Developing an Anti-Phishing Large Language Model: A Focus Group Study on Human, Technological, and Legal Factors

George A. Thomopoulos
Department of Electrical and Computer Engineering
University of Patras
gthomop@upatras.gr

Panagiota Kiortsi
Researcher, Faculty of Law
Aristotle University of Thessaloniki
kiortsip@gmail.com

Damianos Dumi Sigalas
Department of Electrical and Computer Engineering
University of Patras
damianos.dumisigalas@ac.upatras.gr

Dimitris Kosmopoulos
Department of Computer Engineering and Informatics
University of Patras
dkosmo@upatras.gr

Ioannis Igglezakis
Professor, Faculty of Law
Aristotle University of Thessaloniki
iingleza@law.auth.gr
https://orcid.org/0000-0002-5967-4232

Christos Fidas
Department of Electrical and Computer Engineering
University of Patras
fidas@upatras.gr

***Abstract* – The phishing problem poses a significant threat in modern information systems, putting both individuals and businesses at risk of financial and professional harm. Owing to social media's rapid development and widespread appeal, deception of this sort is hurting millions of people and growing more dangerous. The current work, as part of the AILA (Artificial Intelligence-driven Framework and Legal Advice Tools for Phishing Prevention and Mitigation in Information Systems) project, aims to specify and validate an AI-driven multifactor (human, technology and legal) anti-phishing data model, with the implementation of focus group studies. The findings assist to provide human, technology, and legislative user model endpoints that will be identified and discussed for explicit and implicit user modeling, which will guide the development of the corresponding AI-driven user modeling and profiling mechanisms. To this end a Large Language System is planned to be employed.**

Keywords - phishing, focus group, artificial intelligence, large language model

I. INTRODUCTION

Phishing cyberattacks is one of the biggest security risks to contemporary information systems, which may harm end users and service providers in various ways, both financially and professionally. Phishing is the practice of passing off a fraudulent communication as authentic to trick, coerce, or influence targets into unintentionally disclosing personal information to unapproved parties. It refers to social engineering techniques that involve phone calls, emails, texts, or social media posts posing as reliable service providers to trick end users into disclosing sensitive information like credit card numbers, passwords, or pin codes. Due to the popularity and quick growth of social media on the internet, these kinds of deceit are becoming an ever-bigger menace and are harming millions of people [2].

Phishing is a threat that allows an attacker to get personal and sensitive information from possible victims by social engineering tactics and/or other means, either technological or non-technological. This data breach can result in financial or other consequences. While electronic phishing emerged over twenty years ago, comparable methods date back to the 1800s at the latest. Since phishing attempts often target a large number of "victims," the attackers frequently favor successful communication strategies with wide exposure [3]. To carry out the assault, they might be transmitted using various ways like email, SMS (smishing), voice (vishing) and social media (Facebook, Twitter etc.). Usually, the goal is to get the user's login credentials or financial information, also known as identity theft and cat-phishing. While cat-phishing primarily relies on imitating an individual to request possible victims to deliver money to the attacker, identity theft engages acquiring private information such as credit card, tax security and social security numbers and other sensitive information like name, address, date of birth, with the intention of directly benefiting the attacker financially. Successful phishing attacks may have disastrous consequences for sensitive domain companies, including banks, schools, and health care, as well as for individual end users. According to current research, phishing assaults continue to be the most popular and straightforward attacks, with more than two million phishing sites identified as of January 2021 [1].

Leading-edge anti-phishing research is exceptionally monolithic and fragmented and does not examine the phishing problem from a multidisciplinary perspective. Current anti-phishing user models face phishing threats either in terms of human, technology or law but not in a multidisciplinary method. According to Varshney et al. [7] several anti-phishing solutions exist that include phishing detection schemes (either real-time or non-real time) and phishing prevention schemes, and many countries have developed new anti-phishing laws to help prevent phishing threats. Furthermore, Artificial Intelligence (AI) technologies have shown substantial benefits in preventing most phishing assaults, with most organizations securing their systems relying on AI-based cyber security. The information provided to AI systems, as well as the advancement of machine learning technologies, have been shown to have a major impact on preventing assaults and promoting cyber security. This is accomplished by real time system learning, which ensures the safety of the user [6].

979-8-3503-6883-3/24/$31.00 ©2024 IEEE

The convergence of human behavior, technology, and legal frameworks in phishing can be structured into a comprehensive anti-phishing user model. This model aims to facilitate the creation of advanced algorithms, enabling organizations to assess their readiness against phishing attacks and to gauge the vulnerability of their employees and stakeholders. Ultimately, this approach promises significant strides in scientific innovation for preventing phishing and tailoring anti-phishing measures to individuals. This will suggest an AI-driven and human, technology, and legal centered user modelling approach for providing a more comprehensive method to identify vulnerable users for phishing attacks within an organization and proceed to mitigation actions if necessary.

This research explores the convergence of generative AI and phishing, employing focus group analysis to define and validate an AI-driven anti-phishing data model that integrates human, technological, and legal factors. The focus groups aim to discuss and pinpoint user model endpoints related to these factors for both explicit and implicit user modeling. The insights obtained will guide the creation of AI-driven user modeling and profiling mechanisms and will be used to validate the user model based on specific metrics for static and dynamic profiles. The studies involve key stakeholders from participating organizations, such as data protection officers, security experts, lawyers, and system administrators.

This paper structure is the following: in Section 2 we present our research motivation and contribution, Section 3 presents focus groups analysis and methodology followed, Section 4 defines the research questions and presents the implementation of the focus groups and the analysis of results, in Section 5 we discuss the results of the research and outline potential future work and finally Section 6 concludes the paper.

II. Research Motivation and Contribution

A. Related Work

Several papers were found from the literature that have utilized focus groups in phishing research.

Lötter, A. and Futcher, L. [8] used a focus group at Nelson Mandela Metropolitan University's School of Information and Technology to verify a framework developed for identifying phishing attacks, as a viable technique for raising awareness of phishing assaults using email client user interfaces. The eight individuals present were urged to offer comments on the framework, whether it was good or constructive criticism.

Button et al. [9] in their research conducted six focus groups organized completely or partially over the internet, with 48 members of online fraud victims. The focus groups were employed to investigate incidents regarding online fraud methods, perceptions concerning the severity of online fraud offenses, the responsibility of the perpetrator and the crucial factors that should exacerbate or alleviate the seriousness of these crimes.

Williams et al. [10] conducted six focus groups with 32 persons, to explore whether susceptibility of employees to spear phishing is affected by additional factors within the working environment. The researchers examined the various factors that might make an individual more susceptible to certain influences, considering both the person's characteristics and the work environment they are in. This was done by gathering insights from employees about their perceptions of susceptibility within their workplace.

Karagiannopoulos et al. [11] in their research conducted 3 focus groups with total of 12 individuals to enhance comprehension of their perceptions and encounters with cybercrime participants aged over 60 took part in the focus groups, sharing their internet usage experiences and encounters with cybercrime.

In Althobaiti et al. [12] research, 32 participants of Human-Computer Interaction (HCI) experts, security experts, and average users, conducted a set of eight focus group sessions. The goal of this study is to improve the design of a new URL feature report, which will help users make informed decisions about whether a URL is trustworthy or potentially malicious.

Misra et al. [13] conducted a focus group to determine decisive components when designing the user interface of the presented software game (named "Phish Phinder"), to empower users with the skills and know-how to effectively combat phishing attacks, offering a comprehensive education that covers both the underlying principles and practical steps to take to avoid falling victim to phishing scams. The purpose of doing this empirical inquiry was to verify that the consumers' preferences were integrated during the design phase to optimize user involvement within the game.

B. Motivation and Contribution

State-of-the-art anti-phishing research is highly fragmented and monolithic and does not address the problem from a multidisciplinary perspective. In this work we utilize focus groups aiming to implement a multidisciplinary approach that will formalize a novel anti-phishing user model that will incorporate human, technology, and legal factors of phishing and study the subject across the three pillars as shown in Fig. 1.

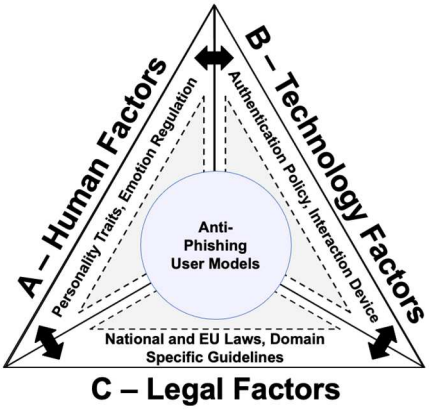

Fig. 1. Anti-phishing multidisciplinary research model

The first pillar refers to the human factor that will gather the static and dynamic metrics of the users while engaged in interaction with the system. The human factor will embrace metrics like age, gender, nationality, personality, emotion level etc.

The second pillar refers to technology factor that will classify the users through machine learning algorithms, based on the applied user authentication policy and their susceptibility to phishing attacks based on the statistical models. The technology factor will embrace metrics like

authentication policy, knowledge-based authentication, interaction device etc.

The third pillar refers to the Legal Factor that will reason about the best-fist recommendation legal advice aiming to mitigate phishing attacks to end-users and service providers. In many countries phishing legislation and laws do not exist and traditional criminal laws of computer crimes and identity theft are used for most of the phishing attacks. Therefore, recommending of specific laws at National or EU level is an important step in the mitigation of phishing attacks, especially considering the dynamic context of cybersecurity legislations in the EU and in Greece, where the legal factor will embrace National and EU laws, domain specific guidelines etc.

This research explores the intersection of generative AI and phishing and utilizes a focus group analysis to specify and validate an AI-driven multifactor (human, technology and legal) anti-phishing data model. The goal of the focus groups is to discuss and identify user model endpoints related to the human, technology and legislative factor for both explicit and implicit user modeling. The insights gathered from these discussions will inform the development of corresponding AI-driven user modeling and profiling mechanisms and will be utilized to validate the user model, according to specific metrics aligned to static and dynamic profiles.

III. Focus Groups Analysis

A. Focus Groups

Focus group is an interview with a group of a pre-defined and specific number of people discussing a certain issue. This research method comprises a focused debate among a limited number of people who share specified qualities, resulting in qualitative data to get insights into the issue of interest. In contrast to traditional group interactions, such as personal meetings or group interviews, which aim to obtain consensus or suggestions, focus groups are organized to examine a variety of perspectives, fetching results from rigorously comparing the data obtained from these groups [15]. A focus group allows for repeated interactions not just between the interviewer and the respondent, but also among all members in the group.

Focus group interviews are used in research to provide insights into respondents' attitudes, feelings, beliefs, experiences, and responses that would be difficult to obtain using alternative methods like questionnaire surveys, one-on-one interviews or observation. Focus groups are especially helpful because of the social gathering and interaction they give, which allows for the disclosure of attitudes, sentiments, and opinions [16]. Furthermore, focus groups are useful when there are fundamental differences among participants whether they are individuals, decision-makers or experts. They are especially useful for investigating certain groups culture and their everyday language, determining the amount of consensus on a given issue.

Focus groups can be used at many times of a study, including the early preparatory stages to define requirements, throughout the study to evaluate or establish a specific program of activities, and after the program has completed to analyze its impact or suggest new paths of inquiry [16]. In terms of scheduling, focus groups are appropriate for introducing a new program or service, providing questions that are difficult to answer in a written survey, or supplementing knowledge gathered from written surveys. Focus groups often generate verbal, open-ended, wide, and qualitative comments. The interaction among participants not only displays their opinions on the world but also discloses the language they use to discuss a problem.

Focus groups are structured around a set of predetermined questions, with several participants adequate to generate rich discussion and ensure active participation without exclusion. While there is no definitive rule, different researchers have suggested varying sizes ranging from 4 to 31 participants [18]. Factors such as group dynamics, participant engagement, and facilitator experience are crucial in determining the optimal size. Larger groups may generate diverse ideas but can be difficult to control and may lead to unequal participation whereas smaller groups encourage active participation but may limit the range of perspectives. Furthermore, researchers encounter a challenge when dealing with a limited pool of participants who are hard to access, while their study design necessitates group discussion on the topic. In such cases, researchers can only gather a small group comprising two to five participants, typically consisting of individuals with advanced expertise in the field [4], which are identified as mini focus groups.

The number of groups can vary from a single meeting to several sessions, with the last session typically lasting 45-60 minutes to 1½-2 hours. Generally, there is a preference for slightly longer sessions, typically between one-and-a-half to three hours, as they allow for more in-depth discussion and information gathering. However, the optimal session length depends on the available time and the need to gather maximum information efficiently [18]. Participants in focus groups must be a representative sample of people whose opinions are important [16]. Ideally, focus groups should be relatively homogeneous to ensure equality of contribution while still allowing for diverse viewpoints. Ultimately, the choice of group size should consider the research objectives and context to facilitate meaningful discussion and generate valuable insights [18]. Participants' homogeneity is critical for maximizing transparency, and they must feel at ease with one another. Furthermore, several groups can be formed to accommodate more people and various opinions.

Every focus group must have a facilitator or moderator who is responsible for providing clear explanations of the group's purpose, making participants feel at ease, facilitating interaction, promoting debate, challenging participants, probing for details, or advancing discussions, and keeping the session focused while avoiding personal opinions. The facilitator's role varies depending on the research approach and objectives. Morgan (1988) [19] suggests that an external facilitator is preferable to avoid researcher bias. Conversely, when focus groups are exploratory, researcher bias is unlikely as their opinions are not fully formed [18]. Nassar-McMillian and Borders (2002) [20] advocate for researcher-led groups, citing the researcher's subject knowledge as beneficial for maintaining focus. Additionally, a co-facilitator may be required to take notes and manage records [16].

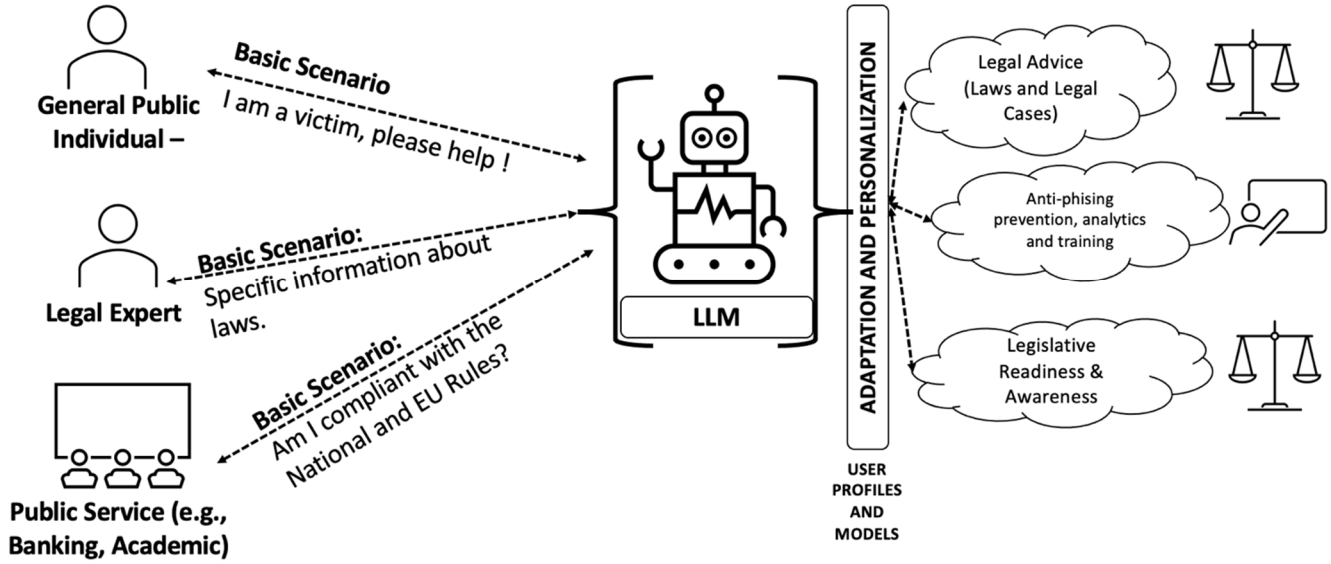

Fig. 2. Focus groups design representation

Sim and Waterfield [17] in their research reffer that focus group research produces different ethical concerns that do not entirely correlate to those posed by one-on-one interviews, and it addresses three important issues: consent, secrecy and anonymity, and risk of damage. In terms of ethics, full information about the purpose and uses of participants' contributions must be provided, any sensitive material with confidentiality must be handled appropriately, confidentiality and anonymization of data from the group must be maintained, and all participants must complete a consent form. Additional actions that should be made to complement and reinforce these concerns closer to the real setting of the focus group include a briefing just prior to the conversation, during the discussion itself, and a debriefing immediately following the focus group.

IV. Methodology and Research Questions

A. Research Questions

Based on the aforementioned research model, we formulated the following research questions related to the focus group research we conducted: *"What are the user model requirements for an AI-driven multidisciplinary (human, technology and legal) anti-phishing data model?"*. By answering this question, we will be able to assist the development of corresponding AI-driven user modeling and profiling mechanisms that will be utilized to validate the user model, according to specific metrics aligned to static and dynamic profiles.

B. Methodology

Background of the study

In the current work, three focus groups were conducted aiming to identify and discuss user model endpoints related to the human factor, technology factor, and legislative factor for both explicit and implicit user modeling. The objective of this task is to identify and confirm the user model by conducting Focus Group studies and the insights gathered from these discussions will be used as input for the development of corresponding AI-driven user modeling and profiling mechanisms. The focus group studies involved key stakeholders from participating organizations, including data protection officers, security experts, lawyers, and system administrators.

The scope of the current study is to implement Focus Groups that will discuss and identify human, technology, and legislative user model endpoints that will allow organizational and individual users to reason about their phishing readiness level and take informed decisions with regards to employees and stakeholders' susceptibility to phishing attacks. Thus, three focus groups were conducted consisting of Individual Users, Legal Experts and Public Services Users, that included participants from general public and representative stakeholders from organizations like data-protection officers, security experts, lawyers and system administrators (Fig. 2).

All three focus groups discussions were conducted online through teleconference. The synchronous online focus groups procedure was preferred versus the in-person focus groups, due to several advantages the first offers. Online focus groups allow researchers to resolve many problems associated with the cost, location, and attendance of busy participants, therefore broadening the availability of possible participants and adding significant flexibility to the interview scheduling process. Internet and teleconference technology has overcome several limitations of in-person focus groups, including the difficulty of reaching and scheduling busy participants, immobility due to financial or physical reasons, unavailability due to different time zones and the difficulty of coordinating to attend at a specific, in-person location. Online groups provide solutions to problems that arise in conventional face-to-face focus groups, with real-time conversations guided by one or more moderators and the participation of a small number of members [21].

a) Study Design and Participants: The process of participant recruitment and group assembly for qualitative academic research is pivotal, influencing the dynamics and quality of interaction within focus groups. A thoughtful approach to sampling and selection is crucial to eliciting diverse opinions, attitudes, and life experiences. Typically, researchers establish a sampling frame, select potential participants, and employ strategies like invitations and

questionnaires to gather demographic information and consent [14]. Balancing group heterogeneity and homogeneity is key; while diversity enriches discussion, too much variation can hinder communication, and excessive homogeneity may limit perspective. Theoretical sampling aids in ensuring a comprehensive range of observations by strategically selecting participants based on pertinent characteristics [16]. Overall, focus groups should encompass diverse backgrounds, views, and experiences to effectively illuminate complex phenomena [5]. Participant identification is paramount, as focus group dynamics play a vital role in generating data, and group composition should align with the research objectives.

The focus group sessions were conducted with 18 total participants to explore perspectives on phishing prevention and mitigation strategies from varied stakeholder groups, as stated in Table I. The first focus group consisted of 8 participants, including practicing attorneys and legal experts from academia. Facilitated by a moderator with expertise in IT law, the discussion centered on the legal dimensions of phishing, with a particular emphasis on capturing the experiences of professionals when supporting victims of phishing attacks. Additionally, participants shared their views on the potential development of an AI system aimed at assisting both legal professionals in their work and their clients.

Subsequently, the second focus group consisted of 5 participants with administrative roles in public organizations, including IT departments of higher education institutions and public health institutions. This session facilitated discussions focused on practical challenges encountered in implementing cybersecurity protocols and strategies to counter potential phishing attacks. The discussion was structured to first capture the opinions of the participants on common phishing techniques encountered in their respective workplaces. Subsequently, the group delved into a discussion on the potential design and usefulness of an AI system aimed at assisting users in recognizing and responding to phishing threats effectively.

The third focus group consisted of 5 participants with diverse backgrounds, potentially including individuals who have been victims of phishing attacks in the past. One participant represented the National Data Protection Authority. Through open dialogue, participants shared personal experiences contributing to a comprehensive understanding of their views on phishing threats and prevention measures. Additionally, this focus group aimed to explore the perceptions of the general public regarding the jurisdictional authority of different entities in addressing cybersecurity challenges. Building upon the insights gathered from the previous two focus groups, discussions centered on the potential utilization of an AI system to assist users in navigating regulatory complexities and accessing relevant support resources effectively. The discussions of the second and third focus groups were moderated by the same facilitator.

TABLE I. FOCUS GROUPS PARTICIPANTS

Focus Group	**Participants**		
	Number	***Age***	***Characteristics***
Public Services	5	35-55	Higher Educational Institutes, Health Organization
Individuals	5	27-60	Victims, DPO
Legal Experts	8	40-60	Practicing attorneys, Legal experts from academia

b) Procedure: All three focus group interviews were held via teleconference and had a duration 50-60 minutes. The focus groups were nanaged by two research team members with the first team member, the principal researcher, acting as the moderator for all groups and the second team member acting as the assistant moderator and helping participants with technical issues. To begin the focus groups the moderator welcomed participants, shared a slide presentation with the questions-topics of conversation and briefly discussed tips to help the conversation. Then he asked each question, probed further to clarify or get more details on responses as appropriate, and facilitated discussion. During the interview, the assistang moderator acted as a note taker captured the main points of the groups' conversations as a backup in the event that the recording failed or corrupted.

All participants were recruited through targeted email invitations, drawn from the professional and academic networks of the facilitators of the discussions. The discussions were conducted in the participants' native language. Additionally, notes were taken by the moderator during the focus groups to capture key points and facilitate analysis. Following the completion of all three focus groups, the recorded discussions were manually transcribed into text for further analysis.

c) Ethics: In terms of ethics, all participants completed a privacy notice and a consent form that provided information regarding the purpose of the research the utilisation of participants' contributions, the confidentiality and anonymization of personal data. The participants consented be video recorded for for evaluation and research purposes. Participants completed a survey with demographic and organizational information. Before each focus group a short briefing was held with information about the procedure and the subject and after the end of it a debriefing finalized the procedure. As part of the ethics procedure, each participant was assigned a reference code for ensuring anonymity.

V. ANALYSIS OF RESULTS

Through the implementation of Focus Group studies that were implemented, human, technology, and legislative user model endpoints will be identified and discussed for explicit and implicit user modeling, which will guide the development of the corresponding AI-driven user modeling and profiling mechanisms. Our aim is, through the focus group procedure, to answer the research question stated in the previous section.

A. Public Services Focus Group

The Public Services Focus Group discussed phishing challenges, emphasizing neglect of verification processes, dependence on the IT department, and susceptibility under high workloads. Concerns were raised about the general scope of phishing attempts not being highly targeted but still posing a threat due to their broad reach. The group

highlighted the necessity for a user-centric system with personalized responses, legal information, and an intuitive user experience to enhance phishing defense mechanisms in organizations. This was illustrated by comments from two group participants.

"So, in the first instance, it doesn't want information, it wants a simple command, throw it [the phishing e-mail], don't throw it, or it's safe. Onwards, they could exist with an interactive process between the user and the system but rather on a second level"

"Yes, I would definitely use some legal information on how to remedy this whole situation and what legal responsibilities I have, because there are definitely legal responsibilities. Aside from the cheating issue"

Additionally, the session emphasized the potential value of an AI system in providing immediate post-attack guidance but raised concerns about its potential overuse and disruptive nature, suggesting a need for careful design to balance engagement and utility, as another participant commented.

"So, this could help a lot in providing ready-made checklists. [.....] That's where the AI system will be very useful because it will have the cold logic and speed to choose the right actions to be taken on a case-by-case basis and that's where AI could help"

B. Individual Users Focus Group

The Individual Users Focus Group discussed personal experiences with phishing attacks, emphasizing the varied nature of attacks and the vulnerability of certain groups like older individuals and those less familiar with technology. Participants highlighted the need for assistance in recognizing and responding to phishing attempts, providing clear guidance post-attack and personalized advice.

"..... I am contacting [an AI system] because I want to be informed in general about what legislative framework is applicable, or I am contacting because something has happened to me, and I am trying to find a solution to my problem"

There was also a focus on legal implications, data privacy, and the importance of offering up-to-date information on phishing tactics. The feedback indicated the necessity for a technically accurate and user-friendly system that ensures privacy and security of user data, incorporating legal advice and awareness of data protection regulations. As stated by three participants

"...it is a fact that an AI system can be used to perform various antiphishing actions, such as for user training, identity verification, anomaly detection, etc."

"Legal information, yes, provided that it would be done with moderation in a very careful way regarding the decisions one could make, in so that one can comprehend the induction the correlation"

"The goal is to ensure full anonymity, security, privacy of information and how long the information will be stored. And the security of the application itself matters. Because the user should be able to feel comfortable giving all this information to a system"

C. Legal Focus Group

The Legal Focus Group involved legal experts discussing National and EU legal framework against phishing, legal advice in order to combat phishing attacks, GDPR compliance, and the utilization of AI in addressing phishing attacks. The group focused on actions victims should take after phishing attempts, emphasizing contacting providers and authorities, preserving evidence, and seeking legal counsel if necessary. The feedback highlighted that it would be of major importance that the AI system could provide advice on fraud prevention and mitigation, specifically concerning the legal grounds for compensation. They also pointed out the lack of concrete information on behalf of Banks/financial institutions regarding phishing. As a participant commented

"I think this gap can be filled by such an artificial intelligence system that will guide the victim in his further actions. It is important that the AI system is fully up to date with the current models applied by the banks"

The group also discussed the need for the AI system to provide guidance on legal protections, emerging crimes, and jurisprudential trends at different judicial degrees.

"... the artificial intelligence system must also work as an aid to the lawyers, so that they are informed about the necessary actions that should be taken based on the latest protection model at a time, but also about the legal protection enjoyed by the victim of a phishing"

The AI system should also inform lawyers regarding the necessary actions to be taken, legal protections, emerging crimes, and jurisprudential trends, contribute to understanding these trends at various levels, and assist in reporting incidents to banks while collecting evidence from victims.

"After providing information to the AI system, [the AI system must inform] a) which provision of the Code applies, b) at the level of civil liability, if the Bank is liable under the legislation for electronic payments and for non-compliance with obligations"

VI. Discussion

A. Discussion on focus groups results

The results of the three focus groups highlighted various key points. The first focus group with stakeholders from public services discussed challenges related to phishing attacks, emphasizing the need for autonomous phishing detection methods accessible to all users and ongoing training programs to enhance the community's ability to respond to threats. The second focus group involving individual users focused on the need for an AI system to assist in recognizing and responding to phishing attempts, providing clear guidance post-attack, and offering personalized advice. The

third focus group with legal experts emphasized the importance of contacting providers in case of fraud, preserving evidence, and seeking legal advice to mitigate phishing attacks effectively.

AI systems can be designed to effectively assist users in recognizing and responding to phishing attempts by providing functionalities like user training, identity verification, and personalized guidance. They should offer clear, actionable advice post-attack, categorize user problems, and provide up-to-date information on phishing tactics to enhance user awareness. Legal implications and data privacy considerations are crucial, emphasizing the need for anonymization of data, careful collection of personal information, and clear guidelines for navigating legal landscapes surrounding phishing incidents. Overall, AI systems should be technically accurate, user-friendly, and ensure privacy and security of user data while incorporating legal advice and awareness of data protection regulations.

Victims of phishing attacks have legal responsibilities such as contacting authorities, preserving evidence, and seeking legal counsel if necessary. AI systems can guide victims by advising on fraud prevention, legal protections, and necessary actions post-attack, contributing to understanding emerging crimes and jurisprudential trends, and assisting in reporting incidents to banks while collecting evidence from victims. It is crucial for AI systems to handle legal information carefully, provide correct recommendations, and ensure the privacy and security of user data. Additionally, victims should file complaints or lawsuits for compensation against providers in fraud cases, with the AI system offering guidance on appropriate actions.

B. LLM specifications from Focus Groups

The focus groups provided valuable insights into developing a user-centric AI system for effective phishing defense mechanisms. Key features necessary for developing a user-centric AI system to enhance phishing defense mechanisms in public services organizations include offering prompt guidance post-phishing attack, categorizing emails as 'spam' or 'not spam,' providing personalized, interactive experiences, and ensuring privacy and security of user data. The AI system should offer details on laws regulating phishing attacks, inform users of their legal responsibilities, and provide up-to-date information on phishing tactics to enhance user awareness. Additionally, the AI system should assist in recognizing and responding to phishing attempts by offering clear, actionable advice, categorizing user problems, and ensuring a technically accurate and user-friendly design.

Common main points across the Public Services Focus Group, the Individual User Focus Group, and the Law Focus Group include the emphasis on the need for continuous training and awareness programs to enhance the community's ability to respond to phishing threats effectively. Additionally, there is a shared focus on the potential value of AI systems in assisting users in recognizing and responding to phishing attempts, providing clear guidance post-attack, and offering personalized advice. Furthermore, all three groups highlighted the importance of legal advice, data privacy considerations, and the necessity of contacting providers and preserving evidence in case of fraud to mitigate phishing attacks successfully.

From the three focus groups results analysis emerged the need to explore the effectiveness of AI systems in raising user awareness among individuals with less technical backgrounds due to the lack of human contact during interactions. Additionally, there is a need in understanding how these systems can effectively synthesize complex legal information and provide accurate recommendations to users in post-attack scenarios. Furthermore, there is a necessity in clarifying the roles and responsibilities of various authorities in handling phishing incidents to improve public awareness and provide clearer guidelines for victims.

C. Future work

This research uses focus groups aiming to implement multidisciplinary research that will formalize a novel and multidisciplinary anti-phishing user model that will incorporate human, technology, and legal factors of phishing. The findings assist to provide human, technology, and legislative user model endpoints that will be identified and discussed for explicit and implicit user modeling, which will guide the development of the corresponding AI-driven user modeling and profiling mechanisms. User Model Specification through AI entails mainly the utilization of Large Language Models (LLMs) in order to structure the specific domain of phishing attacks and the requirements and background of the users involved in that. After careful user modeling and domain modeling it is possible to perform inference and build on top of that the recommendation mechanisms. Through the development of an integrated AI-based anti-phishing software framework we anticipate providing to diverse business domains e.g., e-banking, e-government, e-education a comprehensive multidisciplinary anti-phishing framework that will be open source and will provide groundbreaking advances in the delivery of personalized anti-phishing legal advice and guidance.

To this direction we propose the AILA system (***A****rtificial* ***I****ntelligence-driven Framework and* ***L****egal* ***A****dvice Tools for Phishing Prevention and Mitigation in Information Systems*), that suggests an anti-phishing user model that will design and develop innovative recommendation functions and will provide personalized legal advice to organizations (based on domain factors) and end-users (based on profile). Furthermore, it will demonstrate a multidisciplinary and open-source anti-phishing framework that will implement phishing prevention and mitigation algorithms. At the top of that it will deliver validated anti-phishing adaptation heuristics aiming to facilitate the transfer of scientific knowledge, that will derive from valid real-world case-studies that will be implemented within the participating organizations - higher education domain, to other domains like e-government, e-health or e-banking.

VII. CONCLUSION

This research explores the intersection of generative AI and phishing, employing focus group analysis to define and validate an AI-driven anti-phishing data model that incorporates human, technological, and legal factors. The focus group discussions focused on phishing challenges, the need for user-centric AI systems, legal implications, and data privacy considerations. Participants emphasized the importance of personalized responses, legal information, and an intuitive user experience to enhance phishing defense

mechanisms. Overall, the focus groups yielded valuable information for creating an AI-powered system that prioritizes user needs and effectively thwarts phishing attacks.

The research highlighted the potential value of AI systems in providing post-attack guidance, balancing engagement and utility, and ensuring privacy and security of user data. The results will help establish user model endpoints in human, technological, and legal areas, which will be analyzed for both explicit and implicit user modeling. This will direct the creation of AI-driven mechanisms for user modeling and profiling.

ACKNOWLEDGMENT

This work has been financially supported by the Hellenic Foundation for Research & Innovation (HFRI) under the Basic Research Financing (Horizontal support for all Sciences), National Recovery and Resilience Plan (Greece 2.0) (ΑΔΑ: ΡΨΣΖ46Μ77Γ-Ε3Ο), under the project entitled "Artificial Intelligence-driven Framework and Legal Advice Tools for Phishing Prevention and Mitigation in Information System" (AILA) with Proposal ID 15440.

REFERENCES

[1] George Thomopoulos, Dimitrios Lyras, and Christos Fidas. 2023. Methodologies and Ethical Considerations in Phishing Research: A Comprehensive Review. In Proceedings of the 2nd International Conference of the ACM Greek SIGCHI Chapter (CHIGREECE '23). Association for Computing Machinery, New York, NY, USA, Article 3, 1–10. https://doi.org/10.1145/3609987.3609990

[2] Jari, Mousa. "An Overview of Phishing Victimization: Human Factors, Training and the Role of Emotions." ArXiv abs/2209.11197 (2022): n. pag. DOI 10.5121/csit.2022.121319

[3] Eric Chan-Tin, Loretta Stalans, Spencer Johnston, Daisy Reyes, and Shelia Kennison. 2022. Predicting Phishing Victimization: Roles of Protective and Vulnerable Strategies and Decision-Making Styles. In Fifth International Workshop on Systems and Network Telemetry and Analytics (SNTA '22). Association for Computing Machinery, New York, NY, USA, 35–42. https://doi.org/10.1145/3526064.3534107

[4] Ochieng NT, Wilson K, Derrick CJ, Mukherjee N. The use of focus group discussion methodology: Insights from two decades of application in conservation. Methods Ecol Evol. 2018; 9: 20–32. https://doi.org/10.1111/2041-210X.12860

[5] Richard A. Powell, Helen M. Single, Focus Groups, International Journal for Quality in Health Care, Volume 8, Issue 5, 1996, Pages 499-504, https://doi.org/10.1093/intqhc/8.5.499

[6] Ansari, Meraj Farheen; Sharma, Pawan Kumar; and Dash, Bibhu (2022) "Prevention of Phishing Attacks Using AI-Based Cybersecurity Awareness Training," International Journal of Smart Sensor and Adhoc Network: Vol. 3: Iss. 3, Article 6. DOI: 10.47893/IJSSAN.2022.1221

[7] Gaurav Varshney, Rahul Kumawat, Vijay Varadharajan, Uday Tupakula, Chandranshu Gupta, Anti-phishing: A comprehensive perspective, Expert Systems with Applications, Volume 238, Part F, 2024, 122199, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2023.122199

[8] Lötter, A. and Futcher, L. (2015), "A framework to assist email users in the identification of phishing attacks", Information and Computer Security, Vol. 23 No. 4, pp. 370-381. https://doi.org/10.1108/ICS-10-2014-0070

[9] Button, Mark & McNaughton Nicholls, Carol & Kerr, Jane & Owen, Rachael. (2014). Online frauds: Learning from victims why they fall for these scams. Australian and New Zealand Journal of Criminology. 47. 391-408. 10.1177/0004865814521224.

[10] Emma J. Williams, Joanne Hinds, Adam N. Joinson, Exploring susceptibility to phishing in the workplace, International Journal of Human-Computer Studies, Volume 120, 2018, Pages 1-13, ISSN 1071-5819, https://doi.org/10.1016/j.ijhcs.2018.06.004.

[11] Dr. Vasileios Karagiannopoulos, Dr. Annie Kirby, Shakiba Oftadeh-Moghadam, Dr. Lisa Sugiura, Cybercrime awareness and victimisation in individuals over 60 years: A Portsmouth case study, Computer Law & Security Review, Volume 43, 2021, 105615, ISSN 0267-3649, https://doi.org/10.1016/j.clsr.2021.105615.

[12] Kholoud Althobaiti, Nicole Meng, and Kami Vaniea. 2021. I Don't Need an Expert! Making URL Phishing Features Human Comprehensible. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 695, 1–17. https://doi.org/10.1145/3411764.3445574

[13] Misra, Gaurav, Nalin Asanka Gamagedara Arachchilage and Shlomo Berkovsky. "Phish Phinder: A Game Design Approach to Enhance User Confidence in Mitigating Phishing Attacks." International Symposium on Human Aspects of Information Security and Assurance (2017), DOI 10.48550/arXiv.1710.06064

[14] Andrew Parker & Jonathan Tritter (2006) Focus group method and methodology: current practice and recent debate, International Journal of Research & Method in Education, 29:1, 23-37, DOI: 10.1080/01406720500537304

[15] Haney, J. M., Jacobs, J. L., Barrientos, F., & Furman, S. M. (2022). Lessons Learned and Suitability of Focus Groups in Security Information Workers Research. In HCI for Cybersecurity, Privacy and Trust (pp. 135–153). Springer International Publishing. https://doi.org/10.1007/978-3-031-05563-8_10

[16] A. Gibbs, "Focus Groups," Social Research Update Issue 19, Guilford, 1997. http://sru.soc.surrey.ac.uk/SRU19.html

[17] Sim, J., & Waterfield, J. (2019). Focus group methodology: some ethical challenges. In Quality & Quantity (Vol. 53, Issue 6, pp. 3003–3022). Springer Science and Business Media LLC. https://doi.org/10.1007/s11135-019-00914-5

[18] Masadeh, Mousa. (2012). Focus Group: Reviews and Practices. International Journal of Applied Science and Technology. 2. 63-68.

[19] Morgan, D.L. (1988) Focus Groups as Qualitative Research. Newbury Park, CA:Sage

[20] Nassar-McMillan, S. C., and Borders, L. D. (2002, March) "Use of Focus Groups in Survey Item Development.The Qualitative Report [online] 7 (1). Available from http://www.nova.edu/ssss/QR/QR7-1/nassar.html [11May 2011]

[21] David.W. Stewart & Prem Shamdasani (2017) Online Focus Groups, Journal of Advertising, 46:1, 48-60, DOI: 10.1080/00913367.2016.1252288