with the final product and builds trust, as they become partners in the project. Data scientist D2 illustrated this: *"The more you involve people in developing a product, the more influence they have on what they see, the happier they are because they feel they are part of the process, they feel they have something to say."* Thus, adjustments to the reward system – some of which do not involve financial expenditures at all – may help mitigate the undesirable effects.

Improved Feedback Mechanisms. *To address the communication gaps due to ineffective feedback processes identified in our root cause analysis, it is necessary to enhance the feedback mechanisms.* Feedback on model performance is regularly collected during recurring meetings where data scientists, managers, and security analysts are present. However, data scientists have observed that security analysts tend not to speak up during these meetings. *"I think they are sometimes afraid of telling us that they are not using all of the models,"* said data scientist D3. Instead, they provide more honest feedback in informal settings, such as small talk.

Our study highlighted a common challenge in many organizations where hierarchical dynamics can stifle open and honest feedback [72]. This relates to the concept of psychological safety, which is the state of feeling secure enough in a group setting to take interpersonal risks. Essentially, it means that individuals can voice their opinions, seek assistance, or acknowledge errors without fearing criticism or embarrassment. Psychological safety is crucial for increasing team effectiveness and improving performance [33]. To improve the current feedback mechanisms, the following interventions can be implemented:

- **Feedback culture training:** Conduct workshops focused on building a culture of open feedback. These sessions should showcase instances where feedback has led to significant improvements, demonstrating the value of feedback and encouraging more open communication.
- **Designated feedback mediator:** Appoint designated individuals who can collect feedback from analysts and present it during meetings. This approach helps ensure that feedback is shared without direct confrontation and facilitate more honest and constructive communication.
- **User feedback modules:** If technology permits it, implement feedback modules within AI tools where users can provide real-time feedback directly within the application.
- **Feedback impact communication:** Communicate back to users how their feedback has been utilized to improve the AI models. This transparency can help build trust and show that their input is valued and acted upon.

Strategic Communication and Alignment. Organizational strategy significantly influences how technology, including AI, is developed and implemented [48]. Data scientist D4 described the lack of it: *"Even though [Name Redacted] is super great and supportive of us, analytics is still not a big part of their strategy. So, there is no real push from the top down. They always say, "Yeah, you have my full support, we'll do whatever you need. Just let me know.", but that's not permeated within their teams and their organization."*

Strategic communication and alignment among senior managers are essential for the successful development and deployment of AI in cybersecurity. At this hierarchical level, it is crucial to demystify AI and clearly outline the requirements for its full-scale implementation. Given that senior leaders often lack the time to learn about the intricacies of AI, it is imperative to communicate these requirements effectively. This can ensure that senior managers have a comprehensive understanding of what is needed to support and drive AI initiatives, leading to more informed decision-making and successful integration of AI technologies in the Organization's cybersecurity department.

Clear Process Structure. Inefficient business processes related to model development and deployment have been identified as another barrier to the successful implementation of AI projects [72]. While the current limited number of ongoing projects and running models does not present a bottleneck, this issue must be addressed proactively. If the Organization intends to implement more AI models in the future, ensuring scalability and efficiency in developing and managing these models will be crucial to avoid potential challenges. To address this, it is essential to establish a clearer process structure that provides a systematic approach to developing and deploying AI models.

One practical recommendation emerged from this study is to establish a dedicated sub-team of analysts to focus on machine learning projects. Data scientist D4 shared: *"When you start building that network with people that understand what you do, things kind of snowball and get easier."*

Transparent Processes and Documentation. The interviews indicated that transparency in AI processes is essential for building trust and facilitating informed decision-making. According to Lao et al. [54], providing comprehensive information about AI models can significantly enhance their usability and acceptance. The authors advocate for the use of tools like model cards, fact sheets, and an "About Me" tab to offer detailed insights into various aspects of the models. These tools should include information on model performance, documentation, training data, and other relevant details.

The use of model cards has been advocated as a medium to increase transparency among machine learning developers and users by Mitchell et al. [67]. The authors emphasize the importance of viewing model cards as one of many transparency tools, such as third-party algorithmic audits (quantitative and qualitative), adversarial testing, and more inclusive user feedback mechanisms.

Simplify Explanations. The literature on explainable AI (XAI) is quite limited, particularly when it comes to explanations that are evaluated with real-world users [100]. It remains an open question how to address the actual needs of users for understanding AI diagnoses [60]. While the technical field of XAI has developed multiple tools to explain AI models (e.g., SHAP, LIME), it is still unclear how to select the best one and translate it into suitable UX designs [61].

Kim et al. [51] conducted a mixed-methods study with 20 end-users to understand their XAI needs, uses, and perceptions. They found that users prefer practically useful information that resembles human reasoning rather than technical system details. Our interview results also support this finding – in the context of cybersecurity, simple explanations are more useful. Participants noted that SHAP plots, for example, can be too complex and time-consuming for analysts to interpret.

Amarasinghe et al. [8] studied the effect of explanations on fraud analysts by using SHAP among other XAI toolkits and provided a simplified interface. Analysts were presented with the top 6 features with the highest absolute importance. The importance values were indicated by colors: green – for negative importance (no fraud) and red – for positive importance (fraud). A similar interface design could be applied to security analysis, where analysts must decide whether an event is a real threat or not. By providing pairs of feature names and their importance scores in a clear, color-coded manner, we can make the explanations more accessible and actionable for cybersecurity analysts.

Career Mentorship. Our findings of the anxiety related to career uncertainties, when AI is introduced into the workforce, are consistent with existing research [1]. To address these concerns, it is recommended to establish a career mentorship program that promotes analysts' collaborative behaviors with AI [52]. The program can focus on:

- **Career pathways:** Clearly define career pathways that incorporate AI-related roles. This clarity can help alleviate anxiety and demonstrate the organization's commitment to supporting career growth in the evolving technological landscape [46].
- **Mentorship opportunities:** Pair analysts with experienced mentors (i.e., security analysts experienced with AI) who can guide them through the integration of AI into their work. Mentors can provide insights into how AI can enhance their roles and help them navigate potential career transitions.
- **Early engagement:** Engage employees early in their careers by offering training and development opportunities related to AI. This approach aligns with recommendations from Kong et al. [52], advising employers to hire early-career employees and actively train them to work with AI, thereby cultivating a long-term talent pool.

Note that in the interviews, the leaders from the Organization shared that they are not yet highly concerned about this issue. This is because developing AI models is relatively new in the department, and only several models run in production simultaneously. Therefore, while it may seem premature to discuss future careers with analysts due to the limited integration of AI in cybersecurity practices, research indicates that initial trustworthiness beliefs, or first impressions, can have long-term impacts [21]. Trustworthiness beliefs play an essential role in AI integration as they influence how willing employees are to adopt and utilize AI technologies. If employees trust the technology and the organization's commitment to supporting their career growth alongside AI, they are more likely to embrace and effectively use AI solutions [52].

5 Discussion

The Role of AI in Cyber. As expected, the interviews highlighted the potential of AI in cybersecurity. These observations align with the technological optimism outlined in models such as TAM and UTAUT, which highlight performance expectancy and perceived usefulness as key drivers of adoption, both of which are reflected in our interview data.

Several factors underscore the necessity of AI in cybersecurity. The rapid advancement of digital technologies has changed the nature of cyber threats, creating new challenges for organizations around the world. Traditional cybersecurity measures, once considered sufficient, now struggle to contend with the evolving sophistication and frequency of modern cyberattacks [40].

AI can address these challenges due to its capability to process large amounts of data and identify anomalies, allowing it to facilitate fast threat detection – a crucial feature for minimizing damage and mitigating risks. This capability is particularly important in detecting advanced threats such as ransomware, zero-day exploits, and insider attacks, which require more proactive and adaptive defense mechanisms. Moreover, AI can enhance decision-making by providing security analysts with insights and recommendations based on data-driven analysis. This support is crucial in a field where timely and informed decisions can prevent extensive damage. AI systems can prioritize threats, suggest remediation actions, and predict potential outcomes, thereby aiding analysts in making more effective decisions quickly. While our study hypothesized the benefits of AI for cybersecurity based on expert opinions, the literature provides concrete evidence that leveraging AI tools can significantly strengthen organizations' cybersecurity posture (e.g., see Darktrace's most recent report [30]).

Barriers for Human-Centric AI Implementation. Issues such as data quality, the scarcity of labeled data, and the complexity of developing suitable AI models are well-documented in the literature as well. Hence, it was expected that these obstacles would also present significant barriers within the context of the Organization's cybersecurity operations. However, anticipating these challenges does not diminish their significance. Interviews with data scientists highlighted persistent problems related to data quality, noting that many data structures are designed at the back-end of applications solely to capture data, often without ensuring its quality.

Implementing effective data governance strategies is a foundational requirement for transitioning to an AI-driven cybersecurity department. This supports our application of the STS framework, which emphasizes that technical improvements, such as better data management, must be accompanied by changes to social systems, including clearer stakeholder roles and a shared understanding of data responsibilities. These strategies should not only include robust data quality monitoring and auditing processes but also emphasize the importance of educating stakeholders about the critical role of data in analytics and their responsibilities within this framework.

The finding that AI will serve as a complementary tool rather than a replacement for human analysts is also consistent with the prevailing view in the field. Considering the high-stakes environment of cybersecurity, it was anticipated that participants in this study would hold similar opinions. Despite AI's significant advancements, it remains susceptible to errors due to technological limitations and the complexity of its (sometimes black-box) operations. Consequently, human analysts play an indispensable role in mitigating these flaws. This highlights the need for future initiatives to focus on optimizing the collaboration between humans and AI within this context, rather than concentrating solely on advancing the technology itself. We underscore that even highly accurate AI systems may face adoption challenges due to complex human factors. The promise of AI cannot be fully realized if solutions are not designed with human properties and needs in mind. This aligns with the analysis of our interviews, where human-centric concerns repeatedly surfaced even when discussing technically advanced systems, confirming the IRT perspective that resistance may stem from rational psychological or functional barriers rather than simple reluctance.

Our case study highlights the pivotal role of security analysts in AI development for cybersecurity. These professionals not only identify use cases based on their daily challenges but also assist data scientists in understanding data, developing model features, and testing the resulting AI models. Security experts act as both users and domain experts, making their involvement crucial for the successful adoption of AI in cybersecurity. Ignoring their role and the human factors associated with it can impede the effective integration and utilization of AI tools. Therefore, AI solutions for cybersecurity must be engineered with a clear focus on human-centered design, ensuring that the needs and expertise of security analysts are integral to the development process. This approach can enhance the utilization of AI tools and foster trust and acceptance among the workforce, ultimately leading to more effective cybersecurity practices. These findings validate the assumptions of sociotechnical system theory, particularly the joint optimization of human and technical systems. In this case, involving analysts directly in design processes helped bridge the gap between AI capabilities and real-world usability.

Organizational Culture and AI Adoption. Research consistently underscores the significant influence of organizational culture and senior management support on the adoption of AI technologies. In line with this, our case study confirmed the critical role of these factors. For example, resistance to change can be a significant barrier to the adoption of new technologies, especially in established organizations where existing practices and mindsets are deeply ingrained. In this context, strategic alignment from leadership is essential to overcoming these challenges. This finding reinforces the notion that the journey toward AI integration is not solely a technological challenge but also a matter of effectively managing people and organizational dynamics.

While some level of anxiety regarding job security is anticipated with the introduction of AI technologies, the extent to which these concerns and career path ambiguities emerged as significant barriers was surprising. This indicates a deeper level of fear and resistance among security analysts than initially expected, emphasizing the importance of addressing these concerns proactively to mitigate resistance and foster acceptance of AI within the workforce. The prominence of these concerns reflects barriers consistent with IRT, particularly image and tradition barriers, where fears about role erosion or reputational impact inhibit willingness to engage with AI tools. This means that there is a need for organizations to prioritize people-centric strategies alongside technical solutions. This reinforces the organizational insights derived from the TOE framework, which considers internal culture and leadership as essential environmental and organizational components influencing innovation uptake.

In Section 4.3, we proposed career mentorship as a key strategy. To develop these career mentorship programs, organizations should begin by conducting needs assessments to identify employees' skills, aspirations, and concerns related to AI integration. These insights will allow mentorship programs to be customized, providing targeted guidance and support that addresses the identified needs. Additionally, implementing pilot programs can help gather iterative feedback and refine and adapt the mentorship initiatives to ensure alignment with both organizational objectives and individual goals.

The pronounced lack of trust in AI among security analysts, who often view AI as a "black box", was more significant than anticipated. This distrust implies that the perceived opacity of AI systems is a critical barrier to their adoption in cybersecurity. Interestingly, compared to similar studies (see Section 2.2), our research uniquely identified this issue as a significant challenge in the broader adoption of AI in cybersecurity. As detailed in our root cause analysis, the lack of trust in AI among security analysts is caused by several problems, including the analysts' limited understanding of the output of models and AI in general.

Explainable AI. Machine learning models, as a result of their operation, typically only display the final result of a calculation. Therefore, to understand how the model arrives at this output, security analysts require some form of explanation. Despite the availability of well-established methods such as SHAP and LIME, we were somewhat surprised to find that security analysts tend to prefer more straightforward and often simplified explanations of AI model decisions over more comprehensive technical ones. This finding emphasizes that while tools like SHAP are valuable for explainable AI (XAI), their complexity can be a barrier for cybersecurity professionals. Analysts benefit more from simplified, clear explanations that meet their practical needs, highlighting the need for user-centered AI explanation development and testing. Supporting our findings, Suh et al. [94] reported that XAI methods such as SHAP and LIME, are often perceived as confusing by cybersecurity analysts. A potential solution could be using Large Language Models (LLMs) to translate SHAP outputs into user-friendly narratives. Ali and Kostakos [5] proposed an intrusion detection system using an LLM-based conversational agent, though without user testing to confirm its acceptance. Meanwhile, Zytek et al. [106] demonstrated that LLM-generated natural language explanations consistently outperformed SHAP plots in a user study, suggesting promise in this approach. At the same time, we should admit that LLMs can generate inaccurate explanations ("hallucinations"), emphasizing the importance of their output through evaluation to ensure the reliability. This need for transparent and user-friendly explanations underscores TAM's principle of perceived ease of use, while also pointing to gaps in perceived behavioral control as described in TPB – analysts must feel empowered to interpret and act on AI outputs. This further highlights the need for effective human-AI collaborative solutions.

While it is relatively intuitive that transparency and explanations are essential for building trust, it was less expected that a lack of understanding of AI would also have a significant effect on trust. One argument for why the lack of knowledge of how AI functions significantly affects security analysts' trust in AI lies in their work and the historical context of cybersecurity practices. Unlike many technologies used in daily life without complete understanding, cybersecurity has traditionally relied on rule-based models, which are inherently more straightforward to comprehend. These models operate on clear, predefined rules that analysts can understand, modify, and predict, providing a sense of control and transparency. The shift to AI-driven approaches introduces a level of complexity and opacity that contrasts with the simplicity of rule-based systems. Furthermore, the high-risk nature of cybersecurity decisions underscores the need for a deeper understanding and trust in the tools used. Security analysts are responsible for protecting sensitive data and critical infrastructure, and their decisions can have significant and far-reaching implications. When analysts rely on AI tools that they do not fully understand, they may feel anxious about the potential consequences of their decisions. This anxiety may stem from the fear of unintended outcomes or missing critical threats due to the "black box" nature of AI, where the decision-making process is not transparent.

Feedback Mechanisms. Unexpectedly, existing feedback mechanisms were found to be often ineffective, with analysts hesitant to provide honest feedback in formal settings. This suggests that current processes for collecting and integrating user feedback are not effectively fostering open communication, indicating a need for adjustments. We highlight that these findings emerged through our reflexive thematic analysis, which allowed us to surface nuanced interpersonal, cultural, and organizational dynamics that may be overlooked in more quantitative approaches.

We observe in our case that cultural differences play a significant role in how feedback is perceived and delivered [66]. To overcome these communication barriers within multicultural teams, one potential approach is to organize broader workshops focused on cross-cultural communication and feedback.

While the importance of continuous learning and innovation is well-known, the significant positive impact of events like hackathons on collaboration and practical, hands-on courses was particularly notable. These events promote creative problem-solving and help demystify AI, making it more approachable and understandable for analysts. Encouraging such innovative activities can enhance practical engagement with AI technologies.

6 Limitations

Organizational Context. This research is based on a single-case study of a large European organization, providing in-depth insights into AI adoption within a specific organizational and cultural context. While this focus allowed for a detailed exploration of relevant dynamics, the findings may not be fully generalizable to other organizations, industries, or geographic regions. The specific organizational culture, structure, and practices of the studied company can influence the results, meaning that different organizations might face unique challenges and opportunities not captured in this study.

To increase the generalizability of our findings, future research should consider conducting multiple case studies across a variety of settings and organizational contexts, such as comparing cybersecurity practices in small, medium and large enterprises, or examining different industries like finance, healthcare, and manufacturing.

This would help determine whether the findings hold true across different types of organizations and sectors. However, the alignment of our results with those from comparable studies (see Section 2.2) suggests that the insights gained are likely to be generalizable beyond this immediate context.

Participant Selection and Response Bias. A possible limitation of this study is the selection of participants, which may have introduced response bias. Response bias arises when the chosen participants do not accurately represent the broader sample population or the entire industry [82]. In this case, the cybersecurity analysts who were included in the study were predominantly those who were already involved in AI projects or had some level of engagement with AI technologies. This selection potentially skews the findings, as these analysts may hold more favorable views toward AI, given their direct experience and involvement in its implementation. To mitigate this limitation, the interview questions were designed to encourage participants to reflect not only on their own experiences but also on those of their colleagues. This approach aimed to capture a broader range of perspectives, helping to balance the potential bias introduced by the participants' direct involvement with AI.

Sample Size. The study involved 15 participants, which is considered sufficient for qualitative analysis [42] and larger than in some other studies (see Section 2.2). While this sample size allowed for a sufficient examination of key themes, a larger and more diverse sample could provide additional perspectives and capture further nuances.

Qualitative Methodology. This research primarily employed qualitative methods, which are well-suited for exploring complex, context-specific phenomena like AI adoption in cybersecurity. While qualitative approaches provide rich, detailed data, they can also be influenced by biases such as interviewer influence and participant subjectivity.

Incorporating mixed-method approaches in future studies could strengthen the evidence base by complementing qualitative insights with quantitative data. This combination would validate the findings through numerical data and help capture more diverse perspectives by allowing researchers to explore patterns on a larger scale. For example, while qualitative interviews might uncover nuanced experiences and opinions, quantitative surveys could reach a broader audience, revealing trends and variations across different demographics or organizational roles.

Time Constraints. The six-month duration of this study provided a snapshot of AI adoption at a particular moment in time. While this was sufficient to capture key dynamics, it may not fully reflect the long-term impacts and evolving nature of AI integration in cybersecurity. Longitudinal studies could build on this research by offering deeper insights into how AI adoption influences organizations over time.

Long-term studies that follow the integration of AI into cybersecurity over several years would provide valuable insights into how trust and cross-team collaboration evolve, which can be used to design governance frameworks that ensure AI tools are implemented responsibly, with clear guidelines for transparency, accountability, and ethical use. These studies could also track the career development of security analysts as AI tools become more prevalent, providing a clearer picture of the impact on job roles and career trajectories. Understanding the long-term impact of AI on job roles and collaboration can inform the design of targeted training programs that address specific skill gaps and prepare security analysts for the evolving demands of their roles.

Longitudinal insights can provide empirical evidence on how the interaction between social and technical factors evolves. These can contribute to the refinement of the proposed sociotechnical conceptual model and the development of generalized frameworks that could be applied to other sectors facing similar AI integration challenges.

7 Conclusion

We examined the adoption of AI-based tools in cybersecurity within a large international financial organization using reflexive thematic analysis informed by a Sociotechnical Systems (STS) framework. This approach allowed us to jointly assess technical, human, and organizational dimensions as interdependent elements of a complex ecosystem. By evaluating our findings against established technology adoption models – TAM, TPB, TOE, UTAUT, and IRT – we identified several intertwined barriers: lack of trust in AI, limited domain knowledge among data scientists, ineffective feedback loops, organizational inertia, and job security concerns. Our results demonstrate that technically sound models alone cannot succeed without alignment across social and organizational subsystems, which collectively determine the success or failure of AI adoption.

Building on these insights, we proposed targeted interventions such as foundational AI education, mentorship initiatives, transparent model explanations, feedback system redesign, and strategic alignment between leadership and practitioners. These interventions emphasize the importance of developing AI with cybersecurity professionals rather than merely for them, reinforcing the value of participatory and human-centered design. Future research should expand our work by focusing on multi-case studies across industries and organizational scales, complemented by mixed-method approaches that integrate qualitative and quantitative evidence. Such studies would help validate and refine these findings, advancing the development of AI systems that are not only effective but also trusted, adaptive, and resilient within the evolving cybersecurity landscape.

References

[1] Daron Acemoglu and Pascual Restrepo. 2018. Artificial Intelligence, Automation, and Work. In *The economics of artificial intelligence: An agenda*. University of Chicago Press, 197–236. doi:10.3386/w24196

[2] Icek Ajzen. 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 50, 2 (1991), 179–211. doi:10.1016/0749-5978(91)90020-T

[3] Zeynep Akata, Dan Balliet, Maarten de Rijke, Frank Dignum, Virginia Dignum, Guszti Eiben, Antske Fokkens, Davide Grossi, Koen Hindriks, Holger Hoos, Hayley Hung, Catholijn Jonker, Christof Monz, Mark Neerincx, Frans Oliehoek, Henry Prakken, Stefan Schlobach, Linda van der Gaag, Frank van Harmelen, Herke van Hoof, Birna van Riemsdijk, Aimee van Wynsberghe, Rineke Verbrugge, Bart Verheij, Piek Vossen, and Max Welling. 2020. A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect With Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence. *Computer* 53, 8 (2020), 18–28. doi:10.1109/MC.2020.2996587

[4] Khalifa AL-Dosari, Noora Fetais, and Murat Kucukvar. 2024. Artificial Intelligence and Cyber Defense System for Banking Industry: A Qualitative Study of AI Applications and Challenges. *Cybernetics and Systems* 55, 2 (2024), 302–330. doi:10.1080/01969722.2022.2112539

[5] Tarek Ali and Panos Kostakos. 2023. HuntGPT: Integrating Machine Learning-Based Anomaly Detection and Explainable AI with Large Language Models (LLMs). arXiv:2309.16021 [cs.CR]

[6] Azza Alomary and John Woollard. 2015. How is technology accepted by users? A review of technology acceptance models and theories. In *International Conference on 4E*. https://eprints.soton.ac.uk/382037/

[7] Hamza Alshenqeeti. 2014. Interviewing as a Data Collection Method: A Critical Review. *English Linguistics Research* 3, 1 (2014), 39–45. doi:10.5430/elr.v3n1p39

[8] Kasun Amarasinghe, Kit T. Rodolfa, Sérgio Jesus, Valerie Chen, Vladimir Balayan, Pedro Saleiro, Pedro Bizarro, Ameet Talwalkar, and Rayid Ghani. 2024. On the Importance of Application-Grounded Experimental Design for Evaluating Explainable ML Methods. *Proceedings of the AAAI Conference on Artificial Intelligence* 38, 19 (March 2024), 20921–20929. doi:10.1609/aaai.v38i19.30082

[9] Bjorn Andersen and Tom Fagerhaug. 2006. *Root Cause Analysis*. Quality Press.

[10] Meraj Farheen Ansari, Bibhu Dash, Pawankumar Sharma, and Nikhitha Yathiraju. 2022. The Impact and Limitations of Artificial Intelligence in Cybersecurity: A Literature Review. *International Journal of Advanced Research in Computer and Communication Engineering* 11, 9 (2022), 81–90. doi:10.17148/ijarcce.2022.11912

[11] Bokolo Anthony Jnr. 2023. Examining the adoption of telehealth during public health emergencies based on technology organization environment framework. *Journal of Science and Technology Policy Management* 15, 6 (April 2023), 1311—-1338. doi:10.1108/jstpm-05-2022-0079

[12] ATLAS.ti Scientific Software Development GmbH. 2024. ATLAS.ti. Computer software. https://atlasti.com Version 24.0.0, qualitative data analysis software.

[13] Steve Olusegun Bada. 2015. Constructivism Learning Theory: A Paradigm for Teaching and Learning. *IOSR Journal of Research & Method in Education* 5, 6 (2015), 66–70. https://api.semanticscholar.org/CorpusID:37780480

[14] Richard P. Bagozzi. 2007. The Legacy of the Technology Acceptance Model and a Proposal for a Paradigm Shift. *Journal of the Association for Information Systems* 8, 4 (2007), 244–254. doi:10.17705/1jais.00122

[15] Jeff Baker. 2012. The Technology–Organization–Environment Framework. *Information Systems Theory: Explaining and Predicting Our Digital Society* 1 (2012), 231–245. doi:10.1007/978-1-4419-6108-2_12

[16] Mohan Baruwal Chhetri, Shahroz Tariq, Ronal Singh, Fatemeh Jalalvand, Cecile Paris, and Surya Nepal. 2024. Towards Human-AI Teaming to Mitigate Alert Fatigue in Security Operations Centres. *ACM Transactions on Internet Technology* 24, 3, Article 12 (2024), 22 pages. doi:10.1145/3670009

[17] Andreea Bendovschi. 2015. Cyber-Attacks – Trends, Patterns and Security Countermeasures. *Procedia Economics and Finance* 28 (2015), 24–31. doi:10.1016/S2212-5671(15)01077-1

[18] Robert P. Bostrom and J. Stephen Heinen. 1977. MIS Problems and Failures: A Socio-Technical Perspective. Part I: The Causes. *Management Information Systems Quarterly* 1, 3 (Sept. 1977), 17–32. doi:10.2307/248710

[19] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. doi:10.1080/2159676X.2019.1628806

[20] Bybit. 2025. Bybit Security Incident: Timeline of Events and FAQs. Online: https://learn.bybit.com/this-week-in-bybit/bybit-security-incident-timeline/.

[21] Rachel L. Campagna, Alexandra A. Mislin, Kurt T. Dirks, and Hillary Anger Elfenbein. 2022. The (Mostly) Robust Influence of Initial Trustworthiness Beliefs on Subsequent Behaviors and Perceptions. *Human Relations* 75, 7 (2022), 1383–1411. doi:10.1177/00187267211002905

[22] Leong Chan, Ian Morgan, Hayden Simon, Fares Alshabanat, Devin Ober, James Gentry, David Min, and Renzhi Cao. 2019. Survey of AI in Cybersecurity for Information Technology Management. In *2019 IEEE Technology & Engineering Management Conference (TEMSCON)*. 1–8. doi:10.1109/TEMSCON.2019.8813605

[23] Arijit Chaudhuri and Tasos C. Christofides. 2013. *Indirect Questioning in Sample Surveys.* Springer Science & Business Media. doi:10.1007/978-3-642-36276-7

[24] Shih-Chih Chen, Li Shing-Han, and Li Chien-Yi. 2011. Recent Related Research in Technology Acceptance Model: A Literature Review. *Australian Journal of Business and Management Research* 1, 9 (Dec. 2011), 124–127. doi:10.52283/NSWRCA.AJBMR.20110109A14

[25] Kuo-Ming Chu. 2023. A consumer innovation resistance theory perspective on the advanced driver assistance systems. *Economic Research-Ekonomska Istraživanja* 36, 3 (2023), 2153716. doi:10.1080/1331677X.2022.2153716

[26] Cisco. [n. d.]. Artificial Intelligence in Security. https://www.cisco.com/c/en/us/products/security/artificial-intelligence-ai.html. Accessed: 2024-07-14.

[27] Victoria Clarke and Virginia Braun. 2017. Thematic analysis. *The Journal of Positive Psychology* 12, 3 (2017), 297–298. doi:10.1080/17439760.2016.1262613

[28] Haydee M. Cuevas, Stephen M. Fiore, Barrett S. Caldwell, and Laura Strater. 2007. Augmenting Team Cognition in Human-Automation Teams Performing in Complex Operational Environments. *Aviation, Space, and Environmental Medicine* 78, 5: Supplement (May 2007), B63–B70.

[29] Darktrace. [n. d.]. Cyber AI Glossary - Incident Response. https://darktrace.com/cyber-ai-glossary/incident-response. Accessed: 2024-07-14.

[30] Darktrace. 2024. State of AI Cyber Security 2024. https://www.darktrace.com/resources/state-of-ai-cyber-security-2024-executive-summary

[31] Fred D. Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *Management Information Systems Quarterly* 13, 3 (1989), 319–340. doi:10.2307/249008

[32] Thijs van Ede, Hojjat Aghakhani, Noah Spahn, Riccardo Bortolameotti, Marco Cova, Andrea Continella, Maarten van Steen, Andreas Peter, Christopher Kruegel, and Giovanni Vigna. 2022. DEEPCASE: Semi-Supervised Contextual Analysis of Security Events. In *2022 IEEE Symposium on Security and Privacy (SP)*. 522–539. doi:10.1109/SP46214.2022.9833671

[33] Amy C. Edmondson. 2018. *The Fearless Organization: Creating Psychological Safety in the Workplace for Learning, Innovation, and Growth.* John Wiley & Sons.

[34] Mica R. Endsley. 2017. From Here to Autonomy: Lessons Learned From Human–Automation Research. *Human Factors* 59, 1 (2017), 5–27. doi:10.1177/0018720816681350

[35] Babajide Tolulope Familoni. 2024. Cybersecurity Challenges In The Age Of AI: Theoretical Approaches And Practical Solutions. *Computer Science & IT Research Journal* 5, 3 (Mar. 2024), 703–724. doi:10.51594/csitrj.v5i3.930

[36] Martin Fishbein and Icek Ajzen. 1977. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. *Philosophy and Rhetoric* 10, 2 (1977), 130–132.

[37] Gaston Godin and Gerjo Kok. 1996. The Theory of Planned Behavior: A Review of its Applications to Health-Related Behaviors. *American Journal of Health Promotion* 11, 2 (Nov. 1996), 87–98. doi:10.4278/0890-1171-11.2.87

[38] Steven R. Gomez, Vincent Mancuso, and Diane Staheli. 2019. Considerations for Human-Machine Teaming in Cybersecurity. In *Augmented Cognition.* 153–168. doi:10.1007/978-3-030-22419-6_12

[39] Hari Gonaygunta. 2023. *Factors Influencing the Adoption of Machine Learning Algorithms to Detect Cyber Threats in the Banking Industry.* PhD dissertation. University of the Cumberlands.

[40] Shankha Shubhra Goswami, Surajit Mondal, Rohit Halder, Jibangshu Nayak, and Arnabi Sil. 2024. Exploring the Impact of Artificial Intelligence Integration on Cybersecurity: A Comprehensive Analysis. *Journal of Industrial Intelligence* 2, 2 (2024), 73–93. doi:10.56578/

jii020202
[41] James Gusman. 2023. *A Qualitative Study on the Deployment of Artificial Intelligence and Machine Learning within Cybersecurity for Intelligent Decision Making*. PhD dissertation. Capella University.
[42] Monique Hennink and Bonnie N. Kaiser. 2022. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine* 292 (2022), 114523. doi:10.1016/j.socscimed.2021.114523
[43] Rania Hodhod, Harlie Hardage, Safia Abbas, and Eman Abdullah Aldakheel. 2023. CyberHero: An Adaptive Serious Game to Promote Cybersecurity Awareness. *Electronics* 12, 17 (2023), 3544. doi:10.3390/electronics12173544
[44] Hyperproof Team. 2024. Understanding the Change Healthcare Breach and Its Impact on Security Compliance. Online: https://hyperproof.io/resource/understanding-the-change-healthcare-breach/.
[45] Mohammad Hossein Jarrahi. 2018. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons* 61, 4 (2018), 577–586. doi:10.1016/j.bushor.2018.03.007
[46] Jeeyoon Jeong, Byung-Jik Kim, and Julak Lee. 2024. Navigating AI Transitions: How Coaching Leadership Buffers Against Job Stress and Protects Employee Physical Health. *Frontiers in Public Health* 12 (2024). doi:10.3389/fpubh.2024.1343932
[47] Siqi Jiang, Yuyin Tang, and Jung Chieh Lee. 2022. *A Preliminary Study Exploring the Effects of Artificial Intelligence on Fintech Innovation Resistance*. Atlantis Press, 923–927. doi:10.2991/978-94-6463-036-7_136
[48] Jan Jöhnk, Malte Weißert, and Katrin Wyrtki. 2020. Ready or Not, AI Comes – An Interview Study of Organizational AI Readiness Factors. *Business & Information Systems Engineering* 63, 1 (Dec. 2020), 5–20. doi:10.1007/s12599-020-00676-7
[49] Karl M. Kapp. 2012. *The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education* (1st ed.). Pfeiffer & Company.
[50] Steve Kerr and Glenn Rifkin. 2008. *Reward Systems: Does Yours Measure Up?* Harvard Business Press.
[51] Sunnie S. Y. Kim, Elizabeth Anne Watkins, Olga Russakovsky, Ruth Fong, and Andrés Monroy-Hernández. 2023. "Help Me Help the AI": Understanding How Explainability Can Support Human-AI Interaction. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 250. doi:10.1145/3544548.3581001
[52] Haiyan Kong, Zihan Yin, Yehuda Baruch, and Yue Yuan. 2023. The Impact of Trust in AI on Career Sustainability: The Role of Employee-AI Collaboration and Protean Career Orientation. *Journal of Vocational Behavior* 146 (2023), 103928. doi:10.1016/j.jvb.2023.103928
[53] Karel Kuzmiak. 2017. Blue Team: the Game. https://tulipslab.org/projects/1617/CharlesFirewallGameWebsite/Website/v0.7/index.html Accessed: 2025-10-09.
[54] Vivian Lai, Chacha Chen, Q. Vera Liao, Alison Smith-Renner, and Chenhao Tan. 2021. Towards a Science of Human-AI Decision Making: A Survey of Empirical Studies. arXiv:2112.11471 [cs.AI]
[55] Annette Lareau. 2021. *Listening to People: A Practical Guide to Interviewing, Participant Observation, Data Analysis, and Writing It All Up*. University of Chicago Press.
[56] Jehyun Lee, Farren Tang, Phyo May Thet, Desmond Yeoh, Mitch Rybczynski, and Dinil Mon Divakaran. 2022. SIERRA: Ranking Anomalous Activities in Enterprise Networks. In *2022 IEEE 7th European Symposium on Security and Privacy (EuroS&P)*. IEEE Computer Society, Los Alamitos, CA, USA, 44–59. doi:10.1109/EuroSP53844.2022.00011
[57] Paul Legris, John Ingham, and Pierre Collerette. 2003. Why do people use information technology? A critical review of the technology acceptance model. *Information & management* 40, 3 (2003), 191–204. doi:10.1016/S0378-7206(01)00143-4
[58] Lai-Ying Leong, Teck-Soon Hew, Keng-Boon Ooi, and June Wei. 2020. Predicting mobile wallet resistance: A two-staged structural equation modeling-artificial neural network approach. *International Journal of Information Management* 51 (2020), 102047. doi:10.1016/j.ijinfomgt.2019.102047
[59] Jian-hua Li. 2018. Cyber security meets artificial intelligence: a survey. *Frontiers of Information Technology & Electronic Engineering* 19, 12 (2018), 1462–1474. doi:10.1631/fitee.1800573
[60] Q. Vera Liao, Daniel Gruen, and Sarah Miller. 2020. Questioning the AI: Informing Design Practices for Explainable AI User Experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, 1–15. doi:10.1145/3313831.3376590
[61] Q. Vera Liao, Milena Pribić, Jaesik Han, Sarah Miller, and Daby Sow. 2021. Question-Driven Design Process for Explainable AI User Experiences. arXiv:2104.03483 [cs.HC]
[62] Chia-Lun Lo. 2023. The Attitudes Toward Artificial Intelligent Applications: An Investigation Based on Innovation Resistance Theory. In *Proceedings of the 53rd Annual Meeting of the Western Decision Sciences Institute (WDSI)*. Western Decision Sciences Institute.
[63] Hasan Mahmud, A. K. M. Najmul Islam, Ranjan Kumar Mitra, and Ahmed Rizvan Hasan. 2022. The Impact of Functional and Psychological Barriers on Algorithm Aversion – An IRT Perspective. In *The Role of Digital Technologies in Shaping the Post-Pandemic World*, Savvas Papagiannidis, Eleftherios Alamanos, Suraksha Gupta, Yogesh K. Dwivedi, Matti Mäntymäki, and Ilias O. Pappas (Eds.). Springer International Publishing, 95–108. doi:10.1007/978-3-031-15342-6_8
[64] Stacey L. Malek, Shikhar Sarin, and Christophe Haon. 2020. Extrinsic Rewards, Intrinsic Motivation, and New Product Development Performance. *Journal of Product Innovation Management* 37, 6 (Nov. 2020), 528–551. doi:10.1111/jpim.12554

[65] Yogesh Malhotra and Dennis Galletta. 1999. Extending the technology acceptance model to account for social influence: theoretical bases and empirical validation. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*, Vol. Track1. doi:10.1109/HICSS.1999.772658
[66] Erin Meyer. 2014. *The culture map: Breaking through the invisible boundaries of global business*. Hachette Book Group USA.
[67] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*. Association for Computing Machinery, 220–229. doi:10.1145/3287560.3287596
[68] Nachaat Mohamed. 2023. Current trends in AI and ML for cybersecurity: A state-of-the-art survey. *Cogent Engineering* 10, 2 (2023), 2272358. doi:10.1080/23311916.2023.2272358
[69] Dietmar P. F. Möller. 2023. *Cybersecurity in Digital Transformation*. Springer, 1–70. doi:10.1007/978-3-031-26845-8_1
[70] Scott Mongeau and Andrzej Hajdasinski. 2021. *Cybersecurity Data Science: Best Practices in an Emerging Profession*. Springer International Publishing. doi:10.1007/978-3-030-74896-8
[71] Yupeng Mou, Yixuan Gong, and Zhihua Ding. 2024. Complement or substitute? A study of the impact of artificial intelligence on consumers' resistance. *Marketing Intelligence & Planning* 42, 4 (March 2024), 647–665. doi:10.1108/mip-04-2023-0187
[72] Nadia Nahar, Shurui Zhou, Grace Lewis, and Christian Kästner. 2022. Collaboration challenges in building ML-enabled systems: communication, documentation, engineering, and process. In *Proceedings of the 44th International Conference on Software Engineering (ICSE '22)*. Association for Computing Machinery, 413—-425. doi:10.1145/3510003.3510209
[73] Anton J. Nederhof. 1985. Methods of coping with social desirability bias: A review. *European Journal of Social Psychology* 15, 3 (July 1985), 263–280. doi:10.1002/ejsp.2420150303
[74] Anastasija Nikiforova, Antoine Clarinval, Anneke Zuiderwijk, Daniel Rudmark, Petar Milic, and Katrin Rajamäe-Soosaar. 2024. Innovation Resistance Theory in Action: Unveiling Barriers to Open Government Data Adoption by Public Organizations to Unlock Open Data Innovation. arXiv:2407.10883 [cs.CY]
[75] Anastasija Nikiforova and Anneke Zuiderwijk. 2022. Barriers to Openly Sharing Government Data: Towards an Open Data-adapted Innovation Resistance Theory. In *International Conference on Theory and Practice of Electronic Governance (ICEGOV '22)*. Association for Computing Machinery, 215–220. doi:10.1145/3560107.3560143
[76] NOVA LABS. n.d.. Cybersecurity Lab. https://www.pbs.org/wgbh/nova/labs/lab/cyber/ Accessed: 2025-10-09.
[77] Randolph Nwaiwu and Meret Keeris. 2024. The AI Cyber Security Challenge. https://kpmg.com/nl/en/home/insights/2024/06/ai-cyber-security-challenge.html Accessed: 2024-08-11.
[78] Tiago Oliveira and Maria Fraga Martins. 2011. Literature Review of Information Technology Adoption Models at Firm Level. *Electronic Journal of Information Systems Evaluation* 14, 1 (2011), 110–121.
[79] Rudolph Oosthuizen and Leon Pretorius. 2014. Modelling of Command and Control Agility. In *Proceedings of the International Command and Control Research and Technology Symposium (ICCRTS)*. Alexandria, VA, US.
[80] Saheed Femi Osholake, Chinemelum Umealajekwu, Anthony Edohen, Abiola Olusola Majekodunmi, and Uchenna Evans-Anoruo. 2024. Human-AI Collaborative Security Operations: Optimizing SOC analyst cognitive load through augmented intelligence frameworks. *Iconic Research And Engineering Journals* 8, 6 (2024), 1102–1115.
[81] Raja Parasuraman, Thomas B. Sheridan, and Christopher D. Wickens. 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans* 30, 3 (2000), 286–297. doi:10.1109/3468.844354
[82] Jeongeun Park and Minhye Park. 2016. Qualitative Versus Quantitative Research Methods: Discovery or Justification? *Journal of Marketing Thought* 3, 1 (2016), 1–8.
[83] James C. Paterson. 2023. *Beyond the Five Whys: Root Cause Analysis and Systems Thinking*. Wiley, United Kingdom.
[84] Michael Quinn Patton. 2014. *Qualitative Research & Evaluation Methods: Integrating theory and practice*. Sage publications.
[85] Mlungisi Radebe, Pitso Tsibolane, and Mike Hart. 2022. Perceptions of AI Tools for Cybersecurity in Large Enterprises. In *The EWG-DSS International Conference on Decision Support System Technology*. European Working Group on Decision Support Systems (EWG-DSS).
[86] Sundaresan Ram and Jagdish N. Sheth. 1989. Consumer Resistance to Innovations: The Marketing Problem and its Solutions. *Journal of Consumer Marketing* 6, 2 (Feb. 1989), 5–14. doi:10.1108/eum0000000002542
[87] Nikon Rasumov. 2019. Stop the Bots: Practical Lessons in Machine Learning. https://blog.cloudflare.com/stop-the-bots-practical-lessons-in-machine-learning/. Accessed: 2024-12-30.
[88] Tom-Martijn Roelofs, Eduardo Barbaro, Svetlana Pekarskikh, Katarzyna Orzechowska, Marta Kwapień, Jakub Tyrlik, Dinu Smadu, Michel Van Eeten, and Yury Zhauniarovich. 2024. Finding Harmony in the Noise: Blending Security Alerts for Attack Detection. In *ACM/SIGAPP Symposium on Applied Computing*. 1385–1394. doi:10.1145/3605098.3635981
[89] Saqib Saeed, Salha A. Altamimi, Norah A. Alkayyal, Ebtisam Alshehri, and Dina A. Alabbad. 2023. Digital Transformation and Cybersecurity Challenges for Businesses Resilience: Issues and Recommendations. *Sensors* 23, 15 (2023). doi:10.3390/s23156666
[90] Sujeet Kumar Sharma and Jyoti Kumar Chandel. 2013. Technology acceptance model for the use of learning through websites among students. *International Arab Journal of E-Technology* 3, 1 (2013), 44–49.

[91] Yun Shen, Enrico Mariconti, Pierre Antoine Vervier, and Gianluca Stringhini. 2018. Tiresias: Predicting Security Events Through Deep Learning. In *ACM SIGSAC Conference on Computer and Communications Security (CCS '18)*. Association for Computing Machinery, 592–605. doi:10.1145/3243734.3243811

[92] Ronnie H. Shroff, Christopher C. Deneen, and Eugenia M. W. Ng. 2011. Analysis of the technology acceptance model in examining students' behavioural intention to use an e-portfolio system. *Australasian Journal of Educational Technology* 27, 4 (Aug. 2011). doi:10.14742/ajet.940

[93] Adewale Daniel Sontan and Segun Victor Samuel. 2024. The intersection of Artificial Intelligence and cybersecurity: Challenges and opportunities. *World Journal of Advanced Research and Reviews* 21, 2 (Feb. 2024), 1720–1736. doi:10.30574/wjarr.2024.21.2.0607

[94] Ashley Suh, Harry Li, Caitlin Kenney, Kenneth Alperin, and Steven R. Gomez. 2024. More Questions than Answers? Lessons from Integrating Explainable AI into a Cyber-AI Tool. arXiv:2408.04746 ACM CHI 2024 Workshop on Human-Centered Explainable AI (HCXAI).

[95] Shalini Talwar, Manish Talwar, Puneet Kaur, and Amandeep Dhir. 2020. Consumers' Resistance to Digital Innovations: A Systematic Review and Framework Development. *Australasian Marketing Journal* 28, 4 (2020), 286–299. doi:10.1016/j.ausmj.2020.06.014

[96] Shahroz Tariq, Mohan Baruwal Chhetri, Surya Nepal, and Cecile Paris. 2025. A2C: A modular multi-stage collaborative decision framework for human–AI teams. *Expert Systems with Applications* 282 (2025), 127318. doi:10.1016/j.eswa.2025.127318

[97] Shirley Taylor and Peter A. Todd. 1995. Understanding Information Technology Usage: A Test of Competing Models. *Information Systems Research* 6, 2 (1995), 144–176. doi:10.1287/isre.6.2.144

[98] Louis G. Tornatzky, Mitchell Fleischer, and Alok K. Chakrabarti. 1990. *The Processes of Technological Innovation*. Lexington Books, Lexington, Mass.

[99] Eric L. Trist. 1981. *The evolution of socio-technical systems*. Vol. 2. Ontario Quality of Working Life Centre Toronto.

[100] Jasper van der Waa, Elisabeth Nieuwburg, Anita Cremers, and Mark Neerincx. 2021. Evaluating XAI: A Comparison of Rule-Based and Example-Based Explanations. *Artificial Intelligence* 291 (2021). doi:10.1016/j.artint.2020.103404

[101] Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. 2003. User Acceptance of Information Technology: Toward a Unified View. *Management Information Systems Quarterly* 27, 3 (2003), 425–478. doi:10.2307/30036540

[102] Guy H. Walker, Neville A. Stanton, Paul M. Salmon, and Daniel P. Jenkins. 2008. A review of sociotechnical systems theory: a classic concept for new command and control paradigms. *Theoretical Issues in Ergonomics Science* 9, 6 (2008), 479–499. doi:10.1080/14639220701635470

[103] Yingying Yang, Peng Lu, Yuanyuan Niu, and Guohong Yuan. 2024. Research on Unmanned Smart Hotels Resistance from the Perspective of Innovation Resistance Theory. *Sage Open* 14, 3 (2024). doi:10.1177/21582440241281570

[104] Robert K. Yin. 2009. *Case study research: Design and methods* (4 ed.). Sage.

[105] Angie Zhang and Min Kyung Lee. 2025. Knowledge Workers' Perspectives on AI Training for Responsible AI Use. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 1207. doi:10.1145/3706598.3714100

[106] Alexandra Zytek, Sara Pidò, and Kalyan Veeramachaneni. 2024. LLMs for XAI: Future Directions for Explaining Explanations. arXiv:2405.06064 [cs.AI]

Received 24 April 2025; revised 10 October 2025; accepted 30 December 2025