

New breakthroughs in AI Chatbots and their potential in mental health services

Tin Galijašević
Varaždin General Hospital
Zagreb, Croatia
kopaka@oems.hr (0000-0002-68893170)

Maja Škarić
Neuropsychiatric Hospital "Dr. Ivan
Barbot
Zagreb, Croatia
skaric.maja@gmail.com

Eva Podolski
School of Medicine, University of
Zagreb
Zagreb, Croatia
evapodolski3@gmail.com

Filip Mustać
Department of Psychiatry and
Psychological Medicine
University Hospital Centre Zagreb
Zagreb, Croatia
0000-0003-2851-6183

Martina Matovinović
Department of Internal Medicine,
Division of Endocrinology
University Hospital Centre Zagreb
Zagreb, Croatia
0000-0002-6325-7394

Darko Marčinko
Department of Psychiatry and
Psychological Medicine
University Hospital Centre Zagreb;
School of Medicine, University of
Zagreb

Zagreb Croatia 0000-0001-7688-464

Abstract—Since the COVID-19 pandemic the impact of mental disorders on daily functioning and well-being on the general population and the rising demand for mental health support have been in the limelight. The recent surge in new breakthroughs in artificial intelligence technology has generated interest in the potential to utilize AI in mental health services. New chatbot technologies can utilize these new generations of AI to provide special benefits including accessibility and anonymity and can overcome traditional barriers to therapy. Chatbots facilitate continuous data collection which could enable personalized treatment approaches and improve diagnostic accuracy. While this is promising, the need for human oversight and the importance of ethical data collection and privacy remain an issue. The implementation of AI chatbots into mental health services presents an opportunity to greatly enhance accessibility but careful consideration of their role and limitations is necessary for optimal integration in clinical practice.

Keywords—Artificial intelligence, Machine learning, Psychiatry, Mental Health

I. INTRODUCTION

A mental disorder manifests as a notable disruption in an individual's cognition, emotional regulation or behaviour typically leading to impairment in significant areas of functioning such as school, work and everyday life. It includes a wide range of mental health conditions, with anxiety, depression, bipolar and post-traumatic stress disorder being amongst the most common ones. According to the Institute of Health Metrics and Evaluation 1 in every 8 people around the world lived with mental disorder in 2019. Unfortunately, during Covid-19 pandemic number of people living with anxiety or major depressive disorders increased from 26 % to 28 %. [1] Various individual, familial, communal, and structural elements can simultaneously influence mental wellbeing positively or negatively. While many individuals demonstrate resilience, those confronted with adverse conditions like poverty, violence, disability, and inequality face heightened risks. Protective and risk factors encompass personal psychological and biological traits, such as emotional capabilities and genetic predispositions. Changes in brain structure and/or function often influence these factors. Numerous resources are available to enhance the journey towards mental wellness, including medication, counseling, social support, and education. Counseling encompasses a

range of techniques, from acquiring relaxation skills to restructuring cognitive patterns. Additionally, the acceptance, encouragement, and support from friends, family, and peers play pivotal roles. Education on managing mental health conditions alongside other medical concerns equips individuals with valuable skills and assistance in achieving overall recovery and wellness. Collaborating with a treatment team facilitates the development of a comprehensive recovery plan tailored to individual needs. This plan may incorporate counseling, medication, support groups, educational programs, and other effective strategies, ensuring a holistic approach to wellness. However, nowadays, with the advancement of artificial intelligence and technology, new methods of providing support and assistance to individuals with mental illness are emerging. In this paper, we will examine the history of the development of chatbots, their current potential in clinical practice, their potential issues and shortfalls and share our opinion where further research should be focused on.

II. DEVELOPMENT OF CHATBOTS

Chatbots are computer programs designed to engage in text-based or speech-based conversations with users through interactive interfaces. Chatbots operate fully automatically, interacting with users without human intervention. In addition to their role in customer communication for online shopping, chatbots are increasingly utilized in education and the gaming industry. They serve as a unique human-machine interface, granting users chat-based access to application functions and data. With the ability to launch multiple instances effortlessly, chatbots offer high flexibility, enabling simultaneous conversations with numerous users. This flexibility frees up resources for more complex tasks, making them particularly valuable in various economic domains. The roots of human-computer interaction technology date back to academia half a century ago. In 1964, ELIZA, a programmable natural language processing program developed by Joseph Weizenbaum at the MIT Artificial Intelligence laboratory, emerged as a pioneering example. [2] Acting as a person-centered psychotherapist, ELIZA lacked the ability to comprehend conversation content, yet many users perceived it as intelligent and even formed emotional connections with it. Fast forward fifty years, and the technology that once powered ELIZA is now ubiquitous on smartphones and smart home devices worldwide. The potential of chatbots in mental health is already emerging. The rapid advancement of technology has opened up new

possibilities for healthcare, including clinical psychology and psychotherapy. Research interest in utilizing chatbots for mental health support is on the rise, evidenced by the growing number of studies and online services offered by healthcare providers. They are being utilized in suicide prevention, cognitive-behavioral therapy, and tailored to specific populations such as veterans. [2] Chatbots offer a promising avenue for individuals hesitant to confide in human therapists, potentially improving access to mental health treatment and effectiveness for those reluctant to engage with traditional therapy. Despite considering them "the future of therapy," psychiatric chatbots lack a clear definition and role in clinical practice. In this article we aim to explore the potential for integrating chatbots in clinical practice and present the current understanding of their user base and therapeutic effects in light of recent technological developments.

III. POTENTIAL OF CHATBOTS

Chatbots represent a significant advancement in mental health care, offering numerous benefits across various psychological domains. According to many authors they provide accessible and immediate support, overcoming geographical and financial barriers that often impede traditional therapy access [3], [4], [5], [6]. Studies have recognized that the need for such support to existing mental health service emerged during COVID 19 pandemic due to the shortage of human mental health advisers [4], [7], [8]. A chatbot, powered by artificial intelligence, has endless amounts of time and patience, never forgets what a patient has said [9], can deliver personalized mental health interventions with consistency, anonymity [10] and privacy, judgement free [5] ensuring users can seek help without inconvenience [11]. A study by Fiske found out that using chatbot will reduce embarrassment in asking for specific information or services or feelings of shame when admitting noncompliance with a treatment plan, leading to improved trust and openness between patients and the medical system. Users can self-direct the pace of information- especially important for low-literacy patients [9]. Fiske et al. stated that AI interventions might be particularly well placed for populations living in resourcepoor settings, for reaching high-risk groups such as veterans, or for those who are concerned about the social stigma associated with psychotherapy, patients may respond positively to the fact that the counterpart is not a human therapist[9]. According to Pham, they can also provide support for people with communication difficulties[12]. Haque and Brunn stated in their studies that chatbots offer social and psychological support in situations where realworld human interaction, such as connecting to friends or family members or seeking professional support, is not preferred or possible to achieve [5], [6]. Evidence from Ahmed and Pham suggests chatbots are particularly effective in administering cognitive support using cognitivebehavioural therapy (CBT) principles and exercises [13], behavioural modification, social motivation, attention enhancement, and biofeedback [12]. They are utilized to assist with making psychiatric diagnoses, symptom tracking, prediction of acute disease exacerbations and recovery, and psychoeducation [12]. According to Ray, machine learning methods have shown good diagnostic ability in case of dementia, in differentiation of adults with ADHD from healthy controls, in prediction of depressive symptoms in post-partum. During the COVID times, machine learning methods have been used to detect burn out in health care workers [14]. According to many authors, they utilize interactive messaging to guide users through therapeutic

exercises, deliver emotional wellness coping strategies [15], mindfulness practices, coach individuals through times of emotional distress and encourage mood tracking [12], fostering self-care and awareness, in some cases, they can even connect users to mental health professionals [5]. Besides, Sabour stated that emotional support received from chatbot, in form of chatting partner that can effectively listen to their problems and provide a channel for them to vent their emotions has shown improvement in depressive symptoms, negative affect and insomnia [16]. A review from Ahmed found out there are chatbots with unique feature acting as the one having a mental illness and looking for help. With this technique of the chatbot being the care-receiver it allows the user to self-reflect and reach deeper level of understanding their own vulnerable experience [13]. According to many sources there are numerous conditions and disorders where usage of chatbots have shown symptoms relief. Depression and anxiety being the most common mental disorder around the world take large number of chatbots focusing on their symptom's management [4], [13], [14], [17]. Another potential target for chatbots are patients with schizophrenia, dementia, phobic disorders, addiction, eating disorders and sexual disorders [8], [9], [14], [17], or public speaking anxiety [13]. They even provide improvements in chronic pain management [18] or mental distress in general [16]. A study by Ray found that in geriatric psychiatry- 'companion bot' assist in helping elderly or depressed patients by providing companionship and interaction in aim to reduce loneliness, support older people in maintaining their independence and daily activities reminding them on time to wake up, go to bed, or eat [14]. Authors like Abd-Alrazaq, Fiske and Pham said that chatbots are reported to be effective for patients with autism spectrum disorders through education and therapy, they can teach children social skills and help them with facial recognition and appropriate gaze response [9], [12], [17]. Pham and Ray mentioned another use in form of, so called, avatar therapy are faces that interact with patients with schizophrenia - they challenge the persecutory voice hallucinations they experience and gradually learn to gain control over the distressing voices [12], [14]. According to Haque chatbots are effective in dealing with obesity with interventions for increasing physical activity, achieving relevant weight loss, and improving diet by sending daily check-in reminders [5]. Brunn stated that the integration of chatbots into mental health care also facilitates continuous data collection, enhancing treatment personalization and effectiveness. AI can assist in identifying novel biomarkers which may allow to better identify variation in phenotypes and targeted treatment options [6]. We highlight the anticipated benefits of chatbot implementation from the findings of several researchers in Fig 1.

TABLE I.

Study	Anticipated Benefits	
	<i>Targeted concerns</i>	<i>Structural benefits</i>
Fiske	Depression, anxiety, autism, schizophrenia	Patients experiencing shame, low income areas
Pham	Schizophrenia, children with autism	Cognitive behavioral therapy automation
Haque	Stress, anxiety, depression	Ease of access, judgment-free zone

Brunn	Prediction of psychosis, depression and anxiety	Good for regions with low access to medical care, data collection and analysis
Ahmed	Depression and anxiety	Self-reflective therapy, cognitive behavioral therapy integration
Ray	Psychotic disorders, geriatric patients, sleep disorders, addiction	Low costs, potential method of mental health services in rural and developing countries

Fig. 1. Anticipated benefits in targeting clinical concerns and structural benefits of chatbot implementation

Through machine learning algorithms, chatbots can adapt to individual user responses and progress, tailoring interventions to better meet specific needs over time. However, it's important to acknowledge their limitations, including the need for human oversight in complex cases and ensuring ethical use of user data. According to Jungmann, as diagnostic medium, chatbots are not substitutes for doctors or psychotherapists, they are rather a complementary tool in mental health care, low-cost, low-threshold, and time-efficient support in the diagnosis of mental disorders in adulthood, with needed improvements in field of diagnosis mental disorders in childhood [19].

IV. ISSUES AND SHORTFALLS OF CHATBOTS

Several research groups highlight the need for rigorous consideration before implementing chatbots in clinical practice. Some of the concerns are data security and confidentiality given the private health information that is collected by AI. The lack of professional training clinicians have in AI usage is also a major factor that limits chatbot implementation. Also, the complexity and opacity of AI algorithms is an issue as this limit doctor's ability to understand and review AI behaviour and predictions [9], [20]. While today, especially in psychodynamic type psychotherapy, transference and countertransference and their interpretation are most often relied on as an indispensable part of the psychotherapy process, chatbots probably have less potential to deal with this, but probably over time they will be able to provide support, as well as monitor objective parameters such as previously described the use of mobile applications in people with eating disorders [21]. Fiske et al also highlight the potential in some healthcare systems to justify replacing established services which would increase healthcare disparities. They also note the lack of up-to-date legal and ethical frameworks that could address rapidly advancing technology like AI in an appropriate timeframe - for example it is unclear if AI would have to be legally responsible to inform healthcare providers or authorities if a patient is deemed a threat to themselves or others [9].

McCradden et al. highlight how while well-intentioned, AI implementation may lead to unintended consequences. AI is often perceived as an „expert“ system, potentially overshadowing the expertise of clinicians. This elevation of AI's predictions may inadvertently diminish the importance of clinical judgement and patient experiences. Blind acceptance of AI's superiority risks perpetuating historical inequities faced by individuals with mental illnesses [22]. Monteith et al. note that AI technology remains immature for widespread implementation in other medical fields, for example while over 150 AI products have been approved for usage, current reviews indicate limited evidence of

significantly improved diagnostic outcomes. They also note possible biases which could arise from inaccurate training data which would skew the AIs ability to detect certain pathologies. While Physician attitudes toward AI are generally positive, yet concerns about ethical, legal, and educational gaps persist. Surveys highlight a lack of formal training in AI among physicians, underscoring the need for comprehensive education to foster critical assessment skills. They note automation bias and deskilling of medical professionals where they could over time start to excessively rely on AI generated recommendations [20]. Addressing these problems requires concerted efforts between clinicians and developers to improve data quality, mitigate bias, and enhance physician education in AI. Only through such measures can AI be safely integrated into clinical practice, empowering physicians to leverage its benefits while ensuring patient safety and quality of care.

V. CONCLUSION AND FUTURE WORK

In conclusion, chatbots show great potential as a transformative innovation in the domain of mental health care, with their accessibility, anonymity and end-user oriented design being highlights. Chatbots possess an unparalleled capacity for personalized interventions in many different clinical scenarios ranging from anxiety and depression, to dementia, schizophrenia and others. Their ability to collect user data could be of great use to clinicians by providing them with extensive symptom tracking and alerting them to patients that are in a potential crisis. However, chatbots also face problems with adequate implementation as most clinicians currently lack professional training and extensive knowledge of AI algorithms to safely judge and use the data chatbots could provide. Concerns regarding data security and a lack of a clear legal framework are also major hurdles that need to be overcome. We urge lawmakers and regulatory bodies to recognize the advent of novel AI technologies and create a modern legal framework to allow further research and protect patient's privacy and safety. We believe future work and efforts on this topic should focus on refining and clarifying chatbot algorithms and large scale longitudinal studies to research what clinical entities would be the best target for widespread implementation of this technology and how to adapt this technology to target specific patient populations like young adults, veterans or the elderly. By prioritizing these areas researchers can provide invaluable data to chatbot developers to create even better algorithms which would ultimately allow healthcare providers to improve clinical outcomes and face current challenges in mental healthcare.

ACKNOWLEDGMENT

Conceptualization and design of article: T.G., F.M.; literature search, writing— original draft preparation: F.M., T.G., E.P. and M.Š. supervision: M.M. and D.M.; writing— review and editing: all authors. All authors have read and agreed to the published version of the manuscript.

REFERENCES

- [1] World Health Organization, "Mental Health and COVID-19: Early evidence of the pandemic's impact," World Health Organization, 2022.
Accessed: Mar. 06, 2024. [Online]. Available: <https://www.jstor.org/stable/resrep44578>
- [2] A. N. Vaidyam, H. Wisniewski, J. D. Halamka, M. S. Kashavan, and J. B. Torous, "Chatbots and Conversational Agents in Mental Health: A

- Review of the Psychiatric Landscape,” *Can J Psychiatry*, vol. 64, no. 7, pp. 456–464, Jul. 2019, doi: 10.1177/0706743719828977.
- [3] S. Suharwardy et al., “Feasibility and impact of a mental health chatbot on postpartum mental health: a randomized controlled trial,” *AJOG Glob Rep*, vol. 3, no. 3, p. 100165, Mar. 2023, doi: 10.1016/j.xagr.2023.100165.
 - [4] Y. He et al., “Mental Health Chatbot for Young Adults With Depressive Symptoms During the COVID-19 Pandemic: Single-Blind, Three-Arm Randomized Controlled Trial,” *J Med Internet Res*, vol. 24, no. 11, p. e40719, Nov. 2022, doi: 10.2196/40719.
 - [5] M. D. R. Haque and S. Rubya, “An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews,” *JMIR Mhealth Uhealth*, vol. 11, p. e44838, May 2023, doi: 10.2196/44838.
 - [6] M. Brunn, A. Diefenbacher, P. Courtet, and W. Genieys, “The Future is Knocking: How Artificial Intelligence Will Fundamentally Change Psychiatry,” *Academic Psychiatry*, vol. 44, no. 4, pp. 461–466, Aug. 2020, doi: 10.1007/s40596-020-01243-8.
 - [7] Y. Zhu, R. Wang, and C. Pu, “‘I am chatbot, your virtual mental health adviser.’ What drives citizens’ satisfaction and continuance intention toward mental health chatbots during the COVID-19 pandemic? An empirical study in China,” *Digit Health*, vol. 8, p. 20552076221090031, Mar. 2022, doi: 10.1177/20552076221090031.
 - [8] J. Sun et al., “Artificial intelligence in psychiatry research, diagnosis, and therapy,” *Asian Journal of Psychiatry*, vol. 87, p. 103705, Sep. 2023, doi: 10.1016/j.ajp.2023.103705.
 - [9] A. Fiske, P. Henningsen, and A. Buyx, “Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy,” *Journal of Medical Internet Research*, vol. 21, no. 5, May 2019, doi: 10.2196/13216.
 - [10] I. Hungerbuehler, K. Daley, K. Cavanagh, H. Garcia Claro, and M. Kapps, “Chatbot-Based Assessment of Employees’ Mental Health: Design Process and Pilot Implementation,” *JMIR Form Res*, vol. 5, no. 4, p. e21678, Apr. 2021, doi: 10.2196/21678.
 - [11] G. Sanabria et al., “‘A Great Way to Start the Conversation’: Evidence for the Use of an Adolescent Mental Health Chatbot Navigator for Youth at Risk of HIV and Other STIs,” *J Technol Behav Sci*, pp. 1–10, May 2023, doi: 10.1007/s41347-023-00315-4.
 - [12] K. T. Pham, A. Nabizadeh, and S. Selek, “Artificial Intelligence and Chatbots in Psychiatry,” *Psychiatr Q*, vol. 93, no. 1, pp. 249–253, 2022, doi: 10.1007/s11126-022-09973-8.
 - [13] A. Ahmed et al., “Chatbot features for anxiety and depression: A scoping review,” *Health Informatics J*, vol. 29, no. 1, p. 14604582221146719, 2023, doi: 10.1177/14604582221146719.
 - [14] A. Ray, A. Bhardwaj, Y. K. Malik, S. Singh, and R. Gupta, “Artificial intelligence and Psychiatry: An overview,” *Asian J Psychiatr*, vol. 70, p. 103021, Apr. 2022, doi: 10.1016/j.ajp.2022.103021.
 - [15] C. Grové, “Co-developing a Mental Health and Wellbeing Chatbot With and for Young People,” *Front Psychiatry*, vol. 11, p. 606041, Feb. 2021, doi: 10.3389/fpsy.2020.606041.
 - [16] S. Sabour et al., “A chatbot for mental health support: exploring the impact of Emohaa on reducing mental distress in China,” *Front Digit Health*, vol. 5, p. 1133987, May 2023, doi: 10.3389/fgth.2023.1133987.
 - [17] A. A. Abd-Alrazaq, M. Alajlani, A. Abdallah Alalwan, B. M. Bewick, P. Gardner, and M. Househ, “An overview of the features of chatbots in mental health: A scoping review,” *International Journal of Medical Informatics*, vol. 132, Dec. 2019, Accessed: Mar. 06, 2024. [Online]. Available: <https://eprints.whiterose.ac.uk/151992/>
 - [18] J. D. Piette et al., “Patient-Centered Pain Care Using Artificial Intelligence and Mobile Health Tools: A Randomized Comparative Effectiveness Trial,” *JAMA Internal Medicine*, vol. 182, no. 9, p. 975, Sep. 2022, doi: 10.1001/jamainternmed.2022.3178.
 - [19] S. M. Jungmann, T. Klan, S. Kuhn, and F. Jungmann, “Accuracy of a Chatbot (Ada) in the Diagnosis of Mental Disorders: Comparative Case Study With Lay and Expert Users,” *JMIR Form Res*, vol. 3, no. 4, p. e13863, Oct. 2019, doi: 10.2196/13863.
 - [20] S. Monteith, T. Glenn, J. Geddes, P. C. Whybrow, E. Achtyes, and M. Bauer, “Expectations for Artificial Intelligence (AI) in Psychiatry,” *Curr Psychiatry Rep*, vol. 24, no. 11, pp. 709–721, 2022, doi: 10.1007/s11920-022-01378-5.
 - [21] F. Mustač, T. Galijašević, M. Matovinović, and D. Marčinko, “Smartphone app based psychological interventions for patients with eating disorders,” in *2023 8th International Conference on Smart and Sustainable Technologies (SpliTech)*, Jun. 2023, pp. 1–4, doi: 10.23919/SpliTech58164.2023.10193373.
 - [22] M. McCradden, K. Hui, and D. Z. Buchman, “Evidence, ethics and the promise of artificial intelligence in psychiatry,” *J Med Ethics*, vol. 49, no. 8, pp. 573–579, Aug. 2023, doi: 10.1136/jme-2022-108447.