Potential of AI for User-Centric Cybersecurity in the Financial Sector

Muriel Frank
SnT, University of Luxembourg
muriel.frank@uni.lu

Martin Brennecke
SnT, University of Luxembourg
martin.brennecke@uni.lu

Pol Hölzmer
SnT, University of Luxembourg
pol.hoelzmer@uni.lu

Nadia Pocher
SnT, University of Luxembourg
nadia.pocher@uni.lu

Gilbert Fridgen
SnT, University of Luxembourg
gilbert.fridgen@uni.lu

Abstract

The use of cybersecurity tools powered by artificial intelligence (AI) continues to gain traction in the financial services industry. On the one hand, they can strengthen an organization's technical cybersecurity posture. On the other hand, even if cybercriminals also leverage AI to exploit human weaknesses, there are early indications that AI can help equip the workforce against evolving threats. Based on a structured literature review (SLR) and a Delphi study, this article identifies the most promising end-user-focused use cases in which AI can assist financial institutions in combating cybersecurity threats and gearing their workforce up to thwart cyberattacks. For information security executives and researchers alike, this study provides a first set of general directions on which AI-powered and user-centric tools and solutions to focus on in the near future.

Keywords: Artificial intelligence, behavioral cybersecurity, delphi study, awareness, financial services

1. Introduction

The financial services sector has fully embraced the digital transformation and increasingly relies on emerging technologies in nearly all of its everyday activities (Oladipo et al., 2024). Given the industry's position as a prime target for cybercriminals, this dependency creates new challenges (Fares et al., 2023). In fact, the entire financial sector is facing more sophisticated attacks on a daily basis (Javaheri et al., 2024). Recent figures from 2023 indicate that the financial services industry suffers from the highest number of data breaches now at 27%, followed by healthcare at 20% (White, 2023). To make matters worse, cybercriminals increasingly use artificial intelligence (AI) tools to enhance their nefarious activities (Darem et al., 2023; Guembe et al., 2022; Sen et al., 2022; U.S. Treasury, 2024).

Frequently, such breaches result from a lack of awareness or careless behaviors of their own employees (Javaheri et al., 2024; Krombholz, 2015), posing an existential threat to organizations and their business processes (Guembe et al., 2022). With AI attacks on the rise, traditional technical and non-technical cybersecurity measures are increasingly inadequate, which is why organizations need to equip themselves and their cybersecurity teams accordingly (Guembe et al., 2022). In the literature, we find first confirmation that leveraging AI can benefit organizations by making their security systems more resilient, streamlining threat detection, and supporting professionals in the fight against cybercrime (e.g., Kumar et al., 2023; Taddeo et al., 2019).

Yet, the deployment of AI in financial institutions is still in its infancy (Atkins et al., 2024). One reason is that there is a gap between recognizing the potential of AI and lacking the strategic vision of how to utilize it to their advantage (Fares et al., 2023). Especially as individuals play a critical role in ensuring cybersecurity (Oladipo et al., 2024), a greater focus on the human factor is needed, and with it, a better understanding of how to support their security behaviors with AI. Therefore, our work examines user perspectives on AI-powered cybersecurity practices and explores the following research question:

What are the most promising AI-powered user-centric cybersecurity measures in financial institutions?

Answering our research question involves two steps. First, we use a structured literature review (SLR) to gain an overview of used and proposed user-centric approaches to combat cybersecurity threats (Moher et al., 2009), meaning they target the average employee who does not work in information security and is therefore more vulnerable to cybersecurity criminals. The SLR is based on peer-reviewed literature published in various disciplines, including computer science, information systems, and educational studies. It thus captures

URI: https://hdl.handle.net/10125/109389
978-0-9981331-8-8

cross-disciplinary knowledge of how AI is being used to empower organizations and their workforce to fight cybercrime. Second, we build on a three-round online Delphi study with information technology (IT) and cybersecurity experts working in the financial sector (Brancheau et al., 1996; Okoli & Pawlowski, 2004). The Delphi method is well-suited for dealing with uncertainty (Paliwoda, 1983) inherent in the use of AI (Ransbotham et al., 2017) and thus allows us to explore the potential of 13 AI-based, user-centric cybersecurity approaches for use in the financial sector. As a result, we present a ranking of the most promising AI-powered cybersecurity measures, and thereby, we offer practitioners and academia a more holistic understanding of these measures. This work seems especially important as cybercriminals specifically try to deceive average workers in the financial sector, while at the same time, there are relatively few studies that focus on the human factor (Javaheri et al., 2024).

The remainder is divided into four sections: The background offers an in-depth view of AI and cyber risks in the financial sector (Section 2). Next, Section 3 describes the SLR and its results, which informs our Delphi study described in Section 4. Section 5 discusses our results in light of previous findings and in terms of their practical implications. It also concludes the work.

2. Background

2.1. Financial industry cybersecurity threats

The complex technological environment that facilitates the provision of financial services has long made the industry a prime target for cybercriminals seeking to obtain critical information or disrupt essential services (e.g., Dhashanamoorthi, 2021; Rohmeyer & Bayuk, 2019). In recent years, however, the number and sophistication of threats have further increased (Javaheri et al., 2024; Kumar et al., 2023), with banks and other providers regularly facing ransomware, distributed denial-of-service (DDoS) attacks and espionage (Gulyás & Kiss, 2023; Javaheri et al., 2024). This trend is alarming for these organizations, as cyber incidents affect them as a major operational risk with serious impacts (Rohmeyer & Bayuk, 2019), ranging from financial loss to reputation damages, customer frustration, and litigation (Dhashanamoorthi, 2021; Tariq, 2018). Inherently, they threaten the functioning of the financial system and, depending on their scale, may even contribute to triggering a financial crisis (Gulyás & Kiss, 2023).

Also, in the financial industry, cybersecurity risks exist both in the organization's technology and processes and in the behavior of its personnel (Rohmeyer & Bayuk, 2019). This has led practitioners and researchers alike to closely examine this multifaceted threat landscape. While earlier research investigated technology and data-driven threats (e.g., Gai et al., 2018), the focus is now shifting to user-centric threats. According to Javaheri et al. (2024), financial services face three classes of threats. The first includes threats that arise from technological issues (i.e., misconfigurations or system flaws), such as ransomware attacks (e.g., Keshavarzi & Ghaffary, 2020). The second includes threats caused by human error or misuse, such as insider threats or social engineering (e.g., Krombholz, 2015). The last category is procedure-related, resulting from procedural errors or improper policy implementation.

As technology advances, so do cybercriminals. Today, they increasingly rely on smarter and more autonomous AI techniques, resulting in threats that outwit their predecessors (Guembe et al., 2022; Javaheri et al., 2024). As an example, more and more cybercriminals are finding ways to evade traditional detection systems and to disguise their own actions (Guembe et al., 2022). Looking at the most common AI-driven or AI-layered attacks, malicious actors are using AI to exploit various vulnerabilities, such as password guessing, using intelligent self-learning malware or intelligent target profiling (for a more detailed overview see Guembe et al. (2022)).

2.2. Artificial intelligence and cybersecurity

Generally speaking, AI aims to create intelligent computer programs to solve complex real-world problems (McCarthy, 2004). It has the ability to imitate and outperform the cognitive functions of human beings, i.e., by processing and analyzing vast amounts of data (Kumar et al., 2023). More recently, its latest manifestation, generative AI, has taken off, hyped by practitioners and academics worldwide (Chen et al., 2023; Hilario et al., 2024). Notably, AI is not limited to a specific set of methods (McCarthy, 2004). Rather, it is constrained by the jagged boundaries of technological progress (Dell'Acqua et al., 2023) or in the words of Berente et al. (2021) by the "continually evolving frontier of emerging computing capabilities" (p. 1433).

It is only in the last few years, and therefore comparatively late, that AI has made its way to the cybersecurity field. While security professionals and executives once believed that AI could not help protect organization assets (Chan et al., 2019), they now recognize the transformative power of these technologies to reshape security strategies (Kumar et al., 2023) and change the rules of the game (Michael et al., 2023). In particular, AI-enhanced solutions allow organizations to mitigate cybersecurity threats (Zhang et al., 2022) and increase

the efficiency of both cybersecurity measures (Kaur et al., 2023) and security professionals (Zacharis & Patsakis, 2023). For example, AI-driven customized training programs can tailor content to individual user needs and behaviors, ensuring that training is personalized and effective (Jawhar et al., 2024; Kallonas et al., 2024; Trifonov et al., 2020). Machine learning can continuously improve training content based on user performance, emerging threats, and simulation of cyberattacks, making the training process dynamic and up-to-date (Sen et al., 2022; Zeadally et al., 2020). Generative AI, such as ChatGPT, can help guide employees in learning and improving their security knowledge and skills (Kallonas et al., 2024; Nguyen et al., 2024). Zhang et al. (2022) provide an detailed overview of how AI is being used for cybersecurity. The authors identify four categories in which AI is leveraged, namely abnormal traffic detection, dangerous behavior monitoring, network situation awareness, and user authentication.

More recently, Kaur et al. (2023) systematically mapped AI-enhanced use cases to the respective 5 functions and 23 categories of the well-known cybersecurity framework developed by the National Institute of Standards and Technology (NIST). However, based on their taxonomy, only a small share of identified cybersecurity solutions are user-centric (e.g., adaptive security awareness training), while the majority of solutions are technical in nature. Considering the substantial body of literature suggesting not to underestimate the human element when devising cybersecurity measures (e.g., Crossler et al., 2013), as well as when addressing AI solutions (Kumar et al., 2023), such a predominant focus on the technical side increasingly appears shortsighted. Following the call of Kaur et al. (2023) to focus more on human-AI interaction, this paper takes a user-centered approach to AI-powered cybersecurity solutions and reviews the academic literature to identify cybersecurity measures that are enhanced by AI.

3. Structured literature review

3.1. Procedure

Our research process started with an SLR, as it provides researchers with a solid understanding of the existing literature by collecting relevant studies (Levy & Ellis, 2006). According to Moher et al. (2009), the SLRs begins with an identification phase, followed by a screening phase, before assessing a paper's eligibility and making an inclusion decision to ensure a rigorous and reproducible process.

We started with a keyword search in abstracts and titles (Levy & Ellis, 2006) to identify relevant papers. To do this, the search string had to be as broad, yet as precise, as possible. For this reason, we did not limit the search to cybersecurity measures in the financial sector only, as human influence on cybersecurity is ubiquitous. For a comprehensive yet effective view of how AI can augment user-centric cybersecurity efforts, we used the following search string: *(AI OR artificial intelligence) AND [IT, information, computer, cyber]{1} security) AND (awareness training OR [human, user]{1} [behavio[u]{0,1}r, attitude, engagement, perception]{1})*.

We sourced information from libraries and well-established databases, including *ACM Digital Library*, *AIS eLibrary*, *arXiv.org*, *IEEE Xplore*, *Springer*, *Emerald* and *Elsevier ScienceDirect*. We further enriched the results using AI-enhanced databases, such as Elicit and Dimensions, and by searching Google Scholar. Inclusion criteria were peer-reviewed manuscripts written in English and published between January 2018 and August 2024. We did this to ensure relevance and timeliness. In addition, we removed articles from mega publishers, articles with abstracts only, inaccessible articles, articles with a strong technical rather than behavioral or user focus and articles that were neither empirical nor literature reviews. Since our search term was quite broad, we limited the results per database and library to the top 200.

Initially, we identified a total of 2638 records, 1502 from databases and 1136 from libraries (see Figure 1). After duplicate removal, screening, and applying inclusion and exclusion criteria, the SLR yielded 19 quality records with a majority of empirical papers (more than 80%). The SLR process was meticulously documented to ensure reproducibility, with all steps and decisions recorded for transparency. For data synthesis, we detailed the processes for determining study eligibility, preparing data for presentation, and displaying results. This thorough review laid the foundation for a list of user-centric AI-supported cybersecurity measures for the subsequent Delphi study to refine and validate our findings through expert consensus. Then, using inductive reasoning, we consolidated the articles into broader dimensions, as elaborated in the next section.

3.2. Results

The literature on user-centric AI-powered cybersecurity measures or solutions can be divided into two predominant streams: 1) proactive solutions and 2) reactive solutions. The first stream of research includes studies that apply cybersecurity methods before an incident occurs to prepare or train users and systems in advance, while the second stream focuses on studies that apply cybersecurity measures in response to incidents

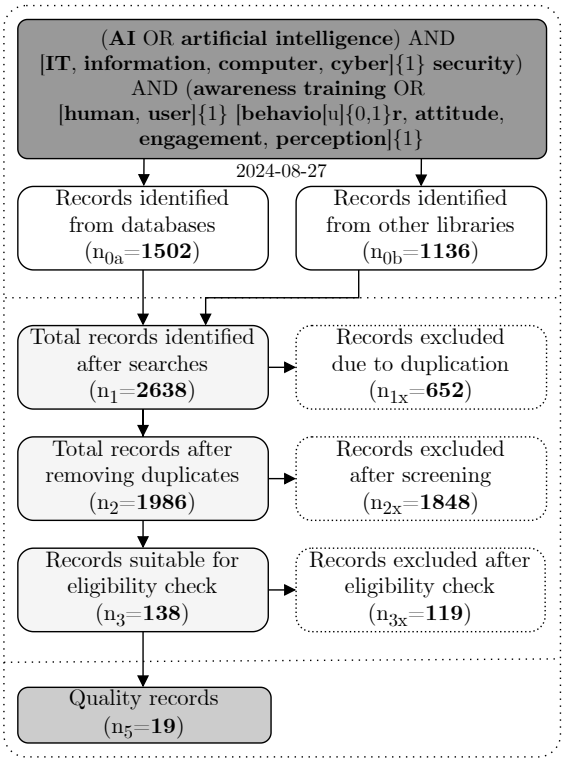

Figure 1. Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) reporting flowchart following Moher et al. (2009).

to mitigate the impact and learn from them (Jawhar et al., 2024). Additionally, these streams can be further subdivided depending on whether or not end users are involved (Espinha Gasiba et al., 2021) (see Tables 1). The following subsections describe various AI-powered user-centric cybersecurity measures, organized by the categories introduced. Notably, these categories are not mutually exclusive and highlight different aspects of how and when user-centric AI-powered solutions can improve cybersecurity.

Proactive user-centric measures. One of the most popular proactive cybersecurity measures that involves people is Security Education Training and Awareness (SETA) programs (Dash & Ansari, 2022). These programs are designed to mitigate the risk of user-caused security incidents. Today, AI-based approaches are increasingly used to analyze large amounts of (real-time) data and feed their findings into the design of innovative awareness programs or security education strategies (Abu-Amara et al., 2021; Tan et al., 2020). Usually, they can automatically take into account user behavior and needs, as well as personal characteristics such as phishing susceptibility, thereby improving the effectiveness of awareness and anti-phishing training (Espinha Gasiba et al., 2021; Zhang et al., 2022). They can also help personalize the learning experience by automatically identifying knowledge gaps and providing targeted interventions to address specific weaknesses, improving overall training outcomes, and ensuring users are well prepared for evolving challenges (Al-Mashhour & Alhogail, 2023; Dash & Ansari, 2022; Kallonas et al., 2024). This may include using AI to simulate real-world attack scenarios to train employees and correct misbehavior (Abu-Amara et al., 2021; Barletta et al., 2023; Espinha Gasiba et al., 2020).

In addition, AI can also help keep security training up to date. It does this by continuously analyzing emerging threats and organizational data, to adjust training based on those insights. As a result, users can be trained on the latest cybersecurity threats and risks, thanks to an AI-based analysis of the threat landscape (Zhang et al., 2022). This makes the training more effective and engaging (Barletta et al., 2023; Trifonov et al., 2020; Zhang et al., 2022).

Literature also sees potential for AI-based gamified awareness training, as it can help automatically analyze player behavior, provide personalized feedback, and dynamically adjust challenges based on user performance (Abu-Amara et al., 2021; Barletta et al., 2023). This contributes to user engagement and to improving learning outcomes (Espinha Gasiba et al., 2021; Jawhar et al., 2024). Finally, another area with potential for AI-based approaches is the design of frameworks and SETA programs based on international standards and best practices (Jawhar et al., 2024; Trifonov et al., 2020). This ensures ongoing cybersecurity resilience and preparedness (Oladipo et al., 2024; Trifonov et al., 2020).

Reactive user-centric measures. The second stream of research focuses on behavioral cybersecurity analytics and monitoring to identify high-risk groups and mitigate insider threats (e.g., Ansari, 2022; Barone IV et al., 2023; Koutsouvelis et al., 2020). For example, the use of AI could improve security measures by continuously analyzing complex behavioral data, identifying subtle patterns that are not apparent to human analysts but may indicate security risks, and providing real-time alerts and targeted interventions (Espinha Gasiba et al., 2021; Jawhar et al., 2024; Zhang et al., 2022). This enables more effective detection and prevention of repetitive misbehavior that increases security vulnerabilities (Espinha Gasiba et al., 2021; Tan et al., 2020). More specifically, AI-based approaches can help identify at-risk groups based on misbehavior for targeted interventions (Ansari, 2022; Nguyen et al., 2024) or to predict phishing susceptibility (Al-Mashhour & Alhogail, 2023; Sharif et al., 2018; Zhang et al., 2022).

	Manual	Automated
Proactive	**Simulation of Cyberattacks**: AI simulates cyberattacks (e.g., phishing attacks) to train users and correct misbehaviors (Abu-Amara et al., 2021; Ansari, 2022; Barletta et al., 2023; Espinha Gasiba et al., 2020). **Adaptive Security Training**: AI helps to customize training modules by analyzing user interactions and performance (Al-Mashhour & Alhogail, 2023; Dash & Ansari, 2022; Kallonas et al., 2024; Kaur et al., 2023; Tan et al., 2020; Trifonov et al., 2020). **Awareness Reminders**: AI helps provide regular security-related updates and refreshers on best practices (El Hajal et al., 2021; Oladipo et al., 2024). **SETA Program Design**: AI helps develop comprehensive cybersecurity education programs based on assessments of international standards and best practices (Jawhar et al., 2024; Trifonov et al., 2020).	**Content Updates for Training Systems**: AI helps update cybersecurity training content based on real-time data sources, incidents, and threat landscape monitoring (Jawhar et al., 2024; Kaur et al., 2023; Tan et al., 2020). **Intelligent Coaching and Challenge Assessment**: AI helps integrate virtual coaches into e.g. security training platforms and automates the scoring of challenges to provide instant feedback and adjust difficulty levels accordingly (Abu-Amara et al., 2021; Dash & Ansari, 2022; El Hajal et al., 2021; Espinha Gasiba et al., 2020, 2021; Nguyen et al., 2024). **Refinement of Awareness Training**: AI analyzes user behavior to refine training content and reduce susceptibility to deception (e.g., phishing click rates) (Barletta et al., 2023; Kaur et al., 2023)
Reactive	**Insider Threat Detection**: AI detects anomalies and suspicious behavioral patterns, identifying and visualizing local threats, and alerts the user (Barone IV et al., 2023; Koutsouvelis et al., 2020; Zhang et al., 2022).	**Targeted Recommendations**: AI analyzes user behavior to identify at-risk groups and makes recommendations for targeted training interventions (Ansari, 2022; Nguyen et al., 2024). **Prediction of Susceptibility to Deception**: AI predicts users' susceptibility to cyber threats based on behavioral factors (Al-Mashhour & Alhogail, 2023; Sharif et al., 2018; Zhang et al., 2022).

Table 1. AI-powered user-centric cybersecurity measures.

4. Delphi study

4.1. Methodology

Following previous studies (e.g., Brancheau et al., 1996; Dhillon et al., 2021), this article reports on a three-round online Delphi survey of IT or information security professional working in the financial services industry and regularly using AI. The use of the Delphi method in the field of information systems (IS) to obtain controlled expert opinions on complex problems is on the rise (Dhillon et al., 2021; Okoli & Pawlowski, 2004). It is an iterative feedback technique that allows researchers to 1) predict or identify key issues and prioritize problems for management action or 2) develop conceptual frameworks or theories (Okoli & Pawlowski, 2004; Schmidt, 1997). In all cases, it is an aid to the structured evaluation of the relative importance of the problems under investigation (Delbecq et al., 1975; Okoli & Pawlowski, 2004) as in subsequent rounds, experts have the opportunity to reevaluate their opinions based on the previous rounds' pooled responses (Brancheau et al., 1996).

This technique dates back to the 1950s when the RAND Corporation conducted a series of linked questionnaires to reach consensus among experts (Dalkey & Helmer, 1963). The typical procedure is as follows: As a first step, participants are either asked to provide a list of items and a rationale for listing them (Schmidt, 1997) or are provided with a list of questions (Brancheau et al., 1996). This list usually contains between six (e.g., Schmidt, 1997) and 21 items (e.g., Brancheau et al., 1996). Their responses are collected, summarized, and played back to the participants to get their feedback on the reordered list (Schmidt, 1997). Unlike other survey tools, Delphi studies do not require a large panel size. However, there is no consensus regarding the most adequate sample size, which typically ranges from ten to 30 participants (Paré et al., 2013). For our study, participants were recruited through Prolific, a platform with a pool of online workers that can be used for robust scientific research (Palan & Schitter, 2018), and through the researchers' networks. Selection criteria included being currently employed, working in IT or IT security or a related field in the financial sector, as well as regularly

working with AI (at least once a week). This resulted in 67 eligible participants on Prolific. 15 additional participants were invited through our network. All of them were assured of the anonymity of their responses.

4.2. Procedure & results

In the first round, we asked participants to rank the ten AI-powered cybersecurity measures identified through the SLR ranging from 1 (= highest priority) to 10 (= lowest priority). We presented measures in random order to avoid bias. We also invited respondents to provide further input on how AI can enhance user-centric cybersecurity measures, which four IS researchers then independently reviewed and categorized. Overall, respondents (n=26) contributed a total of 40 additional suggestions for cybersecurity measures in the financial sector. However, the majority were either already covered (n=20), unclear (n=2) or out of scope (n=14), meaning that the suggestions made were not applicable to user-centric interventions. Finally, we identified those that were novel, not redundant and applicable (n=3). This set included measures such as AI-powered self-monitoring, where AI supports users in analyzing their behavior and in making safe(er) decisions. In the second round, participants from the previous Delphi study received the ranked list as well as the new proposals to provide their input on the updated ranking. They came up with 44 new suggestions for user-centric cybersecurity measures. However, half of them were already in use and the rest were out of scope, which is why we did not raise another request for new ideas. In the third and final round, the participants once again had the opportunity to revise and re-order their rankings. We collected 22 responses, representing 27% of those invited to participate. The majority of participants were male (87.5%), all full-time IT professionals, and located primarily in the United Kingdom (n=12) and the United States (n=6). They ranged in age from 23 to 59. We present the final ranking and mean rankings for the individual rounds in Table 2.

5. Discussion & conclusion

Given the evolving threat landscape driven by AI, this work follows a two-step approach to identify the most promising AI-powered cybersecurity measures that empower the average financial sector employee to act more securely. Using an SLR-approach following the PRISMA-method, we identified ten AI-powered cybersecurity measures. Building on these measures (later expanded to include suggestions from participants), we conducted a three-round Delphi study in which we asked IT and/or cybersecurity experts in the financial sector to rank the different measures based on their (expected) importance in the financial sector. To provide further insight and explore their potential in more detail, each of the top five measures ranked will be discussed below.

Insider threat detection: Recognizing that AI can process massive amounts of data (Zhang et al., 2022), our experts see clear value in AI assisting humans in analyzing user behavior and identifying suspicious behavioral patterns – e.g., based on log files (Koutsouvelis et al., 2020). Throughout our Delphi study, insider threat detection ranked in the top two (second in the first round and first in the following two rounds). This aligns with the financial sector leading all others in terms of insider threat (Aljawarneh & Gupta, 2017; Proofpoint, 2021). Further, when we asked participants for further ideas on how AI can enhance user-centric cybersecurity tools in the financial services sector, suggestions in this category consistently ranked first.

Simulation of real-world attack scenarios: In response to the increase in attack vectors, organizations are beginning to change their approach to training (Abu-Amara et al., 2021; Barletta et al., 2023). Our experts see great potential in such approaches in the financial sector, where AI helps simulate real-life scenarios or phishing attacks, as they can help identify gaps (Sen et al., 2022) but also better tailor future training to users' awareness levels (Espinha Gasiba et al., 2020). They continuously suggested related measures in the idea-collection section – e.g. simulated phishing emails, mock threats, phone calls, etc. What must be considered, however, is that phishing simulations, for example, are subject to legal restrictions (Sutter et al., 2022).

Targeted training recommendations: Previous work has demonstrated how AI can improve the understanding of behavioral cybersecurity measures (Ansari, 2022; Kaur et al., 2023). Not surprisingly, the experts also see potential in analyzing behavioral patterns to identify misbehavior and exploring the factors that predict at-risk groups, which in turn can be used to target them with specific security training later on.

Approaches to predict phishing susceptibility: Phishing susceptibility is a perennial problem plaguing organizations and researchers alike. In 2023, almost 30% of phishing attacks worldwide targeted the financial industry (Statista, 2024). As a result, they spend a great deal of effort analyzing the factors that increase or decrease the likelihood that individuals will fall for phishing (Moody et al., 2017). This is also reflected in the expert opinions of the Delphi study, which ranked the use of machine learning algorithms to help predict phishing susceptibility in the top five. This is made possible by their ability to process vast amounts of data (Zeadally et al., 2020).

AI-powered cybersecurity measure	Strategy	User involvement	R1	R2	R3
Insider threat detection	Reactive	Yes	3.88	3.10	2.81
Simulation of real-world attack scenarios	Proactive	Yes	3.12	3.11	3.60
Targeted training recommendations	Reactive	No	6.27	5.00	5.00
Approaches to predict phishing susceptibility	Reactive	No	4.38	5.30	5.21
Self-monitoring	Proactive	Yes	n.a.	6.00	5.39
Awareness reminders	Proactive	Yes	6.65	6.94	5.75
SETA program design	Proactive	Yes	5.42	5.12	5.75
Adaptive security training	Proactive	Yes	5.85	5.33	5.81
Incident response strategies	Reactive	Yes	n.a.	5.76	5.88
Intelligent coaching and challenge assessment	Proactive	No	6.77	5.71	6.73
Post-incident education	Reactive	Yes	n.a.	6.92	6.88
Refinement of awareness training	Proactive	No	6.42	6.50	7.00
Content updates for training systems	Proactive	No	6.23	7.29	7.47

Table 2. Ranking of AI-powered cybersecurity measures with means for all rounds of the Delphi study. In order of final ranking in R3. Lower means indicate higher perceived potential.

Self-monitoring: Unintentional security misbehavior that allows cybercriminals to infiltrate systems is a major organizational vulnerability (Crossler et al., 2013). This seems to be an important issue for our experts as well, as they suggested in the first round of our Delphi study to add AI-powered tools that allow users to monitor their own behavior and alert them in case of deviation. More specifically, experts wish for an AI assistant that gives them security advice on the go. They envision the assistant helping them browse the web more safely or scanning their email and warning them if it contains a suspicious link. It could also warn them if they download risky applications.

5.1. Contributions & implications

Our main contribution is a thorough forecast of the potential of AI for user-centric cybersecurity measures in the financial sector. We present a total of 13 measures, identified through literature review and experts participating in our Delphi study, that point to user-centric measures able to strengthen organizations' security posture (Oladipo et al., 2024). Based on a three-round online Delphi study, we were able to gather solid predictions from experts in the field (Okoli & Pawlowski, 2004) and thereby expanded our understanding of AI in behavioral cybersecurity. In particular, our results suggest that behavioral cybersecurity in the financial sector should be approached using a mix of reactive and proactive measures. Our experts ranked three reactive and two proactive measures in the top five. This split seems reasonable because preventive measures can help raise awareness and equip people with skills to fight cybercrime (El Hajal et al., 2021), while reactive measures can help correct and prevent repeated misbehavior (Al-Mashhour & Alhogail, 2023).

Since limited access to security data is no longer a limitation (Zeadally et al., 2020), our work is the first to consolidate various applications for AI-powered cybersecurity measures that target the average employee. In this sense, it can be used by other researchers to test and compare their effectiveness. It can also spark ideas about how to facilitate the implementation of such user-centric cybersecurity measures, both within the financial sector and beyond. As noted by Fares et al. (2023), many practitioners lack a vision of how to integrate AI to strengthen their security posture. Therefore, our work also has practical value in guiding practitioners in the financial sector who need to make a decision on what might be promising. Specifically, we provide them with a comprehensive set of measures that IT and security experts consider important for securing financial assets in the future.

5.2. Limitations & outlook

This work is not without critics and, as such, should be evaluated in consideration of potential limitations that could indicate paths for future research. Considering the scope of this paper, we want to highlight a total of three

limitations that give rise to future research opportunities. First, the expert consensus remains a controversial criterion; expert opinions differ, and full consensus is frequently unrealistic (Keeney et al., 2006). However, expert opinions tend to stabilize with a sound research methodology and several rounds. Second, it is worth noting that our study does not capture the entirety of AI-powered user-centric cybersecurity measures in the financial services industry. Instead, it is intended to highlight the potential of technologies such as AI to better equip the workforce in the face of evolving security threats. Future work can, thus, explore additional AI-powered user-centric cybersecurity measures, possibly in other sectors targeted by cybercriminals, such as healthcare, where the cost of data breaches has increased by 53% since 2020 (IBM, 2023). In addition, it would be worthwhile to examine whether the application of these measures would be feasible from a legal and business point of view. Third, given the novelty of (particularly generative) AI in a personal and business context, everyone, including our experts, still has limited experience using AI. Even though we made sure that all of them use AI regularly, the potential of AI needs to be explored further. One promising avenue is to conduct in-depth interviews on the subject matter. Another is to use an idiographic approach, such as Cram et al. (2024), to accompany the implementation of AI-powered cybersecurity measures in the financial services industry.

Acknowledgments

This work was funded by Luxembourg's FNR and PayPal, PEARL grant ref. 13342933/Gilbert Fridgen, and grant ref. NCER22/IS/16570468/NCER-FT (CryptoReg), and supported by Banque et Caisse d'Épargne de l'État (Spuerkeess). For open access purposes, the authors have applied a CC BY 4.0 license to any Author Accepted Manuscript arising from this submission.

References

Abu-Amara, F., Almansoori, R., Alharbi, S., Alharbi, M., & Alshehhi, A. (2021). A novel SETA-based gamification framework to raise cybersecurity awareness. *International Journal of Information Technology*, *13*(6), 2371–2380.

Aljawarneh, S. A., & Gupta, M. (Eds.). (2017). *Online banking security measures and data protection*. IGI Global.

Al-Mashhour, A., & Alhogail, A. (2023). Machine-learning-based user behavior classification for improving security awareness provision. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *14*(8).

Ansari, M. F. (2022). A Quantitative Study of Risk Scores and the Effectiveness of AI-Based Cybersecurity Awareness Training Programs. *International Journal of Smart Sensor and Adhoc Network.*, 1–8.

Atkins, L., Banerjee, S., Boer, M., Craig, L., Greis, J., Hao, G., & Idler, M. (2024). The cyber clock is ticking: Derisking emerging technologies in financial services. Retrieved May 11, 2024, from https://mckinsey.com/capabilities/risk-and-resilience/our-insights/the-cyber-clock-is-ticking-derisking-emerging-technologies-in-financial-services

Barletta, V. S., Calvano, M., Caruso, F., Curci, A., & Piccinno, A. (2023). Serious Games for Cybersecurity: How to Improve Perception and Human Factors. *2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, 1110–1115.

Barone IV, C. R., Mekni, M., & Nassar, M. (2023). Gargoyle Guard: Enhancing cybersecurity with artificial intelligence techniques. *3rd Intelligent Cybersecurity Conference (ICSC)*, 127–132.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS Quarterly*.

Brancheau, J. C., Janz, B. D., & Wetherbe, J. C. (1996). Key issues in information systems management: 1994-95 SIM Delphi results. *MIS Quarterly*, *20*(2).

Chan, L., Morgan, I., Simon, H., Alshabanat, F., Ober, D., Gentry, J., Min, D., & Cao, R. (2019). Survey of AI in cybersecurity for information technology management. *Technology & Engineering Management Conference*.

Chen, B., Wu, Z., & Zhao, R. (2023). From fiction to fact: The growing role of generative AI in business and finance. *Journal of Chinese Economic and Business Studies*, *21*(4), 471–496.

Cram, W. A., D'Arcy, J., & Benlian, A. (2024). Time will tell: The case for an idiographic approach to behavioral cybersecurity research. *MIS Quarterly*, *48*(1), 95–136.

Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M., & Baskerville, R. (2013). Future directions for behavioral information security research. *Computers & Security*, *32*, 90–101.

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*, *9*(3), 458–467.

Darem, A. A., Alhashmi, A. A., Alkhaldi, T. M., Alashjaee, A. M., Alanazi, S. M., & Ebad, S. A. (2023). Cyber threats classifications and countermeasures in banking and financial sector. *IEEE Access*, *11*.

Dash, B., & Ansari, M. F. (2022). An effective cybersecurity awareness training model: First defense of an organizational security strategy. *International Research Journal of Engineering and Technology*.

Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes*. Scott, Foresman & Co.

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality [ssrn:4573321].

Dhashanamoorthi, B. (2021). Artificial intelligence in combating cyber threats in banking and financial services. *International Journal of Science and Research Archive*, *4*(1), 210–216.

Dhillon, G., Smith, K., & Dissanayaka, I. (2021). Information systems security research agenda: Exploring the gap between research and practice. *The Journal of Strategic Information Systems*, *30*(4).

El Hajal, G., Abi Zeid Daou, R., & Ducq, Y. (2021). Human firewall: Cyber awareness using WhatApp AI chatbot. *3rd International Multidisciplinary Conference on Engineering Technology (IMCET)*.

Espinha Gasiba, T., Lechner, U., & Pinto-Albuquerque, M. (2020). Sifu - a cybersecurity awareness platform with challenge assessment and intelligent coach. *Cybersecurity*, *3*(1), 24.

Espinha Gasiba, T., Lechner, U., Pinto-Albuquerque, M., & Porwal, A. (2021). Cybersecurity awareness platform with virtual coach and automated challenge assessment [arXiv:2102.10430].

Fares, O. H., Butt, I., & Lee, S. H. M. (2023). Utilization of artificial intelligence in the banking sector: A systematic literature review. *Journal of Financial Services Marketing*, *28*(4), 835–852.

Gai, K., Qiu, M., & Sun, X. (2018). A survey on FinTech. *Journal of Network and Computer Applications*, *103*, 262–273.

Guembe, B., Azeta, A., Misra, S., Osamor, V. C., Fernandez-Sanz, L., & Pospelova, V. (2022). The emerging threat of AI-driven cyber attacks: A review. *Applied Artificial Intelligence*, *36*(1).

Gulyás, O., & Kiss, G. (2023). Impact of cyber-attacks on the financial institutions. *Procedia Computer Science*, *219*, 84–90.

Hilario, E., Azam, S., Sundaram, J., Imran Mohammed, K., & Shanmugam, B. (2024). Generative AI for pentesting: The good, the bad, the ugly. *International Journal of Information Security*.

IBM. (2023). Cost of a data breach report 2023. Retrieved June 10, 2024, from https://www.ibm.com/reports/data-breach

Javaheri, D., Fahmideh, M., Chizari, H., Lalbakhsh, P., & Hur, J. (2024). Cybersecurity threats in FinTech: A systematic review. *Expert Systems with Applications*, *241*.

Jawhar, S., Miller, J., & Bitar, Z. (2024). AI-driven customized cyber security training and awareness. *3rd International Conference on AI in Cybersecurity*.

Kallonas, C., Piki, A., & Stavrou, E. (2024). Empowering professionals: A generative AI approach to personalized cybersecurity learning. *Global Engineering Education Conference (EDUCON)*.

Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, *97*, 101804.

Keeney, S., Hasson, F., & McKenna, H. (2006). Consulting the oracle: Ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, *53*(2), 205–212.

Keshavarzi, M., & Ghaffary, H. R. (2020). I2CE3: A dedicated and separated attack chain for ransomware offenses as the most infamous cyber extortion. *Computer Science Review*, *36*.

Koutsouvelis, V., Shiaeles, S., Ghita, B., & Bendiab, G. (2020). Detection of insider threats using artificial intelligence and visualisation. *6th Conference on Network Softwarization*, 437–443.

Krombholz, K. (2015). Advanced social engineering attacks. *Journal of Information Security and Applications*, *22*, 113–122.

Kumar, S., Gupta, U., Singh, A. K., & Singh, A. K. (2023). Artificial intelligence: Revolutionizing cyber security in the digital era. *Journal of Computers, Mechanical and Management*, *2*(3), 31–42.

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science: The International Journal of an Emerging Transdiscipline*, *9*, 181–212.

McCarthy, J. (2004). What is artificial intelligence? Retrieved May 9, 2024, from http://jmc.stanford.edu/artificial-intelligence/what-is-ai/

Michael, K., Abbas, R., & Roussos, G. (2023). AI in cybersecurity: The paradox. *IEEE Transactions on Technology and Society*, *4*(2), 104–109.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*.

Moody, G. D., Galletta, D. F., & Dunn, B. K. (2017). Which phish get caught? An exploratory study of individuals' susceptibility to phishing. *European Journal of Information Systems*, *26*(6), 564–584.

Nguyen, Q. H., Wu, T., Nguyen, V., Yuan, X., Xue, J., & Rudolph, C. (2024). Utilizing large language models with human feedback integration for generating dedicated warning for phishing emails. *2nd Workshop on Secure and Trustworthy Deep Learning Systems (SecTL)*.

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information & Management*, *42*(1), 15–29.

Oladipo, J., Okoye, C., Elufioye, O., Falaiye, T., & Nwankwo, E. (2024). Human factors in cybersecurity: Navigating the fintech landscape. *International Journal of Science and Research Archive*.

Palan, S., & Schitter, C. (2018). Prolific.ac - a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.

Paliwoda, S. J. (1983). Predicting the future using Delphi. *Management Decision*, *21*(1), 31–38.

Paré, G., Cameron, A.-F., Poba-Nzaou, P., & Templier, M. (2013). A systematic assessment of rigor in information systems ranking-type Delphi studies. *Information & Management*, *50*(5), 207–217.

Proofpoint. (2021). Managing insider threats in financial services. https://www.proofpoint.com/sites/default/files/e-books/pfpt-uk-eb-managing-insider-threats-in-financial-services.pdf

Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping business with artificial intelligence: Closing the gap between ambition and action. *MIT Sloan Management Review*.

Rohmeyer, P., & Bayuk, J. L. (2019). *Financial cybersecurity risk management: Leadership perspectives and guidance for systems and institutions*. Apress.

Schmidt, R. C. (1997). Managing Delphi surveys using nonparametric statistical techniques. *Decision Sciences*, *28*(3), 763–774.

Sen, R., Heim, G., & Zhu, Q. (2022). Artificial intelligence and machine learning in cybersecurity: Applications, challenges, and opportunities for MIS academics. *Communications of the Association for Information Systems*, *51*(1), 179–209.

Sharif, M., Urakawa, J., Christin, N., Kubota, A., & Yamada, A. (2018). Predicting impending exposure to malicious content from user behavior. *Conference on Computer and Communications Security*.

Statista. (2024). Share of financial phishing attacks worldwide from 2016 to 2023. https://www.statista.com/statistics/1319867/share-of-financial-phishing-attacks/

Sutter, T., Bozkir, A. S., Gehring, B., & Berlich, P. (2022). Avoiding the hook: Influential factors of phishing awareness training on click-rates and a data-driven approach to predict email difficulty perception. *IEEE Access*, *10*, 100540–100565.

Taddeo, M., McCutcheon, T., & Floridi, L. (2019). Trusting artificial intelligence in cybersecurity is a double-edged sword. *Nature Machine Intelligence*.

Tan, Z., Beuran, R., Hasegawa, S., Jiang, W., Zhao, M., & Tan, Y. (2020). Adaptive security awareness training using linked open data datasets. *Education and Information Technologies*, *25*(6), 5235–5259.

Tariq, N. (2018). Impact of cyberattacks on financial institutions. *Journal of Internet Banking and Commerce*, *23*(2).

Trifonov, R., Nakov, O., Manolov, S., Tsochev, G., & Pavlova, G. (2020). Possibilities for improving the quality of cyber security education through application of artificial intelligence methods. *International Conference Automatics and Informatics (ICAI)*.

U.S. Treasury. (2024). Managing artificial intelligence - specific cybersecurity risks in the financial services sector. Retrieved May 11, 2024, from https://home.treasury.gov/system/files/136/Managing-Artificial-Intelligence-Specific-Cybersecurity-Risks-In-The-Financial-Services-Sector.pdf

White, D. (2023). Data breach outlook: Finance surpasses healthcare as most breached industry in 2023. Retrieved May 11, 2024, from https://www.kroll.com/-/media/kroll-images/pdfs/data-breach-outlook-2024.pdf

Zacharis, A., & Patsakis, C. (2023). AiCEF: An AI-assisted cyber exercise content generation framework using named entity recognition. *International Journal of Information Security*, *22*(5), 1333–1354.

Zeadally, S., Adi, E., Baig, Z., & Khan, I. A. (2020). Harnessing artificial intelligence capabilities to improve cybersecurity. *IEEE Access*, *8*.

Zhang, Z., Ning, H., Shi, F., Farha, F., Xu, Y., Xu, J., Zhang, F., & Choo, K.-K. R. (2022). Artificial intelligence in cyber security: Research advances, challenges, and opportunities. *Artificial Intelligence Review*, *55*(2), 1029–1053.