American Journal of Computer Science and Technology
2024; 7(1): 1-8
http://www.sciencepublishinggroup.com/j/ajcst
doi: 10.11648/j.ajcst.20240701.11
ISSN: 2640-0111 (Print); ISSN: 2640-012X (Online)

Securing Well-Being: Exploring Security Protocols and Mitigating Risks in AI-Driven Mental Health Chatbots for Employees

Sourav Banerjee*, Ayushi Agarwal, Ayush Kumar Bar

Datalabs, United We Care, Gurgaon, India

Email address:

souravbanerjee@unitedwecare.com (Sourav Banerjee), ayushi@unitedwecare.com (Ayushi Agarwal), abar@unitedwecare.com (Ayush Kumar Bar)

*Corresponding author

To cite this article:

Sourav Banerjee, Ayushi Agarwal, Ayush Kumar Bar. (2024). Securing Well-Being: Exploring Security Protocols and Mitigating Risks in AI-Driven Mental Health Chatbots for Employees. *American Journal of Computer Science and Technology*, *7*(1), 1-8. https://doi.org/10.11648/j.ajcst.20240701.11

Received: December 18, 2023; **Accepted**: December 29, 2023; **Published**: January 11, 2024

Abstract: In today's workplace, mental health is gaining importance. As a result, AI-powered mental health chatbots have emerged as first-aid solutions to support employees. However, there are concerns regarding privacy and security risks, such as spoofing, tampering, and information disclosure, that need to be addressed for their implementation. The objective of this study is to explore and establish privacy protocols and risk mitigation strategies specifically designed for AI-driven mental health chatbots in corporate environments. These protocols aim to ensure the ethical usage of these chatbots. To achieve this goal, the research analyses aspects of security, including authentication, authorisation, end-to-end encryption (E2EE), compliance with regulations like GDPR (General Data Protection Regulation) along with the new Digital Services Act (DSA) and Data Governance Act (DGA). This analysis combines evaluation with policy review to provide comprehensive insights. The findings highlight strategies that can enhance the security and privacy of interactions with these chatbots. Organisations are incorporating heightened security measures, including the adoption of Two-factor Authentication (2FA) and Multi-Factor Authentication (MFA), integrating end-to-end encryption (E2EE), and employing self-destructing messages. Emphasising the significance of compliance, these measures collectively contribute to a robust security framework. The study underscores the critical importance of maintaining a balance between innovative advancements in AI-driven mental health chatbots and the stringent safeguarding of user data. It concludes that establishing comprehensive privacy protocols is essential for the successful integration of these chatbots into workplace environments. These chatbots, while offering significant avenues for mental health support, necessitate effective handling of privacy and security concerns to ensure ethical usage and efficacy. Future research directions include advancing privacy protection measures, conducting longitudinal impact studies to assess long-term effects, optimising user experience and interface, expanding multilingual and cultural capabilities, and integrating these tools with other wellness programs. Additionally, continual updates to ethical guidelines and compliance with regulatory standards are imperative. Research into leveraging AI advancements for personalised support and understanding the impact on organisational culture will further enhance the effectiveness and acceptance of these mental health solutions in the corporate sector.

Keywords: AI-Driven Mental Health Chatbots, Privacy Protocols, Security Threats, GDPR Compliance, Corporate Mental Health, Risk Mitigation, Data Security

1. Introduction

The recognition of the significance of mental health programs within companies is growing in today's business world. This is because mental health problems have a considerable impact on both employee productivity and overall well-being [15, 12]. In the United States, one out of every five adult's experiences mental illness each year, yet only a third of those individuals who require assistance actually receive it [4]. This gap often leads to presenteeism,

where employees come to work despite struggling with their mental or physical health. Unfortunately, this can have a negative effect on their performance and productivity.

The World Health Organization (WHO) has estimated that depression and anxiety have a significant impact on the economy, resulting in an annual loss of $1 trillion in global productivity [22]. These figures emphasise the pressing requirement for effective mental health programs in work environments. To emphasise the importance of emotional well-being, many employers are actively working to improve the benefits associated with mental health. These improvements encompass stress management initiatives. Assisting in managing invisible conditions such as anxiety and depression [3]. Research has demonstrated that prioritising employees' mental health can greatly enhance productivity since individuals who receive treatment report improved job performance [24]. Additionally, addressing health concerns can contribute to higher rates of employee retention, as a considerable number of workers attribute their decision to leave a job to mental health factors.

Individuals who suffer from severe mental illness often have higher rates of cardiovascular and metabolic disorders, indicating a link between mental and physical well-being [2]. Consequently, it has become essential to implement workplace initiatives that focus on mental health to avoid severe consequences. This research paper showcases how Kareify's application addresses mental health concerns. The app provides a variety of self-help tests and clinical programs that users can take to evaluate their mental health. Based on the results, users can also book appointments with clinical psychologists and psychiatrists, as exemplified in Figure 1 (a) and (b). These examples demonstrate how technology can be used to promote mental health and improve access to care through Mental health Chatbots.

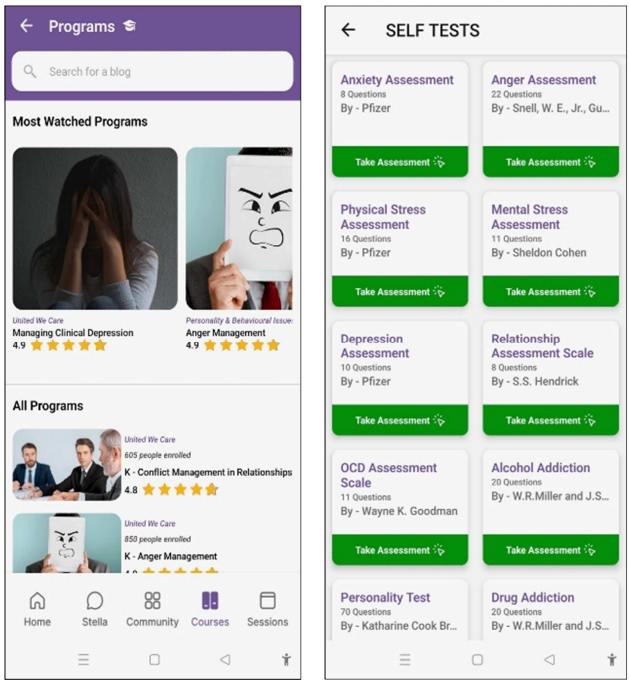

Figure 1. *(a) Self-Help Programs; (b) Self-help Assessments.*

From a research standpoint, it is widely recognised that implementing health programs in the corporate setting is essential for the success of organisations and the well-being of employees. These programs have been proven to increase productivity and job satisfaction in the workplace. Neglected mental health disorders can have a negative impact on employee performance. Research, including a study conducted by Deloitte, suggests that organisations can achieve a return of up to $4.20 for every dollar invested in health initiatives [5, 21]. This return is mainly attributed to absenteeism, improved employee productivity, and lower healthcare costs.

Moreover, mental wellness initiatives play a role in enhancing employee retention rates. When workers feel that their employer genuinely cares about their well-being, they tend to display higher levels of loyalty and dedication, resulting in reduced turnover and related costs. Additionally, companies that prioritise the health of their employees are more likely to gain positive perception from customers, investors and prospective hires, which ultimately contributes to an improved corporate image and reputation.

From a moral and legal standpoint, employers have a responsibility to ensure that the workplace is safe and promotes good mental health. Failing to meet these conditions can lead to repercussions. Essentially, implementing health programs in corporate settings is not only advantageous but also necessary for the overall well-being of both the organisation and its employees.

1.1. Rise of AI-Driven Mental Health Chatbots

The emergence of health chatbots powered by AI indicates a notable change in how mental healthcare is addressed within the business community. These chatbots harness the capabilities of machine learning and artificial intelligence to replicate interactions via text or voice, providing a wide array of services spanning from mental health assistance to customer support.

Notably, voice-based chatbots, integrated into devices like smartphones and smart speakers, require sophisticated speech-to-text translation, which factors like accent and background noise can influence. Text-based chatbots, accessible via various platforms like Messenger, Slack, and web applications, offer a more controlled interaction with features like quick replies and easier handling of complex queries.

The evolution of chatbots from the early days of ELIZA, a program designed to mimic a psychotherapist, to current sophisticated systems using deep neural networks reflects significant technological advancements [1, 13]. These modern chatbots are increasingly adopted for digital interventions in mental health, a field where access to services remains a challenge globally.

AI-driven chatbots in mental health, such as Woebot, have shown promising results in demonstrating their ability to reduce symptoms of mental health conditions like depression effectively, and the most common response to the use of AI mental health chatbots are cheap, easy-to-use and accessible

anytime [23]. However, these chatbots also face challenges, including privacy concerns, limitations in understanding human language nuances, and the ability to handle sensitive topics like suicide prevention.

The study also delves into the Mental Health First Aid applications developed by Kareify, focusing particularly on Stella, an AI-based chatbot trained in Cognitive Behavioural Therapy (CBT), as depicted in Figure 2(a) and (b). Stella is designed to deliver on-demand Mental Health First Aid and has gained considerable adoption, with over 550,000 monthly users regularly accessing it.

Analysis conducted within this research indicates that these applications are notably effective in equipping users with essential resources for managing their mental health concerns. Specifically, the chatbot Stella has been identified as an invaluable asset during crises, equipped with SOS redirection capabilities that ensure users are guided to the appropriate resources swiftly and efficiently.

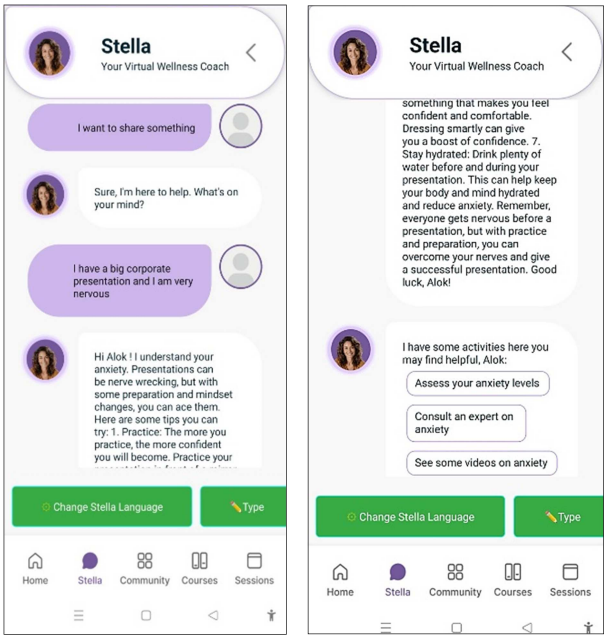

Figure 2. *(a) AI chatbot STELLA; (b) SOS and Emergency Response.*

Despite their challenges, chatbots offer several advantages, including 24/7 availability, cost savings, and anonymity, encouraging more open disclosure of sensitive information. They can be particularly beneficial for people who face barriers to accessing traditional mental health services due to stigma, geographic isolation, or unconventional work shifts.

However, it is imperative to acknowledge the inherent limitations of chatbots in mental health care and offer clear elucidations regarding the capabilities and potential biases of these conversational agents. It becomes paramount to prioritise privacy safeguards, enhance adherence, and maintain a reasonable balance in relying on these digital tools to ensure their safe and ethical utilisation. Mental health professionals assume a pivotal role in this context, serving as crucial guides in both the development and application of these chatbots. Their involvement is instrumental in ensuring that these technological tools complement, rather than substitute, traditional mental health services.

1.2. Overview of the Paper's Focus on Protocols and Risk Mitigation

The central emphasis of this paper lies in the exploration of protocols and strategies to mitigate risks associated with the use of AI-driven mental health chatbots for employees. It seeks to establish comprehensive guidelines for the responsible utilisation of these chatbots, ensuring their alignment with ethical standards and the safeguarding of user privacy. The paper further delves into mechanisms designed to protect against potential risks, encompassing aspects such as the misinterpretation of mental health conditions, concerns related to data security, and the potential over-reliance on chatbot interactions for mental health support. Additionally, it underscores the imperative for continuous monitoring and evaluation of these AI tools to iteratively refine and enhance their efficacy and safety within a corporate environment.

1.3. Privacy Protocols in AI-Driven Chatbots

The utilisation of AI-driven mental health chatbots in the corporate sector raises significant privacy concerns, necessitating stringent protocols to safeguard user data. These chatbots, while beneficial in providing accessible mental health support, encounter challenges in ensuring data security and privacy. The primary security issues revolve around threats and vulnerabilities, categorised under the STRIDE model, which includes risks like spoofing, tampering, and unauthorised information disclosure [6, 13].

A fundamental facet in guaranteeing the security of chatbots involves the meticulous implementation of robust authentication and authorisation procedures. This assumes paramount significance, particularly in instances where chatbots are entrusted with managing sensitive user data, as observed in sectors like banking or healthcare. The authentication process typically entails the utilisation of a combination of methodologies, incorporating biometrics and Two-Factor or Multi-Factor Authentication, to validate user identity rigorously. This multi-layered approach ensures that access to sensitive information is restricted solely to duly authorised individuals, thereby fortifying the overall security architecture.

An additional pivotal factor is the implementation of end-to-end encryption (E2EE), a measure designed to uphold the confidentiality of communication between users and the chatbot, restricting access solely to the intended participants [13]. This becomes particularly crucial in scenarios involving the exchange of sensitive mental health information. While numerous chatbots employ HTTPS for secure data transfer, it is imperative to recognise that only E2EE can comprehensively ensure the privacy of the conversational exchange [13].

Moreover, it is crucial to comply with privacy regulations by the General Data Protection Regulation (GDPR) [10].

GDPR emphasises the need for pseudonymisation and encryption of data, highlighting the significance of safeguarding data in chatbot interactions. Developers of chatbots must guarantee that their systems adhere to these regulations, particularly when it comes to handling, storing and sharing user data. According to the Blueprint for an AI Bill of Rights, systems must undergo testing before they are deployed, with a focus on identifying and minimising risks. Continuous monitoring is also necessary to ensure the safety and effectiveness of these systems, taking into account their intended use and addressing any potential issues that may arise beyond their original purpose. It is essential for these systems to adhere to industry standards in order to maintain their reliability and safety [25].

The introduction of self-destructing messages in chatbots, particularly in financial and healthcare contexts, is another method to enhance privacy [13]. This feature aligns with GDPR's requirement for data minimisation, ensuring that personal data is not stored longer than necessary and should be forgotten/deleted.

2. Security

In the domain of AI-driven mental health chatbots, security considerations assume a pivotal role in safeguarding the confidentiality, integrity, and availability of sensitive data. The incorporation of chatbots in mental health care within corporate settings underscores the imperative for a thorough examination of potential security risks and vulnerabilities. Concurrent with this scrutiny is the essential task of devising robust protocols designed to mitigate these identified risks effectively.

A primary security concern is the potential for unauthorised access to sensitive information. Chatbots, such as ChatGPT and BARD, are trained on extensive datasets, potentially including sensitive personal information. This raises notable concerns about data privacy and the unauthorised use of personal data. The lack of explicit consent from individuals whose data are used in training these models underscores the ethical and legal challenges inherent in deploying these systems. Additionally, the potential for proprietary or copyrighted information to be included in these datasets further complicates the security landscape.

Another important concern revolves around the exposure of sensitive data. When users engage with chatbots, there is a possibility that they might unintentionally disclose information, such as personal health records or business secrets. Such unintended disclosures can result in breaches of data, affecting both personal privacy and corporate confidentiality. Therefore, it is crucial to evaluate the integration of chatbots into systems that handle Protected Health Information (PHI), taking measures to prevent unauthorised access and ensure compliance with data protection regulations [19, 20].

Ensuring transparency in the utilisation of data for training AI models is a matter of concern. The "right to be forgotten" mentioned in Article 17 of the General Data Protection Regulation (GDPR) becomes especially important in this context [7]. Individuals might want to modify or delete their data if it has been used without their consent or is inaccurate. It is crucial to ensure the accuracy and relevance of the data employed in training these models to maintain trust and effectiveness in chatbot systems.

The exploitation of chatbots for malicious purposes, such as phishing and malware creation, represents another significant security threat and raises ethical concerns [16]. The ability of AI systems to generate convincing phishing content or assist in the creation of malware poses a substantial risk to both individuals and organisations. This underscores the need for robust security protocols to detect and prevent such malicious uses.

Lastly, the escalation of security risks is underscored by the increasing reliance on technology. Despite the potential fortification offered by encrypted transmissions and multifactor authentication, the ramifications of a system breach persist as a significant concern. A breach could lead to widespread exposure of sensitive information or the spreading of misleading or biased information, and this emphasises robust security measures in the realm of technological dependence.

2.1. Authentication and Authorization in AI-Driven Mental Health Chatbots

The differentiation between authentication and authorisation is crucial in the security architecture framework of AI-powered health chatbots, particularly in corporate environments. Authentication involves verifying the identity, which becomes crucial when chatbots handle user data, such as in banking or healthcare situations. Common authentication methods include using credentials like usernames, passwords, biometric information and tokens generated by the system. It's worth noting that, in applications where enhanced security is necessary, these tokens are often temporary to ensure a higher level of security by limiting their lifespan.

Moreover, it is crucial to incorporate role-based access control (RBAC) as an aspect [11]. RBAC establishes access rights based on users' roles and responsibilities, ensuring that authorised personnel have the ability to access sensitive information. This approach strengthens the security measures by limiting data accessibility in accordance with defined roles, effectively reducing risks associated with unauthorised usage.

The concept of authorisation goes beyond verifying someone's identity; it also involves granting them specific access to data and services. Let's take the example of a banking chatbot; once a user has successfully proven their identity, authorisation comes into play by determining which account information they are allowed to view. Additional measures like Two Factor Authentication (2FA) and Multi-Factor Authentication (MFA) are implemented to enhance security further. These protocols require users to confirm their identity through channels like email and text messages, providing an extra layer of protection in the authentication process, as shown in Figure 3(a) and (b).

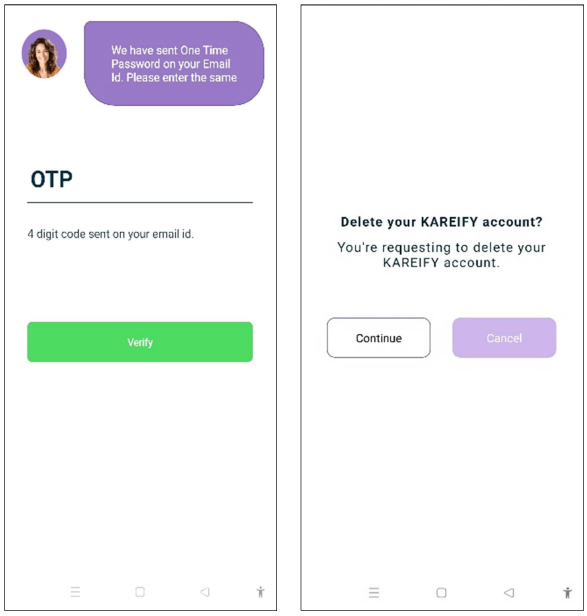

Figure 3. *(a) Two-Factor Authentication; (b) Right to be Forgotten.*

2.2. Security Risks and Mitigation Strategies in Chatbot Communication and Data Protection

The emergence of malicious chatbots, especially within cross-platform messaging apps, has led to novel phishing tactics such as Smishing. These deceptive chatbots pose as legitimate services, aiming to extract personal information from unsuspecting users. Mitigating these attacks hinges on cultivating awareness and refraining from responding to unrecognised services. Additionally, the adoption of temporary tokens within chatbot sessions serves as a preventive measure against unauthorised access to sensitive information, particularly in scenarios where a user's device is lost or compromised. Further, the incorporation of data anonymisation techniques, including de-identification, pseudonymisation, and generalisation, plays a pivotal role in enhancing privacy. These methods involve the removal or alteration of identifiable information, ensuring that individuals cannot be directly identified. This multifaceted approach helps to strengthen defences against potential privacy breaches in the context of malicious chatbots and related security threats [17].

2.3. End-to-End Encryption in Chatbot Conversations

End-to-end encryption (E2EE) plays a role in safeguarding the privacy of our conversations when interacting with chatbots. With E2EE, only the individuals involved in the communication can. Understand the messages. This is achieved by encrypting the data in such a way that only the intended recipient can decrypt it. Implementing E2EE is particularly crucial when it comes to protecting personal information shared during these conversations [14]. While many chatbots on websites use HTTPS for data transfer, it's important to note that HTTPS provides point-to-point encryption rather than end-to-end encryption. This distinction emphasises why it's essential to incorporate E2EE to ensure that our data remains secure throughout its journey.

2.4. GDPR Compliance and Self-Destructing Messages

In adherence to the General Data Protection Regulation (GDPR), specifically referencing Article 32(a) [8] and Article 5(e) [9], organisations are obligated to implement measures such as pseudonymisation and encryption of personal data. Moreover, the adoption of self-destructing messages for the transmission of Sensitive Personally Identifiable Information (PII) aligns with GDPR provisions outlined in Article 17 [7]. This ensures that personal data is retained only for the duration necessary, a crucial element in compliance. This approach holds particular relevance in communications involving financial and healthcare data, where the stringent control of sensitive information retention is imperative [8].

The Digital Services Act (DSA) and the Data Governance Act (DGA) are steps in the European Union's regulations [26], specifically addressing concerns regarding online content and data sharing. The DSA focuses on combating harmful material on platforms requiring proactive removal of things like hate speech, disinformation and terrorist propaganda. It imposes rules on marketplaces and social media platforms while giving national authorities the power to enforce compliance with hefty fines.

At the same time, the Data Governance Act (DGA) is aimed at improving access to and sharing of personal data within the EU [27]. This legislation establishes a framework for data intermediaries that act as trusted third parties facilitating data sharing. Notably, it also introduces data portability rights for businesses, allowing the movement of data across platforms. The overall objective is to increase data availability for research and innovation purposes. These laws highlight the EU's dedication to establishing a more transparent environment while promoting responsible practices in managing data for innovation purposes.

In the United States, President Biden issued a groundbreaking Executive Order to ensure the U.S. leads in AI, emphasising safety, security, privacy, equity, and innovation. The order establishes new standards for AI, requiring safety test sharing, developing rigorous standards, and addressing risks in critical sectors [28]. It prioritises privacy protection, encourages Congress to pass data privacy legislation, and promotes privacy-preserving techniques. The order aims to advance equity, combat discrimination, and ensure fairness in criminal justice. It also focuses on consumer protection and workers' rights and promotes innovation while fostering international collaboration. The government will use AI responsibly, issuing guidance, enhancing procurement, and accelerating AI talent acquisition. The comprehensive strategy reflects a commitment to responsible innovation and continued collaboration with global partners.

2.5. Backend Security in Chatbot Systems

On the backend, the storage and management of user communication data are subject to significant security considerations. Organisations often store these

communications for service analysis and legal protection, with user consent typically embedded in terms and conditions. The data must be stored securely, with regular methods like regular expressions, pattern matching, and entity recognition employed to detect and protect personal data. However, challenges arise with open-domain chatbots, which are less restricted and potentially more vulnerable to security risks.

In addition, the widespread use of chatbots on platforms and the sharing of user information with third-party services highlight the need for a strong privacy policy and security measures [17]. It is crucial to handle the user data collected by these chatbots responsibly, whether it is for enhancing their services or conducting analysis, in order to prevent access and misuse.

In order to tackle these security concerns, it is important to take an approach. This involves enforcing policies to protect data privacy, being transparent about how data is used, and having strong measures in place to safeguard and obtain consent for data. Furthermore, utilising security technologies like anomaly detection and encryption can help minimise the chances of unauthorised access or data breaches. Conducting frequent security audits and compliance checks is also vital for ensuring the reliability of these systems.

3. Discussion

The discussion surrounding the security of AI-driven mental health chatbots in corporate environments highlights a nuanced interplay of technological, ethical, and legal considerations. The imperative for robust authentication and authorisation mechanisms is crucial, especially when these chatbots manage sensitive information, as seen in healthcare applications. The dependence on conventional credentials, coupled with advanced techniques such as biometric identification and temporary token generation, indicates a heightened dedication to user security. However, these measures must construct a delicate balance between user convenience and accessibility, ensuring that heightened security does not hinder the usability of the chatbot or discourage user engagement.

The threat posed by malicious chatbots, exemplified by the rise of Smishing attacks, brings to light the evolving nature of cybersecurity threats in the realm of AI-driven communication. These threats jeopardise not only individual privacy but also corporate confidentiality and integrity [18]. Educating users about the risks associated with engaging with unrecognised chatbots and implementing robust system-level safeguards are crucial steps in mitigating these risks. The utilisation of temporary tokens to guard against unauthorised access in situations involving device loss or theft underscores the necessity for dynamic security measures that can adapt to diverse user contexts. This highlights the importance of a proactive and adaptive security approach to ensure the protection of user data and maintain a secure environment in the face of evolving challenges.

End-to-end encryption (E2EE) plays a role in the security framework of chatbots. Although using HTTPS and SSL/TLS protocols offers a level of security, it is essential to implement E2EE to ensure the privacy of confidential conversations [13]. This is especially important when considering GDPR compliance, where safeguarding data isn't just about security but also a legal requirement. The use of self-deleting messages in line with GDPR's principles of minimising data and limiting retention showcases an approach to protecting information in transient digital communications.

The implementation of security measures in the backend of chatbot systems, essential for the storage and management of communication data, brings forth a range of challenges, particularly in balancing service improvement and compliance requirements with user privacy concerns. To navigate these challenges while maintaining the functionality and analytical value of chatbot systems, GDPR-compliant methods such as data anonymisation and pseudonymisation are employed. These techniques, which include data masking, shuffling, or aggregation, are crucial in protecting user identities and ensuring personal data cannot be linked back to an individual, even in cases of unauthorized access.

The implementation of these privacy-preserving measures involves several critical steps. Initially, it requires the identification and categorisation of sensitive data that needs anonymisation or pseudonymisation. Once identified, appropriate anonymisation techniques are applied to obscure or remove personal identifiers. Regular data audits are conducted to ensure the effectiveness of these methods and compliance with the latest GDPR guidelines. Additionally, when chatbots are integrated with third-party services, it's vital to ensure these partners also adhere to stringent data privacy protocols, often formalised through contractual agreements and regular audits. An essential aspect of this implementation is the continuous update of privacy policies to accurately reflect current data handling practices, thereby maintaining transparency with users about the usage and protection of their data. Staff training and awareness are equally important, ensuring those involved in data processing are fully informed about the importance of data privacy and the specific protocols in place. Finally, a commitment to continuous improvement and adaptation is necessary, particularly in staying updated with technological and regulatory developments, to enhance and refine data protection strategies continually. As chatbots become more integrated across various platforms and interact with an increasing number of third-party services, the imperative for vigilant and evolving privacy policies and security protocols becomes paramount. This proactive approach in privacy management is crucial for maintaining ongoing compliance and safeguarding user trust in chatbot systems.

4. Conclusion

This paper emphasises the task of ensuring the safety of AI driven health chatbots in corporate settings. It is an issue that requires a combination of advanced technological solutions, comprehensive user education, compliance with legal regulations and ethical data handling. As digital

communication and AI technology continue to evolve it becomes increasingly crucial to adapt strategies that safeguard user privacy and security while utilising chatbots for health support in the workplace.

Looking ahead there is potential for research and development in this field from various perspectives. Future initiatives could explore emerging technologies such as blockchain to enhance security and data privacy in chatbot interactions. Considering the impact of international data protection laws on chatbot deployment will require a global outlook that ensures compliance while understanding regional differences in privacy regulations. Collaborations across disciplines will play a role bringing together technologists, mental health professionals, legal experts and ethicists to develop comprehensive chatbot solutions focused on user's needs. A significant challenge lies in finding the balance, between personalisation and privacy by employing innovative data processing methods that prioritise user consent and anonymity. Moreover, it is essential to understand how AI powered chatbots reshape workplace dynamics and integrate with existing health policies. Understanding the elements that influence user trust and acceptance of these technologies is vital for their implementation. It is important to address scalability and accessibility concerns to ensure that these tools can reach a range of people including those with limited access to technology. By exploring these areas future research and development can make significant contributions, to the field improving the functionality and safety of AI powered health chatbots. Taking a perspective is crucial to ensure that technological advancements are made in an ethical and legally compliant manner while also promoting the well-being of users. Ultimately this will shape the landscape of mental health support in workplaces.

Funding

This research is sponsored by United We Care to support the work and provide funds to cover publication costs.

ORCID

0009000559359944 (Sourav Banerjee)
0009000525759907 (Ayushi Agarwal)
0000000330506478 (Ayush Kumar bar)

Acknowledgments

This research was sponsored by United We Care, who provided financial support for the research of this study. Their contribution was instrumental in facilitating the data collection, analysis, and overall success of the research endeavour.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Bassett, C. (2018, February 21). The computational therapeutic: exploring Weizenbaum's ELIZA as a history of the present. AI & SOCIETY, 34(4), 803–812. https://doi.org/10.1007/s00146-018-0825-9

[2] C. (2022, April 26). Heart Disease and Mental Health Disorders | cdc.gov. Centers for Disease Control and Prevention. https://www.cdc.gov/heartdisease/mentalhealth.htm

[3] Catapano, P., Cipolla, S., Sampogna, G., Perris, F., Luciano, M., Catapano, F., & Fiorillo, A. (2023, October 20). Organizational and Individual Interventions for Managing Work-Related Stress in Healthcare Professionals: A Systematic Review. Medicina, 59(10), 1866. https://doi.org/10.3390/medicina59101866

[4] Centers for Disease Control and Prevention. (2023, April 25). About Mental Health. https://www.cdc.gov/mentalhealth/learn/index.htm

[5] Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016, May). Scaling-up treatment of depression and anxiety: a global return on investment analysis. The Lancet Psychiatry, 3(5), 415–424. https://doi.org/10.1016/s2215-0366(16)30024-4

[6] Deng, M., Wuyts, K., Scandariato, R., Preneel, B., & Joosen, W. (2010, November 16). A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. Requirements Engineering, 16(1), 3–32. https://doi.org/10.1007/s00766-010-0115-7

[7] GDPR.eu. (2018, November 14). Art. 17 GDPR - Right to erasure ('right to be forgotten') - GDPR.eu. https://gdpr.eu/article-17-right-to-be-forgotten/

[8] GDPR.eu. (2018, November 14). Art. 32 GDPR - Security of processing - GDPR.eu. https://gdpr.eu/article-32-security-of-processing/

[9] GDPR.eu. (2018, November 14). Art. 5 GDPR - Principles relating to processing of personal data - GDPR.eu. https://gdpr.eu/article-5-how-to-process-personal-data/

[10] GDPR.eu. (2018, November 7). What is GDPR, the EU's new data protection law? - GDPR.eu. https://gdpr.eu/what-is-gdpr/

[11] Grayson, N. R. (2023). Cybersecurity Framework Profile for Electric Vehicle Extreme Fast Charging Infrastructure. https://doi.org/10.6028/nist.ir.8473

[12] Hamberg-van Reenen, H. H., Proper, K. I., & van den Berg, M. (2012, August 3). Worksite mental health interventions: a systematic review of economic evaluations. Occupational and Environmental Medicine, 69(11), 837–845. https://doi.org/10.1136/oemed-2012-100668

[13] Hasal, M., Nowaková, J., Ahmed Saghair, K., Abdulla, H., Snášel, V., & Ogiela, L. (2021, June 3). Chatbots: Security, privacy, data protection, and social aspects. Concurrency and Computation: Practice and Experience, 33(19). https://doi.org/10.1002/cpe.6426

[14] Kaspersky. (2023, April 19). Chatbots are everywhere, but do they pose privacy concerns? www.kaspersky.com. https://www.kaspersky.com/resource-center/preemptive-safety/chatbots

[15] National Institute for Health and Care Excellence. (2022, March 2). Recommendations | Mental wellbeing at work | Guidance | NICE. https://www.nice.org.uk/guidance/ng212/chapter/Recommendations

[16] Nicole Sette, J. C. (2023, March 23). Emerging Chatbot Security Concerns | Kroll. https://www.kroll.com/en/insights/publications/cyber/emerging-chatbot-security-concerns

[17] Sebastian, G. (2023). Privacy and Data Protection in ChatGPT and Other AI Chatbots: Strategies for Securing User Information. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4454761

[18] Tariq, U., Ahmed, I., Bashir, A. K., & Shaukat, K. (2023, April 19). A Critical Cybersecurity Analysis and Future Research Directions for the Internet of Things: A Comprehensive Review. Sensors, 23(8), 4117. https://doi.org/10.3390/s23084117

[19] U.S. Health Resources & Services Administration. (2019, August 2). Guide to Privacy and Security of Health Information. https://www.hrsa.gov/behavioral-health/guide-privacy-and-security-health-information

[20] United States Department of Health and Human Services (HHS). (2009, November 20). Summary of the HIPAA Security Rule. HHS.gov. https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html

[21] World Health Organization. (2016, April 13). Investing in treatment for depression and anxiety leads to fourfold return. https://www.who.int/news/item/13-04-2016-investing-in-treatment-for-depression-and-anxiety-leads-to-fourfold-return

[22] World Health Organization. (2022, September 28). Guidelines on mental health at work. https://www.who.int/publications/i/item/9789240053052

[23] Zagorski, N. (2022, May 1). Popularity of Mental Health Chatbots Grows. Psychiatric News, 57(5). https://doi.org/10.1176/appi.pn.2022.05.4.50

[24] Goetzel, R. Z., Roemer, E. C., Holingue, C., Fallin, M. D., McCleary, K., Eaton, W., Agnew, J., Azocar, F., Ballard, D., Bartlett, J., Braga, M., Conway, H., Crighton, K. A., Frank, R., Jinnett, K., Keller-Greene, D., Rauch, S. M., Safeer, R., Saporito, D., . . . Mattingly, C. R. (2018, April). Mental Health in the Workplace. Journal of Occupational & Environmental Medicine, 60(4), 322–330. https://doi.org/10.1097/jom.0000000000001271

[25] The White House. (2023, November 22). Blueprint for an AI Bill of Rights | OSTP | The White House. Retrieved December 10, 2023, from https://www.whitehouse.gov/ostp/ai-bill-of-rights

[26] European Commission. (2023, December 15). The Digital Services Act package. Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

[27] European Commission. (2023, December 14). European Data Governance Act. Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

[28] House, W. (2023, October 30). FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. The White House. https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/