Social Engineering and Information System Security- A survey on the Necessity of Prevention

Florence SEDES [1] [0000-0002-9273-302X] and Jonathan DEGRACE [2]

[1] University of Toulouse – IRIT, Toulouse, France
[2] IT Cybersecurity Consultant, Dexper, Aix en Provence, France
sedes@irit.fr

Abstract. This survey investigates the persistent and evolving threat of social engineering to information system security, emphasizing the urgency of prevention strategies. With 74% of cyberattacks involving a human component, social engineering exploits cognitive biases, emotional responses, and psychological manipulation to deceive individuals and breach organizational defenses. The integration of AI significantly amplifies these threats, enabling hyper-personalized, scalable, and deceptive attacks through tools like deepfakes and generative language models. This paper categorizes various attack vectors—phishing, smishing, vishing, social phishing, quishing, and cyber-grooming—detailing their mechanisms and psychological underpinnings. It further explores how AI facilitates the entire attack lifecycle, from reconnaissance to exploitation and exit, by automating pretexting, generating tailored content, and adapting in real-time to target behavior. Beyond the technical perspective, the study stresses the human dimension of cybersecurity, highlighting the need for user training, critical thinking, and dynamic awareness programs. While current countermeasures range from AI-driven detection systems to organizational protocols, their effectiveness varies due to usability limitations and adaptation challenges. The paper advocates for transdisciplinary approaches combining social science insights with technical innovation. It concludes by recommending scalable, adaptive, and ethically compliant prevention strategies, calling for continuous evaluation, regulatory alignment, and user empowerment. The findings underline the necessity of shifting users from security vulnerabilities to active defense assets in the face of increasingly sophisticated AI-powered social engineering**.**

Keywords: Social Engineering, Cybersecurity, Phishing, AI, Cognitive Biases, Prevention, Cyberattacks, Human Factors, Training, Detection.

1 Introduction

The management of an organization's essential data is becoming increasingly sensitive with the heightened risk of (cyber) attacks: overall, it is the Information System (IS) that must be protected against all financial, legal, reputational, etc., damage.

Interconnections with other organizations, the increasingly widespread use of the Internet of Things (IoT), and the heterogeneity of Information Systems expand the attack surface usable by cybercrime, fostering its growth:74% of cyberattacks have a human component and rely on social engineering methods, 50% are phishing professional emails. Furthermore, the use of technologies such as generative AI and Machine Learning increases the quality and quantity of cyberattacks.

Social engineering involves the use of human manipulation and fraud techniques through digital tools and is used against organizations and individuals alike. Social engineering attacks thus highlight humans as the “weak link” in an organization’s cyber defense. Depending on the cybersecurity prevention methods employed, the results in terms of resistance to social engineering vary greatly.

Social engineering attacks remain one of the most pervasive threats to digital security, leveraging psychological manipulation to exploit human vulnerabilities. Traditionally associated with phishing emails and impersonation scams, the field has evolved drastically with the advent of AI. Today, sophisticated language models and generative media tools enable attackers to craft hyper-personalized messages, conduct real-time deception, and generate realistic deepfake content.

In this context, it is important to review the various factors that enable the design and, ultimately, the early and effective prevention of cyberattacks, particularly by relying on collaboration between technical expertise and human skills.

This paper explores the contemporary landscape of AI-assisted social engineering and introduces a framework for understanding, detecting, and mitigating these threats within online social networks. By integrating behavior modeling, adversarial detection mechanisms, and contextual user profiling, we aim to contribute a scalable and explainable approach to countering socially engineered threats in a digital-first era.

2 Definition of *Social Engineering*

Social engineering, also sometimes called psychological fraud, involves the use of human manipulation and fraud techniques through digital tools. Its emergence in the cyber realm dates back to the 1950s with the appearance of *phreakers*. These users bypassed the functioning of telephone lines of the time to obtain free communications or services that were then paid for [1].

Several elements make up a *social engineering* attack. In addition to the technical means that enable the attack (emails, SMS, phone calls, espionage methods, etc.), there are elements specific to humans, related to their intrinsic functioning (cognitive and behavioral manipulations, judgment heuristics, cognitive biases, emotions, etc.), on which various *social engineering* techniques used by attackers will rely [2] [3].

3 Attacks through social engineering

According to the *Verizon 2023* report, 74% of cyberattacks have a human component and rely on *social engineering* methods (50% of professional email *phishing*) [4]. Humans are then considered the most fragile element in an organization's cyber defense, and the resistance they develop against these attacks varies greatly, depending on the prevention method used. Recent case studies are periodically published on Social Engineering Attacks, such as:

Examples of Social Engineering Attacks[1]: This article lists 15 famous cyber attacks where social engineering was the predominant factor, highlighting how these attacks exploit human error.

Social Engineering Statistics for 2025[2]: This report explores the rise of social engineering attacks, their alarming statistics, and effective strategies to safeguard organizations against them.

Common Types of Social Engineering Attacks[3]: This article discusses ten common types of social engineering attacks and how to prevent them, emphasizing the importance of understanding hackers' tactics.

The use of AI leads to an increase in the quality and quantity of cyberattacks. Identifying the factors associated with these challenges, by combining technical and human skills, would therefore enable the design and, ultimately, the early prevention of cyberattacks (for example, analysis of weak signals, "low-frequency signals, or even non-apparent, but deduced from information or a fact" [5]).

3.1 Different Forms of Attack

There are several forms of *social engineering* attacks. These can involve both the cyber realm and physical exchanges. Some require direct contact with the victim (for example, *phishing*), while others operate indirectly (for example, the use of fake websites).

[1] https://phoenixnap.com/blog/social-engineering-examples

[2] https://sprinto.com/blog/social-engineering-statistics/

[3] https://www.crowdstrike.com/en-us/cybersecurity-101/social-engineering/types-of-social-engineering-attacks/

- *Direct attacks* refer, for example, to *phishing*, *tailgating* (physical passage with a victim – generally behind them, and sometimes without their knowledge – through a secured gate) or face-to-face attacks; indirect attacks refer to watering hole attacks (infection of a service commonly used by members of a group when a frontal attack is not feasible) or *shoulder surfing* (looking over the victim's shoulder) [2] [7] [8] [9] [10].
- *Social engineering* attacks can rely on the functionalities offered by AI (information search in *Open Source INTelligence* (OSINT) [12], creation and orchestration of content dissemination, improvement of message impact, etc.) to increase their own effectiveness, for example, with the quasi-industrialization of personalized *phishing* attacks targeting a specific victim ("*spear-phishing*") [11].

Let us remind the case of this multinational corporation where the CFO received a voice call from the supposed "CEO" urgently requesting a fund transfer. The voice was convincing, matching tone, cadence, and speaking patterns. It was later revealed to be an AI-generated deep-fake, trained using publicly available video interviews. The company lost over $240,000 before realizing the fraud. This incident exemplifies the growing sophistication of AI-powered social engineering and underscores the need for adaptive and explainable countermeasures.

Among all the potential threats, we will mainly discuss points related to humans (without going into more technical aspects), with *pretexting* attacks (use of fraudulent "pretexts" to trap a victim) that directly aim to obtain one or more actions from the victim.

3.2 Phishing

Phishing (or "spoofing") involves sending fraudulent emails aimed at opening an internet link or a corrupted attachment. The link will, for example, redirect to a counterfeit website to retrieve credentials. The attachment will most often contain a "*malware*", a malicious program such as a virus, which encrypts the organization's data until a ransom is paid ("*ransomware*") [2]. The *phishing* email will often contain a message that plays on fear or, conversely, has an appeal.

3.3 Smishing

Smishing, the smartphone cousin of *phishing*, involves sending a fraudulent SMS aimed at opening a link or a corrupted attachment, always with the goal of retrieving information or depositing *malware* on the victim's phone. It also plays on fear or arouses interest [2].

3.4 Vishing

Vishing, another cousin of *phishing*, is a fraudulent voice call (phone call, call via a social network, etc.). The attacker calls their victims under a false pretext to extract information or get them to perform an action (information, transfer, access, etc.) for their benefit [2].

3.5 Social Phishing

Social phishing involves sending a fraudulent message on a victim's social network, with the aim of opening a link or downloading a corrupted attachment, to retrieve information or deposit *malware* on the victim's device. Sometimes, *social phishing* involves *cyber-grooming* (trapping a child online by adults): a bond of trust is established to later open up to *social phishing* [10] [11] [13].

3.6 Quishing or QRishing

Quishing, a more insidious form of *pretexting*, involves physically covering a legitimate QR code with a fraudulent QR code designed by the attackers. As in *phishing*, the objective is to redirect to a site or download a corrupted attachment to retrieve information or deposit *malware* [6]. The cyberattacker targets the customers, for instance in a restaurant: they replace (or cover) some of the QR codes stuck on the restaurant's tables to allow customers to access the menu and place orders. The customers' credentials and payment information are stolen during the attempt to view the menu.

3.7 Cyber-Grooming

Cyber-grooming ("grooming" or "online child enticement") describes the fact that an adult contacts children or adolescents via the internet to ultimately trap them.

To achieve their goal, the adult can employ several methods, such as pretending to be the same age as their victims, building relationships with their loved ones, pretending to be a friend and a good listener. The ultimate goal is almost systematically to lower the victims' guard to lure them into the net of (cyber)sexual abuse or trafficking [14] [15].

4 How Social Engineering Works

Social engineering attacks use social influence techniques. These exploit the usually functional modes of human behavior to turn them to their disadvantage: social needs (social relationships, trust, respect for hierarchy, societal rules, the need to process information quickly, etc.), judgment heuristics (shortcuts sometimes used to judge or make decisions), cognitive biases, emotions, etc.

We will only provide a global overview, but each of the points mentioned, which relate to social and cognitive psychology as well as the exercise of critical thinking, deserves to be developed (see Figures **1**, **2**). Social psychology allows us to focus on social influence [16] [17] and, more particularly, on behavioral manipulation [18] [19].

Social engineering attacks work by trying to short-circuit our analytical analysis capabilities to mislead the victim.

In *cyber-grooming* attacks, certain characteristics guide the choice of the victim: isolation, discomfort, naivety, geographical proximity, social environment, incompetence of a young victim to defend themselves (while adults, more vigilant, to better know the dangers, better detect attempts at seduction or manipulation) [14]. These characteristics are then exploited: personalization of the appeal of the trap, coercion (implicit or explicit), deception, isolation of the victim, gradual sexualization, etc.

Phishing, *smishing*, or *vishing* attacks (whose structure changes relatively little from one to the other) also sometimes use these strategies, but to a lesser extent.

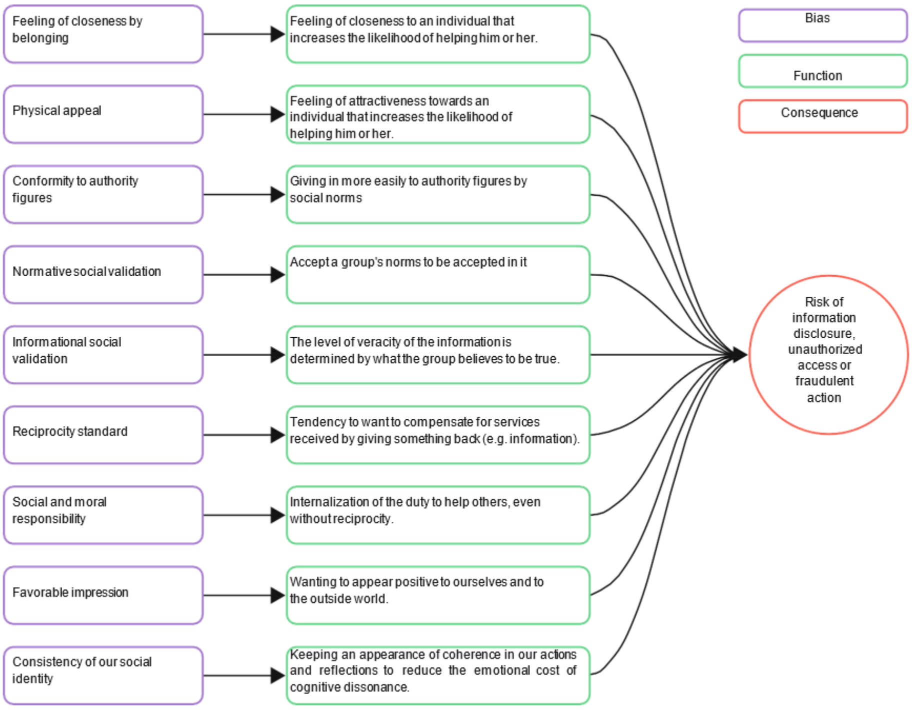
Feeling of closeness by belonging
Feeling of closeness to an individual that increases the likelihood of helping him or her.
Physical appeal
Feeling of attractiveness towards an individual that increases the likelihood of helping him or her.
Conformity to authority figures
Giving in more easily to authority figures by social norms
Normative social validation
Accept a group's norms to be accepted in it
Informational social validation
The level of veracity of the information is determined by what the group believes to be true.
Reciprocity standard
Tendency to want to compensate for services received by giving something back (e.g. information).
Social and moral responsibility
Internalization of the duty to help others, even without reciprocity.
Favorable impression
Wanting to appear positive to ourselves and to the outside world.
Consistency of our social identity
Keeping an appearance of coherence in our actions and reflections to reduce the emotional cost of cognitive dissonance.
Bias
Function
Consequence
Risk of information disclosure, unauthorized access or fraudulent action

Fig. 1 – Some Cognitive Biases and how they function

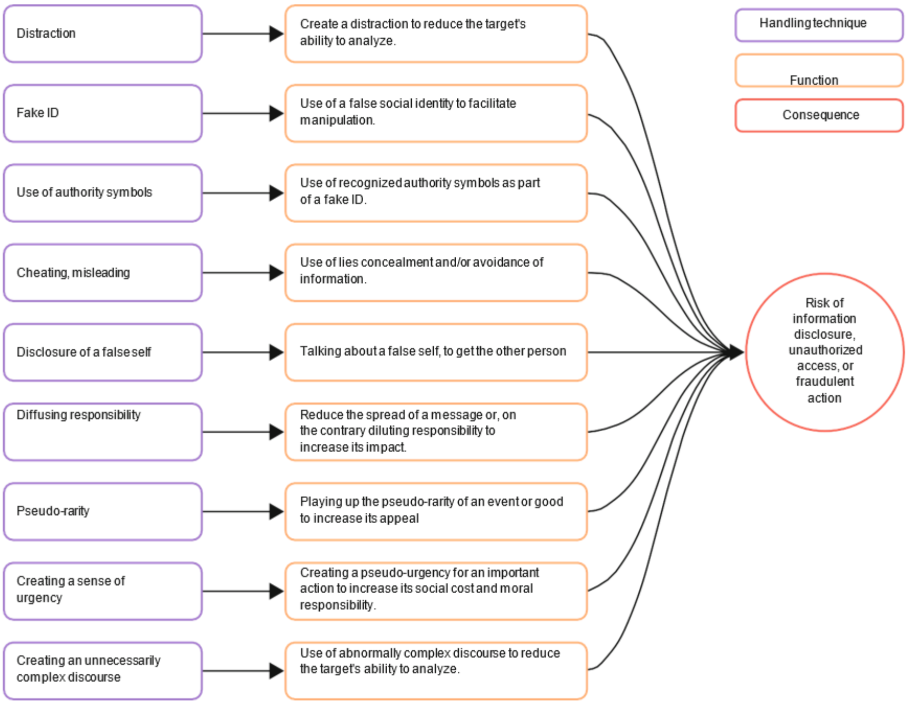
Distraction
Create a distraction to reduce the target's ability to analyze.
Fake ID
Use of a false social identity to facilitate manipulation.
Use of authority symbols
Use of recognized authority symbols as part of a fake ID.
Cheating, misleading
Use of lies concealment and/or avoidance of information.
Disclosure of a false self
Talking about a false self, to get the other person
Diffusing responsibility
Reduce the spread of a message or, on the contrary diluting responsibility to increase its impact.
Pseudo-rarity
Playing up the pseudo-rarity of an event or good to increase its appeal
Creating a sense of urgency
Creating a pseudo-urgency for an important action to increase its social cost and moral responsibility.
Creating an unnecessarily complex discourse
Use of abnormally complex discourse to reduce the target's ability to analyze.
Handling technique
Function
Consequence
Risk of information disclosure, unauthorized access, or fraudulent action

Fig. 2 – Some Manipulation Techniques and how they function

5 Implementing a Social Engineering Attack

According to the cycle described by Mitnick and taken up by Mouton *et al.* (2014), *social engineering* attacks can be broken down into six phases, from their preparation to a possible exit [21]:

- formulation of the attack's goal
- collection and exploitation of available information
- preparation of the attack (definition of the method, the executor, and the scenario)
- start of the attack, creation of the social link, and development of a relationship of trust with the target
- exploitation of the established link to obtain an action from the target, sometimes without their knowledge

- exit from the attack (return to a positive or neutral emotional state to reduce the risk of them raising the alarm, for example).

The number of phases and the respective execution time vary depending on the context and the type of attack chosen. We will illustrate below these different phases in the case of a *cyber-grooming* attack (see Figure **3**).

6 The Use of AI in Social Engineering

Recent advances in generative AI have empowered attackers to create persuasive, context-aware communications at scale. Tools like ChatGPT and LLaMA can mimic human conversational patterns, while deepfake generators clone voices and faces with alarming accuracy. Our framework incorporates AI-generated content detection as a feature, employing text anomaly detection, facial inconsistency scanning, and behavioral pattern mismatches. Furthermore, we discuss the arms race between generative models and detection systems, a human-in-the-loop mechanism remaining mandatory to flag ambiguous cases.

6.1 The Role of AI in Modern Social Engineering Attacks

Contextualizing how AI tools (like LLMs and deepfakes) are being used to automate and scale social engineering techniques (e.g., spear-phishing, impersonation) bridges to our discussion on analyzing online manipulation w.r.t. user's behavior. Explaining how generative AI helps craft tailored phishing or impersonation attempts, referencing the increasing realism of deepfake media, is essential in terms of explanability and acceptability.

Recent developments in artificial intelligence have significantly amplified the effectiveness and scale of social engineering attacks. AI-powered systems can generate highly personalized phishing messages, deepfake audio or video to impersonate trusted individuals, and even simulate real-time conversations using large language models. These capabilities enable attackers to exploit psychological vulnerabilities with greater precision and automation, reducing the effort and expertise required for successful manipulation [35] [36].

For instance, generative AI can scrape and analyze social media data to craft context-aware phishing emails that are nearly indistinguishable from legitimate communications [34]. Additionally, tools leveraging deep learning have been shown to mimic voices and video mannerisms, which can deceive even cautious targets during high-stakes operations [37]. The increasing accessibility of such tools poses a substantial threat to both individual and organizational cybersecurity.

6.2 Forms of AI Used

Social engineering, also sometimes called psychological fraud, involves the use of human manipulation and fraud techniques through digital tools. Its emergence in the cyber realm dates to the 1950s with the appearance of phreakers. These users bypassed the functioning of telephone lines of the time to obtain free communications or services that were then paid for [1].

Several elements make up a social engineering attack. In addition to the technical means that enable the attack (emails, SMS, phone calls, espionage methods, etc.), there are elements specific to humans, related to their intrinsic functioning (cognitive and behavioral manipulations, judgment heuristics, cognitive biases, emotions, etc.), on which various social engineering techniques used by attackers will rely [2] [3].

Social engineering sometimes relies on AI techniques that enable the creation of fraudulent content, data retrieval ("*scraping*"), and analysis, communication using a conversational robot ("*chatbot*"), or even the coordination and evaluation of the different phases of the attack. These are sometimes supported by ML, based on mathematical and statistical approaches to give computers the ability to "learn" from data [22].

Tools like *ChatGPT* or *Bardes* are misused, and illegal tools like *FraudGPT* or *WormGPT* are used [23].

6.3 AI-Enhanced Pretexting

The arrival of generative AI facilitates the generation of falsified content and the organization of mass attacks. These expand the spectrum of potential targets, considerably reducing the cost and potential harm of a *pretexting* attack (*phishing*, *vishing*, etc.).

For example, Microsoft's *VALL-E* is capable of cloning a voice from just three seconds of recording, facilitating the impersonation of a vocal identity during *vishing*.

Other forms of AI automate and personalize *spear-phishing* attacks [23].

6.4 AI in the Different Phases of an Attack

AI techniques are most often used in phases 2 to 5 (see Figure **3**).

For example:

- the target's digital footprint (interests, affiliations, behaviors, etc.) makes it possible to develop the best strategy for the attack (phases 2 and 3) [11]
- generative AI creates specific content and *malware* (phases 4 and 5) [24]

- a *chatbot* establishes bidirectional communication with the victim, adapting in real-time to their reactions to lead them to perform a harmful action for themselves (phases 4 and 5) [11]
- the goal is to obtain information or unauthorized access
- the use of deception is almost systematic
- attacks should be analyzed from the perspective of social and cognitive psychology, based on knowledge and skills that can be developed in the field of critical thinking.
- the use of AI makes it possible to facilitate their industrialization and increase their precision.

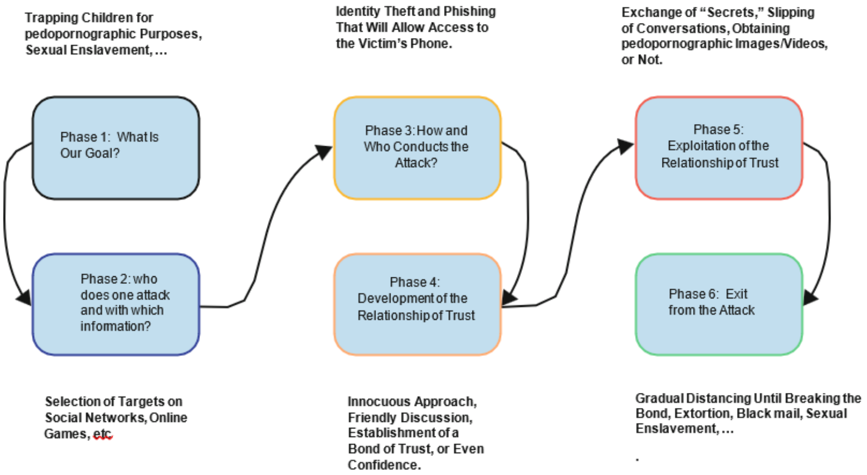
Trapping Children for pedopornographic Purposes, Sexual Enslavement, …
Identity Theft and Phishing That Will Allow Access to the Victim's Phone.
Exchange of "Secrets," Slipping of Conversations, Obtaining pedopornographic Images/Videos, or Not.
Phase 1: What Is Our Goal?
Phase 3: How and Who Conducts the Attack?
Phase 5: Exploitation of the Relationship of Trust
Phase 2: who does one attack and with which information?
Phase 4: Development of the Relationship of Trust
Phase 6: Exit from the Attack
Selection of Targets on Social Networks, Online Games, etc.
Innocuous Approach, Friendly Discussion, Establishment of a Bond of Trust, or Even Confidence.
Gradual Distancing Until Breaking the Bond, Extortion, Black mail, Sexual Enslavement, …

Fig.3 – Phases of a *Social Engineering* Attack. The Example of *Cyber-Grooming*

7 Existing Countermeasures and Effectiveness

7.1 Countermeasures

There is a multitude of countermeasures, with sometimes significant differences in effectiveness, from AI-based detection attempts to user prevention and training campaigns, including posters placed everywhere: prevention messages, preventive *phishing*, explanatory *phishing* "vaccination" (for example, *PhishGuru*), group training, use of *serious games*, etc. [25]. Understanding their advantages and disadvantages (context of use, effectiveness, limits, sustainability over time, etc.) makes it possible to choose the best tools and maintain user engagement.

Below, we will only address countermeasures that involve a human aspect, with their weaknesses and blind spots.

7.2 Decision Trees

There are decision trees to guide users in detecting *social engineering* attacks [24]. However, these are limited to *pretexting* attacks.

It is impossible to apply them to *shoulder surfing*, and under the attacker's pressure, the victim does not have the mental availability to refer to the decision tree in a situation of uncertainty.

7.3 Organizational Processes

For sensitive operations, some cybersecurity policies implement strict processes that rely on other members of the organization or identity control tools [27].

For example, the validation of a transfer order is done by two people who identify themselves and validate the action.

Non-user-friendliness, risk of *bypass*, and complexity of application for individuals or SMEs/craftsmen limit generalization and effectiveness.

7.4 Combining Human and AI Capabilities

Relying on the knowledge and experience of an organization's seasoned users to detect *pretexting* attacks early allows feeding and training a defense AI that is responsible for recognizing these attacks (or similar versions) throughout the organization [26].

The operationalization time of these defense AIs is too long, which complicates their use for individuals or SMEs/craftsmen.

7.5 User Prevention

Prevention based on training upon arrival in the company, once a year, or after failing a *phishing* test is effective but only for a few weeks [29] [30] [31]. According to Junger *et al.* (2017), preventive warning messages have almost no effect, and the increase in cyber defense skills is not automatically correlated with the number of preventive training sessions [29].

Given the costs (of training, immobilization of users and premises, etc.), multiplying sessions is therefore not the most efficient. The most effective preventions over time are those that are playful, provide immediate *feedback*, dynamic exchanges, or micro-training carried out at short intervals [30] [31].

Evaluating users' susceptibility, misinformation, or deception, it is essential to discuss the difficulty of detecting AI-driven social engineering compared to traditional methods, and how detection tools need to evolve alongside generative models.

7.6 Recommendations

Junger *et al.* (2017) [29] and Buller and Junger (2020) [31] recommend, for prevention against *social engineering*: training everyone without exception, regular training, dynamic interventions (illustrations, videos, verbal exchanges, games, etc.), immediate *feedback* for better understanding, and knowing how to decipher a URL.

However, this is not easy when at least **47%** of digital interactions are on smartphones, professional or personal [32], and the methods of combating this consist, for example, of verifying URLs with a mouse pointer. Prevention against *fake news* reduces the risk of a potential compromise of employees (voluntary data leakage, external collaborations, etc.) following an attack on an organization's reputation.

Most of the available training focuses on *phishing* because it is the main attack vector.

It is materially impossible to train users on how to thwart all existing attacks, all the more with the new AI-generated ones that are still unknown. Training must therefore be coupled with organizational measures, such as security policies (password policy, access management, etc.), security processes (two-factor authentication), or technical measures (access restrictions, confidentiality filters, AI-enhanced antivirus, etc.). The developed countermeasures must be subject to a continuous evaluation process to be improved and remain in line with advances in the field and adapt to new threats. A set of recommendations has been proposed by F. Sèdes *et al.* in [33].

Encouraging the mobilization and development of critical thinking is also crucial:

- building a solid general culture
- knowledge of rhetorical forms related to fallacious thinking
- study of cognitive biases and reasons why we adhere to harmful false beliefs, etc.

These knowledge and skills strengthen and enrich prevention against *social engineering* [27].

8 Challenges and Directions

Training users (and why not deploying it on a general population scale) represents a primary challenge in the context of *social engineering* prevention. Research efforts must therefore be continued to produce efficient training tools that can be transposed to less common attack modalities than simple *phishing*.

Understanding the necessary cross-cutting skills is also crucial to enable users to resist attacks that are still unknown. Anticipating future attack vectors will make it possible to cut the grass under the feet of cybercriminals, that means:

- Designing an effective public policy against *social engineering*
- Making existing training more efficient
- Sustaining acquired knowledge and skills
- Training users on different attack vectors.

Beside accuracy, testing and performance assessment, new challenges are opened with broader impact on ethical compliance. The proliferation of AI-enhanced social engineering raises significant issues: while defensive tools grow more sophisticated, so do the offensive capabilities. There is a fine line between proactive surveillance and privacy infringement when analyzing communication behavior. Ethical use of users' data vs. privacy, transparency in algorithmic decisions, and mechanisms for user's consent are crucial to maintaining trust. Our framework advocates for explainability, user control, and compliance with regulatory standards (e.g., GDPR, CCPA) to ensure that defensive strategies uphold ethical principles while maintaining efficacy.

9 Conclusion

Properly trained users could shift from the status of "weak links" to that of valuable assets, capable, for example, of detecting anomalies that are little or not known by digital tools. Including human factors in a cybersecurity policy could increase the overall level of security.

Efforts to combat AI-enhanced social engineering must include real-time anomaly detection, adversarial training of communication platforms, and education campaigns focused on emerging tactics. As the boundary between human and machine-generated deception blurs, it is imperative to adopt adaptive defenses and regulatory frameworks that evolve in step with these technological advancements.

To succeed in this challenge, transdisciplinary collaborations would be essential, between computer science and human and social sciences mainly, as well as inter-organizational collaborations (public, private, large groups, SMEs/VSEs, etc.).

Future research should explore adaptive defenses against AI-driven social engineering, particularly in the context of social platforms and public available personal data, harvested via OSINT tools, personalized content generation and increasingly sophisticated identity spoofing.

References

[1] HATFIELD (J.M.). – *Social engineering in cybersecurity: The evolution of a concept*. Computers & Security, vol. 73, pp. 102–113 (2018).
[2] WANG (Z.), ZHU (H.), SUN (L.) *et al*. – *Social Engineering in Cybersecurity : Effect Mechanisms, Human Vulnerabilities and Attack Methods*. IEEE Access, vol. 9, pp. 11895-11910 (2021). https://doi.org/10.1109/ACCESS.2021.3051633
[3] SEDES (F.), DEGRACE (J.). – *Sécurité dans les SI & social engineering – Un état des lieux,* INFORSID (2024). https://hal.science/hal-04613192/document
[4] VERIZON. – *Rapport d'enquête sur les violations de données* (2024).
https://www.verizon.com/business/resources/reports/dbir/2023/summary-offindings/
[5] ANSOFF (H.I.). – *Managing strategic surprise by response to weak signals*. California management review, 18(2), pp. 21-33 (1975).
[6] RUSSELL (S.J.), NORVIG (P.). – *Artificial Intelligence : A Modern Approach*. Pearson Education, Paris France, 3e éd. (2010).
[7] AFFAN (Y.), RUBIA (F.) *et al*. – *Contemplating social engineering studies and attack scenarios : A review study*. Security and privacy, vol. 2, p. 73 (2019).
[8] PARTY (P.P.), RAJENDRAN (G.). – *Identification and prevention of social engineering attacks on an enterprise*. International Carnahan Conference on Security Technology (ICCST), IEEE, pp. 1-5 (2019). https://ieeexplore.ieee.org/abstract/document/8888441
[9] KAUSHALYA (K.M.H.S.) *et al*. – *An overview of social engineering in the context of information security*. IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS), pp. 1-6 (2018).
[10] DE CASANOVE (O.), SEDES (F.). – *Malicious Human Behaviour in Information System Security: Contribution to a Threat Model for Event Detection Algorithms. Foundations and Practice of Security*. 15th International Symposium, FPS 2022, Ottawa, ON, Canada, December 12-14, pp. 208-220 (2022).
[11] SCHMITT (M.), FLECHAIS (I). – *Digital Deception Generative Artificial Intelligence in Social Engineering and Phishing*. SSRN Electronic Journal, vol. 57, n° 54 (2023).
https://link.springer.com/article/10.1007/s10462-024-10973-2
[12] WANG (Z.), ZHU (H.), SUN (L.). – *Defining Social Engineering in Cybersecurity*. IEEE Access, 8, pp. 85094–85115 (2020). https://www.researchgate.net/publication/341199647_Defining_Social_Engineering_in_Cybersecurity
[13] ALBLADI (S.M.), WEIR (G.R.S.). – *User characteristics that influence judgment of social engineering attacks in social networks*. Human Centrics Computing and Information Sciences, 8:5 (2018). https://doi.org/10.1186/s13673-018-0128-7
[14] WINTERS (G.M.) *et al*. – *Validation of the Sexual Grooming Model of Child Sexual Abusers*. Journal of Child Sexual Abuse, pp. 855-875 (2020).
[15] RINGENBERG (T.R.), SEIGFRIED-SPELLAR (K.C.), RAYZ (J.M.), ROGERS (M.K.). – *A scoping review of child grooming strategies: pre and post internet*. Dans Child Abuse & Neglect, vol. 123 (2022).
[16] BEGUE (L.), DESRICHARD (O.). – *Traité de psychologie sociale*. De Boeck (2024).
[17] CIALDINI (R.). – *Influence et manipulation*. Pocket (2024).
[18] JOULE (R.V.) BEAUVOIS (L.). – *Petit traité de manipulation à l'usage des honnêtes gens* (2024). https://www.education-authentique.org/uploads/PDF-DOC/JBM_Trait%C3% A9_ manipulation_JouleBeauvois.pdf

[19] FALKOWICZ (S.) *et al*. – *Au coeur de l'esprit critique. Petit guide pour déjouer les manipulations*. Eyrolle, pp. 55-98 (2023).
[20] CHO *et al*. – *Effect of personality traits on trust and risk to phishing vulnerability: Modeling and analysis*. Dans IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), IEEE, pp. 7-13 (2016).
[21] MOUTON (F.), MALAN (M.M.), LEENEN (L.) *et al*. – *Social Engineering Attack Framework*. Information Security for South Africa (ISSA), IEEE, pp. 1-9 (2014).
https://ieeexplore.ieee.org/document/6950510
[22] BARRA (V.), CORNUEJOLS (A.), MICLET (L.). – *Apprentissage Artificiel : Concepts et algorithmes*, Eyrolles (2021).
[23] FALADE (P.V.). – *Decoding the Threat Landscape : ChatGPT, FraudGPT, and WormGPT in Social Engineering Attacks*. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 9, n° 5 (2023).
https://ijsrcseit.com/paper/CSEIT2390533.pdf
[24] GUPTA (M.), AKIRI (C.), ARYAL (K.). – *From ChatGPT to ThreatGPT : Impact of Generative AI in Cybersecurity and Privacy*. IEEE Access, vol. 11, pp. 80218-80245 (2023).
https://ieeexplore.ieee.org/document/10198233
[25] SYAFITRI (W.), SHUKUR (Z.) *et al*. – *Social Engineering Attacks Prevention: A Systematic Literature Review*. IEEE Access, 10, pp. 39325 – 39343 (2022).
https://www.researchgate.net/publication/359528837_Social_Engineering_Attacks_Prevention_A_Systematic_Literature_Review
[26] MOUTON (F.), LEENEN (L.) *et al.* – *Social Engineering Attack Detection Model: SEADMv2*. International Conference on Cyberworlds (CW), pp. 216–223 (2015).
https://ieeexplore.ieee.org/document/7398418
[27] SOSAFE. – *Le point sur les menaces et les bonnes pratiques en cybersécurité* (2024).
[28] BURDA (P.), ALLODI (L.), ZANNON (N.). – *Don't Forget the Human: a Crowdsourced Approach to Automate Response and Containment Against Spear Phishing Attacks*. IEEE Euro. Symp. on Security and Privacy Workshops (EuroS&PW), pp. 471–476 (2020).
https://ieeexplore.ieee.org/document/9229829
[29] JUNGER (M.), MONTOYA (L.), OVERINK (F.-J.). – *Priming and warnings are not effective to prevent social engineering attacks*. Computers in Human Behavior, vol. 66, pp. 75-87 (2017).
[30] KUMARAGURU (P.), SHENG (S.), ACQUISTI (A.) *et al*. – *Teaching Johnny Not to Fall for Phish*. Dans ACM Transactions on Internet Technology, 10(2), pp. 1-31 (2010). https://www.researchgate.net/publication/220169843_Teaching_Johnny_not_to_fall_for_phish
[31] BULLEE (J.-W.), JUNGER (M.). – *How effective are social engineering interventions? A meta-analysis* Information & Computer Security, vol. 28, pp. 801-830 (2020).
[32] BAROMÈTRE DU NUMÉRIQUE. – *Enquête sur la diffusion des technologies de l'information et de la communication dans la société française* ARCEP, p. 8 (2022).
[33] DE CASANOVE (O.), LELEU (N.), SEDES (F.). *Applying PDCA to Security, Education, Training and Awareness Programs*. 16th IFIP WG 11.12 International Symposium on Human Aspects of Information Security and Assurance (HAISA 2022), IFIP TC 11 Working Group 12: Human Aspects of Information Security and Assurance, pp. 39-48 (2022).
[34] Whittaker, Z., Leetaru, K., & Goodman, B. (2021). *The Emerging Threat of AI-Powered Social Engineering*. ACM Digital Threats: Research and Practice.
[35] Kumar, A., & Carley, K. M. (2023). *AI-Driven Manipulation on Social Media: A Survey of Techniques and Defenses*. Journal of Artificial Intelligence Research, 76, 451–478.
[36] Brundage, M., et al. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. arXiv:1802.07228.
[37] Vincent, J. (2020). *Deepfake videos could be used to interfere in elections, warn experts*. The Verge.