

Use of ChatGPT and Other AI Technologies in Mental Health Care

Xinyi Huang

Received January 04, 2024

Accepted August 09, 2024

Electronic access October 15, 2024

Since the official launch of ChatGPT in 2022, it has quickly gained attention and brought the topic of artificial intelligence to the forefront. Today, AI is gradually impacting the human lifestyle and is expected to find applications in various fields in the future. One of these promising areas is the use of AI for mental health support, where today's AI technologies have the potential for therapeutic applications in mental health services and will lead to an increasing number of possibilities. This paper explores the possibilities of AI, including ChatGPT, in psychotherapy. It reviews the existing literature on AI in mental health care, examines its therapeutic approaches, assesses the current quality of AI counseling services, and surveys the perspectives and experiences of patients and subjects. In addition, the paper discusses the ethical issues raised by digital mental health and provides recommendations for the associated societal challenges resulting from AI implementation.

Introduction

Positive mental health significantly influences how individuals view their behavior, thoughts, and overall existence. Developing a healthy mindset helps to manage stress, have healthy relationships, and make informed decisions. Mental health, however, is a major concern in contemporary society. In the United States, alarming statistics highlight this point: Mental illness is one of the most prevalent health conditions, affecting more than one in five American adults¹. The World Health Organization's 2022 report shows a 25% spike in anxiety and depression globally in the first year of the 2019 coronavirus disease pandemic². This growing problem has forced many people to focus on mental health issues and desire to find an antidote for this condition.

Addressing mental health issues consists of two main methods: psychotherapy and pharmacotherapy³. In pharmacotherapy, the drug does not completely cure the mental illness but it helps to control the symptoms, so it is usually used in conjunction with psychotherapy. Psychotherapy is a type of treatment that can help people who are experiencing mental health issues. Moreover, psychotherapy can also help to identify the psychological root cause of certain conditions so that the appropriate medication can be prescribed to enable the patient to recover more rapidly. However, people may be turning to new technology, specifically artificial intelligence.

In recent years, the widespread integration of artificial intelligence in various industries has brought about a lifestyle change. The term 'artificial intelligence' was originally coined by computer scientist John McCarthy and is defined as "the science and engineering of making intelligent machines." AI is usually a digital computer that can perform tasks associated with intelligent beings. Training AI usually requires a large amount of relevant data, which is analyzed to form an algorithmic model. In many

areas involving repetitive, detail-oriented tasks, AI can perform even better than humans. Since the development of digital computers in the 1950s, AI has been the subject of much controversy. Proponents believe that AI can perform some tasks better than humans and can improve the quality of human life, while opponents believe that the development of AI carries many risks. The growing popularity of AI has led to the gradual automation of certain jobs, which has led to the replacement of some labor force, especially low-paid workers. One of the fastest-growing and most widely known AI in the market in recent years would be Chatbot, a conversational computer program that can simulate a human conversation by using artificial intelligence and natural language processing (NLP) to understand user questions and respond to those questions automatically⁴.

Among the popular chatbots is ChatGPT, an AI chatbot developed by artificial intelligence research firm OpenAI, which can engage in conversations with users and answer their questions with a high degree of competency. After its launch in November 2022, it attracted more than one million users in just four days. This reflects the tremendous response ChatGPT has received; however, the user community has received mixed reviews of this product and some concerns have been raised about it.

The field of mental health care, similar to other fields, has also been affected by the revolution in digital technology and artificial intelligence. Conversational AI such as ChatGPT and chatbots lead the way, which begs the question of whether AI can replace a counselor's job of conversing with patients. As technology revolutionizes, such conversational AI is indeed reshaping the landscape of mental health care delivery. By leveraging existing diagnostic data as well as analyzing clinician behavior, conversational AI has the potential to transform the traditional approach to psychotherapy.

The present review of literature includes an evaluation of the

general implication of ChatGPT and other AI technologies in mental health-related therapeutic settings. The main elements of this literature include specific chatbots such as ChatGPT's impact on the psychotherapy field, the quality of therapy that artificial chatbot therapies can provide, and the impact of therapy with a chatbot on patient trust and willingness to disclose. The evaluation of these scientific papers is based on a broader question: How can ChatGPT and other AI technologies be used in a therapeutic setting?

Literature Review

Overview of Psychotherapy and AI

Enabling AI to provide a variety of valuable applications in the field of psychotherapy is exactly what it seems to mean; combining AI technology with existing psychological clinical treatment experience to provide users with services in mental health treatment. There are three possible ways in which AI and traditional therapy can be integrated:

1. Human therapist, without AI

In today's traditional psychotherapy, mental health professionals make a diagnosis by gathering information through psychological testing or questionnaires about the patient through a detailed interview. This type of therapy is generally known as conversational therapy or talk therapy. There is a wide range of sessional therapies, the most commonly used of which is Cognitive Behavioral Therapy (CBT), which focuses on identifying and changing negative thinking and thoughts that lead to emotional distress and can help to alleviate mental health disorders such as depression, Post Traumatic Stress Disorder (PTSD) or eating disorders¹.

In addition to this, the counselor provides a secure and comfortable environment for the patients to freely express what is on their minds and release their emotions. The counselor engages in a conversation with the patient by listening to the patient's experience in describing his symptoms to establish a doctor-patient relationship. Finally, the counselor and the patient work together to develop a plan for improvement to achieve the treatment goals set by both parties. Often, after receiving counseling treatment, the patients can gain a deeper understanding of themselves and then use the counselors' advice to improve their problems and become more optimistic about life. The process of psychological counseling relies heavily on the cooperation between the doctor and the patient, with the patient needing a private space to find someone to talk to, and the counselor giving options and suggestions to improve the symptoms.

2. AI therapist, No Human

The development of artificial intelligence has been applied in the field of psychotherapy, also known as AI-driven therapy. AI therapies provides mental health support to a wider range of population, including those who are unable to access traditional treatments for geographic, economic, or social reasons. It can adapt interventions to each person's unique needs and preferences, providing users with a more personalized therapy experience. Its greatest advantage is that it can uplift users throughout the day, keeping an eye on the dynamics of the patient's symptoms at all times. However, AI therapists still have many limitations as they are only in their infancy. Firstly, there may be biases in their training data, leading to possible differences in the treatment recommendations received by different groups. Secondly, regulations and ethical guidelines for AI therapies are still evolving, so privacy policies are still lacking.

A Historical Example: This one-on-one dialog model is reminiscent of the artificial intelligence chat software that has recently become a trend on the Internet. One of the earliest chatbots was developed in 1965 by Rich Alpert; the bot was called ELIZA, a very basic psychotherapist chatbot created to explore communication between humans and machines. It runs a script called DOCTOR, which simulates a psychotherapist of the Rogerian school, which reflects the words of a patient back to the patient's words and responds to questions about the user's input using the rules laid out in the script. The chatbot was one of the first programs to be able to attempt the Turing test, though it only gave the user the illusion of being understood by the program, and did not answer and solve the questions posed by the user. Although ELIZA is not able to have a mutually understandable conversation with the user, it serves as a milestone on the path of AI development that exemplifies how humans have been using AI in conjunction with psychology since the last century. Over time, the field of artificial intelligence has been innovating and evolving until 2022, a big year for AI development⁵.

B Scientific/new Technology Examples: The official launch of the Chat Generative Pre-trained Transformer (ChatGPT) on November 30, 2022, has also brought this Chatbot to the forefront of the conversation. The first GPT model was introduced in 2018 by OpenAI, which is based on the Transformer architecture to improve language understanding through training and then generate coherent text. The 2022 launch of ChatGPT mainly consists of GPT-3.5 and GPT-4, compared to GPT-4 which is "more reliable than GPT-3.5, more creative, and capable of handling more detailed instructions"⁶. Because of the ChatGPT algorithm and its conversational principle of operation, it holds great promise in the medical field. One

example of the use of AI in conjunction with healthcare is the announcement by Microsoft and Epic Systems (one of the largest healthcare software companies in the United States) that they are bringing OpenAI's GPT-4 AI language model to healthcare⁷. Microsoft and Epic expand strategic collaboration with integration of Azure OpenAI Service - Stories. Microsoft News. Retrieved January 3, 2024, from <https://news.microsoft.com/2023/04/17/microsoft-and-epic-expand-strategic-collaboration-with-integration-of-azure-openai-service/>). The purpose of this collaborative research is to use AI for drafting responses to messages sent by healthcare professionals to patients, as well as analyzing medical records while looking for trends. However, in the study, GPT-4 was found to add falsified details to patient notes. This shows that the current ChatGPT has many limitations. Since it operates on the principle of using a language model to generate text, it gives answers to questions that seem to make sense but are wrong⁸. While ChatGPT has an understanding of mental health issues, it does not have a completely accurate solution for these disorders, thus showing that the technology currently needs further improvement and may not be reliable when used in healthcare. Despite current unreliability, when exploring whether ChatGPT and other AI could be viable as personalized therapists in the future, the answer is presumably yes. AI has been slower to develop in the mental health field compared to other medical fields because this field requires practitioners to be more hands-on in their clinical practice to develop a deep emotional bond with patients to observe their emotions to manage their condition.

C Consumer Example: There are also examples of AI being used in the field of psychotherapy in the current market. Woebot is an automated conversational agent chatbot, founded in 2017 by clinical research psychologist Dr. Alison Darcy, that helps users monitor their emotions and learn about themselves. Based on Cognitive Behavioral Therapy (CBT), Interpersonal Psychotherapy (IPT), and Dialectical Behavioral Therapy (DBT), Woebot improves existing emotional problems by using AI and Natural Language Processing to communicate with users and provide them with psychological advice⁹. Studies have shown that Woebot can build a bond of trust with its users within 3 to 5 days¹⁰. When users feel they need support, they can simply open the software Woebot on their phone to help them cope with their daily stress and anxiety. By chatting with the user, it examines the user's daily moods and plots them on a graph to make it easier for the user to understand his or her mood swings, as well as conducting more extended interviews with the user every two weeks. During the program, Woebot develops new treatments based on

previous experiences and conversations with other users, while maintaining complete privacy and security. This AI psychotherapist has the advantage of solving users' problems online 24/7, in addition to temporarily solving the problem of shortage of psychological clinicians and also enabling patients who cannot afford to pay for psychotherapy to get free counseling. However, Woebot does not completely solve the problems that AI therapists have to face. Some users have responded that sometimes Woebot doesn't allow the user to converse freely, but only selects the corresponding option. For example, it will ask a question and the options might be "Awesome!", "Yes" and "Okay", but not "No". This is very frustrating for the user because it doesn't fully reveal what they are thinking².

D Influencing Factors Among Generations: A study have shown that the acceptance of AI-powered mental health virtual assistants also varies widely across generational groups, including Generation X, Generation Y, and Generation Z. The findings suggest that "Gen Y generally having more positive perceptions and stronger intentions, Gen Z closely aligned with them, and Gen X demonstrating." This emphasizes the generational differences in attitudes towards innovative technologies and demonstrates the need to consider different generational preferences for the better development of AI in the psychological field³.

3. Human Therapist-AI Integration

In addition to the use of language models in the applied conversational AI mentioned in the first two examples, big data computational methods are also an important part of making AI applications in psychotherapy. Data statistics in mental health could benefit greatly from AI technology, provided that diagnostic models more suitable for psychotherapy are developed. From the current level of development in AI, data statistics models can be used to collect patient data and a large amount of medical literature, as well as real cases from the past, to summarize the basic linguistic structure of psychotherapy and the corresponding solutions for mental illness¹¹.

Given that psychotherapy is a conversation between the patient and the therapist, the specialties of the therapeutic approach should be found in the utterances used during the therapy. In order to ensure that the data collection process complies with the relevant laws and regulations and privacy policies, researchers collected data on the conversations between patients and therapists in past psychotherapy programs and extracted the emotions that corresponded to the high-frequency questions posed by the therapist and answered by the patient in these conversations. The past cases are digitized and this information is used to collect, analyze, and respond to the data used to "train" the AI. The resulting model can recognize specific content related to the therapy and can provide solutions to the therapist. However, this

method of data collection is not yet perfect. First, it is necessary to ensure the privacy of the patient and to check whether the patient will agree to give his consent to the treatment process to the AI for data training. Secondly, the interaction between the patient and the therapist in the psychotherapy process contains many uncertainties, and the data recorded by the AI cannot include all aspects, so the ability to adapt to the situation during the treatment process is also a problem that AI needs to face in psychotherapy.

As artificial intelligence becomes more prevalent in a therapeutic mental health setting, it is important to understand the psychology of the interaction between human patients and artificial therapists. In addition, it is essential to determine the quality of care provided by AI-only therapists, confirming whether they can improve patient's symptoms and provide accurate treatment options, like CBT. The most important of all is to confirm whether people are comfortable disclosing personal information to these chatbots and to ensure that patients are able to provide consistent information to the AI versus a human therapist, or will have reservations.

Quality of Psychotherapy with Chatbot Therapists

In order for AI to successfully be integrated with current psychotherapy methods, one of two things needs to occur: (1) AI needs to improve outcomes over current therapeutic methods (2) AI therapists need to make users trust them and complete services more efficiently and easily.

1. **What is the current benchmark for success in psychotherapy?** To ensure the development of AI in the field of psychotherapy, it is important to ensure the quality of AI psychotherapy and the user experience. If AI psychotherapy produces very different results from current clinical psychotherapy, or if users do not feel any improvement after experiencing it, then there is a lot of uncertainty and it is not ready to be marketed and utilized by users. Talking therapies in clinical psychotherapy date back to the late 19th and early 20th centuries and have been practiced for over 100 years. Hundreds of clinical trials have been conducted in various forms of talk therapy, and the results of many studies demonstrate that psychotherapy is effective. In one of the most sophisticated statistical analyses on the subject to date, published by psychologists Mary Lee Smith and Gene V. Glass, some 400 studies were reviewed and it was found that among psychiatric patients who received various types of talk therapy, the typical patient outperformed 75 percent of those with similar diagnoses who did not receive treatment. who had a similar diagnosis but did not receive treatment. Now that AI is being integrated with mental health specialty care, and products have been introduced and tested in the marketplace, there is little evidence to support the quality and accuracy of AI psychotherapy. To

test its quality, practice is the best method. After searching through the relevant literature, two of them were found to provide an overview of two relevant trials that have been conducted.

2. What level of success have current AI approaches achieved?

A Woebot Trial 1 In a randomized controlled trial, artificial intelligence for conversationalist psychotherapy was used to provide cognitive-behavioral therapy to young adults who self-identified as having symptoms of depression and anxiety¹⁰. The AI therapeutic program used in this trial was the aforementioned Woebot. To ensure randomization of the trial, the researchers recruited 70 individuals aged 18-28 years from a university community social media site and randomized them to a maximum of 20 controlled sessions over 2 weeks. Participants were divided into two groups; one group used Woebot to engage in a conversational agent interaction using CBT principles, during which users interacted with Woebot and answered questions about their problems. The other portion served as an information-only control group directly using the National Institute of Mental Health eBook, *Depression in College Students*, which provides comprehensive evidence-based information about depression in college students as well as answers to frequently asked questions. All participants completed the 9-item Patient Health Questionnaire (PHQ-9) online 2-3 weeks later, which was used to assess the frequency and severity of participants' depressive symptoms over the past two weeks. The final results showed that participants using Woebot significantly reduced depressive symptoms during the study period while the information control group did not. This trial allows for a preliminary indication that cognitive behavioral therapy with fully automated conversational agents using AI can have a positive impact. However, from reading the literature, this trial is the best evidence of the quality of psychotherapy with AI. The participants in the control group in this trial only read books about depression and did not receive psychotherapy in a practical sense.

B Woebot Trial 2 Another trial explored chatbots' feasibility in providing cognitive behavioral therapy to adolescents suffering from depression and anxiety¹². The trial took place during the COVID-19 pandemic, a period in which the global prevalence of anxiety and depression increased by 25%², and negative psychological symptoms among adolescents escalated to crisis levels. The trial's study population was aged

13 to 17 years and had been newly diagnosed with depression and anxiety in the past 3 months. Symptoms were intervened with the same Woebot that was used in the previous trial except Woebot for Adolescents, which was designed specifically for adolescents. Participants were divided into two groups, one of which received cognitive-behavioral therapy with the Woebot for a 12-week week, and the other group did not receive any interventions for that period. The final results of the trial do not accurately reflect the effectiveness of the AI psychotherapy program due to the insufficient number (less than ten people) of participants and the short duration of the treatment; however, the participants did give positive feedback on the program after their experience. A total of eight (80%) of the participating adolescents agreed or fully agreed with the statement "I liked using the program" and seven (70%) agreed with the statement "It seems possible to treat depression with the program". This suggests that participants were mostly positive and supportive of AI psychotherapists.

3. **What do we need to do to improve AI-based therapy?**

In summary, from the above two trials, it can be seen that the current AI therapist to the user of behavioral cognitive therapy after the results are positive and feasible. Patient after the use of the real improvement, but can not be determined compared to the non-AI conducted by the degree of improvement of the talk therapy. From these two trials, it was found that patients were able to accept the use of AI as a new form of therapy and did not reject it. In recent years there have been very few AI psychotherapy programs on the market for the general public to access, with only Woebot having many users in the major app stores. Therefore there is a lack of data and information on the subject and the conclusions drawn are rather one-dimensional and do not accurately prove the actual quality of AI therapists. To better demonstrate the quality of AI psychotherapy, it should be compared to human talk therapy, so that the current gap between AI and humans can be demonstrated and improved. In addition to this, in order to understand the quality of AI therapists one has to consider whether people are willing to share personal information with these AI therapists. This needs to take into account whether the user is willing to fully disclose their emotions and symptoms when faced with a cell phone screen rather than a real person. Artificially intelligent therapists also need to face some "doctor-patient relationship" challenges. For example, under what circumstances are people willing to share personal information with chatbots, and will conversations with users produce the same results as sharing information with human therapists?

4. **Is there a patient-clinician relationship in the AI-based therapy?** The patient-clinician relationship in talk therapy is critical to patient health and the quality of health-care delivery. Collaboration and cooperation between the physician and the patient are required during the treatment process to ensure the final outcome. The doctor-patient relationship represents a fiduciary relationship, which involves several elements of mutual understanding, trust, loyalty, and respect. Patients with different symptoms or personalities in psychotherapy may also have different outcomes after interacting with the doctor, for example, people with performative personalities may exaggerate their mental illnesses, while people with paranoid and schizoid personalities tend to avoid or not seek treatment¹³. Patient self-disclosure is critical for physicians to gather valid information, which can lead directly to the final diagnosis. Therefore in order to avoid incorrect diagnosis there is a need to build trust between the doctor and the patient so that the patient discloses what is on his mind without reservation. When confronted with an AI psychotherapist, the patient is confronted with an unemotional screen to which the patient needs to reveal his/her heart compared to the original face-to-face emotional sensitization from the doctor, which may lead to a different outcome. As artificially intelligent psychotherapists have begun to be used in several areas of mental health care, more and more people are learning about this technology and are beginning to have concerns. In a Pew Research Center poll, a majority (57%) of patients fear that AI will cause them to lose personal contact with their doctor or therapist. In addition to the technical aspects of AI, patients are also concerned about whether AI can actually solve the "doctor-patient relationship". As mentioned earlier, the patient-doctor relationship is important to patient satisfaction, so patients are now worried about the uncertainty of the final diagnosis for an AI that may not have the patient-doctor relationship. But instead of comparing AI to a human being, we should shift our thinking and weigh the pros and cons of the two. A human therapist will guide the therapeutic relationship in a way that is most conducive to achieving the therapeutic goals. But for an AI therapist, that relationship may be missing, which is the downside of an AI therapist. In the past, the doctor-patient relationship has been cited as the most important step in psychotherapy because talk therapy in the past needed to involve human beings, and human relationships were essential to the therapeutic process, so the establishment of a relationship between the two parties led directly to the final diagnosis. However, this criterion may be different in AI psychotherapy, and therefore should not be stuck to the old criteria that were once in place. Despite the inability to establish a doctor-patient relationship, artificial therapists may be able to achieve the same

proficiency as humans in delivering therapeutic techniques that are at least as effective as human therapists in terms of therapeutic outcomes. In summary, the doctor-patient relationship is not a necessity among AI therapists. The current AI therapists are voluntarily chosen by their clients to serve them, which side by side shows the users' trust in AI, and this trust is also a reflection of the doctor-patient relationship. Even though that is not a comparable trust to doctor-patient relationships, users will gradually adapt over time to develop trust between non-humans.

3. Case study: Willingness to Disclose in AI-based Psychotherapy

Psychologist Sidney Jourard first coined the term "self-disclosure" in 1958, which typically refers to a patient's "statements that reveal personal information about the therapist"¹⁴. This includes emotions, personal experiences, and so on. Trust is promoted when individuals are willing to share their thoughts, feelings, and experiences with others. After the therapist has developed trust with the patient, the patient expresses themselves more authentically, providing much help in subsequent therapy. Similarly, patients' self-disclosure of personal information is critical to successful treatment, and disclosures may include sensitive topics such as their trauma, personal experiences, and thoughts of self-harm. Considering the importance of patient self-disclosure, the following concern is possessed when introducing AI into the field of psychotherapy: if the patient is confronted with an electronic screen when using an AI psychotherapist rather than a therapist in traditional psychotherapy. When the object of disclosure changes from a person to an object, will the patient still disclose all of his inner thoughts to the AI therapist as he has done in the past?

This question already has an answer through trials, and conversational AI has been shown not to diminish patient disclosure. In fact, users are more open to AI therapists than they are to human listeners. In a study started by Gale M. Lucas in 2014, researchers recruited participants to engage in conversational interactions with virtual humans (VHs) and were led to believe that the VHs were controlled by humans or automation. The experiments conclude that of the experiment was that when confronted with a robotic conversation, participants could tell personal or embarrassing things about themselves to an unbiased machine without fear of negative judgment. In contrast, when confronted with a human-controlled VH, participants' responses were less candid than before. In addition to this another trial demonstrated the same results. Tess, a mental health chatbot developed by the X2 Foundation, was used in a trial that found patients improved after treatment with Tess and chose in feedback reports that they preferred chatting with Tess over traditional therapy¹⁵. This indicates that the patients were willing to disclose their emotions to the AI therapist and recognized the

acceptance of this new type of therapy.

In conclusion, people are more willing to disclose aspects of their emotions and experiences to an AI chatbot, instead of sharing it with a real person. This advantage seems like a strong candidate for AI to help bring psychotherapy at scale to a larger group of people.

4. Ethical Issues

As AI gradually develops in the field of mental health, many users are beginning to worry that their private information will be compromised by AI companies. As users pour out their thoughts with AI therapists, it can include a lot of personal user privacy. In today's data-driven world, data breaches occurring on the Internet happen all the time. In the past, it has not been uncommon to see news of user data being utilized by companies to sell or being extorted by hackers. Therefore, it is even more important for the companies involved to work towards achieving responsible clinical implementation. Confidentiality is a part of psychology's code of ethics, and during counseling sessions, psychologists need to make sure that they are more than enough to provide patients with a safe environment where they can easily talk about their private information. To this end, the American Psychological Association has enacted laws to protect patient privacy, and the Health Insurance Portability and Accountability Act has established national standards for the protection of personal medical records and personal health information. Thus, as AI is adapted to counseling therapy, there is a greater need for further research to address the ethical and social issues of the technology.

1. **Examples of current AI therapist privacy policies** When AI is involved in the field of mental health, the importance of privacy protection is particularly acute. The key to training AI therapists is to use real data used in previous counseling sessions and related information from previous patient treatments, which is very private to the patient. Take the example of Woebot, an AI psychotherapist that is already in use in the market today. Woebot Health, the company that started Woebot, declares that user privacy, as well as security of use, is not only their top priority but also the basis of their entire business. In order to gain the trust of their users, they have put a lot of effort into data protection. First and foremost, the company guarantees that all user data is protected as protected health information and complies with all HIPAA requirements. The Health Insurance Portability and Accountability Act (HIPAA) is a federal law enacted in the United States in 1996 to protect sensitive patient health information from disclosure without the patient's consent or knowledge. Compliance with this Act means that Woebot cannot use any protected patient information without the patient's consent. Secondly, they promise that users' conversations with Woebot are

confidential and that staff will only access personal data for their job responsibilities. Users can access their data at any time and delete or save it themselves. When confronted with the query that the AI needs to be trained using user data, Woebot Health stated that they would only use information about how Woebot users use the app and would not use the user's conversational interactions with Woebot. Unless a user uses the service through a clinical program or as part of a study and agrees to share the data, then Woebot will record data about the user's performance, such as responses to questions and emotional trends. These commitments indicate that in terms of data use, the company takes the approach that users share voluntarily and that it does not use users' private data without their permission.

- 2. Transparent regulatory models should be in place for AI therapists** The above examples show that AI needs to be scrutinized at all times. In addition to the commitment to users and compliance with relevant laws, oversight, and scrutiny from multiple parties is important. Political scientist Virginia Eubanks has warned against the unregulated use of data-driven technologies and AI in public services and welfare¹⁶. With the current negative public perception of AI in the logarithmic numbers and a succession of voices opposing the use of big data for AI, only by addressing the issue of trust between AI and its users can AI psychotherapists be widely utilized in the marketplace. Therefore, the implementation of transparent regulation of AI and the inclusion of the public in the regulatory process can increase public trust in the use of big data. Most of the advocates of AI in mental health today are psychiatrists, tech companies, and psychologists, these stakeholders are mainly responsible for the regulation of AI, while patients and their families, who use AI for psychotherapy, are excluded from the scope. And they as direct stakeholders should have the right to regulate the use of AI data. Users should have the right to ownership of the data and be informed about its use.
- 3. Incorrect data may be used for training AI therapist models** The relevant companies should review the data after collecting it, as data containing incorrect information can lead to corresponding errors in the trained AI models. For example, if data collected from therapists is used to build a dialog model, the data will contain the therapist's values, which may be biased statements. There is also the possibility of using data from a specific kind of population to train an AI model, but ending up applying it to a different population, which can lead to biased or inaccurate final results. In order to avoid both of these problems, it is important to ensure that data-specific reviews are conducted in the first place. Only vetted and accurate data should be used when training models with data. The second is to

determine the diversity of data sources and collect representative and diverse data during the training modeling phase. Ensure that the data being trained contains a wide range of demographic and ethnic diversity and can be applied to different populations.

- 4. Emergency situations** Usually, psychotherapists need to take their patients' privacy very seriously. However, in extreme cases, they can also violate confidentiality. When a patient possesses suicidal tendencies or other life-threatening moves, psychotherapists can share private information about them without their consent. Artificial intelligence therapists use the same privacy policies in such cases. Today AI technology is integrated into suicide management and can signal for human assistance if a suicide threat is detected. In 2017, Facebook used a program to automatically detect suicidal content and activate a crisis response plan if it detects that a user is posting content that contains a suicide risk. These programs may include providing users with support resources and crisis hotline information or alerting local emergency responders. The integration of artificial intelligence with mental health care holds great promise for reducing suicide rates. This approach does not undermine the confidentiality constraints of psychotherapy, and the intelligent detection of AI working in tandem with professional intervention can be effective in preventing many impulsive suicidal behaviors.

Discussion

The integration of Artificial Intelligence (AI) with the mental health field has shown great potential, but there are also limitations and challenges. This discussion will explore the future possibilities of AI in mental health treatment.

After reading the literature and researching it was found that artificial intelligence, when trained on large amounts of data, has the potential to make more accurate diagnoses than human therapists. This accuracy stems from the AI's ability to analyze massive amounts of data and identify patterns that human clinicians may have missed. Additionally, AI therapists can provide affordable and accessible mental health support, which allows them to be utilized in areas that lack qualified professionals. And for those who cannot afford traditional therapy, AI can provide a cost-effective alternative to civilianizing mental health care.

Despite these advantages, AI still has many limitations in mental health care. A key issue is the lack of empathy and eye contact in AI systems. Human therapists rely heavily on emotional intelligence and empathy to understand and communicate with their patients. All of today's AIs lack the ability to truly experience and understand human emotions. This limitation can hinder the therapy process as patients may feel misunderstood

or disconnected from the AI therapist.

Another important issue is the ethical implications of using AI in mental health care. AI systems must handle sensitive personal data, which raises concerns about privacy and data security. Companies developing AI therapists must implement strong data protection measures and ensure transparency in data use. Regulatory frameworks are also needed to oversee the ethical use of AI in this regard, prevent misuse and ensure that AI therapists adhere to high standards of care.

In order for AI to contribute positively to mental health, not only must AI algorithms be continually improved to increase their accuracy and reliability. It is also important to enhance interdisciplinary collaboration between AI developers, mental health professionals, and ethicists, which is essential to creating systems that are both effective and ethical. Beyond that, extensive training of human therapists in the use of AI tools can help bridge the gap between technology and human empathy, allowing for a more integrative approach to care.

Limitations

There are several limitations to this literature review. The time frame of the review was limited, which restricted the depth of the analysis. The number of research articles included in the study was also limited; the topic of AI in psychology has only emerged in recent years, and therefore the available literature could be more extensive, which may affect the comprehensiveness of the review. In addition, the research articles included spanned a specific time frame and may have missed relevant studies from other time periods. Future reviews should include a wider range of studies and spend more time analyzing them in order to gain a more comprehensive understanding of the use of AI in mental health.

Conclusion

Artificial intelligence has the potential to revolutionize mental health care by making it more accessible and affordable. However, significant challenges and ethical issues must be addressed to ensure that AI therapists provide high-quality care while protecting patient privacy and data security. Continuing advances in AI technology, coupled with interdisciplinary collaboration and robust ethical frameworks, have far-reaching implications for realizing the full potential of AI in mental health treatment.

References

- 1 *Evidence of Human-Level Bonds Established With a Digital Conversational Agent: Cross-sectional, Retrospective Observational Study*, <https://pubmed.ncbi.nlm.nih.gov/33973854/>, Retrieved August 30, 2023, from.
- 2 *COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide*, [https://www.who.int/news/item/02-03-](https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide)

[2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide](https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide), Retrieved August 15, 2023, from.

- 3 T. Alanzi, Y and Z. NCBI, *AI-Powered Mental Health Virtual Assistants' Acceptance: An Empirical Study on Influencing Factors Among Generations X*, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10753156/>, Retrieved July 7, 2024, from.
- 4 A. Turing, *Framework of Perceptive Artificial Intelligence using Natural Language Processing (P.A.I.N). Artificial Computational Research Society*, http://acors.org/Journal/Papers/Volume2/issue2/vol2_issue2_03.pdf, Retrieved August 16, 2023, from.
- 5 R. Bommasani, *AI Spring? Four Takeaways from Major Releases in Foundation Models*, <https://hai.stanford.edu/news/ai-spring-four-takeaways-major-releases-foundation-models>, Retrieved August 29, 2023, from.
- 6 A. Open, *OpenAI*.
- 7 M. N. Center, *Microsoft News*.
- 8 Bhattacharyya, *Journal of SAARC Psychiatric Federation*.
- 9 Y. Byambasuren and S. Saeb, *NCBI*.
- 10 A. Darcy, *JMIR Mental Health*.
- 11 S. M, *Computational psychotherapy research: scaling up the evaluation of patient-provider interactions*, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4245387/>, Retrieved January 3, 2024, from.
- 12 G. Nicol, *Journal of Medical Internet Research*.
- 13 B. Huntington, *Communication gaffes: a root cause of malpractice claims*, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1201002/>, Retrieved January 3, 2024, from.
- 14 H. E, *Self-disclosure. APA Psycnet*, <https://psycnet.apa.org/record/2002-01390-011>, Retrieved January 3, 2024, from.
- 15 J. A, *Psychological Artificial Intelligence Service, Tess: Delivering On-demand Support to Patients and Their Caregivers: Technical Report*, <https://pubmed.ncbi.nlm.nih.gov/30956924/>, Retrieved January 3, 2024, from.
- 16 V. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, Picador.