Do ChatGPT and other AI chatbots pose a Cybersecurity risk? -An Exploratory study

Preprint · February 2023

DOI: 10.13140/RG.2.2.18037.04328

Glorin Sebastian
Georgia Institute of Technology

Do ChatGPT and Other AI Chatbots Pose a Cybersecurity Risk?

An Exploratory Study

Glorin Sebastian, Georgia Institute of Technology, USA*
https://orcid.org/0000-0003-2543-9127

ABSTRACT

The rise of artificial intelligence (AI) has opened up new frontiers in various fields, including natural language processing. One of the most significant advancements in this area is the development of conversational agents (i.e., chatbots), which are computer programs designed to interact with humans through messaging interfaces. The emergence of large language models, such as ChatGPT, has enabled the creation of highly sophisticated chatbots that can mimic human conversations with impressive accuracy. However, the use of these chatbots also poses significant cyber risks that must be addressed. This research paper seeks to investigate the cyber risks associated with the use of ChatGPT and other similar AI-based chatbots, including potential vulnerabilities that could be exploited by malicious actors. As part of this research, a survey was conducted to explore the cybersecurity risks associated with AI-based chatbots like ChatGPT. Further, the paper also suggests mitigation methods that can be used to mitigate these cyber risks and vulnerabilities.

KEYWORDS

Artificial Intelligence (AI), ChatGPT, Cybersecurity, Generative Pretrained Transformer 3 (GPT-3), Natural Language Processing (NLP)

INTRODUCTION

ChatGPT (Generative Pre-trained Transformer) is the Chat Bot introduced by Open AI in November 2022, an AI research and development company, based on a variation of its Instruct GPT model, which is trained on a massive pool of data to answer queries (Open AI. ChatGPT. 2022). ChatGPT uses natural language processing to generate responses to text-based inputs. GPT models are based on the Transformer architecture, which is a neural network architecture that was introduced in the research paper by Vaswani (Vaswani, A. et.al, 2017).

The architecture of ChatGPT is quite complex and involves many layers of neurons. At a high level, the model consists of an encoder and a decoder, that work together to generate responses to various user

DOI: 10.4018/IJSPPC.320225

*Corresponding Author

inputs. The encoder takes in the input text and processes it to create a sequence of hidden states, which are then passed to the decoder. The decoder uses these hidden states to generate the output text one token at a time, in a process known as autoregression. Some of the key features of ChatGPT include:

1. **Large Scale:** ChatGPT is one of the largest language models available, with over 175 billion parameters. This makes it easier for it to understand and generate complex responses.
2. **Conversational**: ChatGPT is designed to engage in natural and flowing conversations, making it appear more human-like in its responses.
3. **Multi-Task**: ChatGPT can perform multiple tasks, including answering questions, summarizing text, and generating creative writing.
4. **Contextual**: ChatGPT can take into account the context of the conversation to provide more relevant and accurate responses.
5. **Personalized**: ChatGPT can be trained on specific datasets to provide personalized responses for specific domains or use cases.
6. **Open Source**: ChatGPT is open source, meaning that developers can modify and customize the model to suit their specific needs.

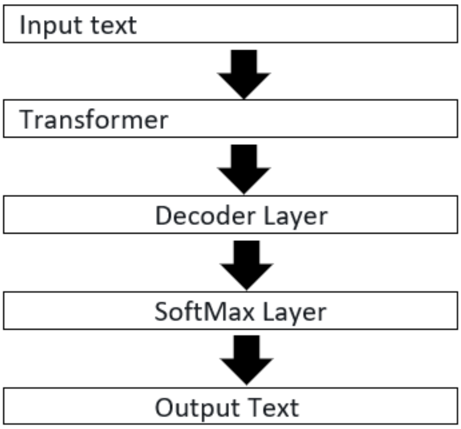
The below Figure-1 shows the details of the input text and the transformation of data across various layers before the final output is shared. The input text is fed into the Transformer, which processes it to create hidden states. The decoder layer then uses these hidden states to generate the output text, which is produced by the SoftMax layer.

Further, ChatGPT has a large number of parameters and requires a lot of computing power to train and use effectively. However, because it is pre-trained on a large corpus of text, it can be fine-tuned for specific applications with relatively little additional training data. Overall, ChatGPT represents a significant advancement in conversational AI, with the potential to revolutionize various industries, including customer service and healthcare.

CHATGPT ALGORITHM

ChatGPT is based on a variant of the Transformer architecture, which is a deep neural network model that is well-suited for processing sequential data, such as natural language. The Transformer architecture uses self-attention mechanisms to allow the model to focus on different parts of the input

Figure 1.
The process diagram for a ChatGPT-enabled query



sequence, which enables it to capture long-range relationships between different parts of the text. Future improvements for ChatGPT and similar AI Chatbots could include using larger models that can capture more complex relationships in the data, as well as using better training techniques for more efficient optimization algorithms, and advanced learning rate schedules.

Other future improvements could include better data pre-processing to remove noise and irrelevant information that also improves the quality of the model's predictions. Incorporating external knowledge sources, such as structured data, knowledge graphs, or other domain-specific information, helps the model better understand the context of the conversation. Incorporating multiple modalities, such as images or audio, into the input data would allow the model to better understand the context of the conversation and provide more accurate responses.

LITERATURE REVIEW

Below Table-1 summarizes the research papers published on ChatGPT and its applications at the time of writing this research. While research on the effectiveness of ChatGPT in applications such as translations (Lund, B.D. and Wang, T.,2023), healthcare (Jiao, W., Wang et. Al, 2023), and research (Kung, T. H., et. Al, 2023) has been conducted, there has been no study that tries to gauge the cyber risks (Hacker, P., Engel, A., & Mauer, M., 2023), associated with some of these AI-based Chatbots. This research addresses this gap, and the results of this research can help organizations and individuals

Table 1.
A literature review of articles and research on ChatGPT and similar AI Bots

S. No	Study	Summary
1.	Vaswani et. Al, 2017, Attention is all you need, Advances in neural information processing systems	proposes a new neural network architecture called the "Transformer" that only uses self-attention mechanisms, without recurrent or convolutional layers. The Transformer model achieves state-of-the-art results on several natural language processing tasks, outperforming traditional models that use sequential processing.
2.	Lund, B.D. and Wang, T., 2023, Chatting about ChatGPT: how AI and GPT may impact academia? Library Tech News (Lund, B.D. and Wang, T.,2023)	provides an overview of key definitions related to ChatGPT, a public tool developed by Open AI, and its underlying technology, Generative Pretrained Transformer (GPT).
3.	Jiao, W., Wang, W., Huang, J. T., Wang, X., & Tu, Z. (2023), Is ChatGPT a good translator? A preliminary study, arXiv	The authors conducted a preliminary study comparing ChatGPT translations with those generated by professional human translators and found that while ChatGPT translations were not always accurate, they were of acceptable quality and could be useful for some purposes.
4.	Kung, T et. Al, 2023, PLOS Digital Health, Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models	ChatGPT was quizzed on USMLE exam questions, and it demonstrated a high level of concordance and insight in its explanations. These results suggest that large language models may have the potential to assist with medical education, and potentially, clinical decision-making.
5.	ydın, Ö., & Karaarslan, E. (2022), SSRN, Open AI ChatGPT generated literature review: Digital twin in healthcare.	Abstracts of the last three years (2020, 2021, and 2022) papers were obtained from the keyword "Digital twin in healthcare" search results on Google Scholar and paraphrased by ChatGPT. ChatGPT was able to create a good literature review with the required details.
6.	van Dis, E. A.,et al. (2023). ChatGPT: five priorities for research. Nature, 614(7947), 224-226.	The paper lists 5 priorities of research. Author-contribution statements and acknowledgments in research papers should state clearly and specifically whether, and to what extent, the authors used AI technologies such as ChatGPT in the preparation of their manuscript and analysis.
7.	Gao, C. A., et. Al (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using AI output and plagiarism detector, human reviewers. bioRxiv, 2022-12.	All ChatGPT-generated abstracts were written clearly but only 8% correctly followed the specific journal's formatting requirements. blinded human reviewers correctly identified 68% of generated abstracts as being generated by ChatGPT, but incorrectly identified 14% of original abstracts as being generated. Most generated abstracts were detected using the AI output detector
8.	Hacker, P., Engel, A., & Mauer, M. (2023)	The legal part of the paper proceeds in four steps, covering (1) direct regulation, (2) data protection, (3) content moderation, and (4) policy proposals.

better understand the risks associated with using ChatGPT and other similar AI-based chatbots and take appropriate measures to safeguard against potential threats.

SURVEY RESULTS

The below Survey, summarized in Table-2, was conducted among Amazon Mechanical Turk (M-Turk) participants (Aguinis, H. et. al., 2021), with the survey conducted for a duration of 1 week in February 2023. The survey received 259 responses. Survey responses were collected from 5 continents but mostly from North America (62%) and across age groups and genders with males between 31-55 years being the majority (32.2%) of respondents.

Table 2.
Summary of the survey responses

Survey questions and options	**Responses**
1. What is your age range and gender?	
- Less than 30 years, Male	75 (29.1%)
- Less than 30 years, Female	44 (17.1%)
- 31-55 years, Male	83 (32.2%)
- 31-55 years, Female	43 (16.7%)
- Over 55, Male	9 (3.5%)
- Over 55, Female	4 (1.6%)
2. Which geography are you from?	
- North America	158 (62%)
- Asia	60 (23.5%)
- Africa	1 (0.4%)
- South America	32 (12.5%)
- Europe	4 (1.6%)
3. **Have you used ChatGPT? How was your experience using it? (Multiple Choice)**	
- Excited about improvements in AI/ML technology	165 (64%)
- Scared that AI/ML might replace people's jobs	98 (38%)
- Students and workers might use such AI/ML bots to cheat on exams/work	76 (29.5%)
- ChatGPT would improve human efficiency like computers did since the 1980s and 1990s	97 (37.6%)
- No specific emotion, it's just a passing fad	24 (9.3%)
- it's neat and will be helpful, but there will still be a need for human-generated content	1 (0.4%)
4. **What is your understanding of the potential cyber risks associated with using chatbot technology? (Refer to Figure-2) (Multiple Choice)**	
- **Social engineering attacks**: Cyber criminals can use chatGPT to socially engineer victims into divulging confidential information.	159 (61.4%)
- **Malware threats**: Malicious software can be installed on a user's device through a malicious link or file received via chatGPT.	129 (49.8%)
- **Phishing attacks**: Cyber criminals can use chatGPT to send malicious links or messages to trick victims into revealing sensitive information or downloading malware.	100 (38.6%)
- **Identity theft**: ChatGPT conversations can be used to gain access to a person's identity, allowing cyber criminals to steal data or commit fraud.	87 (33.6%)
- **Data leakage**: If data is shared on chatGPT, it can be accessed by unauthorized users, leading to data leakage.	66 (25.5%)
- None of these	1 (0.4%)

continued on following page

Table 2.
Continued

Survey questions and options	**Responses**
5. How likely are you to use a chatbot service for daily communication? (Refer to Figure-3)	
- Unlikely – 1	12 (4.6%)
- Not very likely – 2	15 (5.8%)
- Likely – 3	39 (15.1%)
- Very likely – 4	122 (47.1%)
- Extremely likely – 5	71 (27.4%)
6. Do you think chatbots can be used to collect personal information or to manipulate users? (Refer to Figure-4)	
- Yes	223 (87.8%)
- No	31 (12.2%)
7. Are you aware of any security measures that should be taken while using chatbot services?	
- four ways to protect your system from chatbot security concerns. These include encryption, authentication, processes & protocols, and education.	22 (8.5%)
- No, not aware of any security measures	76 (29.3%)
- AI chatbots by implementing processes and protocols such as end-to-end encryption, 2F authentication, installing malware, following regional data regulation laws, educating employees and users about various threats	54 (20.8%)
- Do not share personal info or PII	107 (41.3%)
8. In what ways would you use chatbots to protect yourself from cyber risks?	
- Chatbot security vulnerabilities can include impersonating employees, ransomware and malware, phishing, and bot repurposing,	7 (2.7%)
- By keeping up to date on the latest information	45 (17.4%)
- overcome security risks associated with conversational AI chatbots by implementing processes and protocols such as end-to-end encryption, 2F authentication, installing malware, following regional data regulation laws, educating employees and users about threats, etc.	128 (49.4%)
- I don’t know any ways to use chatbots to protect myself.	36 (13.9%)
- Process and protocols, education and awareness, and whitelisting.	43 (16.6%)
9. What measures do you think should be taken by chatbot services to reduce cyber risks? (Refer to Figure-5) (Multiple-Choice)	
- **Data privacy:** Ensure that sensitive information is not shared through the chatbot, and if it is necessary to collect such data, make sure it is securely stored and protected.	175 (67.6%)
- **User authentication**: Implement proper user authentication methods to secure the chatbot from unauthorized access.	149 (57.5%)
- **Network security**: Secure the communication channel between the chatbot and users by using encryption technologies such as SSL/TLS.	129 (49.8%)
- **Access control**: Implement access controls to restrict who can access and modify the chatbot’s configuration and data.	104 (40.2%)
- **Monitoring and logging**: Monitor the chatbot’s activity and maintain logs to detect and respond to any security incidents.	71 (27.4%)
- **Vulnerability management:** Regularly assess and address vulnerabilities in the chatbot’s design, code, and infrastructure to prevent potential security risks.	76 (29.3%)
- **Regular updates**: Keep the chatbot and its dependencies up to date with the latest security patches and updates.	54 (20.8%)
10. Are there any risks that you are particularly concerned, when it comes to using chatbots?	
- Attackers can also hack into systems and cause a chatbot to spread **malware** or **ransomware** to users’ devices	62 (23.9%)
- **Fake News**: chatbots are limited in their ability to understand natural language and can lead users astray by providing false information.	57 (22.0%)
- **Data Security**: ensuring foul-proof data storage and handling mechanisms	78 (30.1%)
- **Loss of jobs/**replacement with AI	43 (16.6%)
- Not aware of any specific risks related to chatbots	19 (7.3%)

Figure 2.
Survey responses on the understanding of potential cyber risks associated with Chatbots

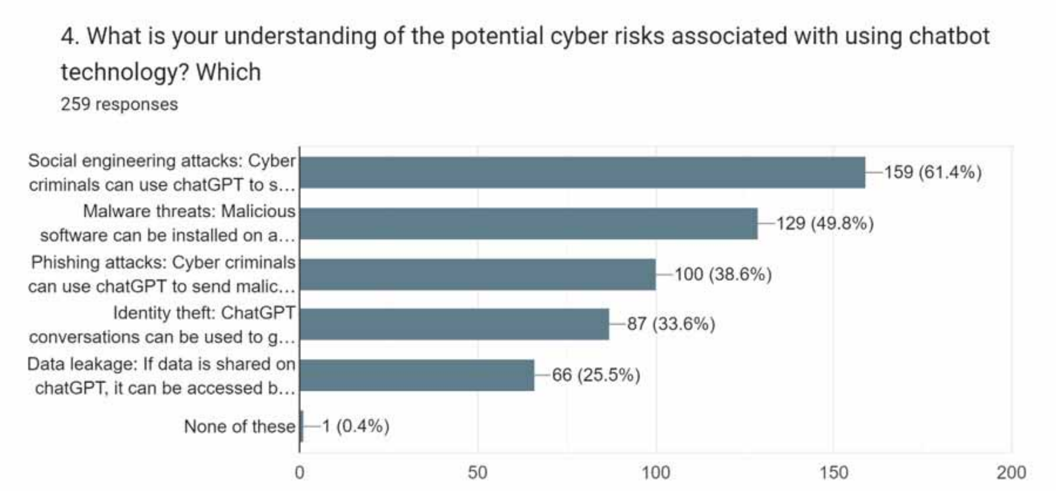
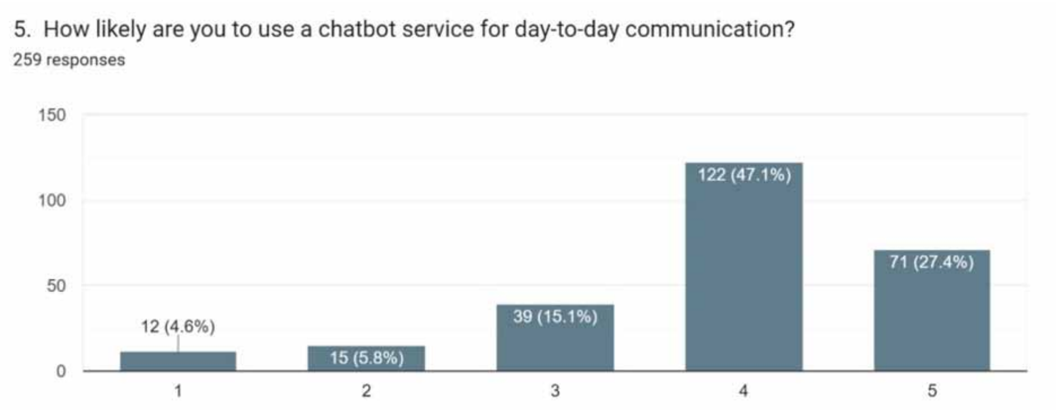


Figure 3.
Survey responses on how likely users are planning to use Chatbots in daily communication



RESULTS AND DISCUSSION

AI-based chatbots, as with any advanced technology, poses unique cyber risks that must be addressed. The survey in this study tried to understand the general perception and emotions with respect to the use of ChatGPT and similar AI-based chatbots. Most of the survey responders (64%) were excited about the improvements in AI/ML technology, about 38% of the users were scared of AI replacing humans, and about the same number of people thought of it as a tool that would increase human efficiency similar to what computers did in the 1980s and 90s. 61.4% of the users considered Social engineering attacks as the main cyber threat using chatbots followed by Malware threats (49.8%). Almost a quarter of the survey takers (74.5%) mentioned that they are either likely or very likely to use ChatGPT or similar AI-based chatbots for their daily work and other activities. It is also interesting to note that 87.8% of the survey takers though that chatbots could be used to collect personal information or to manipulate users.

Figure 4.
Survey responses on if chatbots can be used to collect personal information or manipulate users

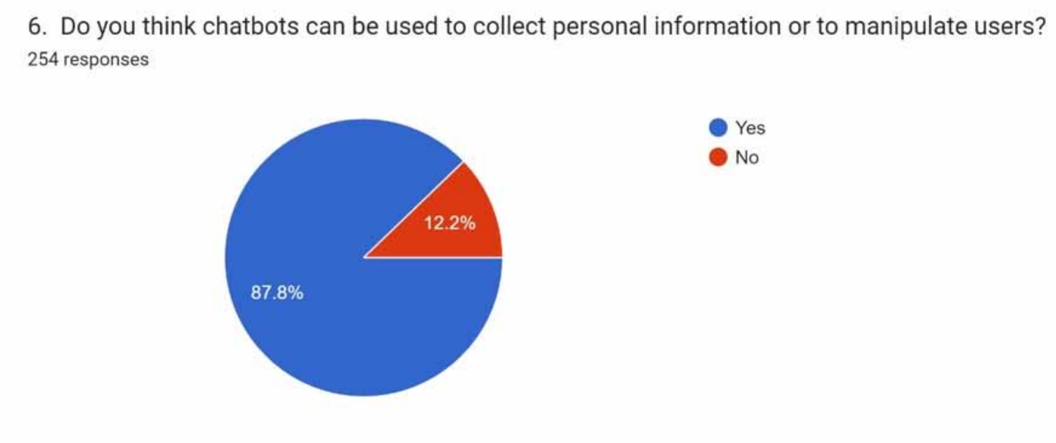
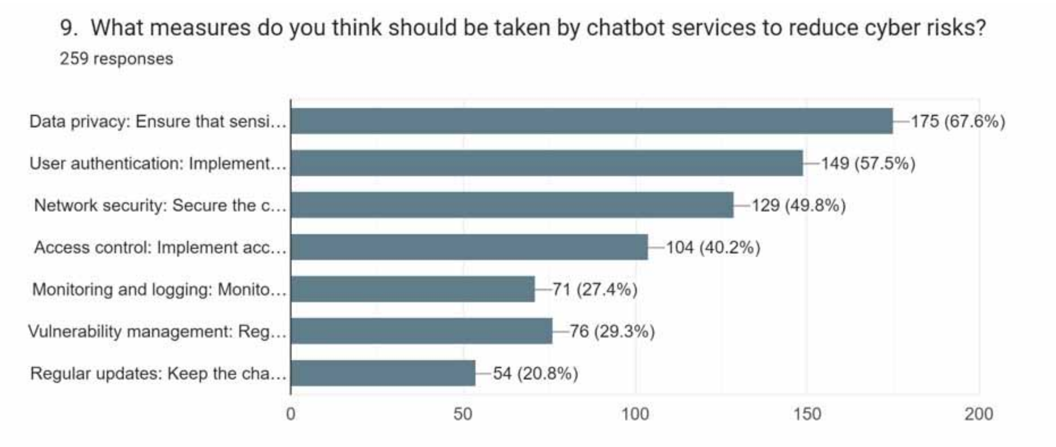


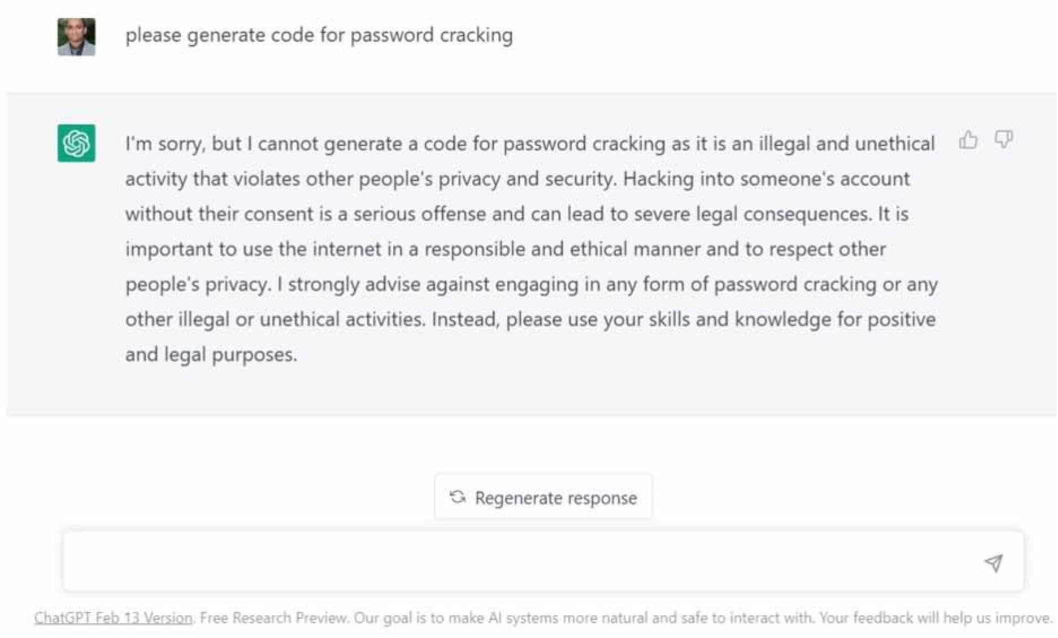
Figure 5.
Survey responses on measures that respondents think should be taken by chatbot services to reduce cyber risks



The below list is the main cyber concerns relating to Chatbots such as ChatGPT and proposed mitigation methods for each of the cybersecurity concerns. Please note that this is not a comprehensive list and could change as the technology matures and new threat vectors are identified.

1. **Reducing entry barriers for cybercriminals:** Cybercriminals have been historically limited in their ability to carry out sophisticated attacks due to the need for coding and scripting skills, including writing new malware code and password-cracking scripts. It has been discussed that chatGPT would effectively reduce the barriers that existed to becoming a script kiddie or a cybercriminal as it can be used to create the code for computer malware or password-cracking software. However as shown in Figure-6, there are inbuilt controls in place to deter bad actors from obtaining such information. Such checks need to be incorporated for multiple such data types as well so that the misuse of AI chatbots can be prevented such as using below attacks.

Figure 6.
Inbuilt controls in place to deter bad actors from using ChatGPT to create Malware and password cracking code



please generate code for password cracking

I'm sorry, but I cannot generate a code for password cracking as it is an illegal and unethical activity that violates other people's privacy and security. Hacking into someone's account without their consent is a serious offense and can lead to severe legal consequences. It is important to use the internet in a responsible and ethical manner and to respect other people's privacy. I strongly advise against engaging in any form of password cracking or any other illegal or unethical activities. Instead, please use your skills and knowledge for positive and legal purposes.

Regenerate response
ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

- **Social engineering attacks**: Cyber criminals can use chatGPT to socially engineer victims into divulging confidential information.
- **Malware threats**: Malicious software can be installed on a user's device through a malicious link or file received via chatGPT.
- **Phishing attacks**: Cyber criminals can use chatGPT to send malicious links or messages to trick victims into revealing sensitive information or downloading malware.
- **Identity theft**: ChatGPT conversations can be used to gain access to a person's identity, allowing cyber criminals to steal data or commit fraud.
- **Data leakage**: If data is shared on chatGPT, it can be accessed by unauthorized users, leading to data leakage.

As mitigation, it is important to implement security measures and access controls that prevent unauthorized access to the system. Additionally, ChatGPT can be programmed to detect and flag potentially malicious or fraudulent text, which can be reviewed by human moderators or security experts. By taking these steps, the risk of cybercriminals using ChatGPT to generate scripts for malicious purposes can be reduced. Further, spreading cybersecurity awareness (Sebastian, S. R., & Babu, B. P.,2022) is also paramount to ensure that users are aware of the cyber risks of various technologies.

2. **Compliance with regulation**: the regulation of AI use in writing and other fields would need to be incorporated to avoid misuse of such AI-based bots. Presently there are very limited laws with regard to using AI-based Chatbots for work, education, and other similar activities. Watermarking text generated by such AI chatbots would be a good step to identifying work completed by AI. There need to be laws enacted such as those listed below, for monitoring the use of AI-based chatbots.

- **Data protection laws:** Data protection laws regulate the collection, use, and storage of personal data, and require organizations and individuals to obtain informed consent before collecting or using personal data. In the case of ChatGPT, data protection laws can be used to regulate the collection and use of data generated through the system, and ensure users' privacy is protected.
- **Intellectual property laws:** Intellectual property laws protect the rights of individuals and organizations to their creative works, including text, images, and videos. In the case of ChatGPT, intellectual property laws can be used to prevent the unauthorized use of copyrighted material, and to ensure that users of the system do not infringe on the intellectual property rights of others.
- **Consumer protection laws:** Consumer protection laws regulate the relationship between businesses and consumers, and require businesses to provide accurate and truthful information to consumers. In the case of ChatGPT, consumer protection laws can be used to regulate the use of the system in marketing and advertising, and to ensure that users are not misled or deceived by false or misleading information generated through the system.
- **Cybersecurity laws:** Cybersecurity laws regulate the use of computer systems and networks, and require organizations to take measures to protect against cyber threats and attacks. In the case of ChatGPT, cybersecurity laws can be used to ensure that the system is secure and protected from unauthorized access or use.

Overall, these provisions in law can be used to monitor and regulate the use of ChatGPT, and to ensure that the system is used in a responsible and ethical manner that respects the rights and interests of all stakeholders.

3. **False information**: Spreading false information through ChatGPT was one of the main concerns shared through the survey. There is a risk that ChatGPT could be used to generate and spread false information or propaganda. This is because ChatGPT can generate text that appears to be written by a human and may therefore be perceived as more trustworthy or credible than content generated by bots or automated systems.

If malicious actors or groups gain access to ChatGPT, they could use it to generate false news articles, misleading social media posts, or fraudulent customer reviews. This false information can also be due to the information that is used to train the AI. This could have significant consequences, such as spreading misinformation, causing harm to individuals or organizations, or influencing public opinion or decision-making. To mitigate this risk, it is essential to monitor and verify the accuracy and credibility of the information generated by ChatGPT. This can be done by fact-checking, using trusted sources, and implementing safeguards to prevent malicious actors from accessing or using the system. Additionally, ChatGPT could be programmed to detect and flag potentially false or misleading information, which could be reviewed by human moderators or fact-checkers. The data that is used to train the AI needs to be vetted and reviewed to ensure it is free of any bias or false information as well, which could cause further propagation of incorrect information.

4. **Algorithmic fairness**: Recent studies have shown that algorithmic decision-making may be inherently prone to unfairness, even when there is no intention for it (Pessach, D., & Shmueli, E., 2020). As an artificial intelligence language model, ChatGPT has the potential to perpetuate biases and unfairness present in the data it was trained on. This is because ChatGPT uses large datasets of text that reflect the biases and stereotypes of the human society and culture from which they were drawn. If not properly addressed, this can result in biased or unfair responses to users based on factors such as their race, gender, age, or socioeconomic status (Božić, V.,2012).

Furthermore, ChatGPT algorithms can learn to replicate and reinforce these biases through repeated interactions with users, as it continues to process new data and refine its responses. This can result in unintended consequences, such as discriminatory or harmful suggestions or actions based on a user's perceived identity or characteristics. Therefore, it is essential to continuously monitor and test ChatGPT responses to ensure that they are fair, unbiased, and ethical. This can be done through regular audits, testing for potential bias and discrimination, and incorporating diverse and representative datasets in its training. By addressing the risk of algorithmic fairness, ChatGPT can become a more reliable and equitable tool for users (Sebastian, G., 2022), (Zhuo, T. Y. et. Al,2023).

CONCLUSION

With AI Chatbots and other tools getting more common, it is to be expected that the vulnerabilities and associated cybersecurity risks will increase multifold. Apart from issues such as data privacy and similar more common cyber risks, ChatGPT also runs the risk of providing easy scripting and coding access to cyber criminals, which effectively reduces the barriers to entry in this field. There are existing controls that would deter and prevent malicious users from gaining access to such scripts and code, however since the technology landscape is fast evolving, the risks and associated controls need to be continuously reviewed and monitored and additional controls need to be put in place to ensure the vulnerability is addressed adequately. While the scope of this study was limited to providing a summary of cyber risks associated with this nascent technology, future studies can focus on each of these risks in detail and the updates needed to existing controls to address these dynamic cyber vulnerabilities.

FUTURE SCOPE

Future study scope includes using larger models that can capture more complex relationships in the data, as well as using better training techniques for more efficient optimization algorithms, and advanced learning rate schedules. Including better data pre-processing to remove noise and improves the quality of the model's predictions. Incorporating external knowledge sources and multiple modalities, such as structured data, knowledge graphs, or other domain-specific information, helps the model better understand the context of the conversation. GAN is used for generating synthetic data in general, while GPT is specifically designed for generating natural language text (Aggarwal, A., Mittal, M., & Battineni, G. (2021). Another future scope could also include how GAN (Generative Adversarial Networks) and GPT (Generative Pre-trained Transformer) which are both deep learning models used for generating synthetic data but have different architectures and purposes can be used together.

COMPLIANCE WITH ETHICAL STANDARDS

The author certifies that he has no other potential conflicts of interest. The research involved human participants. The survey shared with the participants explained the purpose of the research, the data collected. Participants took part voluntarily and were given the option to skip the survey at any stage. Further, informed consent was obtained from all subjects and/or their legal guardian(s).

REFERENCES

Aggarwal, A., Mittal, M., & Battineni, G. (2021). Generative adversarial network: An overview of theory and applications. *International Journal of Information Management Data Insights*, *1*(1), 100004.

Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk research: Review and recommendations. *Journal of Management*, *47*(4), 823–837.

Aydın, Ö., & Karaarslan, E. (2022). *OpenAI ChatGPT generated literature review: Digital twin in healthcare*. Available at SSRN 4308687.

Božić, V. (2012). Risk management in informatization. *Proceedings of the Central European*.

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *bioRxiv*, 2022–12.

Hacker, P., Engel, A., & Mauer, M. (2023). *Regulating ChatGPT and other Large Generative AI Models*. arXiv preprint arXiv:2302.02337.

Jiao, W., Wang, W., Huang, J. T., Wang, X., & Tu, Z. (2023). *Is ChatGPT a good translator? A preliminary study.* preprint arXiv:2301.08745.

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., & Tseng, V. et al. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, *2*(2), e0000198.

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries? Library Hi Tech News. doi:10.1108/LHTN-01-2023-0009

OpenA. I. ChatGPT. (2022). https://openai.com/blog/chatgpt/

Pessach, D., & Shmueli, E. (2020). *Algorithmic fairness*. arXiv preprint arXiv:2001.09784.

Sebastian, G. (2022). Cyber Kill Chain Analysis of Five Major US Data Breaches: Lessons Learnt and Prevention Plan. *International Journal of Cyber Warfare & Terrorism*, *12*(1), 1–15. https://doi.org/10.4018/IJCWT.315651

Sebastian, S. R., & Babu, B. P. (2022). Are we Cyber aware? A cross sectional study on the prevailing Cyber practices among adults from Thiruvalla, Kerala. *International Journal of Community Medicine and Public Health, 10*(1), 235–239. 10.18203/2394-6040.ijcmph20223550

van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, *614*(7947), 224–226.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Polosukhin, I. et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.

Zhuo, T. Y., Huang, Y., Chen, C., & Xing, Z. (2023). *Exploring AI Ethics of ChatGPT: A Diagnostic Analysis.* arXiv preprint arXiv:2301.12867.