# Validation of the Emotion Based Information Security Awareness Measurement

**Anikó Szarvák[1] and Valéria Póser[2]**

[1]Doctoral School of Applied Informatics and Applied Mathematics, Obuda University, Bécsi út 96/B, Budapest, 1034, Hungary, szarvak.aniko@nik.uni-obuda.hu

[2]Biomatics and Applied Artifical Intelligence Institution, John von Neumann Faculty of Informatics, Obuda University, Bécsi út 96/B, Budapest, 1034, Hungary, poser.valeria@nik.uni-obuda

***Abstract:*** *Large Language Models (LLMs) are currently receiving increasing attention, due to their ability to generate natural language output that closely resembles human-written text. ChatGPT was designed with the goal of not being used for illegal activities. However, there are AI tools that do not have such restrictions. As of now, internet users are not prepared for cyberattacks, their activities on the internet are not regulated by Information Security Policies (ISP), and high-level technological protection solutions are not available in their private lives. The use of AI for cyberattacks, fraud, and profit is already present. And defense solutions only follow the attacks, the attack solutions. Therefore, it is essential that users who may come into contact with cyberattacks are prepared. Measuring security awareness based only on a knowledge base, does not necessarily give an accurate picture of a users' real security awareness, due to the rapid obsolescence of the knowledge base. The goal of this article is to supplement the imprecise, knowledge-based measurement, by examining and making measurable relationship obsrvations between cyberattacks and emotions.*

*Keywords: cyberattack; cybersecurity awareness; cybersecurity awareness measurement;*

## 1 Introduction

Nowadays, various cybersecurity attacks are receiving more and more attention. In addition, artificial intelligence (AI) and machine learning (ML) are spreading explosively, which gives attackers new opportunities to deceive unsuspecting users, steal their user identification data, passwords, and access various online service accounts and bank accounts. This article aims to point out that cyberattacks are using increasingly new and sophisticated methods, which the user security

awareness knowledge base can no longer keep up with, which is why the success of an attack may depend on the personal decisions of users. The article points out a possibility of measuring security awareness based on emotions and presents the results of research in this direction.

AI is gaining importance and is beginning to play an increasingly important role in cyberattacks. Because of this, cybersecurity awareness is going to be valued as a first-line of defense. To overview how AI-solutions are used to carry out a cyberattack, the beginning of this article summarizes a few recent cases.

In response to the accelerated development of AI and ML solutions, security awareness and its measurement based solely on knowledge bases are no longer sufficient. Therefore, a new way of measuring security awareness, supplemented by its validation, is considered necessary, where in addition to the information security knowledge base, the emotional reactions of the participants are also taken into account.

## 1.1 AI is a versatile tool

While genAI technologies can be used as part of organizations' cyber defenses, attackers can also use generative AI to carry out more sophisticated cyber attacks.

Attackers with AI models can identify and target individuals or organizations that are most susceptible to phishing attacks. These models can analyze social media posts, emails, and other communication channels to craft persuasive attack messages.

AI used in phishing attacks typically uses a combination of machine learning, natural language processing, deep learning, and reinforcement learning. Natural language processing helps generate personalized, authentic-looking email messages and web pages, and interprets user responses. Reinforcement learning methods allow attackers to continuously refine their attack strategies, responding to user interactions and adapting to defense mechanisms. The application of AI in phishing attacks allows attackers to execute personalized, difficult-to-detect attacks while continuously adapting and learning from their environment [1]

Language models have recently received increasing attention due to their ability to generate natural language output that closely resembles human-written text. ChatGPT was designed with the goal of not being used for illegal activities. However, there are AI tools that do not have such restrictions.

According to an article “a vicious ChatGPT-like AI model is spreading across the dark web, allowing hackers to carry out cyberattacks at an unprecedented scale.” [2]

The emergence of a new artificial intelligence model similar to ChatGPT on the dark web is causing concern among cybersecurity professionals, as this model opens up novel opportunities for cybercriminals. The artificial intelligence model, known

as WormGPT, is designed to create content similar to human-written text that can be used in cybercriminal campaigns.

The long-term consequences of such applications of artificial intelligence are still unclear, but concerns already emerging in the current situation – such as the production of disinformation and misinformation, influencing public opinion, and even impacting political campaigns – pose significant risks for users. [3]

Based on the above citated cases, although the use of AI in cybersecurity also holds great potential, it should not be forgotten that security technologies and solutions will always follow the attacks, as a kind of reaction. In addition, AI-based attacks allow for more targeted, personalized attacks, against which the final line of defense is the human. The likelihood of threats is increasing, especially with regard to the misuse of AI and related cyberattacks. This increase is linked to the proliferation of AI applications and the emergence of ethical issues[4].

### 1.1.1 LLM-Based Phishing and Business Email Compromise

Threat actors exploit Large Language Models (LLMs) to conduct sophisticated phishing and Business Email Compromise (BEC) attacks. They can generate deceptive emails that are contextually relevant, personalized, and grammatically flawless, mimicking legitimate communications from trusted sources, such as company executives or business partners.

The 2023 Verizon Data Breach Investigations Report [5] highlights this trend, noting a near doubling of BEC incidents, which now represent over half of all social engineering attacks. The ease with which LLMs can be used to craft convincing fake emails poses a challenge for traditional security measures.

### 1.1.2 AI-Generated Images

Generative adversarial networks (GANs) [6] and diffusion-based text-to-image models enable the creation of realistic visuals, ranging from human portraits to complex physical scenes. Threat actors, including those aligned with nation-states and independent groups, can craft images that bolster false narratives or impersonate real individuals, helping spread disinformation.

The accessibility of tools like thispersondoesnotexist.com has led to widespread use of generative AI images in information campaigns. Advancements in text-to-image technology promise even broader applications for deceptive content that is challenging to detect.

### 1.1.3 Deepfake

AI technologies can be used to manipulate video footage for disinformation campaigns. This includes customizable AI-generated avatars and deepfake technology that alters existing videos to superimpose faces or mimic voices. Such

manipulations have been used to fabricate narratives or impersonate individuals, enhancing the persuasive impact of false information.

The use of manipulated video footage in cyber operations has been observed since 2021, with instances including AI-generated news presenters and deepfake videos of political figures. In 2024 and beyond, with the launch of the first generative AI tools able to generate novel, photorealistic video footage, that threat will become even more potent. [7]

AI-generated audio and video, while not yet widely adopted, has shown potential for misuse in creating deceptive content. Technologies enabling text-to-voice generation and voice cloning can produce audio tracks that convincingly mimic public figures, making statements that are violent, racist, or otherwise harmful.

These capabilities raise concerns about the ease with which individuals could be impersonated to spread disinformation or incite conflict. Despite the limited use observed so far, tools capable of generating realistic voice recordings from text inputs offer a new vector for social engineering attacks and misinformation campaigns.

### 1.1.4 Improved Reconnaissance

Generative AI enhances reconnaissance efforts, enabling threat actors to analyze and process vast amounts of open-source and proprietary data. [8] With machine learning and data science tools, adversaries can quickly sift through stolen information or publicly available data to identify valuable targets or vulnerabilities. This improves precision in selecting targets.

AI-based tools also assist in refining tradecraft techniques used during reconnaissance. For example, AI can uncover patterns in data that human analysts might overlook, providing insights into more effective ways to approach intelligence gathering or target selection.

### 1.1.5 LLM-Guided Malware Development

LLMs can assist in creating new malware or enhancing existing ones, irrespective of an attacker's technical skill or language proficiency. While there are limitations in LLMs' malware generation capabilities that might require human correction, their contribution to malware creation aids skilled developers and empowers those with lesser expertise.

Reports indicate that financially motivated actors are promoting services on underground forums to bypass LLM restrictions designed to prevent their use in developing and spreading malware, as well as creating deceptive lure materials. Advertisements for LLM services, sales, and API access, as well as LLM-generated code, are increasingly seen on dark web forums. [9]

## 1.2 Impact of the AI-based cyberattacks

Verizon's 2024 Data Breach Investigation Report [10] notes that 68% of attacks are due to human error, including phishing. The World Economic Forum's Global Cybersecurity Outlook 2025 [11] also highlights a sharp increase in phishing and manipulation attacks in 2024, with 42% of organizations reporting such incidents.

Security awareness has indispensable importance for user of the Internet, which should focus more on making them aware of how vulnerable the Internet is, rather than focusing on the various attacks. Explaining the various attacks or keeping the IT knowledge base up to date in the ever-changing IT world is close to impossible. Regulations that support security awareness training in a corporate environment may not keep up with the changes in the IT advancement fast enough.

As of now, internet users are not prepared for cyberattacks, their activities on the internet are not regulated by Information Security Policies (ISP), and high-level technological protection solutions are not available in their private lives.

Raising awareness is essential to reduce the risk and impact of cyberattacks. Instead of restricting the use of technology, end users of technology should be aware of the potential dangers, similar to driving a car or using a weapon. The importance of awareness is becoming increasingly important as technology advances, so much so that, like in physical life, safeguards have been left one step behind. In this situation, risk reduction can be achieved by increasing security awareness.

The use of AI for cyberattacks, fraud, and profit is already present. And defense solutions only follow the attacks, the attack solutions. Therefore, it is essential that users who may come into contact with cyberattacks are prepared. It is a mistake to think that the targets of attacks can only be high-ranking, high-income company executives or government officials. With the help of AI, attackers can determine the people through whom they are more likely to launch a successful cyberattack.

That is why it is of utmost importance to prepare users for the possibility of attacks. The field of cybersecurity, like IT, is an incredibly fast-developing field. It is practically impossible for everyone who comes into contact with IT and the world wide web to have up-to-date knowledge about current attacks and appropriate defense options every minute. In addition, but in parallel, measuring security awareness based only on a knowledge base does not necessarily give an accurate picture of users' real security awareness due to the rapid obsolescence of the knowledge base.

The goal is to supplement the above described, imprecise, knowledge-based measurements, by examining and making measurable the relationship between cyberattack reactions and emotions.

# 2 Emotion-based information security awarenes measurement

The measurement is a multiple-choice test consisting of two parts: the first part is used to record emotions; the second part records the reaction to the cyberattack. Based on the questionnaire survey, the research sought to answer, on the one hand, what emotional reactions are triggered by various, widespread attacks. On the other hand, whether cybersecurity awareness can be measured based on emotional reactions. The aim of the survey was to show what emotional reactions can be experienced in the event of attacks in relation to the KAB [12] model from the point of view of IT security awareness.

The current article used the combined results of the Case and Control groups and repeated the measurement with a group of students majoring in cybersecurity, called Validation group. The students' backgrounds ranged from those working in cybersecurity to those who were just starting out in cybersecurity. The students were at the beginning of their studies. The students' backgrounds were similar to those of the Case and Control groups.

In the present study, the measurement results of the Case and Control groups (both emotional reactions and attack success data) are combined in order to examine the similarities or differences in emotional reactions. The aim of the present study is to compare whether the two completely different groups show similar emotional patterns and reactions to the questions.

Regarding applied methodology, distribution calculation, in this article all distribution results were counted by the summary of individual emotional reactions to the different social engineering scenarios and divided by the summary of the given group results. This method was used to determine the distribution of a given reaction within the entire group. In addition, this method ensures that the results of different emotional reactions within different groups with various number of participants are comparable.

In this article the overall number of emotional reactions are 555, where Case and Control group includes 235 results and Validation group includes 320. The data collection covers multiple years and a wide variety of students, with different knowledgebases in the field of Cybersecurity.

## 2.1 Case and Control group emotional profile review

Case group – is a group of university students, interested in cybersecurity and they are at the beginning of their studies. The second group – hereinafter: Control group – is also a group of university students, highly interested in cybersecurity, even working in this field and they are at the end of their studies. Detailed assessment of their answers was published in "The connection between emotional reactions and successful cyberattacks – a measurement prototype" [13].

It was concluded that the results show that experience and education could reduce but not eliminate completely the threat of a cyberattack. During the education of Control group, the students acquired in-depth knowledge of cyberattacks and they face as many examples as possible. The Control Group is therefore able to identify possible cyberattacks based on typical characteristics of the cyberattack.

Table 1

Answer distribution between Case group and Control group

| | **Case group results** | **Control group results** | **Case group distribution** | **Control group distribution** | **Difference (p)** |
|---|---|---|---|---|---|
| Fear | 8 | 4 | 0.059 | 0.040 | 0.019 |
| Anger | 10 | 39 | 0.074 | 0.390 | 0.316 |
| Indifference | 36 | 25 | 0.267 | 0.250 | 0.017 |
| Surprise | 25 | 8 | 0.185 | 0.080 | 0.105 |
| Contempt | 38 | 21 | 0.281 | 0.210 | 0.071 |
| Joy | 8 | 0 | 0.059 | 0.000 | 0.059 |
| Disgust | 10 | 3 | 0.074 | 0.030 | 0.044 |

The comparison between Case group and Control group in Table 1 shows four significant ($p \geq 0.05$) differences regarding the summary of the emotional profile of the cyberattacks:

- P=0.316 in case of Anger. The Anger emotion was marked in 10 cases in case of Case group, however it was marked in 39 cases in case of Control group. Overall it means a significant difference of the emotional reaction between the two groups and it could be caused by the difference of the acquired knowledge by the Control group members.
- P=0.105 for Surprise, P=0.071 for Contempt and P=0.059 for Joy, all the cases the number of the answers was reduced in Control group in comparison to Case group.

The volume of the given answers between Case group and Control group for Fear (p=0.019), Indifference (p=0.017) and Disgust (p=0.044) were not identified as significantly different.

To have a profile of the emotional based reactions, it was decided to combine Case group and Control group answers to be comparable to the Validation group answers. Due to this action, both groups have same profile regarding the participants, eg. both groups are containing beginner and experienced students in the field of cybersecurity and at the beginning of their studies and with preliminary studies.

## 2.2 Combination of Case and Control group data

To have an overview of the common profile of the Case group and the Control group, their measurement data were combined and evaluated together, hereinafter Case and Control group. The summary of the results of the Case group and Control group is covered in Table 2, where the rows are containing the emotions, the columns are containing the measured cases by the following order: Q1, Phishing on Coronavirus aid, Q2, Access to camera, Q3, Package arrival notification, Q4, Gift notification, Q5, The salary of the teachers.

Table 2

Summary of the results of Case group and Control group

| | Q1 | Q2 | Q3 | Q4 | Q5 | Distribution |
|---|---|---|---|---|---|---|
| Fear | 1 | 6 | 1 | 3 | 1 | 5.11% |
| Anger | 16 | 13 | 10 | 8 | 2 | 20.85% |
| Indifference | 12 | 11 | 7 | 9 | 22 | 25.96% |
| Surprise | 3 | 2 | 9 | 3 | 16 | 14.04% |
| Contempt | 11 | 11 | 17 | 17 | 3 | 25.11% |
| Joy | 0 | 3 | 2 | 2 | 1 | 3.40% |
| Disgust | 4 | 1 | 1 | 5 | 2 | 5.53% |

With this combination of the two groups data allows to compare the results with the Validation group results where the contributors have similar background and knowledge base compared to Case and Control group participants. The attacks were categorized as successful if the participant chose to open or click on a link in their answer as an indication of a successful cyberattack. Distribution of the successful and unsuccessful attacks for Case and Control group is concluded in the following, Table 3:

Table 3

Case and Control group: Distribution between Successful and Unsuccessful Attacks

| | Q1 | Q2 | Q3 | Q4 | Q5 | Total | Distribution |
|---|---|---|---|---|---|---|---|
| Successful attack | 0 | 6 | 3 | 2 | 9 | 20 | 8.51% |
| Unsuccessful attack | 47 | 41 | 44 | 45 | 38 | 215 | 91.49% |

## 2.3 Validation group summary

To create the emotion profile for the measurement, Validation group had been introduced. The subjects of the Validation group are also University students, at the beginning of their studies, but have experience, due to their previous studies or from employment. Therefore, the subjects of Validation group conclude the profile

characteristic of both, Case and Control group. Validation group data collection was equal to Case and Control group data collection and measurement.

Table 4
Summary of the results of Validation group

| | **Q1** | **Q2** | **Q3** | **Q4** | **Q5** | **Distribution** |
|---|---|---|---|---|---|---|
| Fear | 7 | 3 | 1 | 3 | 2 | 5.00% |
| Anger | 8 | 9 | 14 | 14 | 5 | 15.63% |
| Indifference | 19 | 21 | 13 | 12 | 17 | 25.63% |
| Surprise | 4 | 12 | 15 | 1 | 31 | 19.69% |
| Contempt | 18 | 12 | 15 | 22 | 6 | 22.81% |
| Joy | 5 | 1 | 1 | 6 | 2 | 4.69% |
| Disgust | 3 | 6 | 5 | 6 | 1 | 6.56% |

The Validation group results are summarized in Table 4, where the rows are containing the emotions, the columns are containing the measured cases by the following order: Q1, Phishing on Coronavirus aid, Q2, Access to camera, Q3, Package arrival notification, Q4, Gift notification, Q5, The salary of the teachers.

Distribution of the successful and unsuccessful attacks for Case and Control group is concluded in the following, Table 5:

Table 5

Validation group: Distribution between Successful and Unsuccessful Attacks

| | **Q1** | **Q2** | **Q3** | **Q4** | **Q5** | **Total** | **Distribution** |
|---|---|---|---|---|---|---|---|
| Successful attack | 0 | 8 | 2 | 0 | 8 | 18 | 5.65% |
| Unsuccessful attack | 64 | 56 | 62 | 64 | 56 | 302 | 94.35% |

# 3 Comparison of emotional reactions

The total dataset is based on the total 555 collected emotional reactions. The emotional reaction results were compared in three ways: 1, The total emotional reaction distribution difference between Case and Control group and Validation group results, 2, The possibly unsuccessful emotional reaction distribution difference between the two groups and 3, The probably successful attactk's emotional reaction distribution between the two groups.

In this article, it is assumed that the results of the Case and Control groups and the results of the Validation group no significant difference ($p \geq 0.05$) can be identified based on different emotional reactions. If the emotional reaction is not significantly different, then it can be concluded that, in general, the emotional reaction patterns to the different attacks are similar.

The difference between C+C and Validation calculation was performed by summing up all the emotional reactions of the five attack types of the C+C group, a total of 235 measurement values, and the distribution of the given emotional reactions was determined based on all the values. The Validation group calculation and the distribution of emotional reactions were performed following the same principle. In case of Validation group, the total number of measurement was 320. The difference calculation gives the difference between the values of the C+C group and the Validation group, in absolute value.

Table 6

Emotions distribution between Case and Control group and Validation group, total emotional reactions

| | **Case and Control group** | **Validation group** | **Case and Control group** | **Validation group** | **Difference (p)** |
|---|---|---|---|---|---|
| Fear | 12 | 16 | 0.051 | 0.050 | 0.001 |
| Anger | 49 | 50 | 0.209 | 0.156 | 0.052 |
| Indifference | 61 | 82 | 0.260 | 0.256 | 0.003 |
| Surprise | 33 | 63 | 0.140 | 0.197 | 0.056 |
| Contempt | 59 | 73 | 0.251 | 0.228 | 0.023 |
| Joy | 8 | 15 | 0.034 | 0.047 | 0.013 |
| Disgust | 13 | 21 | 0.055 | 0.066 | 0.010 |

The results show that the total emotional distribution of the C+C group and the Validation group shows difference in case of Anger and Surprise. The difference for all other emotional reactions is below 0.05 as it is captured in Table 6. The calculation of the results are based on the distribution of the answers per each group.

The same distribution calculation was performed by summarizing the emotional reaction numbers of the Case and Control group and the Validation group. It can be concluded that emotional reactions overall shows the same patterns for both investigated groups.

## 3.1 Unsuccessful attacks emotional patterns

Separately from the complete emotional pattern calculation, the emotional reactions to attacks where the reaction indicated unsuccessful cyberattack, not opening the linked website, or not opening the attached file, i.e. the attack was considered unsuccessful, were examined, see Table 7.

In case of Fear, Contempt, Disgust, Joy and Indifference, no significant difference can be identified between Case and Control group and Validation group answers, ($p < 0.05$).

In case of Anger and Surprise emotional reactions, the difference is higher than 0.05, therefore significant difference can be identified. The reactions to Anger shows decrease between Case and Control group and Validation group, and the reactions to Surprise shows increase between the two groups. These changes also can be captured in the Total emotional reactions section, see Table 6.

Table 7

Emotional difference between Case, Control and Validation groups for successful attacks

| | **Case and Control group** | **Validation group** | **Case and Control group distribution** | **Validation group distribution** | **Difference (p)** |
|---|---|---|---|---|---|
| Fear | 12 | 14 | 0.056 | 0.046 | 0.009 |
| Anger | 49 | 50 | 0.228 | 0.166 | 0.062 |
| Indifference | 53 | 74 | 0.247 | 0.245 | 0.001 |
| Surprise | 29 | 58 | 0.135 | 0.192 | 0.057 |
| Contempt | 58 | 72 | 0.270 | 0.238 | 0.031 |
| Joy | 2 | 13 | 0.009 | 0.043 | 0.034 |
| Disgust | 12 | 21 | 0.056 | 0.070 | 0.014 |

## 3.2 Successful attacks emotional patterns

Separately from the complete emotional pattern, the emotional reactions to attacks where the reaction indicated successful cyberattack, opening the linked website, or opening the attached file, i.e. the attack was considered successful, were examined.

In the case of hypothetical successful attacks, the emotional reactions do not show the same general pattern as the aggregated emotional reactions described above. In fact, with the exception of a few emotional reactions, such as Anger or Contempt, the emotional reactions show more significant differences between the two groups.

The comparison is calculated as the ratio of the emotion measured for a successful attack in the group to all emotions. The results are concluded in the following table, Table 8.

Based on the measurement, the results shows significant difference ($p \geq 0.05$) in case of Fear ($p=0.111$), Surprise ($p=0.078$) and Joy ($p=0.189$). It worth to comment, that in case of Fear, the Validation group marked it only, regarding to a successful cyberattack.

In case of Disgust, the difference between the Case and Control group and Validation group was equal to the significance limit ($p=0.5$), therefore these are considered significant differences.

Last but not least, in the cases of Anger, Contempt and Indifference, no significant difference can be identified based on the results between the two measurement groups. This leads to the conclusion that in the emotional results of Anger, and Contempt highily likely not leads to a successful cyberattack.

Examining all the emotional reactions, a successful attack can be assumed in 6.85% of the attacks. The difference is not significant when comparing the groups based on the rate of successful attack (Case and Control group has 8.51% and Validation group has 5.63%).

Table 8
Emotional difference between Case, Control and Validation groups for cases of successful attacks

| | Case and Control group | Validation group | Case and Control group distribution | Validation group distribution | Difference (p) |
|---|---|---|---|---|---|
| Fear | 0 | 2 | 0.000 | 0.111 | 0.111 |
| Anger | 0 | 0 | 0.000 | 0.000 | 0.000 |
| Indifference | 8 | 8 | 0.400 | 0.444 | 0.044 |
| Surprise | 4 | 5 | 0.200 | 0.278 | 0.078 |
| Contempt | 1 | 1 | 0.050 | 0.056 | 0.006 |
| Joy | 6 | 2 | 0.300 | 0.111 | 0.189 |
| Disgust | 1 | 0 | 0.050 | 0.000 | 0.050 |

## 3.3 Distribution between Successful and Unsuccesful attacks emotional landscape

Based on the distribution of total emotional reactions between Successful and Unsuccessful attacks, emotional reactions can be linked to the outcome. Table 9 shows the summary of this distribution. By evaluating variability between the two scenarios,it is highly likely that Fear (p=0.002) and Disgust (p=0.038) emotions has no significant impact on the outcome of the attacks.

Table 9
Emotional difference between successful and unsuccessful attacks

| | Successful attack distribution | Unsuccessful attack distribution | Total distribution | Difference (p) |
|---|---|---|---|---|
| Fear | 0.053 | 0.050 | 0.050 | 0.002 |
| Anger | 0.000 | 0.191 | 0.178 | 0.191 |
| Indifference | 0.421 | 0.246 | 0.258 | 0.175 |
| Surprise | 0.237 | 0.168 | 0.173 | 0.069 |
| Contempt | 0.053 | 0.251 | 0.238 | 0.199 |
| Joy | 0.211 | 0.029 | 0.041 | 0.182 |
| Disgust | 0.026 | 0.064 | 0.061 | 0.038 |

In case of Anger (p=0.191), Contempt (p=0.199) it is highy likely that the cyber attack will be unsuccessful.

The positive emotional reactions triggered by the attack like Surprise (p=0.069) and Joy (p=0.182) is highly likely leads to a successful cyber attack. In case of Indifference (p=0.175) invites a more boarder investigation, because this category could point out the lack of knowledge in field of cybersecurity.

# 4 Conclusions

Artificial intelligence that can be used by attackers, creates an opportunity for attackers to pinpoint the most suitable target and carry out these targeted attacks, thus, increasing their success rates. All that is needed for this, is for the attackers to accurately trigger an emotional reaction, with the victum. Artificial intelligence, by having access to huge databases, can help to better understand the target and thus, to design the attack appropriately.

During the evaluation of the data, it can be concluded that no significant difference identified in the aggregated emotional profile between Case and Control group results and Validation group results. This could provide a solid baseline for further measurements and validations.

Significant differences were identified in the case of unsuccessful cyber attack scenarios, related emotional pattern comparison. By evaluating variability between the two scenarios, it is highly likely that Fear and Disgust emotions have no impact on the outcome of the attacks. Anger and Contempt highily are likely not leads to a successful cyberattack. Also, in case of Anger, the conclusion is supported by the capture of the emotional distribution between Case group and Control group, where Anger shows higher values for the Control group, which has more cybersecurity knowledge, thus, the participants have nearly ended their studies.

Also significant differences were identified in the case of successful cyber attack scenarios related emotional pattern comparison. The positive emotional reactions triggered by the attack, such as Surprise and Joy, are highly likely to lead to a successful cyber attack.

By comparing the successful and unsuccesful attack scenarios related emotions, significant differences were identified in case of Indifference (p=0.175). This difference invites a more boarder investigation, because this emotional category suggests that the decision is based on knowledge and not based on other emotions. Therefore, this emotional reaction could point out a lack of knowledge in the field of cybersecurity.

**Further work**

This prototype measurement, currently covers a narrow audience, where the participants are related to the field of Cybersecurity. We can expand the survey for

a wider audience, hopefully to better understand the emotional reactions, in the case of a social engineering attack, and supplement the measurement results to be able to draw a broader baseline, which can be used as a benchmark for basic comparisons or individual measurement of cybersecurity resilience.

## Acknowledgement

Authors gratefully acknowledges the support of the Doctoral School of Applied Informatics and Applied Mathematics on this research.

## References

[1] B. Espinoza, J. Simba, W. Fuertes, E. Benavides, R. Andrade and T. Toulkeridis, "Phishing Attack Detection: A Solution Based on the Typical Machine Learning Modeling Cycle," 2019 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2019, pp. 202-207, doi: 10.1109/CSCI49370.2019.00041.

[2] Tod-Raileanu, G., & Bacivarov, I. (2023). An Overview on Artificial Intelligence Applied in Offensive Cybersecurity. Int'l J. Info. Sec. & Cybercrime, 12, 23.

[3] Polra Victor Falade, Decoding the Threat Landscape : ChatGPT, FraudGPT, and WormGPT in Social Engineering Attacks, Page Number 185-198, Published: 09 Oct 2023

[4] ENISA – Foresight cybersecurity threats for 2030, https://www.enisa.europa.eu/sites/default/files/publications/ENISA%20Foresight%20Cybersecurity%20Threats%20for%202030.pdf, Accessed 2025. 04. 30.

[5] 2023 Data Breach Investigations Report by Verison https://www.verizon.com/about/news/2023-data-breach-investigations-report accessed 2024. 05. 18.

[6] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta and A. A. Bharath, "Generative Adversarial Networks: An Overview," in IEEE Signal Processing Magazine, vol. 35, no. 1, pp. 53-65, Jan. 2018, doi: 10.1109/MSP.2017.2765202.

[7] Valencia A. Jones Artificial Intelligence Enabled – Deepfake Technology, The Emergence of a New Threat, Published by ProQuest LLC, 2020.

[8] Temara, S.: Maximizing penetration testing success with effective reconnaissance techniques using chatgpt, 2023.

[9] Roach, K. Llms for malware offense and defense.2024.

[10] 2024 Data Breach Investigations Report by Verison https://www.verizon.com/business/resources/Te21/reports/2024-dbir-data-breach-investigations-report.pdf Accessed: 2024. 10. 30.

[11] World Economic Forum's Global Cybersecurity Outlook 2025: https://reports.weforum.org/docs/WEF_Global_Cybersecurity_Outlook_2025.pdf Accessed: 2025. 05. 25.

[12] Kruger, H. A., & Kearney, W. D. (2006). A prototype for assessing information security awareness. Computers & security, 25(4), 289-296.

[13] A.Szarvák, V.Póser: The connection between emotions and successful cyber attacks In: ACTA POLYTECHNICA HUNGARICA (1785-8860 1785-8860): 22 5 pp 185-204 (2025)