International Journal of Scientific Research in Multidisciplinary Studies
Vol.**10**, Issue.**3**, pp.**19-30**, March **2024**
E-ISSN: 2454-9312 P-ISSN: 2454-6143
Available online at: www.isroset.org

Research Article

Investigating The Security and Privacy Issues in ChatGPT Usage and Their Impact on Organisational and Individual Security

Polra Victor Falade[1]

[1]Dept. of Cyber Security, Nigerian Defence Academy, Kaduna, Nigeria

**Corresponding Author: pvfalade@nda.edu.ng*

Received: 20/Jan/2024; **Accepted**: 21/Feb/2024; **Published**: 31/Mar/2024

Abstract—Artificial intelligence (AI) technologies, particularly AI-driven chatbots, have become integral to modern communication and task automation. Among these, ChatGPT, developed by OpenAI, has garnered significant attention as a widely accessible natural language processing chatbot. However, concerns regarding security and privacy have accompanied its widespread adoption. This research paper employs blog mining as a methodology to provide a comprehensive analysis of the security and privacy implications surrounding ChatGPT. By extracting insights from relevant blog posts, combined with an examination of existing literature, this study identifies inherent vulnerabilities within the platform and explores potential threats that could exploit these weaknesses. Specifically, the research highlights the risks associated with the mishandling of personal information, the susceptibility to cyber-attacks, and the broader implications for individual privacy and organizational security. Moreover, the paper offers insights into the regulatory landscape governing AI chatbots and suggests future research directions to address these challenges effectively. Ultimately, this study underscores the critical importance of implementing robust security measures and regulatory frameworks to mitigate the risks associated with AI-driven chatbots and ensure their responsible deployment in the digital age.

Keywords— ChatGPT, security issues, Privacy issues, OpenAI, AI, Vulnerabilities, threats

1. Introduction

In recent years, the advent of artificial intelligence (AI) has revolutionized various aspects of human life, offering unparalleled convenience, efficiency, and innovation across diverse domains [1]. One prominent manifestation of AI technology is ChatGPT, an advanced language model developed by OpenAI, capable of generating human-like text responses based on user input [2]. While ChatGPT and similar AI-driven chatbots have garnered widespread acclaim for their ability to streamline communication and automate tasks, their adoption has also raised significant concerns regarding security and privacy [3].

This research paper aims to investigate the multifaceted landscape of security and privacy issues associated with the utilization of ChatGPT, elucidating the potential ramifications for both organizations and individuals. As organizations increasingly integrate AI technologies into their operations, understanding the inherent risks posed by ChatGPT becomes imperative for safeguarding sensitive data, preserving user privacy, and mitigating potential threats. Moreover, elucidating the consequences of these security and privacy challenges is crucial for informing policymakers and stakeholders, facilitating informed decision-making, and fostering the development of robust regulatory frameworks.

Through a comprehensive exploration of the security vulnerabilities, privacy implications, and ethical dilemmas surrounding ChatGPT, this paper seeks to provide valuable insights into the evolving cybersecurity landscape in the era of AI-driven communication. By elucidating the nuances of these issues and their impact on organizational practices and individual behaviour, this research endeavours to contribute to a nuanced understanding of the broader societal implications of AI adoption. Ultimately, by shedding light on the complexities of security and privacy in the context of ChatGPT usage, this paper aims to inform strategies for mitigating risks and fostering responsible AI deployment in an increasingly interconnected digital ecosystem.

The paper is organized as follows: Section 1 introduces the research, highlighting the problem statement, motivation, aims, objectives, and significance. Section 2 discusses related work on security and privacy issues with ChatGPT. Section 3 outlines the research methodology, including data acquisition and analysis procedures. In Section 4, results from the blog mining methodology are presented, along with discussions on the advantages and drawbacks of ChatGPT, its inherent security issues, potential security and privacy threats, and impacts on organizations and individuals. Finally, Section 5 concludes the research, emphasizing its significance and offering directions for future studies.

2. Related Work

ChatGPT, developed by OpenAI, stands as a notable achievement in the realm of AI chatbots, offering natural language responses to a diverse array of queries [1]. While it has gained widespread acclaim and found applications across numerous domains, its remarkable capabilities have also sparked significant scholarly discourse surrounding cybersecurity concerns [1]. These discussions predominantly centre on the potential risks and vulnerabilities inherent in its utilization [1].

Concerns regarding ChatGPT primarily revolve around the potential for information leakage and privacy breaches [4]. This apprehension stems from its ability to process and generate vast amounts of data, driven by sophisticated deep-learning algorithms [3]. Such algorithms, while powerful, can inadvertently uncover patterns in data that were not intended for disclosure, raising questions about the system's security and privacy. Unauthorized access to ChatGPT could lead to data breaches and privacy infringements, prompting the need for robust security protocols and controls to mitigate these risks. Despite these challenges, proponents argue that the benefits of ChatGPT outweigh its potential drawbacks [4].

ChatGPT operates through a combination of generative and retrieval methods, harnessing deep learning algorithms and extensive training data to provide well-crafted responses [5]. As a member of the Generative Pre-trained Transformer (GPT) family [2], ChatGPT has undergone fine-tuning via supervised and reinforcement learning, rendering it versatile in addressing a broad spectrum of tasks and topics [6].

The cybersecurity threats associated with ChatGPT are diverse and concerning. Its capability to generate various forms of malicious content, including malware code, phishing emails, macros, and zero-day viruses, poses significant risks [4]. Such capabilities empower cybercriminals to craft sophisticated and linguistically adept content, enhancing the effectiveness of their attacks while making detection challenging [3], [4].

Instances of data leaks and vulnerabilities linked to ChatGPT have underscored the platform's security challenges, with incidents such as "CVE-2023-28858" exposing personal information [4], [5]. These occurrences highlight the critical need for robust security measures to safeguard against unauthorized access and data exposure.

Moreover, ChatGPT's susceptibility to providing unreliable responses poses the risk of misinformation dissemination, prompting some jurisdictions to impose restrictions to mitigate potential misuse [4]. Additionally, the platform faces the challenge of bad actors utilizing jailbreaking techniques to bypass security measures and access malicious content, further compromising its integrity and security [3], [4]. Furthermore, ChatGPT's capacity to generate deep fake text complicates efforts to discern between original and fake content, adding another layer of complexity to content verification [4].

While existing research has provided valuable insights into the security and privacy challenges associated with AI chatbots like ChatGPT, there remains a need for further investigation into emerging threats, evolving regulatory frameworks, and best practices for mitigating risks. By building upon the foundational knowledge established in prior research, this study seeks to contribute to a comprehensive understanding of the security and privacy implications of ChatGPT usage, informing strategies for responsible AI deployment and governance in the digital age.

3. Research Methodology

In this section, we present the methodological framework employed in our study, encompassing the systematic processes of data collection, screening, extraction, and subsequent analysis.

Our research utilizes a blog mining methodology to gather data concerning the utilization of ChatGPT in cybercrime. Blog mining is a systematic approach utilized for extracting information from publicly available blogs and online content [3]. It serves as a crucial means for obtaining threat intelligence, providing valuable insights into the dynamic trends, strategies, and methodologies employed by malicious actors. Through blog mining, we aim to illuminate the security and privacy issues associated with the use of ChatGPT.

Given the rapid pace of technological advancement within the ChatGPT domain, industry experts and cybersecurity analysts frequently disseminate their insights and concerns through blog posts. As such, these blog entries serve as a rich source of information for gaining a nuanced understanding of the multifaceted dimensions associated with ChatGPT.

However, it is important to admit the inherent limitations of blog mining. These include the absence of rigorous peer-reviewed scrutiny typical of scholarly journals and the susceptibility to personal biases and subjective interpretations. To address these limitations, our methodological approach integrates blog mining with an exhaustive review of academic literature. This comprehensive approach ensures a more robust and nuanced analysis of the subject matter under investigation.

3.1 Data Acquisition

To undertake an extensive data collection procedure, we utilized the Google search engine specifically for blog search, a platform engineered to aggregate publicly accessible content from various blogs on the internet. Our search query revolved around the keyword "security issues and privacy risks associated with using ChatGPT." This search was executed on December 6, 2023, yielding 31,000 search results within a mere 0.26 seconds. To refine the results and exclusively include blogs, we opted for the “News” category and sorted the search results by recency and relevance. This process retrieved approximately 59 blogs, predominantly from the year 2023.

3.2 Blog Screening and Content Analysis

Figure 1 depicts the screening process utilized during the blog mining endeavour. Each of the 59 identified blogs underwent meticulous manual examination to pinpoint content pertinent to the research objectives. Blogs were chosen based on their relevance to the research questions and the credibility of their sources.

Inclusion criteria were established to encompass blogs discussing the security issues and privacy risks associated with using ChatGPT. Ultimately, blogs meeting these criteria were included in the study, while others were excluded for various reasons, such as straying beyond the scope of the primary research topic, requiring subscription access, or falling into categories such as reports. Additionally, certain blogs only provided cursory mentions of ChatGPT without directly addressing the research questions.

After inclusion, the identified blogs underwent a detailed analysis, during which data relevant to the research questions were meticulously extracted. For each blog, details on different facets of ChatGPT were gathered, encompassing its benefits, limitations, security issues, privacy risks and other relevant information.

Following this extraction process, the findings were systematically categorized into relevant themes. This thematic classification facilitated a more organized presentation of the extracted data, allowing for coherent clusters of information to be formed. Such categorization aided in gaining clearer insights into the diverse roles and ramifications of ChatGPT within the realms of cybercrime and cybersecurity.

Figure 1: The blog screening process

4. Results and Discussion

This section unveils the results of the blog mining process, categorizing the findings into essential subheadings such as ChatGPT's advantages and drawbacks, the inherent security concerns associated with ChatGPT, the potential security and privacy vulnerabilities that could exploit these concerns, and the resultant impact of these security and privacy threats.

Figure 2 visually represents the outcomes of the blog mining screening procedure. Out of the 59 blogs initially identified, 34 were considered pertinent to this research and were consequently incorporated into the study, while 25 were excluded due to their inability to meet the designated inclusion criteria.

Through the blog mining process, the data relevant to the research were grouped into four major themes which are advantages and drawbacks of ChatGPT, security issues inherent in ChatGPT, security and privacy threats that can exploit ChatGPT security issues, and impact of security issues from ChatGPT usage. The four identified themes are discussed in the following subsections.

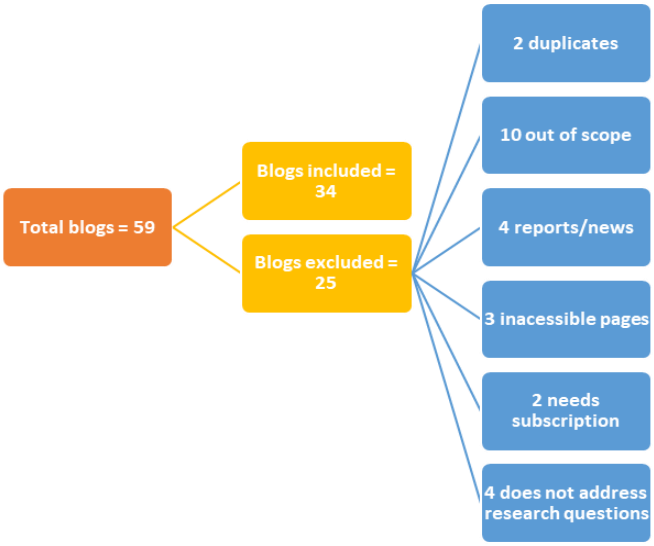

Figure 2: the result of the blog mining process

4.1 Advantages and Drawbacks of ChatGPT

ChatGPT stands out as the pioneering "natural language processing" chatbot driven by artificial intelligence, widely accessible to users. It possesses the capability to engage in human-like conversations and create diverse content formats, including emails, books, song lyrics, and application code. As is typical with novel technologies, generative AI models like ChatGPT present a dual-edged sword, offering both advantages and potential risks [7]. The benefits of using ChatGPT are as follows [8] [9]:

1) Enhanced Operational Efficiency

ChatGPT, being an AI-driven chatbot, excels in handling routine and repetitive tasks with precision and speed. By automating these tasks, ChatGPT allows human employees to allocate their time and energy towards more complex and strategic responsibilities within the organization. This not only increases overall productivity but also ensures that human expertise is utilized where it's most valuable.

2) Cost Savings

One of the major advantages of employing ChatGPT is its cost-effectiveness. Compared to hiring and training additional

staff to handle various tasks, implementing AI chatbots like ChatGPT typically proves to be a more economical solution for organizations. This is particularly beneficial for businesses operating on tight budgets or looking to optimize their resource allocation.

3) Improved Content Quality
ChatGPT serves as a valuable tool for writers and content creators by assisting in refining grammatical errors, enhancing contextual accuracy, and generating creative ideas. By leveraging ChatGPT's language processing capabilities, writers can refine their content, ensuring higher quality and greater engagement from their audience.

4) Educational Support and Training
ChatGPT's ability to provide clear explanations on complex topics positions it as an effective virtual tutor or educational assistant. Users can seek guidance, clarification, and supplementary learning materials from ChatGPT, thereby enhancing their understanding and proficiency in various subject matters. Additionally, ChatGPT can provide instructional guides and interactive learning experiences tailored to individual user needs, further augmenting its educational utility.

5) Expedited Response Time
ChatGPT's real-time responsiveness significantly reduces wait times for users seeking assistance or information. This quick turnaround ensures that users receive prompt and efficient support, thereby enhancing user satisfaction and overall experience with the service.

6) Continuous Availability
Unlike human customer service representatives who operate within specific working hours, ChatGPT is available round-the-clock, providing uninterrupted assistance and support to users across different time zones. This ensures consistent service availability and accessibility, regardless of the time of day or night.

7) Multilingual Communication
ChatGPT's proficiency in multiple languages and its ability to provide accurate translations cater to the linguistic diversity of global audiences. This ensures effective communication and engagement with users from diverse cultural backgrounds, thereby fostering inclusivity and accessibility in communication channels.

8) Personalized Interaction
Through its adaptive algorithms, ChatGPT’s responses are tailored towards users' preferences, behaviour patterns, and previous interactions. This personalized approach not only enhances user engagement but also creates a more immersive and tailored experience for each user, thereby strengthening user satisfaction and loyalty.

9) Scalability
ChatGPT's scalability enables it to handle a large volume of user interactions simultaneously, making it suitable for applications with high user engagement and dynamic scalability requirements. This ensures that the system remains responsive and efficient, even during periods of peak demand.

10) Natural Language Understanding
ChatGPT's advanced natural language processing capabilities enable it to comprehend and generate human-like text, making it adept at tasks such as content generation, answering questions, engaging in conversations, and providing explanations. This enhances its versatility and utility across various applications and use cases.

11) Digital Accessibility
ChatGPT's text-based interface makes it accessible to individuals with disabilities, offering an intuitive communication medium that transcends barriers posed by conventional interfaces. This promotes inclusivity and equal access to information and services in digital environments, thereby fostering a more inclusive and equitable society.

Despite the numerous benefits ChatGPT offer to both individuals and organizations, ChatGPT is not without drawbacks. The limitations of using ChatGPT are outlined as follows [8]:

1) Limitation in Language Comprehension
ChatGPT's understanding of human language is constrained by its training to generate responses based solely on input. Consequently, its replies may lack depth and fail to offer genuine insights into queries.

2) Temporal Knowledge Constraints
Due to its training data ending in 2021, ChatGPT may provide inaccurate or outdated information, particularly concerning events or data beyond this timeframe. Additionally, its responses may be flawed when it fails to grasp the nuances of a query.

3) The naturalness of Responses
ChatGPT's predictive text generation approach may result in responses that sound mechanical and devoid of natural language flow. Its tendency to overuse certain words like "the" or "and" necessitates human review and editing to enhance coherence and readability.

4) Lack of Citations and Analysis
ChatGPT's responses lack proper sourcing and critical analysis of data or statistics. While it may present statistical information, it fails to offer contextual commentary or insights into their significance within the given topic.

5) Inability to Detect Sarcasm and Irony
ChatGPT's reliance on a dataset of text renders it incapable of discerning nuances such as sarcasm or irony, leading to potential misinterpretations of user input.

6) Rigidity in Response Structure
ChatGPT's inability to shift focus or address multiple facets of a question within a single response hampers its adaptability and versatility. For instance, it may fixate on a specific aspect of a query, overlooking other relevant dimensions or inquiries.

4.2 Security Issues Inherent in ChatGPT

In the realm of AI-driven chatbots, ChatGPT has emerged as a prominent player in the tech landscape [10]. However, alongside its widespread adoption come concerns regarding data privacy and security. A fundamental question arises: are there security and privacy issues associated with the use of ChatGPT?

Developed by OpenAI, ChatGPT operates by collecting specific information from its users, which can be broadly categorized into three main types: account information, browser data, and chatbot interactions. Account information encompasses the details provided by users during the account creation process, while browser data includes information such as IP address and location. Additionally, ChatGPT stores the interactions users have with the chatbot itself [10].

These privacy concerns have prompted actions from regulatory bodies, with Italy's data protection agency, Il Garante per la protezione dei dati personali, initially blocking access to ChatGPT within the country [11] due to GDPR-related apprehensions. Subsequently, OpenAI implemented changes aimed at addressing these privacy concerns, leading to the reinstatement of ChatGPT's availability in Italy. However, the swift approval raised questions regarding the thorough examination of core GDPR issues, prompting a new investigation by Polish authorities [12].

Table 1 provides the inherent security challenges embedded within the infrastructure of ChatGPT, which cybercriminal adversaries can potentially exploit, thus precipitating various forms of cyber-attacks. These identified security issues, by their very nature, represent vulnerabilities within the system, thereby necessitating critical attention and remedial action to fortify the platform against potential breaches.

Table 1: security issues inherent in ChatGPT

Issues	References
Plagiarism issues	[13] [8] [14]
Copyright issues	[7] [13] [15]
Biased information issues	[10] [16] [17] [18] [8]
Inaccurate output issues	[17] [19] [18] [20] [15]
Misinformation issues	[21]
Model performance issues	[22]
Storage of User's data issues	[23] [24] [22] [25] [26] [27] [28]
Third-party service providers issues	[27] [29]
Using of input to refine output Issues	[7] [17] [20]
ChatGPT Plugin issues	[30]

1) Plagiarism issues

Utilizing AI tools for text generation can potentially expose businesses to the risk of plagiarism, which stands as a significant concern within the realm of content creation. Although AI technology facilitates rapid content generation, it cannot inherently produce unique text. Consequently, if businesses employ AI-generated content without manual intervention or editing, they face the peril of inadvertently producing duplicate content that may be flagged for plagiarism [13].

The implications of plagiarism extend beyond the corporate sphere and reverberate within educational institutions, particularly concerning students' assignments, projects, and academic submissions. Central to the pedagogical ethos is the cultivation of students' independent learning and the manifestation of their comprehension through original interpretations and insights. However, the advent of AI tools like Chat GPT introduces a novel challenge whereby students may be tempted to utilize AI-generated content in their work, thereby circumventing the essence of academic integrity [14]. This not only raises concerns regarding plagiarism but also engenders ethical dilemmas surrounding the authenticity and credibility of academic outputs.

2) Copyright issues

Utilizing ChatGPT for generating written content poses significant copyright concerns, particularly when the AI draws inspiration from copyrighted material, including licensed open-source resources. For instance, if ChatGPT is trained on open-source code libraries and subsequently reproduces similar code in response to user queries, companies incorporating such code into their products risk violating unfriendly Open Source Software (OSS) licenses [7].

Moreover, AI-powered image, audio, and video generators may inadvertently infringe upon copyright laws by sourcing data from internet content without proper authorization. This could potentially expose users to legal liabilities stemming from the unauthorized use of copyrighted material [13].

Additionally, the data provided to ChatGPT may contain copyrighted material, trade secrets, or confidential information. As ChatGPT generates responses without the explicit consent of data owners, end users must assess whether they possess the requisite rights to utilize or publish such material. Furthermore, compliance with geographical laws and regulatory requirements is paramount when leveraging data sourced from the AI bot [20].

There's also the risk that ChatGPT and the content it produces may be deemed derivative works of copyrighted materials used during its training. This could result in allegations of infringement, particularly if the generated content bears substantial similarity to the copyrighted training data.

Moreover, unresolved legal questions surround the entitlement of legal protection for documents or code generated with or without ChatGPT's assistance. While the United States Copyright Office currently doesn't extend copyright protection to non-human authored works, content produced by ChatGPT from protected publications may be deemed infringing.

Employers using content generated by AI tools may expose themselves to legal risks, as the output may contain portions of protected data, such as trademarked or copyrighted material. Identifying and mitigating potential infringement risks becomes challenging, especially considering that similar content may be generated for multiple users.

Lastly, violating the terms of service by using ChatGPT in the development of other AI systems could jeopardize future AI development endeavours, potentially impacting the company's standing in the AI space.

3) Biased information issues

ChatGPT's training heavily relies on user input and was initially trained using internet data available until 2021. The quality and characteristics of this training data significantly shape the generated content. Given the prevalence of fake and biased information on the internet, AI tools such as ChatGPT are prone to replicating biases inherent in the training data, including instances of discrimination [16].

As ChatGPT is trained on a vast corpus of human-authored content spanning various regions and periods, it inevitably inherits the biases prevalent in human communication. Consequently, the model may inadvertently perpetuate biases present in the real world, including those related to race, gender, and other characteristics.

The data accumulated by ChatGPT through its interactions may unintentionally capture biases present in these conversations. This data could potentially be utilized to develop AI systems that perpetuate discriminatory practices, targeting individuals or groups based on their race, gender, or other attributes [18]. Such practices are ethically unacceptable and can inflict harm on affected individuals or communities [10].

Instances have been observed where ChatGPT has generated responses containing discriminatory content. Efforts are underway by the company to address and mitigate these occurrences. However, the inherent biases ingrained in the model's training data pose ongoing challenges in achieving unbiased and equitable outputs.

4) Inaccurate output issues

One of the significant apprehensions surrounding AI bots pertains to their potential to provide inaccurate information. Instances may arise where the content generated by AI systems diverges from the original context or perspective, resulting in outputs that are inaccurate, incomplete, or biased. This reliance on AI-generated output as factual can inadvertently lead end-users to utilize incorrect information [20].

Despite its impressive capabilities, ChatGPT is not immune to producing inaccurate results. In certain scenarios, such as drafting sections of a legal brief, ChatGPT may cite irrelevant or non-existent cases, or struggle with basic computational tasks, yielding incorrect outcomes. OpenAI acknowledges these limitations and often issues warnings about the possibility of generating inaccurate information [18]. Additionally, ChatGPT may exhibit gaps in knowledge regarding world events post-2021, further contributing to potential inaccuracies.

Furthermore, accuracy concerns extend to other AI tools as well. OpenAI's terms of use acknowledge situations where outputs may not accurately reflect real people, places, or facts. Moreover, there is a possibility that AI tools could generate highly plausible falsehoods, as their performance is contingent upon the quality of the training dataset.

The extent of these risks depends on the use case. While the risks may be lower in scenarios where reviewers can readily identify and correct errors in ChatGPT's outputs, they escalate when reviewers are unable to discern inaccuracies or when no review process is in place. For instance, the risk associated with using ChatGPT to summarize news stories internally differs from relying on it to generate critical code for essential operations within a company's information systems.

However, limited visibility into the datasets used for training, particularly in the case of newer iterations like GPT-4, exacerbates these concerns surrounding accuracy and reliability.

5) Model performance issues

Although ChatGPT offers a free platform for creating code to address cybersecurity challenges such as phishing detection, spam filtering, and malware analysis, the efficacy of the solutions produced hinges on the quality of the training data and the architecture formulated by the AI Chatbot [22].

It's crucial to recognize that the effectiveness of the generated code may fall short in combatting various types of malwares or detecting network intrusions. As a result, relying solely on ChatGPT-generated solutions without adequate validation and testing can expose your system to vulnerabilities, particularly if there isn't a contingency plan or backup solution in place during the development phase [22].

In essence, while ChatGPT presents a valuable resource for cybersecurity development, it's imperative to supplement its outputs with rigorous testing, and validation, and possibly incorporate off-the-shelf solutions to fortify your system's resilience against cyber threats.

6) Misinformation Issues

Similar to other AI language models, ChatGPT relies on the data it was trained on. Despite its impressive performance, there exists a risk of misinformation stemming from two primary factors [21]. Firstly, it may have been trained on biased or erroneous data, potentially perpetuating inaccuracies. Secondly, as artificial intelligence, ChatGPT lacks real-world experience and consciousness, thus it cannot validate the truthfulness of its generated content beyond the scope of its training data. In an era where misinformation, false information, and fake news propagate rapidly across online platforms, the potential implications of ChatGPT producing inaccurate content are concerning [21].

7) Storage of User's data issues

ChatGPT processes user input, feedback, and files to generate content and maintains chat history [28] for 30 days [22]. ChatGPT maintains records of user conversations for accessibility across various devices via the OpenAI platform [24]. This extended retention period represents a significant

window of opportunity for potential security threats associated with ChatGPT to manifest. Privacy experts have criticized ChatGPT for its extensive collection of internet data, including personal and stolen data, without proper consent [23]. This data collection raises concerns about data leakage, as ChatGPT gathers a vast amount of user data, including IP addresses, browser details, site interactions, and browsing activities over time [26]. Users may not have the option to use masked email addresses or passwords for additional safety.

8) Third-party service providers issues

ChatGPT's privacy policy indicates that the platform collects significant users' personal data, including IP addresses, browser information, and browsing activities across various websites. This data may be shared with third-party entities, raising concerns about the potential dissemination of sensitive information to a wide audience, including competitors, customers, regulators, and board members. OpenAI reserves the right to disclose users' personal information to third parties without prior notice [27] [29].

9) Using input to refine output Issues

Due to the nature of AI systems to aggregate information, there is a substantial possibility that personally identifiable information (PII) could be utilized by ChatGPT to generate outputs for end users [7]. Despite warnings from ChatGPT regarding the input of sensitive or personal information, there remains a level of risk associated with the sharing of such data. This risk encompasses potential data breaches, inadvertent data storage, or the misuse of information [20].

Furthermore, there is concern that confidential information inputted into ChatGPT could be retained and used for future responses [17]. For example, an individual could upload sensitive corporate documents or datasets to ChatGPT, intending to obtain refined outputs. However, this confidential information may subsequently be accessible to other users with similar queries, including competitors seeking strategic insights.

10) ChatGPT Plugin issues

Johann Rehberger, a red team director at Electronic Arts and security researcher, has highlighted concerns with ChatGPT's plugins, documenting potential vulnerabilities in his spare time. These plugins could potentially facilitate the theft of chat history, acquisition of personal information, and remote execution of code on users' machines. Rehberger has engaged with plugin developers privately to address these issues and has reached out to OpenAI on multiple occasions regarding these concerns [30].

4.3 Security and Privacy Threats That Can Exploit ChatGPT Security Issues

In this section, we delve into the examination of Table 2, which meticulously outlines the myriad security and privacy threats poised to exploit the previously delineated security vulnerabilities inherent in ChatGPT. Through a comprehensive analysis, we elucidate the multifaceted nature of these threats, shedding light on their potential implications and ramifications.

Table 2: security and privacy threats

Security and privacy threats	References
Data Breach	[13] [10] [21]
Misuse of data	[10] [21]
Data theft and fraud	[22]
Malicious use	[22] [17] [26] [31]
Privacy invasion	[10] [8] [15]
Potential for Censorship	[10]
Data leakage	[32] [7] [22] [33] [26]
Sensitive data exposure	[34] [35] [36] [27]
Model poisoning attacks	[13]
Fake customer support scam	[22]
ChatGPT download	[24]
Malicious ChatGPT apps	[37]
Stolen ChatGPT user's credentials	[38] [28]

1) Data Breach

The storage of data by ChatGPT introduces a notable vulnerability to potential data breaches. Unauthorized access by hackers or malicious entities could compromise the stored data, thereby endangering sensitive user information. Such breaches may entail the exposure of personal conversations, confidential particulars, or any other data exchanged with the AI [10].

2) Misuse of data

There exists a risk that OpenAI or third-party service providers may engage in the misuse of collected data. This potential misuse could extend beyond the intended purpose of enhancing the AI model and encompass activities such as targeted advertising or training other AI models. Of particular concern is the possibility that users may not have provided explicit consent for such secondary uses, thereby raising ethical and privacy considerations regarding data handling practices [10].

3) Data theft and fraud

ChatGPT relies on an open-source large learning model that allows users to make modifications as needed. While this flexibility is crucial for AI training and development, it also introduces vulnerability to potential data theft. Hackers may exploit this openness to gain unauthorized access to users' chat history, thereby exposing the platform to various forms of fraud [22].

Cybercriminals can leverage any gleaned information to target users, including but not limited to email addresses, physical addresses, and code snippets. This susceptibility to data exploitation underscores the importance of robust security measures to safeguard against unauthorized access and mitigate the risk of malicious activities.

4) Malicious use

Utilizing the AI Chatbot requires users to sign up with their name and email address, creating a potential vulnerability if this information falls into the hands of hackers. Access to a database containing millions of ChatGPT users could enable hackers to execute targeted social engineering attacks,

exploiting personal information obtained through the signup process.

Moreover, the advent of AI generation introduces novel extensions to the overall attack surface, presenting new avenues for exploitation by malicious actors. Through generative AI, attackers can develop sophisticated forms of malware, phishing schemes, and other cyber threats that evade traditional defence mechanisms [26]. Such attacks pose significant consequences, including data breaches, financial losses, and reputational damage [22].

Indeed, ChatGPT's capabilities could be exploited by attackers to deceive and target users and their devices. Malicious actors may leverage the AI platform to craft convincing messages or responses aimed at tricking individuals into divulging sensitive information or executing harmful actions on their computers. This underscores the importance of exercising caution and vigilance when interacting with AI-driven chatbots like ChatGPT, as well as implementing robust cybersecurity measures to mitigate the risk of falling victim to such attacks [22].

5) Privacy Invasion

The storage of user data by ChatGPT raises concerns regarding the invasion of privacy. Users may reasonably expect their interactions with ChatGPT to remain confidential, and the retention of data has the potential to undermine this trust. Furthermore, the storage of user data can facilitate the tracking of online behaviour and activities, leading to heightened privacy apprehensions among users [10].

Despite any improvements made to OpenAI's privacy policies in response to data breaches, there may still be challenges in fully complying with the General Data Protection Regulation (GDPR), which mandates stringent data protection measures in Europe. It is plausible that OpenAI collected personal information during the initial training of ChatGPT. While data laws in the United States may offer less definitive protection, European regulations safeguard personal data irrespective of whether it is publicly disclosed or kept privately [15].

6) Potential for Censorship

There exists a concern that the data collected by ChatGPT could be exploited for censorship objectives, potentially involving the identification and suppression of content that criticizes OpenAI or its affiliates. Such actions could impede free expression and curtail the diversity of perspectives in online discourse, thereby raising pertinent issues related to content moderation and freedom of speech [10].

To mitigate these risks, it is imperative to implement meticulous data handling practices, bolster security measures, and establish transparent policies aimed at safeguarding user privacy and promoting the responsible utilization of data. These measures are essential for upholding the principles of free expression and ensuring the preservation of diverse viewpoints within online conversations [10].

7) Data leakage

The extraction of sensitive information by ChatGPT highlights the broader issue of user privacy in the era of AI. As AI models handle increasingly personal data, the industry must prioritize robust privacy measures. Developers should implement encryption and data anonymization techniques to mitigate the risks associated with potential breaches [32].

Data leakage is a significant risk associated with ChatGPT, as any sensitive third-party or internal company information entered into the chatbot becomes part of its data model and may be shared with others who ask relevant questions [7]. Unauthorized disclosure of confidential information into ChatGPT may violate an organization's security policies and put users at risk of exposing personal details, business secrets, website code, or other sensitive information.

While ChatGPT saves user chats for their benefit, users must be cautious about the information they share, ensuring compliance with laws and the platform's terms of use. Personal details and sensitive data should not be included in chats to mitigate privacy risks and safeguard against potential data breaches.

8) Sensitive data exposure

One notable security risk associated with ChatGPT is the potential exposure of sensitive data [39]. If users utilize the publicly available version of ChatGPT within a professional setting, there is a risk of inadvertently inputting confidential information related to their organization or business. This version of ChatGPT utilizes the data provided by users to enhance its learning and responsiveness to future queries [22].

While ChatGPT emphasizes warnings against providing sensitive information, there remains uncertainty regarding the tool's compliance with international privacy laws and the effectiveness of controls in place to protect personal data [7]. The re-use of personal data by generative AI systems for additional purposes poses risks of misuse and reputational harm, potentially undermining trust and breaching privacy commitments to employees, customers, and partners.

It is crucial for users to carefully consider the confidentiality risks associated with interacting with ChatGPT. While AI may enhance output quality and streamline processes, the inadvertent disclosure of sensitive information and the potential for misuse underscore the importance of understanding the terms of use and privacy policies governing AI interactions.

9) Model poisoning attacks

AI tools like ChatGPT and similar chatbots are particularly susceptible to model poisoning, attacks, and data breaches due to their open-source nature. Model poisoning refers to the manipulation of an AI model's training data to introduce malicious inputs or biases, ultimately compromising the model's performance and security.

The open-source nature of these AI tools exposes them to various security risks, including unauthorized access to training data, injection of harmful data or commands, and exploitation of vulnerabilities in the model architecture. Hackers and malicious actors can exploit these weaknesses to manipulate the behaviour of the AI system, leading to undesirable outcomes such as generating misleading information, leaking sensitive data, or facilitating cyberattacks.

Additionally, the collaborative and decentralized nature of open-source development can further exacerbate the vulnerability to model poisoning. With multiple contributors and frequent updates to the AI model, monitoring and detecting malicious activities become challenging, increasing the likelihood of successful attacks.

To mitigate the risk of model poisoning, developers of AI tools must implement robust security measures, including rigorous data validation processes, encryption of sensitive information, access controls to restrict unauthorized access, and regular security audits to identify and patch vulnerabilities. Furthermore, fostering a culture of cybersecurity awareness and collaboration within the open-source community can help enhance the overall resilience of AI models against malicious attacks and data breaches.

10) Fake customer support scam

For individuals who have recently registered for ChatGPT but have yet to initiate its utilization, there often arises a plethora of inquiries regarding its operational mechanics. In such instances, it is common for users to resort to popular platforms such as Slack, Discord, Quora, or Facebook, seeking guidance from proficient users within these communities [22].

However, this recourse presents inherent vulnerabilities, as it exposes individuals to potential cyber threats. Malicious actors, assuming the guise of seasoned experts or representatives affiliated with spurious ChatGPT-related entities, may exploit this scenario to coax users into divulging sensitive information. This can manifest in the form of duplicitous requests for login credentials or personal data, executed under the guise of assisting [22]. Subsequently, such malevolent actions facilitate the perpetration of diverse cyber offences, thus underscoring the critical importance of exercising caution and discernment when engaging in online interactions related to AI technologies.

11) ChatGPT Download

Currently, it is advisable to exercise caution if encountering offers to download ChatGPT from sources other than OpenAI's official website or authorized channels. OpenAI does not provide the option to download ChatGPT as a standalone application for Android or iPhone devices. Instead, the service is readily accessible through desktop or mobile browsers via OpenAI's official website [24].

Instances of unauthorized downloads or purported mobile applications may pose risks to users, as they could potentially be conduits for scams or deceptive practices. Given the popularity of such services, they often attract the attention of individuals seeking to exploit unsuspecting users. Therefore, users should remain vigilant and refrain from engaging with unauthorized sources to mitigate the likelihood of falling victim to fraudulent activities [24].

12) Malicious ChatGPT apps

Investigations have unveiled a surge in counterfeit ChatGPT applications. It is imperative to underscore that there is currently no official ChatGPT application available for download; the service is exclusively accessible via the official webpage.

We have identified several dubious ChatGPT applications, with some being malicious Trojan viruses and others categorized as Potentially Unwanted Applications (PUAs). While PUAs do not meet the criteria for malware, they still present a lower yet notable risk to security and privacy, often resulting in sluggish device performance or the intrusive display of unexpected advertisements [37].

Trojan applications possess the capability to implant malicious software onto devices, potentially leading to data encryption or compromise. These insidious malware types pose a significant threat to users, their data, and their devices. Notably, the ChatGPT trojan applications we have encountered are not novel but have been rebranded under the guise of new identities. Previously, they masqueraded as purported "premium" applications for platforms like YouTube and Netflix, which were also fraudulent [37].

The "ChatGPT" trojan app clandestinely enrols its victims in various premium services through SMS billing fraud. Conversely, the "AI photo" trojan app is affiliated with the Spynote malware, notorious for its ability to pilfer files, SMS messages, call logs, and contact lists from targeted devices [37].

Furthermore, researchers have identified numerous dubious ChatGPT websites and pages sporting counterfeit URLs such as "chat-gpt-pc[.]online." These malicious platforms falsely offer a download of ChatGPT, despite the service being solely browser-based. Their underlying motive is to infect unsuspecting users' computers with malware, an outcome users should vigilantly guard against [37].

13) Stolen user's credentials

Check Point Research (CPR) has highlighted growing concerns regarding ChatGPT's impact on cybersecurity, particularly noting an uptick in the trade of stolen ChatGPT Premium accounts. These stolen accounts grant cybercriminals unrestricted access to ChatGPT, bypassing OpenAI's geofencing restrictions. This surge in stolen accounts reflects a broader trend in the underground hacking community, where the market for account takeovers (ATOs) thrives, traditionally focusing on various online services. CPR has observed increased activity surrounding stolen ChatGPT accounts since March 2023, with cyber criminals engaging in activities such as leaking credentials, trading stolen premium

accounts, and utilizing brute force and checker tools to gain unauthorized access [38]. Additionally, there is a rise in dedicated services offering ChatGPT premium accounts, potentially leveraging stolen payment cards for acquisition.

The escalation in the trade of stolen ChatGPT accounts can be attributed to several factors. Geofencing restrictions imposed by ChatGPT have prompted cybercriminals to seek workarounds, including exploiting the ChatGPT API and premium accounts. Consequently, there is a growing demand for stolen ChatGPT accounts, particularly premium ones, within the underground market. Meanwhile, ongoing discussions surrounding ChatGPT's privacy concerns, evidenced by recent bans in Italy and potential considerations in Germany, further underscore the gravity of the situation. Of particular concern is the privacy risk posed by ChatGPT's storage of user queries, potentially exposing personal and corporate information to unauthorized access [38].

The trade of stolen ChatGPT accounts exploits users' tendency to reuse passwords across multiple platforms. Malicious actors leverage this behaviour by employing account checker software to match stolen email and password combinations, facilitating unauthorized access to online platforms. Such account takeovers constitute a severe breach of security, allowing cybercriminals to assume control of accounts without the account holder's consent [28].

4.4 Impact of Security Issues from ChatGPT Usage

The section outlines the impact of security and privacy issues associated with the use of ChatGPT on both individuals and organizations.

1) Subject to liabilities

The adoption of AI tools by businesses may expose them to additional liabilities, necessitating careful examination of the terms of use associated with such tools. For instance, the utilization of ChatGPT entails potential obligations for users to defend and indemnify OpenAI against any claims, losses, or expenses, including legal fees, arising from or connected to the use of the tool. Moreover, users of ChatGPT are typically bound by agreements mandating arbitration and waiving class action provisions [15]. It's important to note that these terms of use can be modified unilaterally by the creator of the AI tool, presenting an additional layer of uncertainty for users. Therefore, businesses must conduct thorough reviews of the terms of use and understand the implications before integrating AI tools into their operations.

2) Reputation damage

In the event of a security breach compromising ChatGPT's security measures, confidential content including input and output generated by the chatbot, which the organization may have been contractually or legally obligated to safeguard, could be exposed. This breach could be attributed to the organization, potentially leading to reputational damage [7]. The regulatory agency's attention to such occurrences follows several notable instances where the chatbot provided inaccurate information, posing reputational risks to individuals. Customers who receive work products created with the aid of ChatGPT may allege that the failure to disclose this fact constitutes false and deceptive practices [20]. To mitigate such claims, organizations should consider implementing affirmative disclosure and consent mechanisms or obtaining appropriate waivers from customers. This proactive approach can help ensure transparency and mitigate legal risks associated with the use of AI technologies like ChatGPT.

3) Legal implications

ChatGPT, as a third-party system, assimilates data into its repository. Even in the absence of a security breach, sharing confidential customer or partner information with ChatGPT may constitute a violation of agreements with these stakeholders. This is because organizations are often bound by contractual or legal obligations to safeguard such information. Ownership of the code generated by ChatGPT poses another complexity. According to the terms of service, the output produced by ChatGPT belongs to the individual or service that supplied the input [7]. However, complications arise when this output contains legally protected data obtained from input provided by other sources. Resolving ownership disputes in such cases can be intricate and may require legal intervention [7].

4) AI holding human's job

Indeed, the emergence of AI such as ChatGPT passing significant milestones like the bar exam marks just the initial stages of AI encroaching on human job domains [8]. Occupations particularly susceptible to disruption by AI include graphic design, writing, and accounting. The successful completion of the bar exam by a later version of ChatGPT underscores the potential for AI to significantly alter the composition of the workforce in the foreseeable future [18].

5) Impact on the education system

as ChatGPT and similar AI models gain traction in assisting with writing tasks, the necessity for students to acquire traditional writing skills comes into question. This raises existential concerns regarding the future of writing education. Given that students are increasingly turning to ChatGPT for essay assistance, educators are confronted with the pressing need to address this reality [18].

In response, educational institutions are swiftly reassessing their policies to determine whether students should be permitted to leverage AI tools for academic assignments. The impact extends beyond English-based subjects; ChatGPT's capabilities encompass aiding in various tasks such as brainstorming, summarizing, and deriving analytical insights. As such, the educational landscape faces a fundamental shift in how writing skills are cultivated and assessed.

6) Impact on human behaviour and social norm

ChatGPT and similar AI technologies have made our lives easier by automating tasks and providing instant information. However, there is a risk of over-reliance on such technologies. This dependency could lead to a loss of critical thinking skills, decreased human interaction, and a potential

loss. The proliferation of ChatGPT and similar AI technologies has undoubtedly streamlined various aspects of our lives, offering automation and instantaneous access to information. However, this convenience comes with its own set of risks, chief among them being the peril of excessive reliance on such technologies. This dependency has the potential to erode critical thinking skills, diminish human interaction, and compromise privacy [21].

The ramifications of this risk on users are multifaceted. Firstly, an overreliance on AI like ChatGPT may precipitate a substantial shift in human behaviour and societal norms. This could manifest in stunted personal growth, diminished social interactions, and a diminished capacity to navigate daily tasks without the aid of these technologies [21].

7) Impact on decision making

The potential for biased information generated through ChatGPT to influence people's decisions negatively is a significant concern. When individuals base their decisions on biased or inaccurate information, it can lead to adverse outcomes and perpetuate harmful stereotypes or prejudices. Moreover, entrusting critical decisions to AI systems, including ChatGPT, poses a notable hazard [21]. In the event of an error or technical malfunction, the consequences can be severe and widespread. Therefore, while AI undoubtedly offers various benefits, users must remain aware of its limitations and exercise caution in their reliance upon it [21]. Additionally, an overreliance on AI for critical decisions may result in unforeseen consequences if the AI encounters errors or experiences technical failures. Thus, a balanced approach that acknowledges both the advantages and risks of AI technology is essential for informed decision-making.

8) It can mislead developers to malicious package

Attackers have identified a concerning vulnerability in ChatGPT, whereby the AI's tendency to provide false information can be exploited to disseminate malicious code packages. This poses a significant threat to the software supply chain, as it enables the infiltration of malicious code and Trojans into legitimate applications and code repositories like npm, PyPI, GitHub, and others [40].

Through the utilization of "AI package hallucinations," threat actors can craft ChatGPT-recommended code packages that appear legitimate but contain malicious content. Developers may unknowingly download these packages while using the chatbot, subsequently integrating them into widely-used software [40].

Furthermore, developers themselves are susceptible to the allure of ChatGPT, often turning to the AI platform for coding solutions instead of traditional sources like Stack Overflow. This trend creates a prime opportunity for attackers to exploit [40].

As has been observed throughout history, any technology that gains rapid popularity inevitably attracts malicious actors seeking to exploit it for their gain. ChatGPT serves as a real-time example of this phenomenon, highlighting the urgent need for robust security measures to safeguard against such threats.

Young and emerging developers may encounter challenges in distinguishing between malicious and erroneous code packages.

9) Financial loss

The mishandling of personal information acquired via ChatGPT can result in various repercussions, ranging from minor inconveniences to significant financial losses. Unauthorized access or mismanagement of this data may lead to financial harm, highlighting the imperative need to prioritize the protection of sensitive information. Implementing stringent security measures is paramount to mitigate the risk of these detrimental consequences [21].

10) Identity theft

The potential for identity theft stemming from ChatGPT usage is a serious concern, as it could result in innocent individuals becoming victims of malicious activities [21]. If user identities are compromised, perpetrators may exploit this information to perpetrate various harmful actions, posing significant risks to individuals' safety and well-being. It underscores the importance of safeguarding personal information and adopting stringent security measures to mitigate the threat of identity theft and its associated consequences.

5. Conclusion and Future Scope

In conclusion, this research has shed light on the multifaceted landscape of security and privacy concerns surrounding ChatGPT, an AI-driven chatbot. Through meticulous examination and analysis, we have identified inherent security vulnerabilities within the platform, as well as the potential threats that could exploit these weaknesses. From the mishandling of personal information to the risk of cyber-attacks, the implications of these security issues are far-reaching and could have significant ramifications for individuals and organizations alike. As such, it is imperative for stakeholders to prioritize the implementation of robust security measures and regulatory frameworks to mitigate these risks effectively.

Suggestions for Future Work: Moving forward, there are several avenues for future research to explore in greater depth. Firstly, further investigation is warranted into the specific techniques and methodologies employed by cybercriminals to exploit vulnerabilities within ChatGPT. Understanding these tactics can inform the development of more targeted and effective security defences. Additionally, longitudinal studies could be conducted to track the evolution of security threats and privacy concerns in AI-driven chatbots over time, allowing for the identification of emerging trends and patterns. Furthermore, research efforts should be directed towards developing comprehensive guidelines and best practices for organizations to enhance the security and

privacy of AI chatbot systems. Finally, collaboration between academia, industry, and regulatory bodies is essential to foster a holistic approach to addressing these challenges and ensuring the responsible deployment of AI technologies in the future.

Data Availability
All data used are publicly available
Conflict of Interest
None

Funding Source
None

Authors' Contributions
The entirety of the work was undertaken solely by the author.

References

[1] S. Addington, "ChatGPT: Cyber Security Threats and Countermeasures," pp.**1–12, 2023.**
[2] M. Alawida, S. Mejri, A. Mehmood, B. Chikhaoui, and O. I. Abiodun, "A Comprehensive Study of ChatGPT : Advancements, Limitations, and Ethical Considerations in Natural Language Processing and Cybersecurity," *Inf. J.*, **2023.**
[3] P. V. Falade, "Decoding the Threat Landscape : ChatGPT, FraudGPT, and WormGPT in Social Engineering Attacks," *Int. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, October **2023**, doi: 10.32628/CSEIT2390533.
[4] A. Qammar, H. Wang, J. Ding, A. Naouri, M. Daneshmand, and H. Ning, "Chatbots to ChatGPT in a Cybersecurity Space: Evolution, Vulnerabilities, Attacks, Challenges, and Future Recommendations," *J. Latex Cl. Files*, Vol.**14**, No.**8**, pp.**1–17, 2023.**
[5] G. Sebastian, "Do ChatGPT and Other AI Chatbots Pose a Cybersecurity Risk ? An Exploratory Study," *Int. J. Secur. Priv. Pervasive Comput.*, Vol.**15**, No.**1**, pp.**1–11, 2023**, doi: 10.4018/IJSPPC.320225.
[6] M. M. Mijwil, M. Aljanabi, and A. H. Ali, "ChatGPT : Exploring the Role of Cybersecurity in the Protection of Medical Information Intro to ChatGPT," *Mesopotamian J. Cybersecurity*, vol. 2023, pp.**18–21, 2023.**
[7] J. McGinnis, "Is ChatGPT Safe for Organizations to Use?," **2023.**
[8] A. Hetler, "What is generative AI? Everything you need to know," **2023.**
[9] M. V. Jaworski and C. H. Patel, "Your Employees Are Using ChatGPT and Other LLMs: Risks and Legal Implications of ChatGPT in the Workplace," **2023.**
[10] P. Chanda, "Does ChatGPT Save Data? What You Need to Know (2023)," **2023.**
[11] C. Mauran, "ChatGPT rolls out important privacy options," **2023.**
[12] G. Moody, "Poland Opens GDPR Investigation into ChatGPT and OpenAI amid Mounting Privacy Concerns," **2023.**
[13] R. Neubauer, "Do Free AI Tools Pose a Security Risk to Your Business?" **2023.**
[14] M. Malik, "Technical and Legal Risks of ChatGPT: How prepared are we with Laws on AI?" **2023.**
[15] M. T. Renaud, M. McConihe, and N. Liu, "Nothing for free – the real costs of ChatGPT," **2023.**
[16] S. S. Mustufa, "ChatGPT and Data Privacy," **2023.**
[17] D. C and P. J, "ChatGPT and large language models: what's the risk?" **2023.**
[18] G. WU, "8 Big Problems With OpenAI's ChatGPT," **2023.**
[19] C. Zakrzewski, "FTC investigates OpenAI over data leak and ChatGPT's inaccuracy," **2023.**
[20] M. K. Nagothu, "Integrating ChatGPT & Generative AI Within Cybersecurity Best Practices," **2023.**
[21] K. Gilani, "Is ChatGPT Safe? 3 Hidden Risks and 5 Pro Tips (July 2023)," **2023.**
[22] Aparna, "ChatGPT: Analyzing the Security Risks and Ensuring User Safety," **2023.**
[23] Aljazeera, "OpenAI launches business version of ChatGPT after blowback over privacy," **2023.**
[24] F. L. Somoye, "Is ChatGPT safe? Security and privacy risks considered," **2023.**
[25] M. Dixit, "ChatGPT's incognito mode? Users can disable chat history to protect data," **2023.**
[26] T. Jackson, "Exploring The Security Risks Of Generative AI," **2023.**
[27] C. Castro, "ChatGPT: a privacy nightmare?" **2023.**
[28] A. Zacharakos, "ChatGPT users at risk for credential theft," **2023.**
[29] J. D. Neuburger, "ChatGPT Risks and the Need for Corporate Policies," **2023.**
[30] M. Burgess, "ChatGPT Has a Plugin Problem," **2023.**
[31] P. Wagenseil, "Security risks of ChatGPT and other AI text generators," **2023.**
[32] A. News, "How Googlers cracked OpenAI's ChatGPT with a single word," **2023.**
[33] M. Jackson, "Why ChatGPT is a security concern for your organization (even if you don't use it)," **2023.**
[34] G. Moody, "ChatGPT Is a Privacy Disaster Waiting To Happen, Find Out Why," **2023.**
[35] C. Rodrigues, "The Risks of Using ChatGPT to Write Client-Side Code," **2023.**
[36] C. Thorbecke, "Don't tell anything to a chatbot you want to keep private," **2023.**
[37] TREND, "4M Accounts Compromised in #LilyCollinsHack: Fake ChatGPT Apps & Websites Alert," **2023.**
[38] C. P. Team, "New ChatGPT4.0 Concerns: A Market for Stolen Premium Accounts," **2023.**
[39] P. V. Falade and P. O. Momoh, "Evaluating the Permissions of Monitoring Mobile Applications for Remote Employees : Analysing the Impact on Employer Trust and Employee Privacy Concerns," *Int. J. Sci. Res. Comput. Sci. Eng.*, February, Vol.**12**, pp.**42–52, 2024.**
[40] E. Montalbano, "ChatGPT Hallucinations Open Developers to Supply Chain Malware Attacks," **2023.**

AUTHOR PROFILE

Polra Victor Falade- holds a B.Tech in Computer Science with a specialization in Cyber Security, which she earned from the Federal University of Technology Minna, Niger State, Nigeria in 2016. Subsequently, pursued an MSc in Information Security from the University of Surrey, UK, graduating in 2021. These educational experiences have equipped her with a comprehensive understanding of cybersecurity principles and best practices. Currently, she is serving as an Assistant Lecturer in the Department of Cyber Security at the Nigerian Defence Academy (NDA) in Kaduna, Nigeria. In this role, she has been actively involved in educating future cybersecurity professionals, fostering a culture of cybersecurity awareness, and conducting research in the field. Her commitment to academic excellence is reflected in her continuous pursuit of knowledge and dedication to her students. Furthermore, she is a professional member of the Cyber Security Expert Association of Nigeria (CSEAN), which has provided her with valuable networking opportunities and a platform to stay updated with the latest developments in the cybersecurity domain. Also, a member of Internet Society, Nigeria Chapter. Her primary passion lies in research and academic writing, particularly in the areas of Information Security, AI Security, Privacy and cybersecurity-related research.

