Fall 2022

Artificial Intelligence's Impact on Social Engineering Attacks

Sowjanya Manyam
Governors State University

ARTIFICIAL INTELLIGENCE'S IMPACT ON SOCIAL ENGINEERING ATTACKS

By

Sowjanya Manyam

B.Tech., KMM Institute of Technology & Sciences, 2015

P.G.D.M., International School of Management Excellence, 2017

THESIS

Submitted in partial fulfillment of the requirements

For the Degree of Master of Science,

With a Major in Information Technology

Governors State University

University Park, IL 60484

2022

Table of Contents

Abstract

This research paper aims to explore the concept of social engineering attacks and the impact of artificial intelligence on them. Security threats posed by Social Engineering have escalated significantly in recent years. Despite the availability of advanced security software and hardware mechanisms, a vulnerability still exists in the organization's or individual's defense system. In this paper we look at types of social engineering attacks and the basic techniques used by attackers will be described. The primary areas of study are how AI impacts social engineering and is used to detect and prevent social engineering attacks. The application of automated systems is rapidly growing in every lifestyle we imagine – social media, merchandise apps, driverless cars, and cybersecurity companies. Even though AI has improved cybersecurity, it is giving cybercriminals a position to unleash advanced attacks. The employment of chatbots is rising. Chances are we have had an interaction with a Chatbot already, it may well be on Facebook Messenger. Unfortunately, many of us do not realize that we are talking to a bot. This paper also discusses the concepts of voice spoofing, deep fakes and automated social engineering.

Keywords: Social Engineering Attacks, Artificial Intelligence, Voice Spoofing, Deep Fake Recognition, Chatbot

1. Introduction

Today, the computerized world is growing exponentially. Advances in digital communication technology have made communication between people more accessible and faster. However, personal, and sensitive information may be available online through social networks and online services that lack security measures to protect that information. Communication systems are vulnerable to attacks and can be easily penetrated by malicious users using social engineering attacks *[kyoun & Janabi]*. People see cyberspace as a new way of communication, business, information, and entertainment. However, this change in thinking raises serious concerns about the security and privacy of users on the Internet. Despite strict security measures, numerous online attacks are conducted using application design vulnerabilities, fraud, or advanced technical methods. Among these methods, scamming was even before the advent of computers and the internet. In terms of cybersecurity, scamming or phishing are classified as social engineering attacks. Attacks that exploit human vulnerabilities are classified as SE attacks. SE attacks are conducted using human weaknesses such as deception, persuasion, manipulation, or influence.

Artificial intelligence usage is fleetly adding in every aspect of life we can imagine, like social networks, retail operations, independent buses, and of course, cybersecurity companies *[kyoun & Janabi, 2021]*. While artificial intelligence improves cybersecurity, it also provides cybercriminals with an advantage in performing sophisticated attacks. The use of chatbots is increasing. If artificial intelligence can be used to educate bots to communicate friendly with humans, the same technology could be used to conduct cyber-attacks, one of the biggest pitfalls is social reengineering. This paper explains the theory of social engineering, its attacks, the impact

of artificial intelligence on them and how to avoid being the victim of social engineering attacks triggered by Artificial Intelligence.

1.1 Goal

My thesis presents a literature review on Social Engineering, its attack strategies, and types. Discussion of how Artificial Intelligence impacts Social Engineering. An overview of some real-life examples of AI-based social engineering attacks against organizations. The study I conducted and conclusions I reached indicate that it is impossible to stop social engineering attacks from happening; however, we can take countermeasures to prevent any individual or organization from inadvertently falling into such traps, and the main goal is to explore the possibilities of overcoming such attacks in the future.

1.2 Methodology

The methodology used in my research work is literature review which includes qualitative analysis of Social Engineering Attacks its types and the impact of Artificial Intelligence on them. I have employed a secondary exploration approach to gather information about the content. I have explored a pile of journals and websites for reference and described the social engineering attacks which are taking place in the artificial intelligence period. This paper gives a detailed description of what and how social engineering attacks are conducted by hackers and how they are carried out with the help of Artificial Intelligence like using chatbots and we will look at results to how can we overcome it. It's largely insolvable to fully terminate the social engineering attacks, but we will look at a few scenarios where we can avoid not being a victim of Artificial Intelligence triggering Social Engineering Attacks by detecting them. In terms of research paper figures and conclusion derivative, I will ensure that my exploration adheres to ethical morals.

2. Background and Related Work

Social engineering is a collection of various psychological approaches and deceptive methods aimed at obtaining credential information about a person with fraudulent sources. They do this to obtain sensitive information like usernames/passwords, confidential personal information, bank card numbers and anything else that can cause personal or financial harm. The term "social engineering" comes from the field of hacking. Hackers often look for vulnerabilities in computer systems to get them *[Ryabchuk, 2019]*.

Social engineering is an attempt to trick someone into revealing information, for example, a password to attack systems or networks. Social engineering attacks are carried out by targeting the weakest link and by manipulating the victims. Social engineering requires the victim to maintain an asymmetric information relationship with the attacker, who uses this asymmetry to establish technocratic control over the victim *[Ryabchuk, 2019]*. A social engineering attacker is someone who wants to access sensitive information or money. When the attacker manipulates the victim, it will notify the victim's vengeful target, causing inconvenience to be avoided. Successful social engineering attacks depend on manipulating the target or inciting them to reveal personal information. Social engineering attacks have turned into phone calls, emails, and face-to-face interactions and now even artificial intelligence by using chatbots. Social engineering attack methods include phishing, social engineering attacks against the Internet social unity or social networks, automated social engineering, and semantic attacks. With the growing digital era of information technology, different types of social engineering are also developing. Thanks to artificial intelligence, social engineering attacks are scarier than ever. Artificial intelligence can enable cybercriminals to create highly targeted links, websites, emails, and social media posts that people can click on. Criminals use machine learning and predictive analytics to find targets and

create manipulative messages. In this research paper, we'll look at the types of social engineering attacks and the impact of AI on them.

2.1 Social Engineering Attacks

Social engineering attacks concentrate on the attacker's use of persuasion and trust. When they encounter these tactics, people are more likely to do things which they would not else be suitable to do. Social engineering attackers quest victims to gain nonpublic information that can be used for specific purposes or vended on the black market and dark web. Social engineering attacks differ from each other, they partake a common pattern with analogous stages. The common pattern consists of four phases (1) collecting information about the victim;(2) developing a relationship with the victim;(3) using available information and launching an attack and (4) exit without a trace. The figure1*[Salahdine & Kaabouch, 2019]* below depicts above mentioned four stages of social engineering attacks

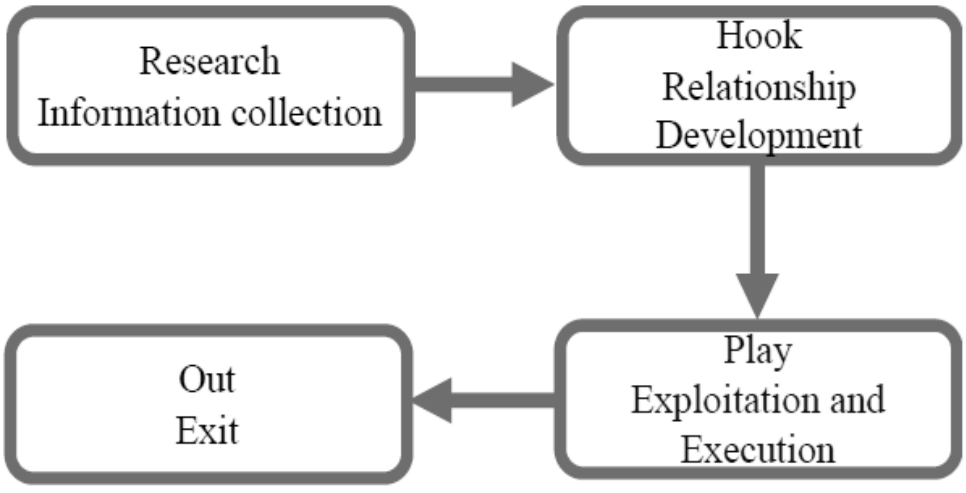

Figure1: Stages of Social Engineering Attacks [Salahdine & Kaabouch, 2019]

During phase one, also known as information gathering, the attacker selects a victim based on specific requirements. In the hook phase, the attacker begins to gain the victim's trust through direct communication or email communication. During the execution phase, the attacker emotionally affects the victim by providing confidential information or making security vulnerabilities. In the exit phase, the attacker leaves without a trace *[Salahdine & Kaabouch, 2019].*

2.2 Social Engineering Attacks Classification

Social engineering attacks into two categories: Human based, and System based. In human-based attacks, the attacker performs the attack himself by interacting with the victim to gather information. Therefore, they can affect a limited number of targets or victims. Software-based attacks are launched using devices such as systems, mobile devices, or artificial intelligence to obtain information from targets. with software-based attackers can attack multiple victims within seconds. The below figure2 *[Salahdine & Kaabouch, 2019]* depicts the classification of social engineering attacks.

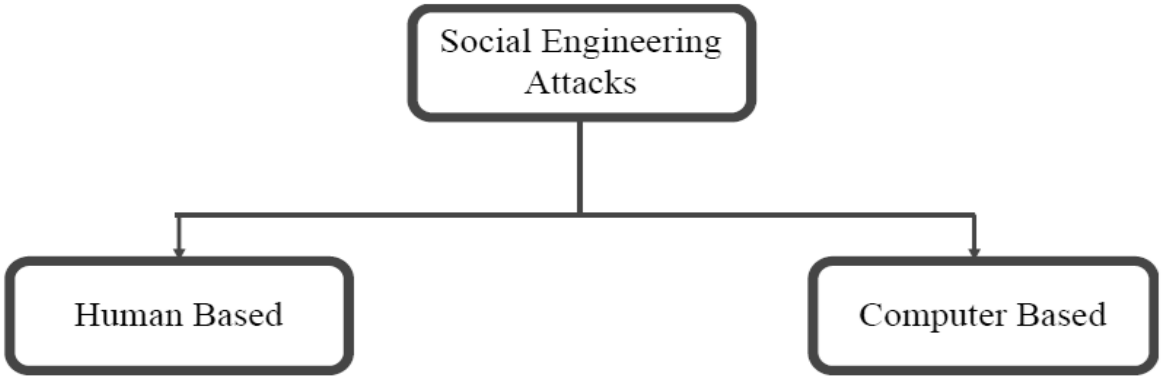

Figure 2: Social Engineering Attacks Classification [Salahdine & Kaabouch, 2019].

Social engineering attacks can also be divided into three categories depending on how the attack is carried out: social, technical, and physical attacks, as shown in figure3*[Salahdine & Kaabouch, 2019]*.

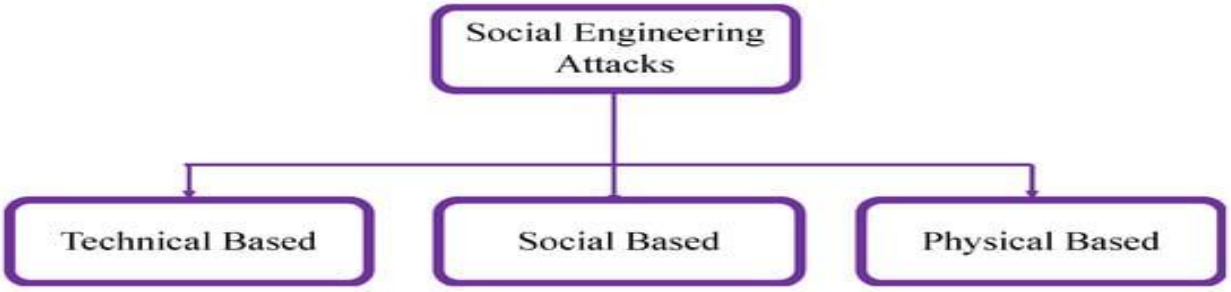

Figure 3: Social Engineering Attacks Classification [Salahdine & Kaabouch, 2019].

Social attacks are carried out through relationships established to play on the psychology and emotions of the victims. These attacks are the most dangerous and effective as they involve human interactions. Examples of such attacks are bait and spear phishing. Technical attacks are carried out over the internet via social media and websites and collect requested information such as passwords, credit card information and security questions *[Salahdine & Kaabouch, 2019]*. Nowadays it is happening with the help of artificial intelligence. Physical attacks refer to physical actions taken by an attacker to gather information about a target. An example of such attacks is the search for valuable documents in trash cans, which is called dumpster diving.

Social engineering attacks can be divided into distinct categories based on different perspectives. By analyzing the various existing classifications of social engineering attacks, we can divide these attacks into two broad categories, direct and indirect. Attacks classified in the first category use direct contact between the attacker and the victim to conduct the attack. They refer to attacks conducted through physical or visual contact or vocal interactions. They may also require the attacker to be present at the victim's workplace to conduct the attack. Examples of such attacks

include physical access, shoulder surfing, container diving, social engineering by phone, fraud, technical support call, and theft of important documents. Attacks classified in the indirect category do not require the presence of an attacker to conduct the attack. The attack can be initiated remotely via malware transmitted via email or SMS attachments. Examples of such attacks are phishing, fake software, pop-ups, ransomware, SMS hunting, online social engineering, and reverse social engineering *[Salahdine & Kaabouch, 2019]*. Figure 4*[Salahdine & Kaabouch, 2019]* depicts the examples of above-mentioned attacks.

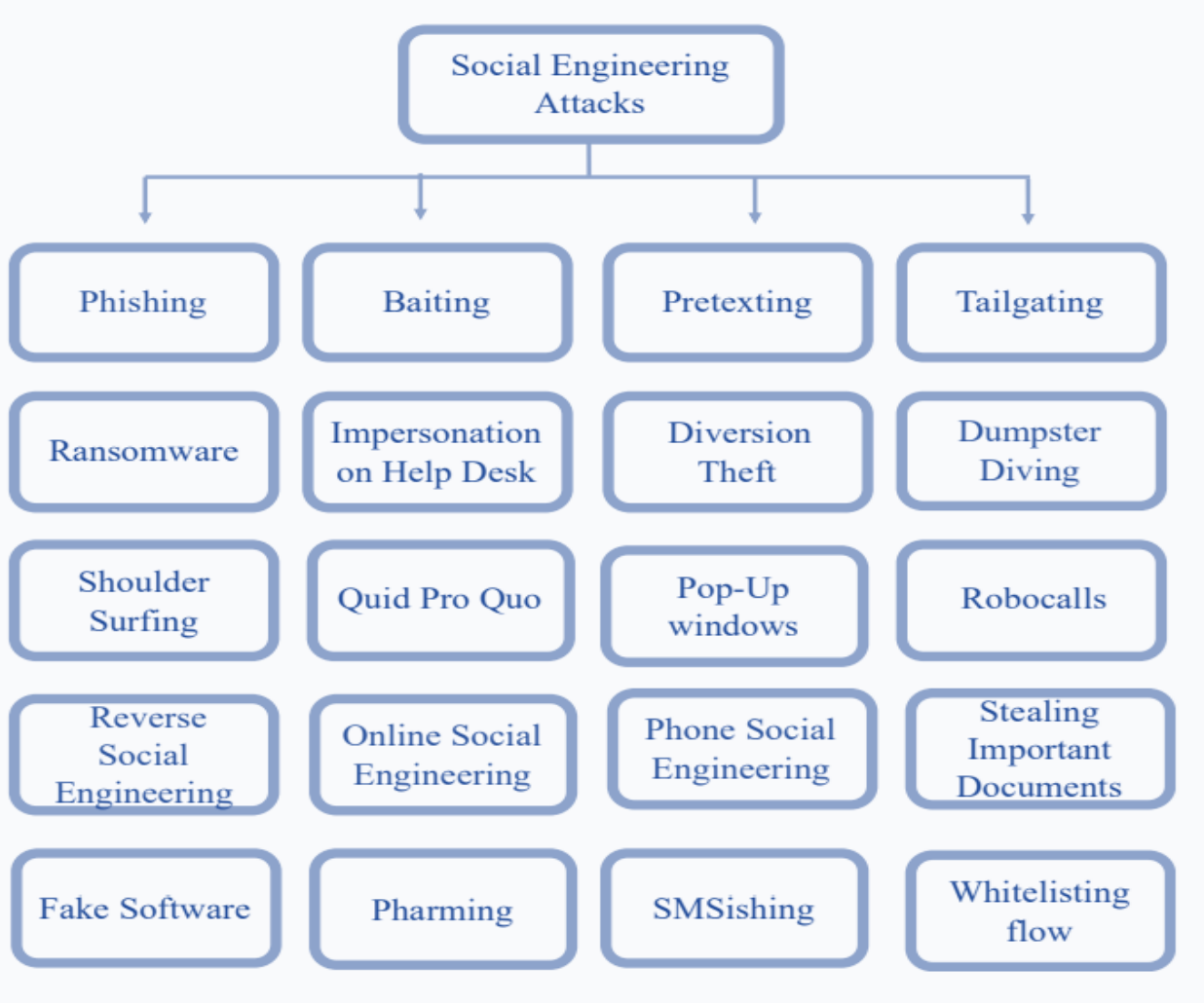

Figure 4: Social Engineering attack types of Examples [Salahdine & Kaabouch, 2019]

3 Social Engineering Attacks Illustration

3.1 Phishing Attacks

The core element of a phishing attack is a message sent via email, social media, or other electronic means of communication. The phisher can use public resources, especially social media, to gather essential information about the victim's personal and professional experiences. These resources are used to collect information such as the potential victim's first and last name, location and e-mail address, interests, and activities. The phisher can use this information to create a credible fake message. Typically, the email the victim receives appears to be from a known person or organization. The attacks are conducted through malicious attachments or links to malicious websites. Attackers often create fake websites that appear to belong to a trusted organization, such as the victim's bank, workplace, or university. Attackers try to collect confidential information such as usernames and passwords or payment information through these websites *[Salahdine & Kaabouch, 2019]*.

3.1.1 Phishing Attack Types

There are diverse types of phishing which are as follows Email phishing, Malware Phishing, Spear phishing, Whaling, Smishing, Vishing, Angler Phishing.

Email Phishing

The most ordinary form of phishing, this type of attack uses tactics such as fake hyperlinks to lure email recipients into sharing their personal information. Attackers often pretend to be a major account provider like Microsoft or Google or even a coworker.

Malware Phishing

Another common approach to phishing, this type of attack, is to attach malware to an email (such as a resume or bank statement) that appears to be a trusted attachment. In some cases, opening a malware attachment can cause entire information systems to fail.

Spear Phishing

While most phishing attacks cover a wide net, spear phishing targets specific individuals using information gathered through research about their work and social lives. These attacks are highly personalized, making them particularly effective at bypassing basic cybersecurity.

Whaling

When bad actors target the "big hitman" like a business manager or celebrity, it's called whaling. These scammers often conduct a thorough study of their targets to find the right moment to steal login credentials or other sensitive information. If you have a lot to lose, attacking whalers have a lot to gain.

Vishing and Smishing

Vishing is voice phishing where an attacker poses as legitimate source and tricks targets to gather credential information. Smishing is SMS phishing where attackers use similar tricks used in email phishing and vishing, but the only difference is this attack is carried out via SMS.

3.2 Pretexting

Pretexting attacks consist of creating false and believable scenarios for the theft of the victim's personal data. They rely on pretexts that make the victim believe and trust the aggressor. The attack is conducted by phone, email, or physical means [2]. Attackers are publishing information in co-worker’s directories, public websites, or conferences Field meeting to carry out

the attack. The pretext may be an offer to provide services or find employment, ask for personal information, help a friend access something or win the lottery.

3.3 Baiting

Baiting is a type of social engineering attack where the scammer uses false promises to lure the victim into a trap that could steal personal and financial information or cause malware on the system. The trap can take the form of a malicious attachment with a seductive name.

The most ordinary form of baiting is that it uses physical methods to spread malware. For example, attackers pull malware-infected flash drives into visible areas where they are confident potential victims will see them. When the victim inserts a flash drive into a computer at work or at home, the malware automatically installs itself on the system. Bait scams are also available online in the form of attractive advertisements that redirect users to malicious websites or encourage users to download a malware-infected app.

3.4 Tailgating Attacks

Tailgating is considered a psychological manipulation in which attackers unknowingly complicate employees in their crimes. It is an information security trick designed to allow attackers access to restricted areas and information by deceiving authorized individuals, such as spear phishing or phishing, including whaling.

Although tailgating and piggybacking attacks are often used interchangeably, it's important to remember that the two have distinct differences. Monitoring attacks are attacks in which an attacker follows an unsuspecting user to gain access to an unauthorized domain. Conversely, in a piggybacking attack, an employee or former employee knowingly provides access to the protected environment to an unauthorized person as part of a coordinated attack.

3.5 Ransomware Attack

Ransomware is malware which is designed to deny a user or firm access to files on their computer. By cracking these files and demanding a rescue payment for the decryption key, cyberattacks place associations in a position where paying the rescue is the easiest and cheapest way to recapture access to those files. Some variants have added fresh functionality similar as data theft to give farther incitement for ransomware victims to pay the rescue. Ransomware has snappily become the most important and visible type of malware. Recent ransomware attacks have impacted hospitals capability to give pivotal services, crippled public services in metropolitan cities, and caused considerable damage to different firms *[Salahdine & Kaabouch, 2019]*.

3.6 Fake Software Attacks

Fake software attacks, also known as fake sites, rely on fake sites to trick victims into believing they are known and trustworthy sites or software. The victim enters real login information on a fake website that gives the attacker the victim's credentials to use on a legitimate website.

For example, access to online banking accounts. An example of such threats is tab nabbing that uses a fake website that pretends to be the login page for a popular website frequently visited by the victim, such as online banking, Facebook, or Twitter. Victims enter their login information when: focusing on something else. A malicious user exploits the victims' trust in these websites and gains access to their credentials.

3.7 Reverse Social Engineering Attacks

Reverse social engineering attackers claim to have fixed the network issue. It includes three main steps: cause a problem such as network blocking; advertising that the attacker is the only

person to solve this problem; Solve the problem by taking the information you want and leaving it without it being detected.

3.8 Pop Up Windows

Pop-up attacks refer to windows that appear on the victim's screen and inform about the loss of connection. The user responds by re-entering their login credentials, causing the pre-installed malware to run in a windowed view. This program transmits remotely the attacker's login information. For example, pop-ups can be random pop-up warning messages about online advertisements designed to trick the victim into clicking on the window. Pop-ups can also be fake messages that warn you that a virus has been detected on the victim's computer. A pop-up will prompt the victim to download and install a recommended antivirus to protect the computer. There can also be false alerts that your computer's memory is full and needs to be scanned and cleaned to gain more space. The victim panics and reacts quickly to solve the problem malware transferred in popup.

3.9 Robocall Attacks

Robocall attacks have emerged recently as mass calls from computers to specific individuals with known phone numbers. Their targets are mobile, home and work phones. robocall is a device or computer program that automatically dials a list of phone numbers to deliver recorded messages. It relies on the Voice over Internet protocol to provide various VoIP functions.

like interactive voice response and text-to-speech. These calls may be for the provision or sale of services or for problem resolution. Helping solve tax problems is a well-known example of an offensive that has gained momentum in recent years. When the victim answers the call, the phone

number is stored in the attacker's database. Even after these connections are blocked, attackers' systems continue to call from other numbers.

3.10 Other Attacks

There are many other types of attacks that can be summarized as follows:

- Impersonation in Help Desk attacks: The attacker pretends to be someone in authority or authority calling a company employee and the helpdesk to request information or services.
- Dumpster Diving attacks: collect sensitive documents from company or company garbage waste equipment such as old computer equipment, drives, CDs, and DVDs.
- Quid Pro Quo Attacks: Trap attacks that offer free victim seduction services. Require exchange of information in exchange for a service or product.
- Diversion theft attacks: send a shipping company to a courier or Pack to the selected location.
- Shoulder surfing attacks: These are the attacks in which the victim is followed while entering a password or Sensitive information.
- Crucial Document Theft Attacks - This includes stealing files from someone's desktop for personal interests.
- Online social engineering attacks: attacker pretends to be network administrator company and prompts for a username and password.
- Pharming attacks: An attacker steals traffic from a specific site by redirecting it to a specific site. another fake website from which the transmitted information can be obtained. This attack works by hacking the Domain Name System (DNS) server and the Internet protocol (IP) address of the host and server which takes advantage of all vulnerabilities.

4. Artificial Intelligence

Artificial Intelligence is the most admired technology which has become more prominent in today's era. Artificial Intelligence is omnipresent in our essence, from reading our emails to getting directions to music or movie recommendations. Artificial intelligence is now part of our life, whether we are apprehensive of it or not.

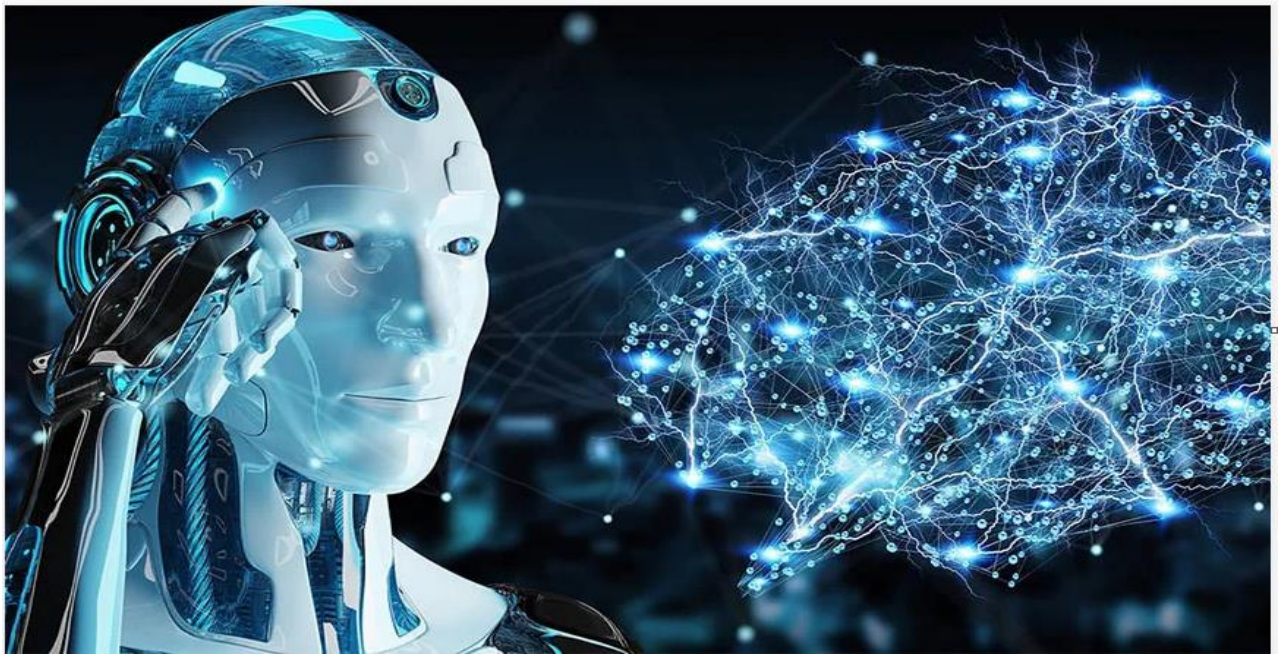

Figure 5: Artificial Intelligence

Artificial intelligence is eternal. From autonomous cars, dictation tools, translation applications, predictive analytics, and application tracking, as well as retail vehicles such as smart shelves and cars, to applications that assist people with disabilities, AI can be a powerful component in great technology products and services. Every time we open our Facebook, surf Google, buy an Amazon recommendation, or book a trip online, artificial intelligence lurks in the background. There are also popular artificial intelligence applications that help us. There are several artificial intelligence programs out there now, including Apple's Siri, Amazon's Alexa, and Google's assistant. Or pop-

up features on websites that ask for answers to constantly asked questions. On the other hand, it can also be misused maliciously.

5. Impact of Artificial Intelligence on Social Engineering

My initial reaction when I got to know about social engineering was shock at seeing how attackers exploit people. The ease with which sensitive information could be obtained astounded me. In today's world, social engineering attackers have evolved and are powered by future tools such as artificial intelligence, allowing them to exploit human psychology and gain access to data and systems. The impact of artificial intelligence on social engineering will be discussed here. When talking about the impact of AI let us consider the following: Threats Detection, Increase Authentication Level, Lower Cost and Resources for security implementation which falls under positive side. When coming to negative side a few new tools from AI such as Voice Spoofing, Deepfakes, Automated Social Engineer Bots will be discussed along with how they are affecting social engineering.

We will look in detail at Voice Spoofing, Deepfakes, Automated Social Engineer Bots.

5.1 Voice Spoofing/Cloning

Voice cloning usually requires you to collect hours of recorded speech to create a dataset and then use the dataset to train your new voice model. But not anymore. The new AI offers an extraordinary set of real-time voice cloning tools, such as Google's artificial intelligence that allows anyone to clone their voice from just five seconds of a voice sample. Users enter a short speech sample, and only the model trained for playback can deliver text-to-speech in immediately sampled speech style. According to research, our brain does not register a significant difference between real and artificial sounds. In fact, it is more difficult for our brain to distinguish between

fake sounds and fake images. These AI systems now require only a small amount of voice to train a working artificial voice that mimics a person's speech style and tone. The likelihood of abuse increases. So far, scientists have not been able to pinpoint the neural differentiation of how the brain can distinguish between what is true and what is false. Consider how artificial voices can be used in an interview, news, or press conference to make listeners believe they are listening to a government official or company CEO.

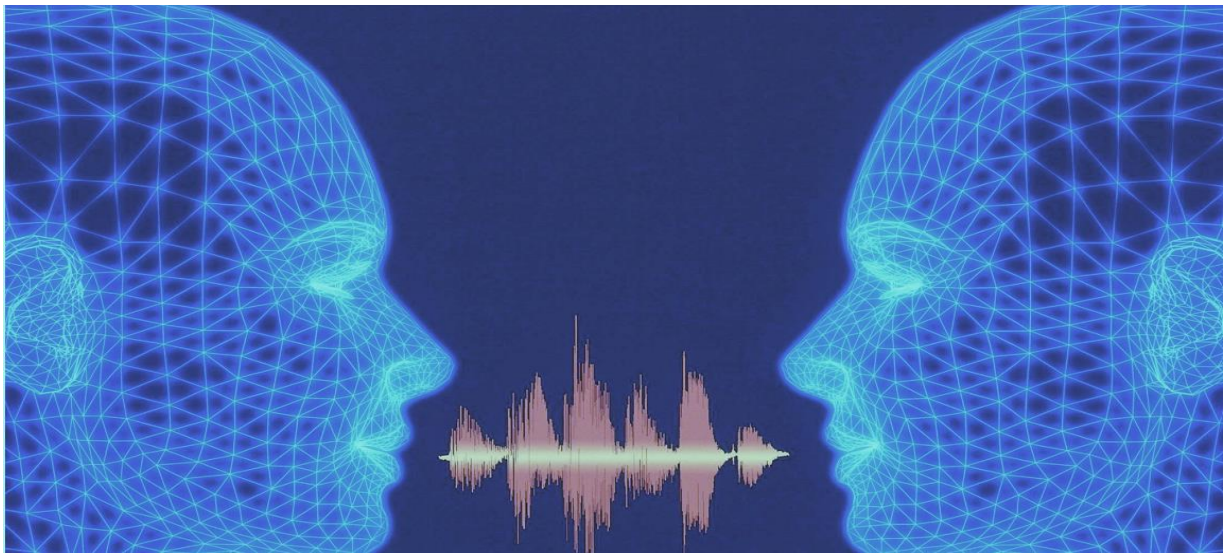

Figure 6: Voice Spoofing/Cloning

Voice Spoofing/Cloning is the process of creating a digital copy of someone's voice. Though the cybercriminals and social engineering attackers use it for malicious purposes there are several advantages to consider. One needs to make a training or marketing video and let us assume a company’s CEO just does not have the time or opportunity to be in front of the camera. With voice cloning, one can easily create high-quality audio and the business leader speaks directly to the audience without wasting time recording audio. Or you have a customer service center and need more engaging and personal interactions with customers. Voice cloning can also help with this, by providing authentic human interaction that seems to come from the agent himself. Now anyone

can clone their voice, thanks to artificial intelligence-assisted voice technology. Cybercriminals are getting smarter and incorporate voice chat into their attacks to lure victims to download malware or hand over sensitive data. The use of voice-based social engineering to justify attacks by threat groups is nothing new. But cybercriminals are exploring more innovative ways to integrate sound into attack chains. For example, cybercriminals set up call centers to integrate with the malware distribution process, while scammers are reinventing the wheel using voice phishing tactics to persuade victims to hand over sensitive data.

Figure 7: Voice spoofing/cloning

In 2019, attackers used voice-generated artificial intelligence to impersonate a CEO on the phone and convinced a UK company CEO to transfer $243,000 to an attacker's-controlled account. That is what happened to Bill Gates, whose voice was cloned by Facebook engineers, without his approval. Voice cloning is already used for fraud. In 2019, scammers cloned the voice of the

president and successfully tricked him into transferring substantial sums of money. Similar crimes were uncovered using the same technology. We will look at detailed description of examples how social engineering attackers cloned voice and lured lump sum of amount from victims *[Brewster,2021]*.

5.2 Deepfakes

A deepfake is a video or image manipulated by artificial intelligence to make you believe something that is not true. While most people use deepfake technology to create memes, bad actors use it to spread disinformation more widely and faster than ever before. For example, you can create people who do not really exist, or show real people doing and saying things they do not. Deepfake makes it incredibly easy to create highly confusing audio-visual content, meaning they can be used for good or for bad. Deepfake is a way to manipulate images. Deepfakes technology can seamlessly link anyone in the world to a video or photo they have never taken. Currently, the core technology that creates deep fraud is artificial intelligence. Deepfake technology using artificial intelligence offers greater possibilities, but also increases the scale of manipulation and interference by bad actors. The main component of deepfake is machine learning, which makes it possible to generate deepfakes much faster at a lower cost. To make a fake video, the creator first trains a neural network of hours of authentic video of the person to provide a realistic "understanding" of how they look from different angles and under different lights. They then combined a trained mesh with computer graphics techniques to superimpose one person over another actor.

Figure 8: Images which depict Real and AI Generated deep fake[google]

Despite the use of AI to make the process faster, it still takes time for the composite to be convincingly placed in a fictional situation. Deepfake videos are created by swapping a person's face with another with the help of a facial recognition algorithm and a deep learning computer network. Fake faces that look almost identical to real faces can be generated with AI deep learning models. There is almost no difference between an AI-generated face and a real face these days because these programs have become so powerful over the years. Deepfakes are tricky to detect. They can manipulate and threaten individuals and companies. Executive members can protect themselves and their companies by becoming familiar with the technology. Some of the most popular deepfake apps are faceapp, Speakpic, wombo.

5.3 AI based automated social engineering bots

A bot is a program that recreates human movement by connecting with frameworks or other users. Bots are mechanized to do specific errands/tasks and interactions and can frequently run without human help. They start an immense measure of the traffic on the web, and there are both good and bad bots. Ordinarily, they often perform repetitive activities to automate tasks. Bots, also called web bots or robots, are involved across numerous enterprises for online client care,

booking, and that is just the beginning. Explicitly in the domain of network safety, bots often support detection and response platforms to reduce the need for cybersecurity specialists in a workforce shortage. A chatbot is a rules-based computer program that simulates human interaction with end users through a chat interface. In other words, the chatbot can talk to you like a real person, ask questions and answer questions according to predefined rules and logic. Good bots often surf the Internet to suit our needs and requirements. For example, Google bots help catalog online content so our search results can be faster and more optimized. On the other hand, chatbots are a suitable alternative for customer service as they interact with users to observe and serve appropriately. On the other hand, bad or malicious bots can be programmed to hack user accounts and steal data, infect computers with dangerous viruses or malware, or send constant spam that leads to a site crash. Cybercriminals use malicious bots to take control of one computer and connect it to another to create a network of "zombie computers" known as botnets, which can then launch large-scale cyberattacks, thus shutting down Internet users completely.

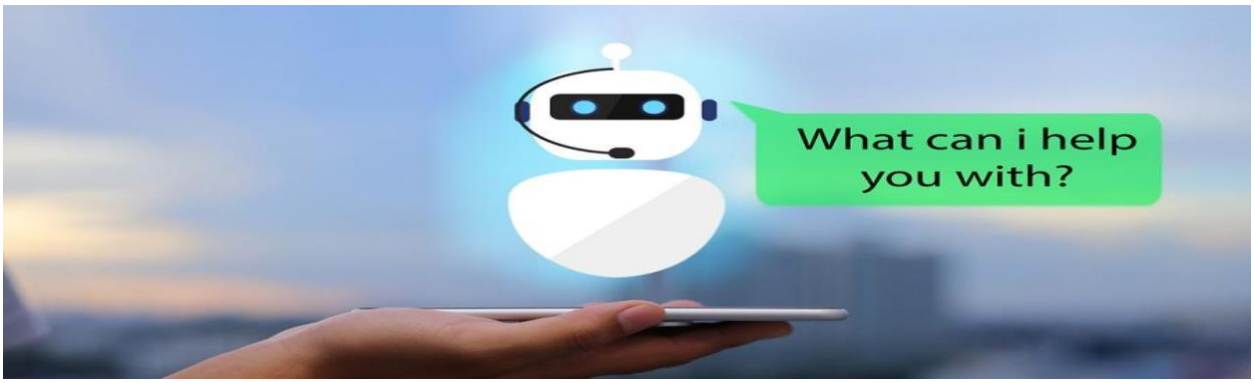

Figure 9: Artificial intelligence Chatbot

There are some of the best AI chatbots used across various industries. Watson Assistant, Rulai, Inbenta, LivePerson, Bold360.

The following example illustrates how threat actors used voice cloning/spoofing, deepfakes, and automated social engineering bots to exploit their victims or organizations *[Theertharaja, n.d.]*.

6. Real life scenarios of Social Engineering attacks

6.1 Voice cloning/Spoofing

Example1: Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case. As discussed earlier in our paper will look in detail at how cybercriminals attempted voice cloning with the help of Artificial intelligence. Cybercriminals impersonated a UK business owner using artificial intelligence (AI) and voice technology, resulting in the illegal transfer of $243,000. According to the Wall Street Journal, an unidentified hacker group allegedly used AI-enabled software to impersonate the voice of a well-known corporate executive to deceive his subordinate, the CEO of a UK-based energy subsidiary. The CEO was then persuaded to execute transactions under the guise of necessary money intended for the company's German parent.

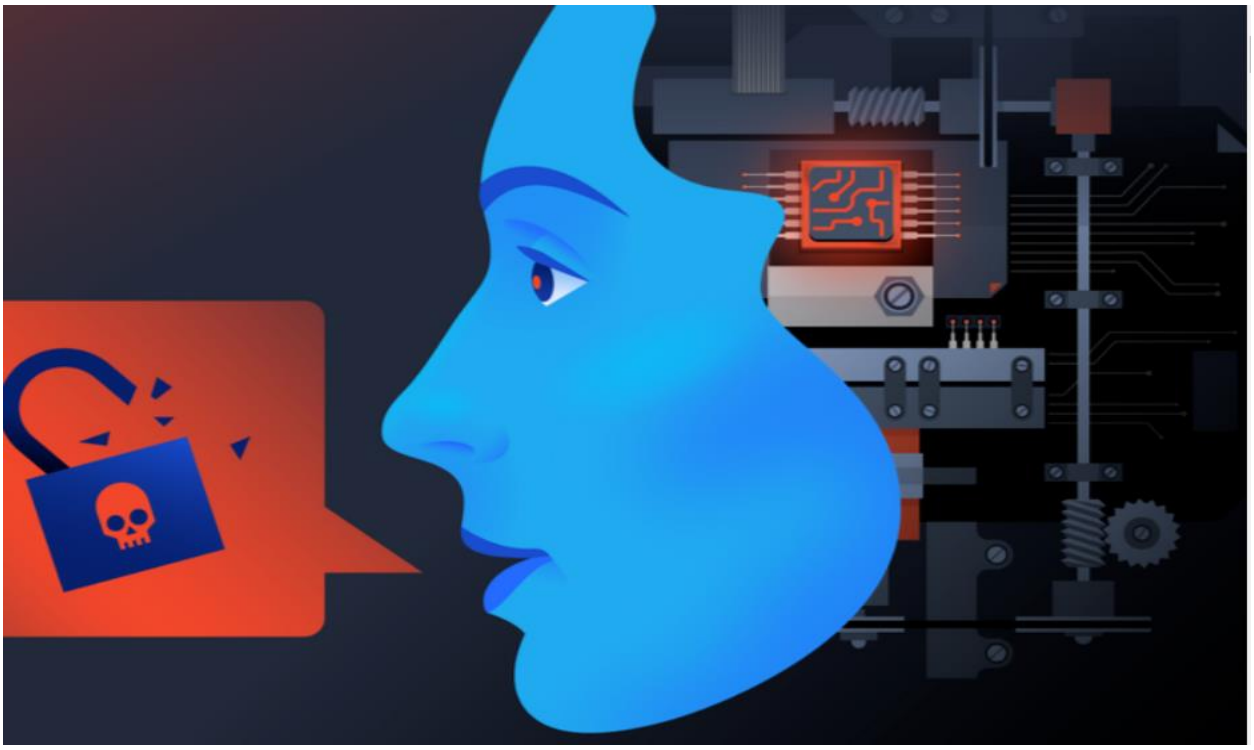

Figure 10: AI based voice cloning

The fraudsters allegedly called the CEO in the UK and demanded a transfer to a Hungarian supplier. They contacted him once more, this time posing as the CEO of the parent firm, to reassure him that they would pay back the transfer. Before any refund monies had surfaced, the CEO was then contacted a third time to request an additional urgent transfer. The CEO then became suspicious and refused to make any additional payments. However, the money that was sent to Hungary was quickly transferred to Mexico and other places, and law enforcement is still seeking for culprits. "It requires fewer recordings to fake voices. These are beginning to become even simpler to produce as computational power rises, which paints a terrifying future."*[Brewster,2021]*

Example2: Another such scenario is where despite warnings regarding the use of the modern technology by cybercriminals, AI voice cloning is utilized in a massive crime that Dubai investigators are looking into. Early in the year 2020, a bank manager in Hong Kong got a call from a man whose voice he recognized a director at a business he had previously spoken with. The director had wonderful news: His company was set to complete an acquisition, and he needed the bank to approve certain transfers worth $35 million. The bank manager could see emails from the director and Martin Zelner, verifying what money needed to transfer where, in his inbox because he had employed Zelner as a lawyer to organize the operations. The bank manager started the transfers because he thought everything looked authentic.

According to a court document obtained by Forbes, the U.A.E. has asked American investigators for assistance in locating $400,000 in stolen funds that were transferred into American-based accounts held by Centennial Bank. He did not realize that he had been the victim of an elaborate scam in which fraudsters had impersonated the director's speech using "deep voice" technology. The United Arab Emirates (UAE), which is looking into the crime since it involved domestic

businesses, thinks it was a complex conspiracy involving at least seventeen people that transferred the stolen funds to accounts around the world. It is only the second known instance of fraudsters allegedly using voice-shaping tools to commit a heist, but it seems to have been much more successful than the first, according to the Wall Street Journal, in which fraudsters attempted to steal $240,000 in 2019 by impersonating the CEO of a U.K.-based energy company using the technology *[Brewster 2021]*.

6.2 Deepfakes

What a sight it would be to sit around a roundtable with Robert Downey Jr., Tom Cruise, George Lucas, Jeff Goldblum, and Ewan McGregor! But it was a deep fake. As Netflix and other streaming services compete for viewers, the entertainment sector is currently undergoing a seismic transformation. Taking this into consideration, Collider created this hilarious deepfake featuring the faces of Tom Cruise, Robert Downey, Jr., George Lucas, Ewan McGregor, and Jeff Goldblum debating streaming and the future of film. Both Barack Obama's PSA and the Donald Trump mocking Belgium for sticking with the Paris Climate Agreement were posted by Buzzfeed. The 21st century's response to photoshopped photographs and films is these fantastic deepfake examples. Artificial intelligence deep learning technology is used in synthetic media, or "deepfakes," to replace an existing person in a picture or video with a different person.

2018 saw the release of a deepfake of former president Barack Obama by BuzzFeed Video. You couldn't tell the video was false because the deepfake flawlessly and correctly imitated his voice and mannerisms. The identity of the individual who was portraying the former president was made known toward the end of the video by actor Jordan Peele.

Despite its message that it's crucial for individuals to use the internet responsibly in a time when misinformation is common, the satirical film generated ethical questions among viewers. [*Six Real life deepfake examples*,2022]

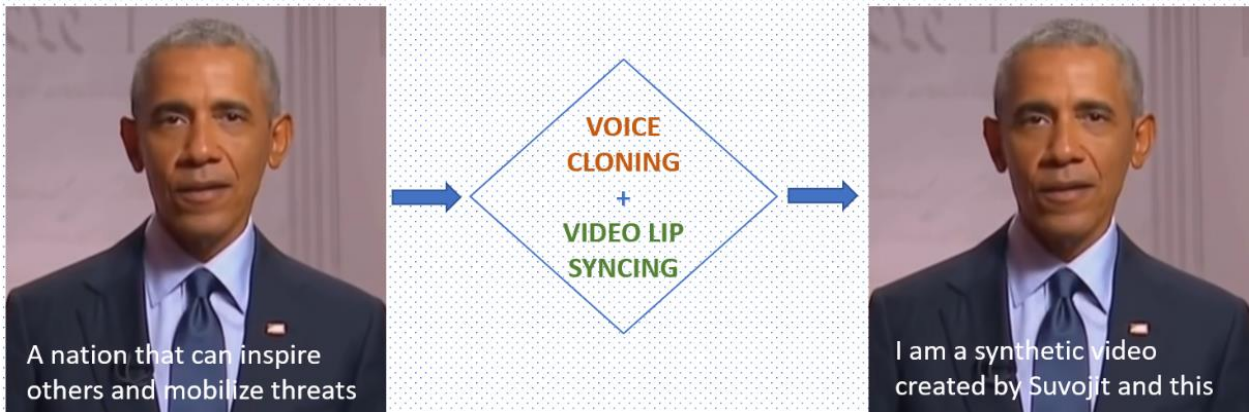

Figure 11: Creating Deepfake

The fact that popular celebrities have a lot of photographs readily available online for AI to train on and learn from is one factor contributing to the widespread application of deepfake technology in these individuals. But as technology advances and more photographs and data are put online every day by average people, it will not be long until a deepfake is used on real people. Deepface creation is already accessible to the public thanks to AI tools like FaceApp and DeepFaceLab.

6.3 Automated Social engineering bots

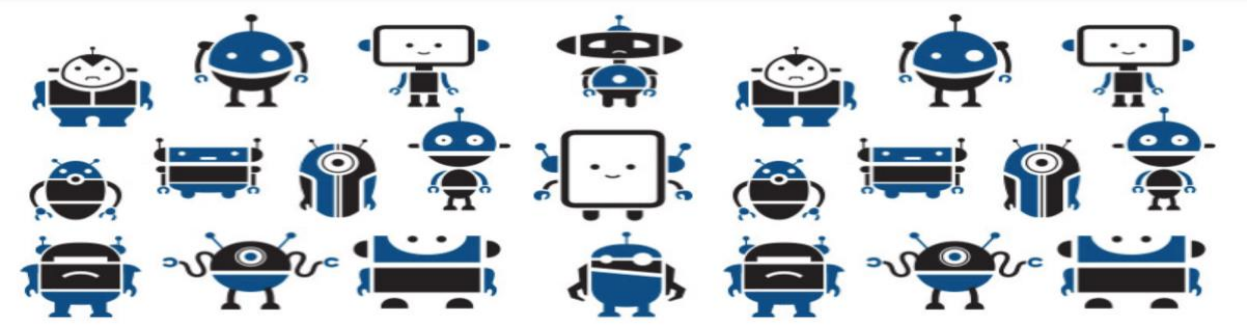

Figure 12: AI based automated social engineering bots

Malicious chatbots are used in a new Phishing Attack to Perform Real-Time Social Engineering. Trustwave researchers have seen a phishing campaign that uses a chatbot to give the fraud more credibility. By starting a discussion and guiding the victim through the process, the chatbot on a legitimate website persuades the user to visit the phishing site. According to the study, integrating chatbots adds an interactive element to a website. Because it increases the site's user interest and engagement, this frequently leads to a greater conversion rate. This is what the people behind this phishing attempt are hoping to take advantage of. Along with impersonating the target brand on the phishing email and website, the chatbot-like element gradually leads the victim to the real phishing sites. The phishing website also appears more trustworthy thanks to the addition of mobile OTP and CAPTCHA pages. An email warning the recipient that their Facebook page has broken community standards gives them 48 hours to appeal the judgment before their page is destroyed. This is the first step in a phishing attempt. The user is allegedly given the opportunity to remedy the issue at Facebook's support center, and is advised to click on an "Appeal Now" button to get there. When the victim clicks that button, a chatbot impersonating a Facebook customer service representative will start a conversation in Messenger. The chatbot's Facebook page is a typical company page with no followers and no posts. A notice stating that the profile is "Very responsive to messages" would be visible if a victim checked the profile, indicating that it is being utilized. The victim will receive an "Appeal Now" button from the chatbot on Messenger, which directs users to a website posing as a "Facebook Support Inbox" even though the URL is not on Facebook's domain. Additionally, as Trustwave points out, the case number on that page does not correspond to the one that the chatbot earlier supplied. Nevertheless, those details are not likely to alert worried people to the fraud. The primary phishing page, which is displayed below, asks users to submit their email address, complete name, page name, and phone number to appeal

the decision to delete their page. Once this information has been input and the "Submit" button has been clicked, a pop-up window demanding the account password will appear. Following that, a POST request is used to send all the data to the threat actor's database. Finally, a bogus 2FA website is accessed and the victim is instructed to enter the OTP they received through SMS on the supplied phone number. That page will accept anything; thus, its sole purpose is to give the entire procedure a false air of validity. Following the verification, the victims arrive on a genuine Facebook page with copyright and intellectual property policies that are purportedly pertinent to the user's breach. Since the phishing attack is automated, the actual use of the stolen credentials may occur at a later stage. Therefore, the threat actors must cultivate in the victims' minds a false sense of validity to postpone any remediation activities for the breach. To automate the theft of credentials and scale up their operations without expending a lot of time or money, threat actors are increasingly using chatbots in phishing assaults. Since many websites include AI and chatbots in their help pages, it can be more difficult to spot these frauds when they occur. When support cases are opened, these scams can appear to be normal *[Phishing websites use chatbots to steal information, 2022]*.

7. Graphical Representation

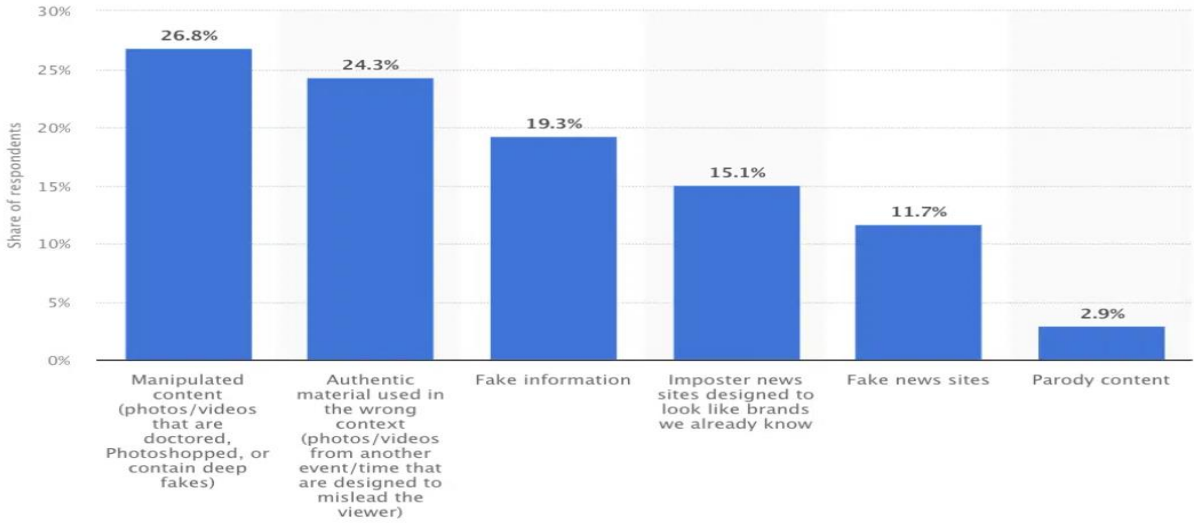

Figure 13: Graphical representation of Deepfakes

As in today's environment, when fake news spreads like wildfire, tools like Deepfake stir up disorder and dissatisfaction in both the individual and societal levels. As shown in the figures above, in the US, 26% of all false news is produced using photo editing software, such as Photoshop and Deepfakes.

8. Detection and Precautionary measures

Deepfake

We have examined the effect artificial intelligence has on social engineering so far in this work. We will now examine how to recognize such attacks and what safety precautions to take to be safe from them. There are few ways to detect deepfake videos.

- Videos where the person never blinks, blinks too frequently, or blinks in an unnatural way are the result of current deepfakes' struggles to convincingly animate faces.
- Keep an eye out for facial issues, skin, or hair issues, and faces that seem blurrier than the surroundings in which they are situated. Abnormally soft appearance of the focus.
- Is the lighting artificial looking? Since the lighting in the target video is far better than that in the model clips used to create the fake video, deepfake algorithms frequently keep the lighting from those clips.
- If the video was fabricated but the original audio was not, it might not seem like the voice belongs to the person.

Voice Cloning/Spoofing

To identify when hackers are attempting to spoof/clone a voice, a new solution known as Void (Voice liveness detection) can be embedded in a smartphone or voice assistant software. Void works by identifying the differences in spectral power between a live human voice and a voice replayed through a speaker. If you have any cause to think the caller is replying to your dialogue with pre-recorded clips, ask the same question again and listen to the answer to determine if it differs from the previous response. Ask open-ended questions that will be challenging to respond to using a script. The ideal posture for such a call would be for the caller to adopt a more powerful, authoritative, and demanding tone in discussion to direct the conversation and have more control over what they need to "play." If people start asking questions and veering off topic, this plan will backfire. So, if you are unsure, try posing a divergent question.

Automated Social Engineering Chatbots

Never give out any personal information to anyone. This goes for names, usernames, email addresses, passwords, PINs, and any other information that could be used to identify you. Such questions are never asked of clients by bank people or automated callers. Avoid taking calls from unknown numbers, and if you do and a caller asks you for personal information, hang up right once. Take it slowly. Scammers frequently use a false feeling of urgency to coerce you into providing your personal information. Hang up or say you will call back later if someone tries to pressure you into making a choice. Call the company's main number that they are claiming to represent next.

9. Future Agenda

This project's major goal is to describe how social engineering attacks are affected by artificial intelligence. It focuses exclusively on the usage of AI in the attacks. The impact has

advantages and disadvantages. This blog post will discuss recent social engineering assaults that have taken place and how AI is attempting to find a solution. This blog will explore recent attacks, that involved manipulation and control of social media. This essay will examine why these attack methods' capacity to blend into existing systems makes them so potent. Additionally, a quick overview of how to spot the symptoms and address the issue will be included.

10. Conclusion

This study aims to educate readers about the impact social engineering is having due to artificial intelligence. Everyone in the present era needs to be aware of these recent technologies and the ways in which they might be utilized maliciously. By including these topics in routine cybersecurity education, we can all be safer and more secure in the face of new forms of criminality. For instance, the voice transfer AI fraud victim might have used other means to authenticate the caller and changed the outcome if they were aware that anyone's voice could be copied live over the phone. Like this, if people knew how realistic and straightforward deepfake technology has gotten, everyone would have a healthy level of skepticism, especially those watching political films. Use antivirus software, avoid posting privileged information on social media sites, make complicated passwords, avoid clicking on dodgy links, and be wary of mailing list subscriptions. Businesses should teach staff members the fundamentals of information security. Most significant, though, is that you should approach the "news" you find online with a fair amount of skepticism. As we indicated in our thesis paper, it is highly unlikely that social engineering attacks will never occur, but if we are aware of all the methods and have safeguards in place, we may avoid becoming a victim.

References

Brewster, T. (2021, Oct 14*.). Fraudsters Cloned Company Director's Voice In $35 Million BankHeist, Police Find*. Forbes. https://www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/?sh=b25d20575591

Foley, J. (2022). *19 deepfakes examples that terrified and amused the internet.* Futureplc. https://www.creativebloq.com/features/deepfake-examples

Gayan, A. (2021, Mar 28). *Impact of AI on Social Engineering.* Medium. https://medium.com/unpackai/impact-of-ai-on-social-engineering-e9bd763a77db

Hatfield, J.M. (2018, Mar.). *Social engineering in cybersecurity: The evolution of a concept.* Computer Secure, *73*, 102-113, - https://www.researchgate.net/publication/341199647_Defining_Social_Engineering_in_Cybersecurity

King, T. C., Aggarwal, N., Taddeo, M., & Floridi, L. (2020). *Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions.* Science and engineering ethics, *26*(1), 89-120. https://link.springer.com/article/10.1007/s11948-018-00081-0

Koyun, A. & Janabi, E. A. (2017, June.). *Social Engineering Attacks.* Journal of Multidisciplinary Engineering Science and Technology, *4*(6), 7533-7538. https://www.jmest.org/wp-content/uploads/JMESTN42352270.pdf

Phishing websites use chatbots to steal information(n.a.). (2022, Mar 23). Phishing Tackle. November 9,2022. Retrieved from https://phishingtackle.com/articles/phishing-websites-use-chatbots-to-steal-information/

Ryabchuk, N., Grishko, N., Grishko, V., Rudenko, A., Petryk, V., Bapiyev, I., & Fedushko, S. (2019). *Artificial Intelligence Technologies Using in Social Engineering Attacks*. Semantic Scholar. https://www.semanticscholar.org/paper/Artificial-Intelligence-Technologies-Using-in-Ryabchuk-Grishko/ac7be08160bb86c6c1a9349ea7d2bb9f7e8be12d

Salahdine, F. & Kaabouch, N. (2019, April.). *Social engineering attacks: A survey* Future Internet, *11*(4), 89; https://doi.org/10.3390/fi11040089

Six Real life deepfake examples. (2022, Nov 9). Q5id proven identity management. November 9, 2022, Retrieved from https://q5id.com/blog/6-real-life-deepfake-examples

Social Engineering. (2022). Carnegie Mellon University. https://www.cmu.edu/iso/aware/dont-take-the-bait/social-engineering.html

Theertharaja, B.(n.d.) *Advanced Automated Social Engineering Bots: The High Tide of Social Engineering Bots and the Scammers Riding Them*. CloudSEK. https://cloudsek.com/advanced-automated-social-engineering-bots-the-high-tide-of-social-engineering-bots-and-the-scammers-riding-them/