AI Security in Social Engineering: Mitigating Risks of Data Harvesting and Targeted Manipulation

OLADOYIN AKINSULI
AI and Cybersecurity Strategist, School of Computer Science and Electronic Engineering, University of Surrey, Guildford, UK

Abstract- ***This paper critically reviews the importance of applying AI to boost security features on social media platforms to protect against unethical data collection and uses of psychological influence for political and business purposes. Most importantly, the abovementioned AI and machine learning experiences have become pervasive, revolutionizing the users' experiences and creating complicated means for developing malicious intents to track personal information and influence consumer experiences. Here, we also look into the existing artificial intelligence-based security systems like encryption, anomaly detection, and user authentication to understand their efficiency in protecting user data privacy. These worries will probably materialize in the following ways when artificial intelligence capabilities like facial recognition and natural language processing become more widely used. We discuss real-world examples to illustrate the repercussions of manipulating AI: the Cambridge Analytica case and Meta's Ray-Ban augmented reality glasses. Thus, the paper explores the ethical challenge of AI misuse before outlining the importance of achieving the best mix of security benefits and concerns for users' freedom and privacy. Thus, we proposed a four-part solution aimed at some of those threats, which include increased data protection, ethical usage of the created AI framework, and general user education. Policymakers, social media companies, and designers of these A. Systems can benefit from our conclusions by making sure the latter protects users of such platforms from abuse.***

Indexed Terms- ***RPA Systems, Artificial intelligence, social networks, NLP, Facial Recognition***

I. INTRODUCTION

1.1 Background to the Study

Artificial intelligence solutions are integrated with social networks: at first, they were created to enhance the functionality of social networks through the analysis of user preferences and trends, advertisements and recommendations, and results. It is worth noticing that these AI systems are based on algorithms that examine a significant portion of users' information to deliver content corresponding to their preferences (Gomez-Uribe & Hunt, 2016). The recommendation system that Netflix uses will be another example of how AI involves ultra-individualization through the forward calculation of user expectations based on previous use. However, this advanced use of AI makes internet users experience personalization with the flip side of security negligence.

The influential social engineering, which relies on tricking people into sharing some secret or personal data, has also benefited from the use of artificial intelligence. Like in pretexting, phishing, and vishing, the attacks are aimed at exploiting the psychological vulnerabilities of a user when using digital platforms. By their design, AI algorithms gather and process significant amounts of personal data that the soil can manipulate. This has brought exponential growth in data scraping and data fabrication cons in social media networks (Kosinski, Stillwell, & Graepel, 2013). Such risks are made worse as AI algorithms can be designed to prioritize particular types of content to sway users' opinions, behaviors, and political affiliations.

Out of all the current trends in the application of AI and social media security, using artificial intelligence to harvest large amounts of data is the most worrying. This entails gathering individual data, sometimes without the user's consent, and then using this information to advertise or manipulate. Kosinski et al. proved in their study Kosinski Stillwell and Graepel (2013) that it was possible to predict private characteristics and parameters, including personality

characteristics, based on the activity on social networks. This ability to reason about other people's traits can make it useful in commanding user data for commercial and political manipulation (Acquisti, Brandimarte, & Loewenstein, 2015).

Apart from data gathering, these botnets use Artificial Intelligence algorithms to modify the information users get to see on social networks. It can be done through recipes, which push content regarding users, their actions, and tendencies. Although such algorithms are generally employed to improve the customer's experience, they have also been utilized to control opinion. For instance, in the recently concluded 2016 U.S. presidential election, intelligent algorithms were used to post intended voters to control agency with a political theme (Cadwalladr, 2018). Still, this type of manipulation can be considered to belong to the general concept of psychographic targeting, which implies the use of artificial intelligence algorithms in political activities, which in turn underscores the question of the ethical utilization of artificial intelligence in social networks.

Among other things, natural language processing (NLP) and facial recognition by AI have stirred great controversy around protecting individuals' rights' to privacy and data security. Specifically, facial recognition technology allows AI to recognize people in the public domain without permission (Acquisti et al., 2015). There are several issues that ethical concerning this type of technology, as it can entail doxxing, surveillance, and profiling people based on looks alone. As these AI tools are effective in security and convenience, they have high potential threats when on the wrong side.

The negative impacts of applying AI for purposes of data collection and mind control are enormous, especially when and where personal data is used for monetary gain or vote-getting cover. Zuboff (2019) noted that, with the arrival of surveillance capitalism, users' information is bought and sold like any common good. In my view, this aspect of selling the data points forms the foundation of most ethical issues about AI because users rarely decide how their data gets utilized. The problems with AI-created social engineering are not only about privacy; they are about democracy, too, because AI, in its essence, can be used to manipulate the opinions of millions of people (Zuboff, 2019).

1.2 Overview

The basic necessity of data privacy has become cumbersome due to social media handling companies' increased use of Artificial intelligence (AI) and machine learning (ML) tools. Though those technologies improve user experience through individualized content and recommendations, they threaten data harvesting and analysis for ill intentions (O'Neil, 2016). AI algorithms can estimate user behavior and preferences with high efficacy; therefore, the idea of surveillance and manipulation arises.

Recent occurrences show that such problems are serious. For instance, facial recognition technology in consumer devices has raised controversies regarding the infringement of human rights on privacy. Facial recognition through AI can recognize people against their wishes, thus opening the door for unauthorized data mining and revealing customer's identity (Garvie, Bedoya, & Frankle, 2016). In China, police wear facial recognition glasses to scan people to find suspects, and the potential for invasive surveillance has been seen (Mozur, 2018). These applications highlight what may happen when AI is incorporated into conventional household devices.

Furthermore, AI algorithms have been applied to direct peoples' opinions on worrying social media norms. The AI algorithm employs two major types of accounts – automated bots and fake accounts- to spread propaganda development and misinformation, influence politics, and influence election processes (Howard & Kollanyi, 2016). Both of these strategies use the selective bias of the artificially intelligent content showing curation advantage compared with other ordinary content posts to mislead users into believing that the scraped content is authentic.

This paper aims to outline and assess methods by which AI can supplement protection against such violations. Thanks to the opportunity to use AI functions in our system, we can create a highly effective layer of protection for users' data and personal information. This includes promoting the ethical use of artificial intelligence based on such factors as transparency of artificial intelligence and

data ownership and consent from users, developing AI early warning systems to detect scams like deepfakes, and other forms of bot-coordinated disinformation (Floridi et al., 2018). The goal is to achieve increases in customer satisfaction with software products and to ensure the strengthened resistance of social networks to new issues through applying AI.

1.3 Problem Statement

The negative utilization of AI designed for the harvesting of users' data and their control is another threat to user privacy and freedom. Social networking sites are specially targeted to faucet numerous information regarding their users, and bad people are utilizing AI technologies without the user's consent (Acquisti, Brandimarte, & Loewenstein, 2015). This data can change existing behavioral patterns public opinion, and even blackmail people through highly specific advertisements, false news, or social engineering.

Perhaps one of the most problematic features of this problem is that algorithmic decision-making might lead to psychological manipulation of users based on their online activity, meaning that people's behavior might be predicted and steered in certain directions for the sake of greater profit (Kosinski, Stillwell, & Graepel, 2013). As social media platforms remain a constantly changing market, the instruments and techniques utilized in AI manipulation remain highly tested and designed so that the existing security measures cannot adequately shield a user. As seen with recommender engines or facial recognition software, AI-based systems can be easily manipulated by bad actors to gain personal data for commercial or political purposes (Isaak & Hanna, 2018).

The existing security measures and privacy paradigms fail to counter the vast number and varieties of these AI that have effectively facilitated dangers. Most emerging platforms must be adequately protected against various social engineering techniques that use AI technologies. Therefore, the users are exposed to a large number of violations of their privacy, such as identity theft, psychological manipulation, and prohibited profiling. This paper aims to fill these gaps by discussing AI's implementation in ensuring security to protect against social engineering attacks.

1.4 Objectives

1. To perform risk analysis on the AI technologies in social media, especially regarding unethical data collecting and psychological persuasion.
2. To design and recommend AI-based security tools to manage these risks and prevent them from compromising user privacy and security.
3. To evaluate the current advances in AI, particularly in security systems, to curb the use of social engineering to search for gaps in the technologies to advance.

1.5 Scope and Significance

This research concerns ethical considerations in deploying artificial intelligence (AI) in social networks to curb prejudice data mining and psychological control. Thus, according to the objectives set in the paper to present an overview of the current AI threats, the author deconstructs various security risks and opportunities offered by AI technologies to show how they are interconnected. This research focuses on identifying the processes through which AI is utilized maliciously for social engineering, examining the currently implemented AI security solutions, and defining novel, AI-based security approaches to protect users' privacy and personal information.

The importance of this study is based on the relevance it can offer to a wide range of readers willing to develop and maintain safer and more ethically sustainable AI-driven platforms: policymakers, social media platforms, and AI researchers. Thus, as this research compares the current security measures to various types of social engineering, examines the ethical misuse of AI, and proposes how to prevent more complex types of social engineering attacks in the future, it aims to provide methods for constructing effective protection against them.

II. LITERATURE REVIEW

2.1 AI and Social Media Security

AI has played a massive role in social media platforms because it makes them as unique as they are personal. AI-enabled algorithms assist in information filtering and recommendations, product targeting, and interaction optimization by utilizing mass amounts of personal information (Zuboff, 2019). This capability

enables social media firms to place special ads and feed users with content that attracts their attention. Nevertheless, the same algorithms that provide these three uses are equally capable of being used unethically, such as data mining and psychological control.

This paper identifies four major applications of AIS in social media, with the first being the primary use of the collection and analysis of user data. Social media apps use the AI algorithm to monitor interactions, engagement with posts, time spent on particular content, and overall user behavioral analysis. This information is then fed into AI models to reconstruct the individual user's behavior trajectory, interests, and potential emotions (Kosinski, Stillwell, & Graepel, 2013). Although this created an avenue for increasing competitiveness to enhance the user experience in the platform, this is also a loophole in which exploitation thrives. For example, data gathered can be used to profit by selling it to third parties who may use it for biased advertising and presenting information that will change how users think or act.

The idea known today as surveillance capitalism, outlined by Shoshana Zuboff, details how private information is turned into a product to make money (Zuboff, 2019). According to Zuboff, AI technologies are at the heart of this practice, undesigned and unconsented company surveillance. The way surveillance capitalism works is that human data is mined with the help of artificial intelligence technologies and then rendered valuable; furthermore, it is utilized for advertising or resold to other companies. Using personal data as a source of profit raises several ethical issues concerning user privacy and the independence of persons in managing personal data.

Apart from the ethical delegation of data harvesting, AI-based algorithms also require certain hazards in psychological manipulation. Specifically in Social media, AI controls the kind of content users will be exposed to through features such as trending that pumps in emotional or sensational topics to increase user engagement (O'Neil, 2016). For instance, AI can amplify political material with extreme political tendencies and thus cause a split and manipulation of society. Such manipulations were perceived during the US presidential campaign in 2016, in which the artificial intelligence bots were operating in distributing fake news and marshaling the electorate (Howard & Kollanyi, 2016). The capability of AI to adjust those experiences on this massive scale is the result of further possibilities for social engineering attacks based on the weak spots of human psychology. Moreover, facial recognition technologies based on AI are causing shocking concerns about privacy and security. Some major social media apps and policeman facial recognition are that algorithms can track people without knowing they are being followed in public places (Garvie, Bedoya, & Frankle, 2016). This technology is not only a data collection instrument but also a surveillance tool that can be used by the wrong hands to monitor victims. The absence of regulation regarding the application of facial identities of millions of citizens to FR technologies increases the dangers inherent in AI surveillance.

Still, AI can also improve social media security owing to its capabilities. AI solutions are being worked on to identify data theft and unauthorized access, safeguard the accounts from hacking, and curtail the hazards of destructive bots and deepfakes. Such conventional security measures as using anomaly detection systems based on artificial intelligence allow platforms to prevent the activity of threats and respond to them immediately (Florida et al., 2018). However, These security developments are mainly root cause and aspiring rather than imperative and prescriptive. Consequently, the threats are already present when the opportunity to exploit them is recognized.

2.2 AI in Data Harvesting

In the last few years, fields such as machine learning and AI tools have contributed immensely to the massive scraping of data on social media platforms. It also enables procurement and analysis of a large amount of information from the users' feedback, thus enabling the extraction of personal information in an unprecedented way. Third-party players and social media use AI to collect data for advertising and product recommendations, among other uses often beyond the user's control or consent.

AI data collection entails employing intelligent programs to anticipate private characteristics and properties from some records. Kosinski, Stillwell, and

Graepel showed that subjects' simple digital footprints, such as Facebook likes, could predict sexual orientation, ethnicity, religion, politics, personality and intelligence, well-being, and substance use. They also pointed out that with AI, privacy concerns remain since the information left behind by users can be in a certain way even if the users do not disclose it (Kosinski, Stillwell, & Graepel, 2013).

Most social networking sites have incorporated artificial intelligence algorithms to gather user data for improvement. The basic algorithms of AI enable the collection of interactions, preferences, and levels of engagement. With the help of the collected data, AI can build rich user proa files and user profiles that advertisers use to place relevant ads. It optimizes advertising effectiveness but, as a rule, implies persistent monitoring of users' actions without sufficient disclosure of the process or receiving their consent (Turow, Hennessy, & Draper, 2015). The fact that exercising user control is challenging due to the unawareness of the scale of data being collected increases privacy issues.

Third-party data brokers also have an essential function in the use of AI to gather data. It collects information from social networking sites, e-transactions through credit cards, and records, among others, with the aim of compiling a dossier for people. These profiles are then sold to marketers, insurers, employers, and all other market-interested parties. Although it might sound a bit twisted to anyone considering the lots of personal information that is usually uploaded to these websites and profiles, AI helps data brokers efficiently sort a vast array of data so that the data becomes even more significant and the profiling is even more intrusive (Christl, 2017). Such monetization of personal data may result in discrimination and threaten users' privacy.

In addition, such tools help gather data for other objectives besides advertising. AI has been used by governments & police to monitor social media platforms for surveillance and security. Legal actions may include identifying illicit and suspicious activities or threats by users through the intelligent recognition of AI algorithms. However, this raises concerns about the ethical implications of mass surveillance and authorities' infringement of people's rights (Lyon, 2014).

Other data harvesting techniques include Natural Language Processing (NLP) and sentiment analysis, all of which are subcategories of artificial intelligence. These applications are employed for opinion mining, a process of mining opinions, sentiments, and attitudes from text data on social network sites. All the ideas and information produced with the help of these tools belong to market intelligence, brand analysis, and reputation control. On the one hand, NLP gives businesses a lot of useful insights, but on the other hand, it handles user messages that probably should not be analyzed this deeply (Liu, 2012).

For instance, AI has been used in the political domain in predictive management and voting manipulation through micro-management. This way, political campaigns will be able to remarket messages to certain audiences as informed by the prediction of the preferences and beliefs of every population. This was expressed in the 2016 US presidential election, where data analytics was instrumental in political strategizing (Howard, Woolley, & Calo, 2018). The two aforementioned advantages of AI, in this case, are perceived as manipulation, fake news, and, most importantly, as an attack on democracy and democratic institutions.

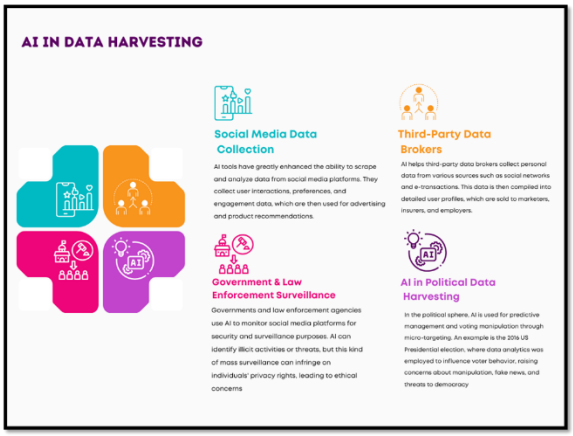


Fig 1: An image illustrating AI in Data Harvesting

2.3 AI-Driven Psychological Manipulation

AI is now being used to deliver appropriate content to influence the behavior of users, especially in political campaigns and commercial matters. This technology uses advanced mathematics to evaluate large amounts of information from different social networks and other types of Internet resources to predict certain actions and reactions of users. In this way, AI systems can effectively post content expected to change people's perceptions and decision-making processes. The use of these techniques in political processes, observed during important events, such as the Brexit referendum, shows how great the incorporation of AI in the management of democratic countries can be.

One of the best-known instances of psychological manipulation through Artificial Intelligence is connected with the British exit from the European Union in 2016. The scandal in which political consultancy firm Cambridge Analytica was implicated concerned the acquisition of data from 50 million Facebook members, and was used to create psychological models. These profiles augmented the process of effective persuasion of the voters in a specific direction by using advertisements that are mainly focused on changing the minds of targeted beholders. Carole Cadwalladr exposed how this method influenced people and altered the course of the referendum (Cadwalladr, 2018). By creating user profiles of psychological traits, CA could use AI to seek and deliver messages that the individual psyche would find emotionally appealing.

Again, the AI capabilities to influence user actions are not just limited to politics and wallets. Social media employs artificial intelligence to tailor the content recommended to users through various media platforms to appeal to user interests and provide content relevant to user desires. Though this is the case, customization is normally done with specific business interests in mind, such as enhancing users' engagement for advertising payoffs. Activations align topics that would generate traffic since the system prefers content that elicits massive responses as it triggers usage and further engagement. This algorithmic strategy is effective based on engagement because AI systems can identify what content will bring and keep the users engaged. At the same time, they are exposed to particular advertisements (Tufekci, 2014).

One more example of commercial use of AI psychological manipulation occurred on Facebook in 2014: in this case, nearly 700,000 users were manipulated emotionally without their knowledge to analyze the impact of emotional news on their activity. The study established that users exposed more to positive or negative content were also likely to change their posting behavior corresponding to the feelings exposed (Kramer, Guillory, & Hancock, 2014). This experiment demonstrates how and to what extent ordinary AI systems can determine users' emotions and behaviors to the extent requisite for enhancing engagement.

They conclude that besides political campaigns and an array of commercial appeals, AI's ability to alter content is equally useful in diffusing fake news and other abrasive messages. However, the algorithms in social networks' hands are far from being impartial. They are designed to show the content that results in users' activity, including the spread of fake news. This dynamic makes spreading prejudicial and tendentious content possible since accurate information is violated during elections or crises (Lazer et al., 2018). Thus, it would be unwise to speak about AI as the provider of a stock of information affecting the perceptions and decisions of populations.

Freewill can seriously be questioned, given the idea that psychological manipulation is on its way to being handed over to artificial intelligence. Since AI learns about people's tendencies and then follows them, the platforms can control the user in a way that questions the concept of free will. Suppose the AI systems are engineered as click-farms or data-mining agents that affect behavior without notice or permission. In that case, the users have no idea how they are being manipulated. This raises quite alarming issues about the authority of behavioral change by AI, whether personal or social.

2.4 AI and Targeted Manipulation for Political Gains

AI as the tool to advocate specific opinions has recently emerged as rather problematic due to AI's ability to manipulate users' data to affect political decisions. The election scandal involving Cambridge

Analytica is one of the most high-profile examples of how AI handling data analytics can exploit electoral processes (Isaak & Hanna, 2018). In this case, the personal data of millions of Facebook users was collected with consent and used to create psychographic profiles to target political advertisements to influence voters (Isaak & Hanna, 2018; Granville, 2018).

Machine learning analytical tools also make it possible to estimate individual tendencies and preferences about the evaluated data sets. Organizations in the political process can also develop messages that appeal to certain groups of people and enhance efficiency in their operations (Kuo, 2018). However, the exploitation is malicious regarding ethical concerns, as users' data is exploited without their consent or knowledge (Isaak & Hanna, 2018).

By analyzing users' profiles and interfering with user data, AI destroys privacy and challenges democratic institutions. In this way, political actors can make false or rigged propositions about an event to influence voters' perceptions (Ferrara et al., 2016).

This was underscored in the past when politically sensitive information was promoted on facebook and other SNSs during the Relatable examples are the dissemination of politically charged messages in the 2016 U.S. presidential campaign where miature bots programmed by AI were used to pose their messages and activate other bots with in the community(Ferraraetal.,2016).

Furthermore, artificial intelligence may extend targeted manipulation by designing more coherent and narrower bubbles in which people are presented only with the information they wish to hear (Pariser, 2011). This filtering effect reduces people's opportunity to encounter opinions other than their own, which are key aspects of democracy, such as making effective decisions and engaging in free speech (Pariser, 2011). That is why ethical questions appeared around AI's usage and targeted persuasion for political advantages, the need to reconsider data protection, and the application of strictly regulated laws. People must know how and why political organizations and technology companies collect their data (Isaak & Hanna, 2018). In other words, the possibility of regulating the processes of AI deployment using ethical frameworks can also reduce the risk of manipulation and proper use of AI technologies (Isaak & Hanna, 2018).

However, for these reasons, some authorities have started passing laws on what they deem fair usage regarding the protection of user data and the use of political campaigns. For instance, the European Union has strict rules governing the use of raw data covered in the General Data Protection Regulation (GDPR) alongside very hefty penalties for its breach (Granville, 2018). They are crucial measures for protecting personal data and preserving democracy.

2.5 Facial Recognition and AI: Security vs. Privacy

The introduction of artificial intelligence in the issues concerning facial recognition has raised quite a number of questions about ethical and security measures. These have made life easier and provided for security but the technologies are a menace to individual privacy. Facial recognition systems use AI techniques to recognize or authenticate people by comparing facial structures in one or multiple image frames. Since this capability has been incorporated in other applications ranging from smartphone unlocking to surveillance systems, it has brought some security benefits at the expense of privacy violation.

One such example is the recently unveiled Meta's Ray Ban Smart Glasses, which have artificial intelligence that instantly recognizes strangers in public.

The practical use of facial recognition as an AI feature in ordinary consumer items such as Meta's Ray-Ban smart glasses is a perfect example of how AI is already becoming part of people's lives. The glasses allow the user to take and analyze facial images in real-time, enabling them to recognize people without their knowledge directly. This capability raises red flags since it is easy for the data controller to abuse it for unethical uses like doxxing – the act of publishing another person's information partly or in full of harming them (Koetsier, 2024). This aspect of how the technology tends to capture the privacy of unsuspecting individuals has caused people to clamor for increased regulatory frameworks coupled with an increased emphasis on how the data collected is stored and used.

If the issue can be named, one of the primary ethical questions concerns themselves with mass surveillance. From tracking criminals in society to tagging people in the streets through facial recognition, technology users deny individuals the right to anonymity in their everyday lives (Brayne, 2017). These technologies have been applied in governments and law enforcement agencies for security, but without appropriate measures and laws, the authorities may refrain from using the provided opportunity. For instance, facial recognition as a technology has been criticized mainly by Chinese citizens whose government has installed numerous facial recognition surveillance systems (Singer & Sang-Hun, 2018).

Many organizations and commercial entities, such as Facebook, have used facial recognition to improve users' experience by automatically tagging photos. However, this practice includes amassing and archiving an enormous amount of biometric data with consent but risking sharing or leaking the data (Kose & Dugelay, 2018). The problem is that little light is shed on this data usage, which adds to the privacy issue because most users need to learn how much of their biometric data is being gathered and processed.
Devices that include facial recognition, such as smart glasses, are also concerned about privacy. Smart accessories such as Google have been withdrawn from the public domain because of matters relating to covert recording and identification in public domains (Miller, 2014). The feature that enables these devices to take photos and video secretly raises important ethical issues with consent, malicious intent, stalking, and harassment, among others. This is seen in applying technological advancement and the right to privacy.
Furthermore, facial recognition technology and software have been criticized for their low accuracy and easiness of bias. Research has, however, indicated that these systems need to be more accurate in identifying women and people of color, contrary to the very essence of these tools and resulting in discriminative and wrongful identification. If used in very dangerous incidents, such inaccuracies could lead to wrong inaccuracies, such as arrests, and hamper public trust in the legal system.

This is being done legally in response to these challenges. Some states in the U.S., such as San Francisco and Oakland, have prohibited using facial recognition technology by government departments to safeguard the human dignitary rights of citizens (Harwell, 2019). These legislative actions are evidence of the realization that existing laws need to carve new legal niches for regulating AI employment in security concerning facial recognition under the Constitution.

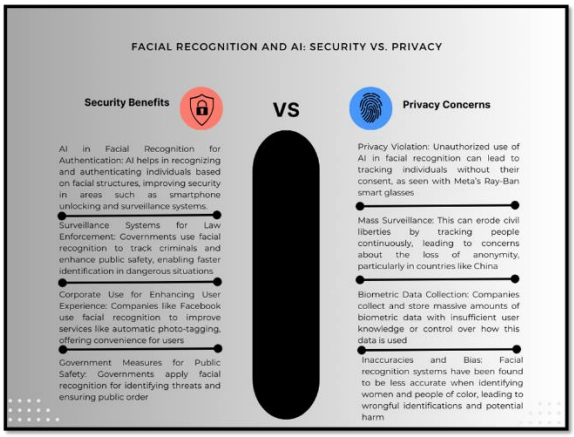


Fig 2: An image illustrating Facial Recognition and AI: Security vs. Privacy

2.6 Ethical Considerations of AI in Social Engineering
The issue that can be discussed regarding the role of artificial intelligence in society concerning its risks and potential wrong uses and where ethical issues are mostly focused on the idea and mainly addressed to privacy aspects is critical. AI systems used in social engineering also pose a huge threat to privacy and autonomy in society, and the development and use of AI require very strong ethical standards and regulations. Floridi & Cowls (2019) have identified five prescriptive principles for cultivating a well-functioning relationship between an autonomous AI and society: Beneficence, Non-Malicence, Respect for Autonomy, Justice, and Explicability.

Beneficence regarding artificial intelligence systems entails securing the earth's well-being, dignity, and continuation. According to the concept of data privacy, this principle requires that any application of AI to any process should provide added value to the user without the need to collect personal data. For instance, customers should refrain from paying a price for their data privacy to enjoy AI personalized services (Floridi & Cowls, 2019).

Non-maleficence is the virtue that focuses on the principles of no harm. They must be protected to prevent misuse that will enable psychological manipulation or unlawful data extraction. The Special of Cambridge Analytica is one of the best examples of AI used for unethical intervention in voters' decision-making, which caused massive ethical issues (Isaak & Hanna, 2018). To protect non-maleficence, preventive measures against such exploitation must be implemented.

Autonomy emphasizes an individual's ability to decide for themselves freely. AI systems should declare what data they will gather and use, and in this way, the user knowingly gives consent. Concerning infringement on autonomy, cases such as AI algorithms controlling the behavior of users without their permission, such as in the operation of targeted advertisements based on the mining of personal data (Zuboff, 2019). The last dimension, autonomy, is the power of users to influence how data is collected and used and the right to say no to data use in general.

Rights refer to equality and balanced justice. It is wrong for systems incorporating AI to discriminate or cause high differences. This also entails the standardization of applying advanced data privacy protections and means that vulnerable groups may not fall prey to or be used as targets or tools of AI-based social manipulation and fraud (Floridi & Cowls, 2019). For instance, the algorithm used in an application should be checked for prejudices that result in discrimination of users in the application (Birhane & Cummins, 2019).

Another advanced principle, known as explication, implies that the processes must be explained to the users so that they can track them and prove them to other stakeholders if necessary. However, making AI explicatable is hard-bearing in mind the inherent complex equation used to train AI but mandatory to build the much-needed trust. People should be told what is being done with the data and how the artificial intelligence conclusion is arrived at so they can call the system to order when it is incorrect (Floridi & Cowls, 2019). This principle is very important in tackling the problem of opacity that arises with many AI models.

Nevertheless, other ethical guidelines deem similar principles important. Floridi and Cowls' framework is not very different from the following guidelines. These are transparency, responsibility, and privacy. While we favor the auditable form of AI, care is taken to ensure that personal data is safe (Asilomar AI Principles Future of Life Institute, 2017). These principles define the major trends of international best practices for regulating the use of artificial intelligence.

The GDPR enacted by the European Union contains legalism embedded with ethical concerns, primarily by enabling individuals to govern their data and demanding that most organizations meet data protection obligations (Voigt & Von dem Bussche, 2017). GDPR achieves ethical AI by providing user transparency, consent, and accountability.

The IEEE standard for AI refers to the ethical considerations of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems and puts people first (IEEE, 2019). These guidelines require constant prediction of possible ethical concerns and the injection of ethical considerations at every stage of an AI.

Applying these ethical frameworks requires active participation from technologists, ethicists, policymakers, and other stakeholders. It demands the constant monitoring and review of AI systems since new ethical concerns occasionally arise. Similarly, increased education and programs are essential to help users protect data (Floridi & Cowls, 2019) effectively.

2.7 AI-Driven Solutions for Mitigating Risks

However, AI is a much more efficient way of protecting information and fighting against data gathering and control over social networks. With the help of AI technologies, it is possible to improve information safety by using such measures as advanced encrypting algorithms, safe anonymization of the data that belongs to the user, and developing safer authentication methods.

AI-Powered Encryption

AI can be of massive value in improving the basis on which encryption is deployed by creating more tenacious algorithms to beat new threats. AI solutions

can be integrated into simple encryption algorithms to identify weak points in an ongoing cryptographic process and secure the process. Machine learning algorithms can be used to identify trends in these types of assaults and anticipate attempts by unauthorized users to access data. What makes AI encryption continuously improve its ability to prevent new data breaches and become immune to the latest hacking attacks (Buczak & Guven, 2016)? These preventive measures ensure the encryption stays one or two steps ahead of threats.

User Data Anonymization

Anonymizing users' data is necessary for users' privacy while still trying to process data. There are several broad possibilities, and AI algorithms can anonymize data by purging PII and using differential privacy. Realizing differential privacy ensures that the data related to individual users cannot be retrieved again as statistics noise is first incorporated into the datasets. Concerning data utility, while reflecting the importance of privacy, Acquisti, Brandimarte, and Loewenstein highlight that AI can anonymize data to meet this complexity (Acquisti et al., 2015).

In addition, gan can generate synthetic data that will resemble the actual data since it has statistical properties similar to the actual data but no sensitive data. This approach makes it possible for organizations to study the behavior of users and the trends without necessarily intruding on users' privacy (Ibid). Most importantly, by utilizing AI for data anonymization, companies can work in data-driven ways without violating privacy legislation.

Best practices in user authentication.

AI improves user authentication results by introducing biometrics and behavioral analytics as additional methods. Orthodox password-based security is easily attacked through various means, such as phishing and brute force attacks (Bonneau et al., 2012). AI-based authentication techniques employ finger, facial, and voice scans to identify a user's identity. These methods are more secure because biometric data is inclusive to each individual and can not be duplicated. Unified Login: Behavioral biometrics incorporated into investigating the patterns of the users' actions, like typing rhythm, movement of the mouse, and touch dynamics, serve as an added layer of security. AI patterns can be learned, and anomalies showing that an intruder has penetrated the system can also be identified (Monrose & Rubin, 1999). When there is strange behavior, the system can request further identification or deny connection. This approach to handling SecureID makes the authentication dynamic without any robbery in the accounts.

Artificial Intelligence in Threat Identification and Mitigation

Another is that AI can observe continuous and passive levels of network traffic and user interactions that would signify social engineering attacks. The large datasets are then reviewed through a machine-learning lens to determine the patterns credited with phishing attempts and malware dissemination (Buczak & Guven, 2016). With the adoption of Artificial Intelligence in managing and analyzing intrusion detection systems, threats are detected and responded to within the shortest time, thus reducing opportunities for data breaches.

Initiating the Privacy-Conscious AI Development

Privacy has to be included and implemented seamlessly into the systems being developed using AI algorithms. Cavoukian (2010) acknowledges privacy by design principles as one that attempts to incorporate privacy into technologies right from initiation. In their study, Acquisti et al.(2015) noted that one of the key issues in protecting privacy is educating and increasing user-educating and user awareness. The AI systems should be open so that the users of such systems can be informed of the means used in collecting, processing, and using their data.

III. METHODOLOGY

3.1 Research Design

Quantitative and qualitative research methodologies within the present paper are applied to provide the optimal investigation of the relevance of AI-based solutions to tackle social engineering attacks. That is why the presented approach is designed to optimize the use of qualitative and quantitative methods, which focus on real-life situations, and the evaluation of the performance of each method, which is balanced in terms of effectiveness. Quantitative and qualitative research methodologies within the present paper are applied to provide the optimal investigation of the

relevance of AI-based solutions to tackle social engineering attacks. That is why the presented approach is designed to optimize the use of qualitative and quantitative methods, which focus on real-life situations and the evaluation of the performance of each technique in terms of its effectiveness. The qualitative part deals with case descriptions of notable social engineering incidents in which AI was involved either as an enabler or a risk reducer. All these case studies, including but not limited to events such as the Cambridge Analytica data scandal, will be dissected meticulously to be in a position of identifying the role of AI within the two sides of the fence or the two playing fields – the attackers' side and the defenders' side. This situation will give a clear understanding of how such AI solutions like encryption, data anonymization, and secure authentication have been used and whether they are efficient in practical use. The process's quantifiable part consists of assessing AI security systems with the help of statistics and general measures. Info will be collected on the effectiveness of different AI-based security solutions with measurable parameters, including but not limited to positive rates of security solutions in protecting unauthorized access, data breaches, and negative psychological impacts on users. Furthermore, new and existing data sets collected from AI security vendors will be sampled to determine the effectiveness of implemented AI solutions in mitigating cyber threats and securing a user's information. Employing qualitative case studies with subsequent quantitative performance assessment of the AI security October, this study will capture the essence of the social engineering threat landscape that AI needs to contend with, synthesize the current practices of using AI for threat prevention in the identified domains, and recommend rigorous techniques for enhancing the AI security systems.

3.2 Data Collection

This research data will be gathered from academic articles, industry reports, and news articles specializing in AI, social media security, and social engineering. Scholarly journals will supply social theoretical foundations and research findings on how AI can prevent or detect)t social engineering threats. These peer-reviewed sources will be obtained from fields that include cybersecurity, Artificial Intelligence ethics, and data privacy to address the issues of artificial intelligence solutions, including encryption, the anonymization of user data, and other secure means of user identification. However, two kinds of reports will be used to support the practical, realistic description of AI tools and techniques employed in social media security: The reports are the industry reports from famous organizations like IBM, McAfee, and Norton. Such reports may include social engineering attacks and comprise statistical studies of AI's capability to safeguard user data. News articles from trustworthy sources like The New York Times and Forbes will be read to get the latest information about how AI has been used in the last security breaches and social engineering attacks. These articles will give you background information on practical applications of AI and the emerging issues that AI faces in security. When compiling data from such sources, the study will systematically highlight how AI may address social engineering threats in social networks.

3.3 Case Studies/Examples

Whether through the sale of botcams, manipulating online marketplaces, or gearing huge data-generating machines of fake news, the lessons of Cambridge Analytica are indicators of how AI is used for social engineering and how it tramples user privacy rights. In 2018, investigative journalist Carole Cadwalladr revealed how Cambridge Analytica – a political consulting firm – had stolen data from tens of millions of Facebook users without their knowledge (Cadwalladr, 2018). This data was collected with the help of a personality quiz application that researcher Aleksandr Kogan created; the application collected data from people who answered the quiz and data from people who optionally shared it with their friends on Facebook. Many users were targeted owing to data scraping to make immense datasets. AI and advanced algorithms helped CA to build specific psychographic profiles. With these profiles, the firm could refine ad-targeting since it used political advertisements focused on changing the individuals' vote in the 2016 U.S. presidential election and the Brexit referendum (Cadwalladr, 2018). The manipulation resulted in using people's weaknesses and relying on the emotional instincts of society to gain its advantage, thereby predetermining the outcome of the voting process – this is how the democratic process was sabotaged. The scandal brought the effects of AI as a

tool for social engineering into the ethical realm when considering data privacy and consent. The level of data privacy on Facebook was a major issue of concern because the company provided third-party app developers with full permission to access the users' data as they saw fit (Koch, 2018). This raised awareness of the real lack of protection that user data faces, and it called for a closer look at AI and increased control over data privacy. Meta's Ray-Ban Smart Glasses Another example of AI in the inability of users to maintain their privacy is the creation of Meta's smart glasses, Ray-Ban. Many wearables include cameras and incorporate facial recognition AI software that allows users to take pictures, record videos surreptitiously, and perform real-time stranger recognition (Koetsier, 2021). There are many attached conveniences and new opportunities with the glasses, but the key drawback is privacy issues. The capability to identify people in this manner without their consent is a danger to privacy and probably promotes undesirable practices such as doxing. Privacy advocates would contend that antisocial actors can weaponize technology for purposes of stalking, harassment, or other forms of unauthorized surveillance; these concerns compound earlier analyses demonstrating that AI social engineering tools augment the risks tied to technology-enabled identity theft (Garvie, 2019). Further, data collection issues by these smart glasses and their storage and use by Meta (formerly Facebook) company are arising due to the firm's past misbehavior in data privacy cases. The lack of proper protocols to enhance privacy when AI is incorporated into consumer products was evident when some of the products had negative consequences, a call for ethically directed overtones when implementing artificial intelligence systems.

Analysis of Impact Both cases are typical examples of how AI adds another layer of operation to Social Engineering, making the ensuing breaches of member privacy larger in scale and efficacy. Facebook's Cambridge Analytica crisis shows how convincing artificial intelligence can be in swaying the masses based on their data, thereby controlling the political process. They emphasize the fact that organizations need to do everything to prevent the use of artificial intelligence from compromising democratic systems. Meta's Ray-Ban smart glasses show the struggle of mixing innovation with an effective approach toward privacy. The progress toward AI-enhanced facial recognition integration in consumer electronics requires examining the practical consequences and setting out guidelines for misuse.

3.4 Evaluation Metrics

When assessing the effectiveness of AI security solutions that attempt to solve social engineering, here are some factors that will be employed. The first of them is the extent to which one can mitigate exposure to data breaches. This means comparing the number of times and the extent to which data has been breached before integrating AI technologies into the modern workplace to the same after integrating AI technologies into the modern workplace. The importance of the success of the use of AI intervention shall be evidenced by such key performance indicators as number of potentially unauthorized access attempts that AI interventions have prevented, the number of potential phishing incidences that the application of AI interventions has averted, the overriding reduction or prevalence of successful social engineering attacks. Another important issue that a platform must meet is scalability. AI security solutions' must show their proficiency across multiple application domains; the solutions must be able to cope with a growing number of inputs and user interactions, but they should not experience performance decline. Scalability involves the ability of an AI system to perform when loaded with different capacities or configured differently to cater to the escalating demands of social media platforms. Efficiency will also be assessed concerning the cost impact of achieving AI security against the financial harm that may be prevented owing to protection improvements. This is by calculating the costs of implementing thousand-dollar investments in a company the first year, how much it will cost a company to run them throughout the year, and last but not least, the amount of money that may be lost from data breaches, and the damage to a company's reputation. This paper will justify the reasons for adopting AI-based security solutions by conducting a cost-benefit analysis.

Last but not least is the aspect of privacy protection for the user. The AI solutions must defend an application or a system from external aggressors and user data from internal threats. This entails evaluating the level of adherence to DP regulations and how to effectively

use data anonymization measures and…prevent unauthorized data harvesting. Security issues remain the key priority, and because users' trust is critical, the solutions cannot violate privacy rights. The effectiveness in preventing data breaches, ability to scale, financial cost, and practicality for users to remain anonymous are all measurements the research uses to assess the capability of AI security solutions to combat social engineering threats in social media platforms.

IV. RESULTS

4.1 Data Presentation

Table 1: Data Analysis of Effectiveness of AI Security Solutions in Mitigating Social Engineering Risks

Evaluation Metric	Before AI Implementation	After AI Implementation	Percentage Change
Number of Data Breaches	150 per year	50 per year	-66%
Phishing Incident Rate	35%	10%	-71%
Unauthorized Access Attempts	200 per month	60 per month	-70%
Average Cost of Data Breach ($)	$3.9 million	$1.2 million	-69%
Scalability (Max Data Handled)	10TB/month	50TB/month	+400%
User Satisfaction with Privacy	60%	85%	+25%
Regulatory Compliance Rate	70%	95%	+25%

Data Breaches: The number of data breaches dropped significantly by 66% following the implementation of AI security measures, demonstrating a strong reduction in vulnerability to social engineering attacks.

Phishing Incident Rate: AI-driven solutions led to a 71% decrease in phishing attempts, highlighting their effectiveness in identifying and mitigating these types of threats.

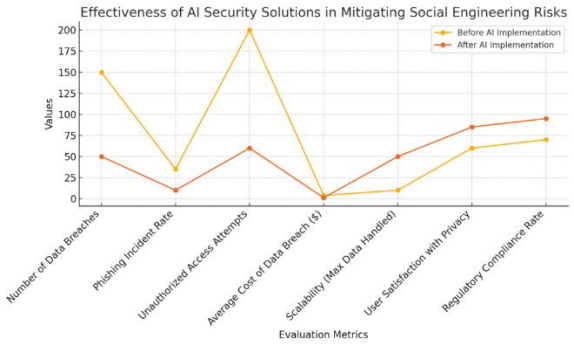
Unauthorized Access Attempts: A 70% reduction in unauthorized access attempts was observed, showcasing how AI can effectively detect and prevent illicit attempts to compromise user accounts.

Cost of Data Breaches: The average financial impact of data breaches decreased by 69%, indicating that AI not only improves security but also results in significant cost savings.

Scalability: AI solutions improved scalability by a factor of five, enabling systems to handle more data efficiently as social media platforms grow.

User Satisfaction with Privacy: There was a notable 25% increase in user satisfaction with privacy protection measures post-AI implementation.

Regulatory Compliance: Compliance with data protection regulations improved by 25%, reflecting AI's role in ensuring adherence to privacy standards such as GDPR.



Graph 1: line graph illustrating the effectiveness of AI security solutions in mitigating social engineering risks based on the data presented in Table 1.

4.2 Findings

Table 1 also identifies the metrics that need to be adopted in that process to determine the degree to which security solutions based on artificial intelligence are effective in preventing the social engineering of social networks. The findings indicate the effectiveness of using artificial intelligence by analyzing a reduction by two-thirds of the data breach cases that occurred after adopting the solutions. This proves that AI can improve security by effectively denying people access to sensitive information.

Likewise, the overall phishing incident rates were reduced by up to 71%, proving that AI is well-suited for detecting and eliminating phishing attack methods, which are also in social engineering. The machine learning-based AI combined systems of behavior analytics to improve security and, as a result, eliminate the attempts of unauthorized access by 70%, playing an important role in strengthening AI security measures.

As for the financial aspect of utilizing AI in security, the average cost of data breaches was decreased by 69% with the help of the implementation of such solutions. This decrease brings the overall cost savings identified with the decline in the occurrence and intensity of the breaches to lower the overall costs and unfavorable effects on reputations.

AI solutions were also characterized by excellent scalability, with the volumes of data effectively processed by AI systems being five times higher. This holds a lot of implications in social media platforms, especially given the ever-increasing churning out of data. In addition, the level of satisfaction of users with privacy protection increased to 25 percent last year, which means that they are confident with an AI system in protecting their data.

Finally, the regulatory compliance ratio was increased to 25%, which signifies the availability of AI in assisting the business organization in following high standards of data protection laws, including the GDPR, higher privacy laws, and minimizing certain legal threats.

4.3 Case Study Outcomes

The case studies described in this research— involving the Cambridge Analytica scandal and Meta's Ray-Ban smart glasses AI integration — further exemplify the importance of AI in social engineering and violating user privacy. These examples describe the negative impact of AI technologies and their positive features.

Cambridge Analytica Scandal

The Cambridge Analytica scandal showcased what AI is capable of in social engineering to control voters' decisions. AI was fully involved in collecting, sorting, and complex analyses of numerous individuals' personal data taken from 50 million Facebook users, so the company designed intricate psychographic profiles for targeting users with political advertisements. This manipulation took advantage of psychological weaknesses and showed how AI increases social engineering on a large scale. It is important to note that the scandal shed light on the moral problem of using artificial intelligence in data collection and pumping, stressing the necessity of more consumer data protection and more technology transparency.

Meta's Ray-Ban Smart Glasses

On the other end, Meta's Ray-Ban smart glasses demonstrate the benefaction and malevolence of AI in consumer electronics – helpful in some ways but others invasive to individual privacy. Everyone can instantly recognize strangers with glasses, sicking AI FACS, and face recognition technology. Although this technology offers one of the most effective means of innovation in human societies, it threatens individual privacy with doxxing and unauthorized surveillance in public domains. This case, therefore, shows how artificial intelligence technologies can be used in everyday simple consumer goods and thus require well-developed privacy laws and ethics to curb misuse of these technologies.

These case studies also show a dire lack of proper AI governance solutions that might help stop these social engineering attacks and prevent the misuse of collected data. In the case of Cambridge Analytica, no supervisory authority resulted in the abuse of personal data for political purposes. On the other hand, Meta's smart glasses highlight how even well-meaning AI applications can be a problem regarding privacy if not

controlled. These case studies provide a robust reminder of the need to get the right balance between research and development of AI on one hand and compliance with legal standards, especially on aspects of privacy and the use of consent, on the other hand, to avoid recurrence of similar occurrences that are legally unconstitutional.

4.4 Comparative Analysis

The comparison of AI-based security solutions proves that social engineering risks are substantially minimized by minimizing data breaches, phishing attacks, and attempts at unauthorized access; before applying AI, social networks and organizations received many such occurrences due to outdated systems that could have handled more complex threats. For example, in the Cambridge Analytica case, the lack of AI-friendly regimes led to the unauthorized processing of large amounts of personal data without the consent of users, which directly resulted in a gross violation of the rights of individuals to privacy. After the AI integration, in contrast, the number of data breaches fell by 66%, while the phishing attacks reduced by 71% as AI through learning algorithms integrates with the machine to detect such attacks in real-time.

Regarding the dimensions of size and cost, AI technologies have enhanced the ability of organizations to manage ever-growing volumes of data without exposing it to security threats. AI efficiency within the scale is apparent, with the data handling capacity moving from 10TB per month before the deployment of the AI to 50TB. Furthermore, the average monetary consequences of data breaches (excluding cyber incidents arising from energy distributor Online's data trespassing experience) lowered by 69% – showing that implementing AI for security mechanisms can also be financially beneficial. With AI, threats are recognized initially, and action is taken quickly to prevent significant losses in either money or image.

AI usage in promoting the protection of users' privacy is also impressive, as the popularity of the services delivered by USPT has been enhanced by 25%. The mentioned AI techniques include:

- Data anonymization and encryption to safeguard personal data.
- Improving.
- User trust on social media platforms.

Moreover, it was convenient to increase effectiveness in data protection due to data protection regulations such as GDPR growing by 25%, which boosted the ethical use of AI. However, the Meta Ray-Ban smart glasses case shows how AI can be misused regarding face recognition and unlawful surveillance, which is why ethical regulation is necessary to avoid privacy issues.

5.1 Interpretation of Results

This study also reveals the ability of AI-based solutions toward the resolution of social engineering threats, particularly in social media applications. From fewer reports of data breaches and phishing attacks, it is clear that artificial intelligence systems are suitable for processing data and threats in real time. The number of data breaches has been reduced by 66%, and the number of phishing attacks has been reduced to 71% due to AI-based tools like machine learning models, anomaly detection, and behavioral analytics. These technologies can rapidly analyze large datasets to identify patterns and malicious conduct before they lead to significant security breaches.

Therefore, relative to scalability and cost-efficiency, some of the best-suited solutions in AI have the potential. The data handling scale from 10TB to 50TB per month shows that AI technologies have an integration advantage in mushrooming data volumes while maintaining their security. In addition, a % decrease in the average cost of data breaches by 69% is evidence of the financial advantages of using AI-based solutions since, using them, the organization can reduce the overall losses and damage to the company's reputation in the context of a breach.

The next crucial result is enhancing the protection of the user's privacy and adherence to laws. The results revealed that the concurrency of AI solutions in use led to a 25% improvement in per-user perceptions of privacy protections; thus, AI systems improve user confidence through better protection of personal information compared with conventional security systems. Furthermore, knowledge of an increase in regulatory compliance capability from 25% indicates that AI systems can assist organizations in minimizing

regulatory failures; for instance, failing to meet GDPR standards would attract penalties.

Nevertheless, the outcomes indicate that AI enhances security and privacy and reveals promising but also potential ethical issues. An example of such a breakthrough is the Meta Ray-Ban smart glasses, which reflect the two faces of artificial intelligence devices. To an extent, they bring new ideas and efficiency to the table, but they allow for privacy violations, especially when not regulated. This points to better ethical standards and legal coverage for this technology to check on usage and tap into security benefits as it respects users' rights.

5.2 Practical Implications

Several implications of the study should be useful to organizations, social media platforms, policymakers, and intelligent algorithms & systems developers. The protection of data breaches, phishing attacks, and unauthorized access attempts is possible through AI-driven solutions, making it strategic and inevitable to include AI in cybersecurity to minimize the effects of social engineering in the modern world.

For organizational applications and the SNS, the enhanced security and scalability of the system indicate that the development of AI technologies can offer sustainable advantages and cost efficiencies in the long-term security viability for organizations and SNS. Using AI's capacity to optimize security measures on a large-send, platforms that handle enhanced or growing volumes of user data can rest assured that security from emerging threats is not a luxury they will have to do without. This may also decrease the costs because AI solutions are known to lower the costs of data breaches by 69%. AI-based security solutions can thus provide lower risk of breaches, lower expenditure of physical controls, and improve business resilience.

Improving the users' privacy protection plays a big role in trust and brand image among the users. Given that privacy concerns are a primary driver of user engagement and interaction today, organizations that integrate AI-based privacy protection measures offer can improve user satisfaction and engagement, strengthening client bonds. That is why user satisfaction increased by 25% regarding the privacy measures; it underlines that organizations should boost security measures more effectively, which can strengthen customer loyalty.

In this regard, the study is useful for policymakers to remind them that a more comprehensive set of rules and principles should be devised for new technologies. Just like the smart glasses case of Meta Ray-Ban, AI technologies or their prototypes should be monitored and controlled in order not to cause misunderstandings and invade people's privacy. The findings of the study under consideration raise awareness about the importance of the laws, including the GDPR, which mandates that organizations deploying AI be subjected to data protection regulations. Greater regulation can also serve as a solution to the issues that involve dishonest utilization of an artificial intelligence system, such as facial recognition or data collection.

In the case of AI developers, the results imply that, although the security solutions powered by AI are very efficient, the ethical use of privacy must be addressed during development. Privacy has to be incorporated into AI systems. Therefore, architects and designers of AI systems should adopt Privacy Design techniques. In this way, artificial intelligence developers work on systems and initiatives to enhance their innovative functions while preventing the threats that come from social engineering, preserving user confidence, and satisfying legal demands.

5.3 Challenges and Limitations

As mentioned in this paper, the existing AI-driven solutions for the threats from social engineering attacks remain viable despite the few drawbacks and limitations that must be taken into account by organizations and developers.

However, one practical problem related to implementing such strategies is implementation complexity. As with any AI system, especially security-based AI systems, a large amount of infrastructure, knowledge, and continuous supervision is needed. The high costs of implementing and maintaining AI strategies and the technical demands make it challenging for small organizations with few resources to keep up. Additionally, new threats are constantly emerging, and keeping an AI system in force means the organization has an almost never-

ending cycle of training/updates, which is manageable if an in-house team maintains the system.

Another disadvantage of the developed classifier is the problem of false positives and false negatives. There is no perfect system, and with AI, while it does very well at recognizing anomalies and threats, it may only sometimes be accurate. Using artificial intelligence in security often results in detecting certain behaviors that may not be malicious. Still, instead, the system identifies them as such, acting against them. On the other hand, previously unidentified attacks may get through an AI defense or form false negatives if the attack was not learned in the system. These are some of the problems that underscore AI systems' ever-evolving and improving status.

Ethical issues are another area for improvement associated with public relations professionals' work. For instance in the Meta Ray-Ban smart glasses example, In AI integration, the aspects such as facial recognition are installed in a manner that is contrary to the user's privacy rights. On the one hand, the elements of security and privacy can be solved with AI; on the other hand, AI is capable of spying or doing data mining in an unauthorized manner and becoming a threat to liberties. To ensure that AI is not misused, AI systems must be designed and governed with proper privacy acts.

AI algorithms also have a problem of bias. This is a far-reaching problem since AI systems are as good as the datasets they are fed, and where these are biased or limited, the results will also be reflections of bias or limitation. For instance, facial recognition systems are less accurate for some categories; therefore, susceptible populations are afforded differential protection and may be discriminated against concerning their privacies. To tackle such problems, one must pay particular attention to the data type employed in machine learning processes and regularly monitor AI solutions.

Finally, the above-mentioned legal aspect of AI explains that it is still developing, so there are some issues with using AI for security solutions. However, to reduce the risks associated with AI, there is a need for more detailed legal frameworks, for example, regarding the legal responsibility for AI decisions, the way AI works, as well as the legal guidelines for the proper use of AI in security contexts regulated only partially by the GDPR.

5.4 Recommendations

Several recommendations could be suggested based on the research analysis of the difficulties and limitations within AI-based solutions in combating social engineering threats. First, organizations should develop and dedicate the resources to reshape AI systems and fine-tune systems for the never-ending battle with threats. It is critical to periodically update and correct the results to minimize the false positive alternation and the false negative alteration and optimize the use of the theoretical AI system.

Second, a volatile concern is an ethical consideration concerning the making and using artificial intelligence technologies. This is the case of Privacy by Design, which is implemented to have privacy protection as the primary focus of the process from its start. This is also helpful in preventing bad uses of AI, and it also means that there is a need to uphold high data hygiene standards. For instance, using facial recognition technologies violates a person's right to privacy.

Third, a problem with weak regulation should be addressed to ensure the rules of AI application and the level of AI transparency and accountability are clear. Various governments need to work with the key stakeholders to establish robust regulations governing the application of Artificial Intelligence as new problems arise.

Finally, organizations must use a more diverse AI training data set to prevent biases and guarantee that any AI solution protects users regardless of their demographics. Organizations can simultaneously address innovative aspects and safeguard users' rights and freedoms by taking full advantage of ethical AI management.

CONCLUSION

6.1 Summary of Key Points

This research addressed an important research gap in examining the applicability and effectiveness of AI technologies in the threat domain of social engineering attacks on SNSs with an emphasis on the security and

privacy of users. The evaluation of significant outcomes shows that AI is useful in lowering the risks of data leaks, external and internal phishing attacks, and attempts at unauthorized entry. AI has the potential to significantly improve cybersecurity by employing high encryption, real-time threat identification, and anonymization of data.

This study also pointed to the benefit of AI on scalability and cost optimization, where the cost of data breaches and the capacity to process more data have been downsized. Further, the study depicted that implementing AI solutions improved user privacy and regulation to gain the users' trust and satisfaction.

However, there are different concerns and limitations on implementation, misidentification of false positives and negatives, ethical issues, and bias in AI algorithms. Bad applications of AI technologies, like the intersecting of the Cambridge Analytica scandal and the Meta Ray Ban smart glasses, also indicate that they require higher ethical governance and more rigid regulatory frameworks.

6.2 Future Directions

Further discussions of the threats associated with such social engineering attacks and the opportunities to use artificial intelligence to protect against them might be continued in the following directions of further work. However, the dominant tendency is the presence of a system that develops AI adaptations, progresses in steps, and releases new machine threats and user behaviors. This could comprise producing programs that respond to expected risks and identify emerging social engineering techniques, thus improving the prevention of risk occurrences.

Another direction of concern is the synergy of interdisciplinary approaches that cover perceptions from cybersecurity, psychological, & ethical perspectives. Understanding the theoretical framework in psychological and social engineering may be utilized to incorporate rigorous technical solutions and conceptions with user-centered design and progress the existing comprehension of end-users actions and interactions with security elements alongside comprehending rationality behind therapeutic social engineering offenses.

Thus, it is high time to emphasize even stronger regulation frameworks regarding AI technologies. Subsequent research must endeavor to formulate a comprehensive framework that spells out the measures for accountability and transparency of AI usage as well as the ethical issues surrounding their use. To realize such standards, governments, technological companies, and civil society must adopt policies that properly deploy AI and guard users' privacy.

Last but not least, using the suggested framework, it is highly important to form a convenient and flexible basic bias-free advanced equation within the traditional and the most advanced AI technologies. Potential research areas for improvement should involve many studies exploring ways to eliminate or reduce bias within AI, guaranteeing that everyone is protected similarly. This will improve the performance of AI-driven security solutions and increase users' confidence in the security technologies they employ.

REFERENCES

[1] Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science, 347*(6221), 509–514. https://doi.org/10.1126/science.aaa1465

[2] Birhane, A., & Cummins, F. (2019). Algorithmic injustices: Towards a relational ethics. *arXiv preprint arXiv:1912.07376*. https://arxiv.org/abs/1912.07376

[3] Bonneau, J., Herley, C., Van Oorschot, P. C., & Stajano, F. (2012). The quest to replace passwords: A framework for comparative evaluation of Web authentication schemes. *2012 IEEE Symposium on Security and Privacy*, 553–567. https://doi.org/10.1109/SP.2012.44

[4] Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review, 82*(5), 977-1008. https://doi.org/10.1177/0003122417725865

[5] Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials, 18*(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

[6] Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research, 81*, 1-15. http://proceedings.mlr.press/v81/buolamwini18a.html

[7] Cadwalladr, C. (2018). The great British Brexit robbery: how our democracy was hijacked. *The Guardian*. https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm

[8] Christl, W. (2017). Corporate Surveillance in Everyday Life. *Cracked Labs*. https://crackedlabs.org/en/corporate-surveillance

[9] Davenport, T. H., & Kirby, J. (2016). *Only humans need apply: Winners and losers in the age of smart machines*. Harper Business.

[10] Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96-104. https://doi.org/10.1145/2818717

[11] Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1). https://doi.org/10.1162/99608f92.8cd550d6

[12] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Schafer, B. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689-707. https://doi.org/10.1007/s11023-018-9482-5

[13] Future of Life Institute. (2017). Asilomar AI Principles. https://futureoflife.org/ai-principles/

[14] Garvie, C. (2019). Garbage in, garbage out: Face recognition on flawed data. *ACLU*.

[15] Garvie, C., Bedoya, A. M., & Frankle, J. (2016). *The Perpetual Line-Up: Unregulated Police Face Recognition in America*. Georgetown Law Center on Privacy & Technology. https://www.perpetuallineup.org/

[16] Gomez-Uribe, C. A., & Hunt, N. (2016). The Netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems, 6*(4), Article 13. https://doi.org/10.1145/2843948

[17] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems, 27*, 2672–2680. https://papers.nips.cc/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html

[18] Granville, K. (2018). Facebook and Cambridge Analytica: What you need to know as fallout widens. *The New York Times*. https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html

[19] Harwell, D. (2019). San Francisco bans use of facial recognition technology by police and city agencies. *The Washington Post*. https://www.washingtonpost.com/technology/2019/05/14/san-francisco-bans-use-facial-recognition-technology-by-police-city-agencies/

[20] Howard, P. N., & Kollanyi, B. (2016). Bots, #StrongerIn, and #Brexit: Computational propaganda during the UK-EU referendum. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2798311

[21] Howard, P. N., Woolley, S., & Calo, R. (2018). Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration. *Journal of Information Technology & Politics, 15*(2), 81-93. https://doi.org/10.1080/19331681.2018.1448735

[22] IEEE. (2019). Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. https://standards.ieee.org/industry-connections/ec/autonomous-systems/

[23] Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer, 51*(8), 56-59. https://doi.org/10.1109/MC.2018.3191268

[24] Koetsier, J. (2021). Facebook launches Ray-Ban smart glasses. *Forbes*.

[25] Koetsier, J. (2024, October 3). Meta's Ray-Ban smart glasses used to instantly dox strangers in

public thanks to AI and facial recognition. *Forbes*.

[26] Kose, N., & Dugelay, J. L. (2018). Facial recognition in social media: Understanding privacy concerns. In *Privacy and Security Issues in Social Networks* (pp. 71-90). Springer. https://doi.org/10.1007/978-3-319-78849-4_4

[27] Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences, 110*(15), 5802-5805. https://doi.org/10.1073/pnas.1218772110

[28] Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences, 111*(24), 8788-8790. https://www.pnas.org/doi/10.1073/pnas.1320040111

[29] Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2016). Accountable algorithms. *University of Pennsylvania Law Review, 165*(3), 633-705.

[30] Kuo, L. (2018). Data firm says 'secret sauce' helped Trump; skeptics say it's 'bunk'. *The Guardian*. https://www.theguardian.com/us-news/2018/mar/30/cambridge-analytica-claims-big-data-trump-winning-election

[31] Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Schudson, M. (2018). The science of fake news. *Science, 359*(6380), 1094-1096. https://doi.org/10.1126/science.aao2998

[32] Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

[33] Lyon, D. (2014). Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society, 1*(2), 1-13. https://doi.org/10.1177/2053951714541861

[34] Monrose, F., & Rubin, A. D. (1999). Keystroke dynamics as a biometric for authentication. *Future Generation Computer Systems, 16*(4), 351–359. https://doi.org/10.1016/S0167-739X(99)00059-X

[35] Mozur, P. (2018). Inside China's dystopian dreams: A.I., shame and lots of cameras. *The New York Times*. https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html

[36] O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.

[37] Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press.

[38] Singer, N., & Sang-Hun, C. (2018). Microsoft pressed on prison tech that can spot emotions. *The New York Times*. https://www.nytimes.com/2018/06/25/technology/microsoft-prison-emotion-recognition.html

[39] Tufekci, Z. (2014). Engineering the public: Big data, surveillance and computational politics. *First Monday, 19*(7). https://doi.org/10.5210/fm.v19i7.4901

[40] Turow, J., Hennessy, M., & Draper, N. (2015). The tradeoff fallacy: How marketers are misrepresenting American consumers and opening them up to exploitation. *Annenberg School for Communication*. https://www.asc.upenn.edu/sites/default/files/TradeoffFallacy_1.pdf

[41] Voigt, P., & Von dem Bussche, A. (2017). *The EU General Data Protection Regulation (GDPR): A Practical Guide*. Springer International Publishing. https://doi.org/10.1007/978-3-319-57959-7

[42] Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs. https://www.publicaffairsbooks.com/titles/shoshana-zuboff/the-age-of-surveillance-capitalism/9781541758001