IEAI Technical University of Munich
School of Social Science and Technology
Institute for Ethics in Artificial Intelligence

Research Brief – October 2023

Minds and Machines: Understanding the Ethical Risks and Challenges of Using AI in Mental Healthcare Practice

By Daniel Burger & Auxane Boch

Mental health, as a foundation of well-being and quality of life, has gained a considerable importance in public discourse since the outbreak of COVID-19. Many individuals have experienced social isolation and/or economic hardship, often resulting in negative and sometimes harmful feelings and emotions. Simultaneously, there has been a decrease in the number of mental health professionals. A question remains if AI can address this bottleneck by ethically supporting mental health care efforts, and if yes, how? In this research brief, we will introduce current opportunities and future challenges associated with the use of AI in mental health.

The World Health Organization (WHO) defines a person's health as the interaction of physical, social, and mental factors: *"Health is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity. The enjoyment of the highest attainable standard of health is one of the fundamental rights of every human being without distinction of race, religion, political belief, economic or social condition"* (WHO, 2019, p. 1).

Mental health has played an increasingly significant role in modern times. For example, in Germany, an average sick leave due to mental illness rose from 27.3 days (2006) to 39.2 days (2021) (Statista, 2022). Furthermore, it is noteworthy that although mental disorders account for only 6.8 percent of work incapacity cases in Germany, they account for 18.8 percent of work incapacity days (Knieps & Pfaff, 2022). It is, therefore, not surprising that health insurers in Germany take mental illnesses very seriously (Techniker Krankenkasse, 2022; Knieps & Pfaff, 2022).

Given these facts, developing a social awareness of mental health and specific mental illnesses is paramount. Efforts to reach those goals are driven in particular by prominent role models. For example, Selena Gomez, Emma Stone and Buzz Aldrin went public with their psychological problems (Lopez, 2019; Uhlmann, 2018).

Considering the global attention and concern for mental health, researchers and practitioners are increasingly searching for tools to manage these conditions and their public health implications. A question has emerged; how can artificial intelligence (AI) assist in managing mental health issues?

The use of AI to provide mental support has already long existed. In 1966, Joseph Weizenbaum developed ELIZA, a speech analysis program capable of leading a conversation (Weizenbaum, 1966). Building upon work like Weizenbaum’s, this research brief considers current opportunities and future challenges associated with using AI in the mental health field.

Source Title Page Image: Wellness Tech AI Health Analysis. Source: Markus Winkler

Given the global attention and concern for mental health, researchers and practitioners are increasingly searching for tools to manage these conditions and the public health implications.

Critical Mental Health Conditions

First, we consider two widespread mental health conditions that need innovative approaches to diagnosis and treatment.

Major Depressive Disorder

Depression is a well-known example of a mental disorder and is considered to be a serious illness. A generic form of depression is Major Depressive Disorder (MDD), which accounted for 10.8% of cases of lifetime prevalence of depression in a survey of 30 different countries (Lim et al. 2018). In the US, the percentage of lifetime prevalence of depression was even higher at 20.6% (Hasin et al., 2018). Globally, MDD affects 322 million people (WHO, 2017). Depression in adolescence is especially a problem; 8% of adolescents suffering from MDD commit suicide (Bostwick and Pankratz, 2000; Hawton, 2014; Korten et al., 2012; Nock et al., 2008; Phillips and Cheng, 2012).

According to ICD-11, MDD is a period in which a person exhibits either a depressed underlying mood or decreased interest in all activities, occurring almost daily for at least two weeks and accompanied by other accompanying symptoms such as difficulty concentrating, a sense of worthlessness, hopelessness, suicidal thoughts and behaviours[1], insomnia or difficulty falling asleep, loss of appetite, fatigue, decreased drive, or tiredness to be considered a depressive episode (World Health Organization, 2022).

[1]In this context, the term suicidal behavior refers to suicidal ideation, planned, attempted and committed suicide (Posner et al., 2007; Sveticic and De Leo, 2012).

After a diagnosis of MDD, three different treatment options are available: psychotherapy, psychopharmacotherapy and a combination of both (Klein & Klein, 2021).

Schizophrenia

Another example of a mental illness is schizophrenic disorder or schizophrenia. Schizophrenia has a prevalence of about one percent (Saha et al., 2005). Yet, this mental disorder incurs more than $150 million in healthcare costs in the United States (Cloutier et al., 2016). For most patients, the disorder emerges in young adulthood (McCutcheon et al., 2020). The early onset and subsequent long-term treatment, and the impact on social and professional relationships explain why such a rare mental disorder can incur such large costs (Andreasen, 2010). At the same time, schizophrenia is associated with a shortened life expectancy of approximately 15 years and a suicide rate of 4 to 13 percent during the lifetime of patients (Hjorthøj et al., 2017; Siris, 2001).

According to the 11th International Classification of Diseases (ICD-11) proposed by the World Health Organization, schizophrenia is a disorder of mental modalities that includes thinking, perception, self-awareness, cognitive abilities, volition, affective feelings, and behavior (World Health Organization, 2022). The ICD-11 states core symptoms include delusions, thought disorders, and hallucinations (World Health Organization, 2022).

Similar to the treatment of MDD, the treatment of schizophrenia involves a choice of pharmacological, psychotherapeutic treatment or a combination of both methods.

AI in Mental Health - Opportunities & Chances

As noted above, mental disorders such as depression and schizophrenia are widespread around the world and represent a significant burden for the people affected as well as for society as a whole. AI technology can potentially add value, especially in the context of diagnosis, detection and treatment of mental disorders, for example, by diagnosing diseases faster and more accurately or by offering individualized treatment options (Alowais et al., 2023).

It is important to highlight that many of these applications hold great promise. However, proper standardization of ethical processes in the life cycle of AI in mental healthcare still must be clearly defined. The following section will discuss the possibilities and opportunities of using AI in mental health. The arguments are presented by application area: (1) diagnosis and detection and (2) treatment and support.

AI technology can potentially add value, especially in the context of diagnosis, detection and treatment of mental disorders…

Diagnosis & Detection

AI can already assist in diagnosing by using imaging techniques to detect diseases at preliminary stages of development based on characteristic features or anomalies, and thus significantly help both patients and medical professionals (Elavarasan & Pugazhendhi, 2020). It is possible to identify a risk for depressive disorder, especially in individuals who do not have the opportunity to see a physician or specialist (Cook et al., 2016) or to have a schizophrenic disorder identified through a written text (Wu et al., 2012).

It is possible for treating physicians to be assisted through these options by receiving a second opinion from AI during the examination of a patient. This not only saves time during diagnosis but also allows for increased precision. Studies by Niu et al. (2019) and Qureshi et al. (2019) demonstrated that their models could distinguish between patients with schizophrenia and healthy subjects with over 90 percent accuracy.

In addition to helping medical staff with AI, patients also benefit from using algorithms during the diagnosis and detection of mental disorders phase. A recent study by the Association of Statutory Health Insurance Physicians of Bavaria in 2021 showed that 97 days or 13.9 weeks pass between the first consultation contact and the start of psychotherapeutic treatment for a mental disorder such as MDD or schizophrenia (Ritter-Rupp et al., 2023). AI has the potential to speed up the diagnostic process, which spans from the initial contact to the diagnosis or start of treatment.

Machine learning can also aid in the prediction of susceptibility to depressive disorders. Based on the analysis of previously collected data (Sato et al. 2015), examined activity in specific brain regions can detect a depressive disorder based on aberrant results earlier than conventional methods. Other studies focused on early detection of depressive disorder using EEG data (Zhang et al., 2012; Acharya et al., 2018; Sharma et al., 2021; Sharma et al., 2022; Yan et al., 2022). In this regard, different studies used different AI methods. Nevertheless, all studies were able to detect individuals with depressive disorders with a high accuracy of over 90% (Zhang et al., 2012; Acharya et al., 2018; Sharma et al., 2021; Sharma et al., 2022; Yan et al., 2022).

…AI has the potential to speed up the diagnostic process, which spans from the initial contact to the diagnosis or start of treatment.

Treatment & Support

AI offers benefits not only for diagnostic purposes but also serves to improve the treatment of mental disorders and provide support during the treatment process (e.g., Van Breda et al., 2016; Yang et al., 2018).

Using a support vector machine (SVM) algorithm, a supervised machine learning method, can significantly increase the likelihood that a patient will respond to a given medication, thereby greatly improving therapeutic treatment outcomes (Ye et al., 2016). For example, Masychev et al. (2020) used AI to recommend specific initial treatments to patients. This not only avoided treatment delays but also selectively limited drug use only to an appropriate patient type (Masychev et al., 2020).

Masychev et al. (2021) succeeded in recommending treatment for a specific patient subset by using a machine-learning algorithm based on eight distinguishing features. The SVM classification of patients was able to achieve 85.7 percent accuracy, which is not only a benefit for supporting clinical psychiatry, but also provides opportunities for further improvement, e.g., by combining data and individual scores such as personality, inventory scores or blood tests (Masychev et al., 2021).

Similarly, Kessler et al. (2016) succeeded in developing a model that can predict the chronicity and severity of a depressive disorder over a longer period by interviewing MDD patients and subsequently analyzing the results with machine learning algorithms. Patients, in particular, can benefit from this analysis, as the results can be used to identify chronic depression at an early stage and treat it accordingly, for example, by providing more intensive care for long-term MDD patients (Kessler et al., 2016).

Intelligent chatbots are considered to be another promising option for supporting the care and treatment of mental disorders (Romanovskyi et al., 2021; Klos et al., 2021). Examples of AI-based chatbots for this purpose in the context of depressive disorder include Woebot, Tess and Wysa (Klos et al., 2021). While Woebot uses a cognitive-behavioral approach to reduce depressive symptomatology, Tess applies various therapy techniques, such as emotion-focused therapy, to provide support (Klos et al., 2021; Greenberg, 2017). Wysa, on the other hand, also uses a behavioral therapy approach but continues to rely on behavioral reinforcement and mindfulness techniques (Inkster et al., 2018).

In this regard, the opportunities that smart chatbots open up for patients' mental health are far-reaching. For example, they can provide access to an entry point to therapies for many people (Fitzpatrick et al., 2017; Patel et al., 2019; Dosovitsky et al., 2020). Moreover, AI thus offers a cost-effective alternative and is a scalable tool that can further expand the range of existing treatments (Fulmer et al., 2018). For example, Liu et al. (2022) demonstrated in their study that patients who used chatbots as a self-help medium could reduce the severity of their depressive disorder, as measured by their Patient Health Questionnaire-9 (PHQ-9) score, within 16 weeks. Regardless, AI for mental health applications still needs to be considered in terms of risks to ensure ethical development and deployment in the long term.

Takeaways

The benefits of using AI to diagnose and support the treatment of mental disorders vary. Medical professionals could benefit from AI through chatbots, virtual therapists or other assistance systems available around the clock and everywhere. They could thus conceal the limited availability of psychotherapists in terms of time or location. At the same time, this argument also describes a benefit for the patient, who could thus avoid long waiting lists and receive help more quickly. Furthermore, patients would benefit from an early detection of mental disorders, which could be treated accordingly.

In addition, the studies and experiments conducted so far reveal further possibilities for optimizing the use of AI in mental health. For example, combining various data that form the basis for the studies conducted makes it possible to further improve the AI-based model and applications, e.g., chatbots. Similarly, different machine learning methods that analyze the underlying data could lead to different insights. Therefore, it is necessary to specify which machine learning method should be applied to which problem to further improve AI results. Next, the risks to consider in implementing such technologies will be discussed.

AI in Mental Health - Challenges & Risks

While the possibilities and opportunities for AI use in medical application areas of diagnosis and detection, as well as treatment and support, and in mental healthcare are numerous and specifically applied, the challenges and risks are of a more general nature.

Sometimes, the most important prerequisite and biggest challenge for using AI is that the application must comply with a previously established ethical framework. For example, Floridi et al. (2018) elaborated on the AI4People approach, which consists of five different principles and twenty accompanying recommendations to create value for society through AI. The five principles can be divided into beneficence, non-maleficence, autonomy, justice, and explicability to structure this ethical approach (Floridi et al., 2018). At the same time, the principles also protect users, individuals and the public, as the stated goal of Floridi et al. (2018) is to move from theory to practice in using AI.

If AI-based applications are made available without first being checked for ethically acceptable use, serious consequences are likely. For example, the suicide of a man in Belgium caused a sensation when it was learned that he took his life after communicating with an anthropomorphic chatbot (Atillah et al., 2023). The company has since implemented safeguards so this does not happen again, but this example highlights the need for clear guardrails when it comes to using chatbots, and even more in the case of vulnerable populations such as mental health patients.

… the suicide of a man in Belgium caused a sensation when it was learned that he took his life after communicating with an anthropomorphic chatbot.

Non-maleficence

The principle of beneficence is considered to have been discussed in the previous section, where the benefits of such tools were outlined. However, the ethical principle of "Non-maleficence", or the "do no harm" condition (Floridi et al., 2018, p. 697), looks at the flip side of this concept. The central aspect of the principle is that no harm may be inflicted upon the individual or society by using AI.

Risks that violate the principle of 'non-maleficence' include biased data, stigmatization of patients and automation bias- "the tendency to use automated cues as a heuristic substitute for attentive information seeking and processing" (Lyell & Coiera, 2016)'. In particular, bias in data poses a significant risk to diagnosis and treatment. Given the risk of false positive diagnoses of mental disorders, it is considered that this not only does not help the patient, but also actually harms them and thereby violates the 'non-maleficence' principle. Further, the possible misuse of diagnosis technologies as surveillance tools, such as in the work context, would clearly violate this principle (APA, 2023).

Furthermore, AI models may also tend to reproduce bias and discrimination when trained with dissimilar data. This can equally lead to misdiagnosis or unequal treatment. Moreover, the example given above is a clear violation of the do no harm principle in which the misbehavior and unsupervised behavior of a conversational AI led to the suicide of an individual. This leads to further questions regarding the intersection of autonomy and non-maleficence principles.

On the other hand, the use of AI in mental health requires access to extremely sensitive data, including, for instance, direct personal conversations, text messages and medical records. Privacy and ethical issues regarding access to this data are critical. It is important to ensure that patient privacy is maintained and that data is not misused. For example, incidents of personal data misuse, such as in the case of the 2018 Facebook Cambridge Analytical scandal, could result in drastic consequences in healthcare settings such as discrimination or stigmatization of patients (Lovejoy et al., 2019).

At the same time, this is considered a major challenge to precisely define future access to AI-based applications and models in order to clarify legal and privacy aspects.

Justice

Within the complex principle of justice, "equity" encompasses several aspects, such as maintaining solidarity and promoting prosperity (Floridi et al., 2018). However, it is also important that the principle has a valid basis that provides guidance and recommendations for action on aspects such as fairness and equality (European Group on Ethics in Science and New Technologies, 2018).

The lack of a legal framework creates a risk to using AI in mental health. Legal requirements related to privacy are seen as a challenge (Lovejoy, 2019), as are violations of attributes such as fairness and equality (Ueda et al., 2023) and a lack of clarity about responsibilities in the event of errors (Choudhury & Asan, 2022).

With respect to the principle of "fairness", for example, the European Group on Ethics in Science and New Technologies warns of the current dangers of data processing, particularly biased and discriminatory datasets, as these datasets serve as the basis for AI (European Group on Ethics in Science and New Technologies, 2018). Biased data, for example, occurs when gender, race, age or ethnicity are not taken into account during data collection and, consequently, are very likely to lead to discriminatory results (Timmons et al., 2022).

The European Commission therefore states in its objectives that it is in the "interest to preserve the EU's technological leadership and to ensure that Europeans can benefit from new technologies developed and functioning according to Union values, fundamental rights and principles" (European Commission, 2021, p. n.a.). From a legal perspective, the implementation of AI in mental health is a multifactorial weighing of the benefits. It must be ensured that AI creates added value for people and society without violating applicable requirements, rules and laws.

The European Commission is in the process of drafting the legal framework that will set the rules for the use of AI and the future development of intelligent support systems (Feingold, 2023). To prevent risk and harm to the public legal good (protected by existing and applicable European Union law), the EU AI Act requires detailed documentation requirements, training and additional monitoring activities for AI systems (Mökander et al., 2021).

It must be ensured that AI creates added value for people and society without violating applicable requirements, rules and laws.

Autonomy

The basic understanding of "autonomy" encompasses the extent of an independent decision. Accordingly, this principle is characterized by the clear distinction of the decision-making authority between humans and machines. In the presence of AI, the intrinsic value of a human decision must be protected to minimize the risks of outsourcing decisions to machines (Floridi et al., 2018).

Human interaction and empathy are invaluable in mental health care. Although AI systems can simulate human conversations, the capacity for true empathy and compassion is still limited. Therefore, it is a considerable challenge to establish consistency between decision support provided by machine learning algorithms and the existing working methods of psychiatrists and therapists regarding the principle of ‘Autonomy’ (Koutsouleris et al., 2022).

…AI is not yet ready to make independent decisions but should instead support healthcare professionals as an assistive system.

From a scientific perspective, the possibility and existence of accurate AI-based models means that the previous boundaries of medical expertise must first be reexamined (Asan et al., 2020). For clinical management, this means adaptations, for example, in the form of human experts who have both medical understanding and the appropriate knowledge of how AI works (Rauseo-Ricupero et al., 2021). Wisniewski et al. (2020) refer to individuals who function at the interface between technology and clinicians or between technology and patients as digital navigators within the treatment ecosystem.

In addition, patient trust in AI systems must be strengthened to build a foundation for the widespread adoption of AI in mental health care. Many people are skeptical of machines when it comes to their mental health (Minerva & Giubilini, 2023). It is important to promote acceptance of AI-driven solutions and ensure that patients feel comfortable and understood.

At this point, it is important to clarify that AI certainly benefits healthcare professionals' everyday work. Nevertheless, AI is not yet ready to make independent decisions but should instead support healthcare professionals as an assistive system (McKendrick & Thurai, 2022).

Explicability

The guiding principle of "explicability" is focused on enabling the ethical principles described above through traceability and establishing accountability (Floridi et al., 2018). More precisely, it focuses on “how AI works (intelligibility) and who is responsible for the way AI works (accountability), in combination with an open research data management system” (Hermann & Hermann, 2021, pp. 29). In the context of AI applications, transparency is also considered important. Transparency in this context means that the input has a logical relationship to the results or that the input provides a clinically meaningful interpretation of the results (Joyce et al., 2023).

'Explicability' presents itself as a fundamental challenge. Non-transparent systems and methods lead to a black-box problem that makes the use of AI-based systems unjustifiable from the point of view of explicability (Wadden, 2021). Moreover, AI, through its non-transparent functioning, reveals a vulnerability that raises doubts about its results (Durstewitz et al., 2019; Chandler et al., 2019). Further, the need for explanation relates to the most basic rights of patients: to understand which treatment or therapy are being proposed to them.

The “need for explanation” also manifests itself in literature. In their paper, Boch et al. (2023) combine AI and bioethics in a theoretical framework to emphasize the prioritization of patient interests, insisting on the need for informed consent in the context of care, in addition to calling for the establishment of regulations and basic standards to ensure patients safety.

Even more explicitly, Fiske et al. (2020) demand that there be absolute clarity about “when and how informed consent must be obtained, as well as best practices for dealing with issues of vulnerability, manipulation, coercion and privacy in patients’ care” (Fiske et al., 2020, p. 214). This emphasizes how important informed consent is in practice. By actively consenting to treatment through a declaration of consent, the patient acts autonomously and independently. The patient's declaration of consent should, therefore, be viewed as an interface between the principles of 'Explicability' and 'Autonomy'.

In addition, there is a clear ethical need for the accountability path to be well defined in the case of AI malfunctioning and possibly causing harm to a patient. Thus, transparency on who is responsible for what should also be at the heart of AI in mental health development processes (Boch et al., 2023).

Transparent and explicable AI systems are considered to be crucial to building the previously addressed trust in AI in mental health. Therefore, in terms of 'explicability', it is recommended that therapists and software developers cooperate under ethical principles. Approaches such as 'ethics-by-design' could establish the framework for cooperation and thus focus on basic ethical principles already in the development and conception stage of AI applications.

Conclusion

The psychiatric and psychological diagnosis and treatment of mental disorders is experiencing promising development through AI, but technological progress also creates challenges that must be overcome. Appropriate solutions have to be developed, especially regarding implementing ethical principles like non-maleficence, justice, autonomy and explicability. It is advisable that representatives from the responsible areas cooperate to develop a strategy to enable the use of AI in mental health. This includes software developers, therapists and representatives from politics and academia. Further, patients and their direct care ecosystem should be involved in such developments, considering all stakeholders. Indeed, patient feedback should be sought as early as possible to identify and address further issues.

> ***The psychiatric and psychological diagnosis and treatment of mental disorders is experiencing promising development through AI, but technological progress also creates challenges that must be overcome.***

Should a functional integration of technology into medicine succeed, weighty opportunities will arise from the use of AI in mental health, such as the permanent availability of diagnostic or prognostic services, as well as the relief of medical professionals during the psychological and psychiatric care of patients (Satiani et al., 2018). In addition, regarding new treatments, continuing to collect data to better treat conditions such as depression through algorithm-based diagnosis assistance and treatment recommendations seems to be a promising path (Daniel et al., 2021).

References

Acharya, U. R., Faust, O., Hagiwara, Y., Tan, J. H., Adeli, H., & Subha, D. P. (2018). Automated EEG-based screening of depression using deep convolutional neural network. Computer Methods and Programs in Biomedicine, 161, pp.103–113. https://doi.org/10.1016/j.cmpb.2018.04.012.

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A., Almohareb, S. N., Aldairem, A., Alrashed, M., Saleh, K. B., Badreldin, H. A., Yami, M. S. A., Harbi, S. A. & Albekairy, A. M. (2023). Revolutionizing Healthcare: The role of Artificial intelligence in clinical practice. BMC Medical Education, 23(1). https://doi.org/10.1186/s12909-023-04698-z.

Andreasen, N. C. (2010). The lifetime trajectory of schizophrenia and the concept of neurodevelopment. Dialogues in Clinical Neuroscience,12(3), pp.409–415.https://doi.org/10.31887/dcns.2010.12.3/nandreasen.

American Psychological Association [APA]. (2023). Worries about artificial intelligence, surveillance at work may be connected to poor mental health. https://www.apa.org. https://www.apa.org/news/press/releases/2023/09/artificial-intelligence-poor-mental-health.

Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. Journal of Medical Internet Research, 22(6), p.e15154. https://doi.org/10.2196/15154.

Atillah, I. E. (2023). KI-Chatbot wird für Selbstmord eines Familienvaters in Belgien verantwortlich gemacht. euronews. https://de.euronews.com/next/2023/04/02/chatbot-eliza-ki-selbstmord-belgien (Accessed on 17 October 2023).

Boch, A., Ryan, S., Kriebitz, A., Amugongo, L. M., & Lütge, C. (2023). Beyond the Metal Flesh: Understanding the Intersection between Bio-and AI Ethics for Robotics in Healthcare. Robotics, 12(4), 110.

Bostwick, J., & Pankratz, V. S. (2000). Affective Disorders and Suicide Risk: A Reexamination. American Journal of Psychiatry, 157(12), pp. 1925–1932. https://doi.org/10.1176/appi.ajp.157.12.1925.

Chandler, C., Foltz, P. W., & Elvevåg, B. (2019). Using Machine learning in psychiatry: the need to establish a framework that nurtures trustworthiness. Schizophrenia Bulletin. https://doi.org/10.1093/schbul/sbz105.

Choudhury, A., & Asan, O. (2022). Impact of accountability, training, and human factors on the use of artificial intelligence in healthcare: Exploring the perceptions of healthcare practitioners in the US. Human Factors in Healthcare, 2, 100021. https://doi.org/10.1016/j.hfh.2022.100021.

Cloutier, M., Aigbogun, M. S., Guerin, A., Nitulescu, R., Ramanakumar, A. V., Kamat, S. S., DeLucia, M., Duffy, R. A., Legacy, S. N., Henderson, C., François, C., & Wu, E. Q. (2016). The Economic Burden of Schizophrenia in the United States in 2013.The Journal of Clinical Psychiatry,77(06), pp. 764–771.https://doi.org/10.4088/jcp.15m10278.

Cook, B. I., Progovac, A. M., Chen, P., Mullin, B., Hou, S. S., & Baca-García, E. (2016). Novel Use of Natural Language Processing (NLP) to Predict Suicidal Ideation and Psychiatric Symptoms in a Text-Based Mental Health Intervention in Madrid. Computational and Mathematical Methods in Medicine, 2016, pp.1–8. https://doi.org/10.1155/2016/8708434.

Daniel, O., Sharon, R., & Tepper, S. J. (2021). A device review of Relivion®: an external combined occipital and trigeminal neurostimulation (eCOT-NS) system for self-administered treatment of migraine and major depressive disorder. Expert Review of Medical Devices, 18(4), pp. 333–342. https://doi.org/10.1080/17434440.2021.1908122.

Dosovitsky, G., Pineda, B. S., Jacobson, N. C., Chang, C., Escoredo, M., & Bunge, E. L. (2020). Artificial Intelligence Chatbot for Depression: Descriptive Studyof Usage. JMIR Formative Research, 4(11), p. e17065. https://doi.org/10.2196/17065.

Durstewitz, D., Koppe, G., & Meyer-Lindenberg, A. (2019). Deep neural networks in psychiatry. Molecular Psychiatry, 24(11), 1583–1598. https://doi.org/10.1038/s41380-019-0365-9.

Elavarasan, R. M., & Pugazhendhi, R. (2020). Restructured society and environment: A reviewon potential technological strategies to control the COVID-19 pandemic. Science of the Total Environment, 725, p. 138858. https://doi.org/10.1016/j.scitotenv.2020.138858.

European Commission (2021). “Laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts”. European Commission. Retrieved April 4, 2021, from https://eurlex.europa.eu/resource.html?uri=cellar:e0649735a37211eb958501aa75ed71a1.0001.02/DOC1&format=PDF (Accessed on 31 August 2023).

European Group on Ethics in Science and New Technologies (EGE). (2018). Statement on artificial intelligence, robotics and “autonomous” systems. https://doi.org/10.2777/531856.

Feingold, S. (2023). The EU’s Artificial Intelligence Act, explained. World Economic Forum. Retrieved May 3, 2023, from https://www.weforum.org/agenda/2023/03/the-european-union-s-ai-act-explained/ (Accessed on 29 September 2023).

Fiske, A., Henningsen, P., & Buyx, A. (2020). The implications of embodied artificial intelligence in mental healthcare for digital wellbeing. Ethics of Digital Well-Being: A Multidisciplinary Approach, 207-219.

Fitzpatrick, K., Darcy, A. M., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults with Symptoms of Depression and Anxiety Using a FullyAutomated Conversational Agent (Woebot): A Randomized Controlled Trial. JMIR Mental Health, 4(2), p. e19. https://doi.org/10.2196/mental.7785.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines, 28(4), pp. 689–707. https://doi.org/10.1007/s11023-018-9482-5.

Friedrich, M. W. (2017). Depression Is the Leading Cause of Disability Around the World.JAMA,317(15), p. 1517.https://doi.org/10.1001/jama.2017.3826.
Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial. JMIR Mental Health, 5(4), p. e64. https://doi.org/10.2196/mental.9782.
Greenberg, L. S. (2017). Emotion-focused therapy. In American Psychological Association eBooks. https://doi.org/10.1037/15971-000.
Hasin, D. S., Sarvet, A. L., Meyers, J. L., Saha, T. D., Ruan, W. J., Stohl, M., & Grant, B. F. (2018). Epidemiology of Adult DSM-5 Major Depressive Disorder and Its Specifiers in the United States. JAMA Psychiatry, 75(4), p. 336.https://doi.org/10.1001/jamapsychiatry.2017.4602.
Hawton, K. (2014). Suicide prevention: a complex global challenge. The Lancet Psychiatry, 1(1), pp. 2–3. https://doi.org/10.1016/s2215-0366(14)70240-8.
Hermann, E. & Hermann, G. (2021). Artificial Intelligence in Research and Development for Sustainability: The Centrality of Explicability and Research Data Management. AI and Ethics, 2(1), 29–33. https://doi.org/10.1007/s43681-021-00114-8.
Hjorthøj, C., Stürup, A. E., McGrath, J. J., & Nordentoft, M. (2017). Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. The Lancet Psychiatry,4(4), pp.295–301.https://doi.org/10.1016/s2215-0366(17)30078-0.
Inkster, B., Sarda, S., & Subramanian, V. (2018). An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. Jmir Mhealth and Uhealth, 6(11), p.e12106. https://doi.org/10.2196/12106.
Joyce, D. W., Kormilitzin, A., Smith, K. & Cipriani, A. (2023). Explainable artificial intelligence for mental health through transparency and interpretability for understandability. npj digital medicine, 6(1). https://doi.org/10.1038/s41746-023-00751-9.
Kessler, R., Van Loo, H. M., Wardenaar, K. J., Bossarte, R. M., Brenner, L. A., Cai, T. T., Ebert, D. S., Hwang, I., Li, J. C., Jonge, D. W. W., Nierenberg, A. A., Petukhova, M., Rosellini, A. J., Sampson, N. A., Schoevers, R. A., Wilcox, M. A., & Zaslavsky, A. (2016). Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. Molecular Psychiatry, 21(10), pp. 1366–1371. https://doi.org/10.1038/mp.2015.198.
Klein, J.P. & Klein, E.M. (2021). Mein Leitfaden Psychiatrie, Springer eBooks. https://doi.org/10.1007/978-3-662-60445-8.
Klos, M. C., Escoredo, M., Joerin, A., Lemos, V. N., Rauws, M., & Bunge, E. L. (2021). Artificial Intelligence–Based Chatbot for Anxiety and Depression in University Students: Pilot Randomized Controlled Trial. JMIR Formative Research, 5(8), p. e20678. https://doi.org/10.2196/20678.
Knieps, F., & Pfaff, H. (Eds.). (2022). BKK Gesundheitsreport 2022. BKK Dachverband. https://www.bkk-dachverband.de/fileadmin/user_upload/BKK_Gesundheitsreport_2022.pdf(Accessedon5 September 2023).
Korten, N., Comijs, H. C., Lamers, F., & Penninx, B. W. (2012). Early and late onset depression in young and middle-aged adults: Differential symptomatology, characteristics and risk factors? Journal of Affective Disorders, 138(3), pp. 259–267. https://doi.org/10.1016/j.jad.2012.01.042.
Koutsouleris, N., Hauser, T. U., Skvortsova, V., & De Choudhury, M. (2022). From promise to practice towards the realisation of AI-informed mental health care. The Lancet Digital Health, 4(11), e829–e840. https://doi.org/10.1016/s2589-7500(22)00153-4.
Lim, G., Tam, W. W., Lu, Y., Ho, C. S., Zhang, M. W. B., & Ho, R. C. (2018). Prevalence of Depression in the Community from 30 Countries between 1994 and 2014. Scientific Reports, 8(1). https://doi.org/10.1038/s41598-018-21243-x.
Liu, H., Peng, H., Song, X., Xu, C., & Zhang, R. (2022). Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness. Internet Interventions, 27, p. 100495. https://doi.org/10.1016/j.invent.2022.100495.
Liu, Q., He, H., Yang, J., Feng, X., Zhao, F., & Lyu, J. (2020). Changes in the global burden of depression from 1990 to 2017: Findings from the Global Burden of Disease study. Journal of Psychiatric Research,126, pp. 134–140. https://doi.org/10.1016/j.jpsychires.2019.08.002.
Lopez, J. (2019). Selena Gomez, Emma Stone, Kendall Jenner, and Adele: Why even the richest & most famous have discussed mental health issues – adolescent growth. https://adolescentgrowth.com/selena-gomez-why-even-the-most-famous-have-discussed-mental-health-issues/ (Accessed on 12 October 2023).
Lovejoy, C. M., Buch, V., & Maruthappu, M. (2019). Technology and mental health: The role of artificial intelligence. European Psychiatry, 55, pp. 1–3. https://doi.org/10.1016/j.eurpsy.2018.08.004.
Lyell, D., & Coiera, E. (2016). Automation bias and verification complexity: a systematic review. Journal of the American Medical Informatics Association, 24(2), pp. 423–431. https://doi.org/10.1093/jamia/ocw105.
Masychev, K., Ciprian, C., Ravan, M., Manimaran, A., & Deshmukh, A. (2020). Quantitative biomarkers to predict response to clozapine treatment using resting EEG data. Schizophrenia Research, 223, pp. 289–296. https://doi.org/10.1016/j.schres.2020.08.017.
Masychev, K., Ciprian, C., Ravan, M., Reilly, J., & MacCrimmon, D. J. (2021). Advanced signal processing methods for characterization of schizophrenia. IEEE Transactions on Biomedical Engineering, 68(4), pp. 1123–1130. https://doi.org/10.1109/tbme.2020.3011842.
McCutcheon, R. A., Marques, T. R., & Howes, O. D. (2020). Schizophrenia—An Overview. JAMA Psychiatry, 77(2), p. 201. https://doi.org/10.1001/jamapsychiatry.2019.3360.
McKendrick, J. & Thurai A. (2022). AI isn't ready to make unsupervised decisions. Harvard Business Review. https://hbr.org/2022/09/ai-isnt-ready-to-make-unsupervised-decisions (Accessed on 17 October 2023).
Minerva, F. & Giubilini, A. (2023). Is AI the future of mental healthcare? Topoi-an International Review of Philosophy, 42(3), 809–817. https://doi.org/10.1007/s11245-023-09932-3.

Mökander, J., Axente, M., Casolari, F., & Floridi, L. (2021). Conformity Assessments and Post-market Monitoring: A Guide to the Role of Auditing in the Proposed European AI Regulation. Minds and Machines, 32(2), 241–268. https://doi.org/10.1007/s11023-021-09577-4.

Niu, H., Yang, P., Chen, H., Hao, R., Dong, S., Shi, Y., Chen, X., Yan, H., Zhang, Y., Chen, Y., Jiang, F., Yang, T., & Guo, Y. (2019). Comprehensive functional annotation of susceptibility SNPs prioritized 10 genes for schizophrenia. Translational Psychiatry, 9(1). https://doi.org/10.1038/s41398-019-0398-5.

Nock, M. K., Borges, G., Bromet, E. J., Cha, C. B., Kessler, R. C., & Lee, S. (2008). Suicide and Suicidal Behavior. Epidemiologic Reviews, 30(1), pp. 133–154. https://doi.org/10.1093/epirev/mxn002.

Patel, F., Thakore, R., Nandwani, I., & Bharti, S. K. (2019). Combating Depression in Students using an Intelligent ChatBot: A Cognitive Behavioral Therapy. IEEE India Conference. https://doi.org/10.1109/indicon47234.2019.9030346.

Phillips, M., & Cheng, H. (2012). The changing global face of suicide. The Lancet, 379(9834), pp. 2318–2319. https://doi.org/10.1016/s0140-6736(12)60913-1.

Posner, K., Oquendo, M. A., Gould, M. S., Stanley, B., & Davies, M. G. (2007). Columbia Classification Algorithm of Suicide Assessment (C-CASA): Classification of Suicidal Events in the FDA's Pediatric Suicidal Risk Analysis of Antidepressants.American Journal of Psychiatry,164(7), pp. 1035–1043.https://doi.org/10.1176/ajp.2007.164.7.1035.

Qureshi, M. N. I., Oh, J., & Lee, B. (2019). 3D-CNN based discrimination of schizophrenia using resting-state fMRI. Artificial Intelligence in Medicine, 98, pp. 10–17. https://doi.org/10.1016/j.artmed.2019.06.003.

Rauseo-Ricupero, N., Henson, P., Agate-Mays, M., & Torous, J. (2021). Case studies from the digital clinic: integrating digital phenotyping and clinical practice into today's world. International Review of Psychiatry, 33(4), 394–403. https://doi.org/10.1080/09540261.2020.1859465.

Ritter-Rupp, C., Fett, S., Pfeifer, A., & Tauscher, M. (2023). Analyse der Wartezeiten in der Psychotherapie in Bayern. Zenodo (CERN European Organization for Nuclear Research). https://doi.org/10.5281/zenodo.7599322.

Romanovskyi, O., Pidbutska, N., & Knysh, A. (2021). Elomia Chatbot: The Effectiveness of Artificial Intelligence in the Fight for Mental Health. COLINS. Retrieved April 22-23, 2021, from http://ceur-ws.org/Vol-2870/paper89.pdf(Accessed on 29 September2023).

Saha, S., Chant, D., Welham, J., & McGrath, J. J. (2005). A Systematic Review of the Prevalence of Schizophrenia. PLOS Medicine, 2(5), p. e141. https://doi.org/10.1371/journal.pmed.0020141.

Satiani, A., Niedermier, J., Satiani, B., & Svendsen, D. P. (2018). Projected Workforce of Psychiatrists in the United States: A Population Analysis. Psychiatric Services, 69(6), pp. 710–713. https://doi.org/10.1176/appi.ps.201700344.

Sato, J. R., Moll, J., Green, S. M., Deakin, J.F.W., Thomaz, C. E., & Zahn, R. (2015). Machine learning algorithm accurately detects fMRI signature of vulnerability to major depression. Psychiatry Research: Neuroimaging, 233(2), pp. 289–291. https://doi.org/10.1016/j.pscychresns.2015.07.001.

Sharma, G., Joshi, A., & Pilli, E. S. (2021). An Automated MDD Detection System based on Machine Learning Methods in Smart Connected Healthcare. 2021 IEEE International Symposium on Smart Electronic Systems (iSES). https://doi.org/10.1109/ises52644.2021.00019.

Sharma, G., Joshi, A., & Pilli, E. S. (2022). DepML: An Efficient Machine Learning-Based MDD Detection System in IoMT Framework. SN Computer Science, 3(5). https://doi.org/10.1007/s42979-022-01250-6.

Siris, S. G. (2001). Suicide and schizophrenia. Journal of Psychopharmacology, 15(2), pp. 127–135. https://doi.org/10.1177/026988110101500209.

Statista. (2022). Psychische Erkrankungen - Durchschnittliche AU-Dauer. https://de.statista.com/statistik/daten/studie/845/umfrage/dauer-von-arbeitsunfaehigkeit-aufgrund-von-psychischen-erkrankungen/(Accessed on 29 September 2023).

Sveticic, J., & De Leo, D. (2012). The hypothesis of a continuum in suicidality: a discussion on its validity and practical implications. Mental Illness,4(2), pp. 73–78. https://doi.org/10.4081/mi.2012.e15.

Techniker Krankenkasse. (2022). Zwei Jahre Coronapandemie: Wie geht es Deutschlands Beschäftigten? Teil 2. Gesundheitsreport 2022. https://www.tk.de/resource/blob/2130932/3432a2d7c9f827e38b1dee99779bb826/gesundheitsreport-2022-data.pdf (Accessed on 12 September 2023).

Timmons, A. C., Duong, J. B., Fiallo, N. S., Lee, T., Vo, H. P. Q., Ahle, M. W., Comer, J. S., Brewer, L. C., Frazier, S. L. & Chaspari, T. (2022). A call to action on assessing and mitigating bias in artificial intelligence applications for mental health. Perspectives on Psychological Science, 174569162211344. https://doi.org/10.1177/17456916221134490.

Ueda, D., Kakinuma, T., Fujita, S., Kamagata, K., Fushimi, Y., Ito, R., Matsui, Y., Nozaki, T., Nakaura, T., Fujima, N., Tatsugami, F., Yanagawa, M., Hirata, K., Yamada, A., Tsuboyama, T., Kawamura, M., Fujioka, T., & Naganawa, S. (2023). Fairness of artificial intelligence in healthcare: review and recommendations. Japanese Journal of Radiology. https://doi.org/10.1007/s11604-023-01474-3.

Uhlmann, B. (2018). Berühmte Depressionspatienten: „Wenn mein schwarzer Hund zurückkehrt ". Süddeutsche Zeitung. Retrieved June 24, 2023, from https://www.sueddeutsche.de/gesundheit/beruehmte-depressions-patienten-wenn-mein-schwarzer-hund-zurueckkehrt-1.2519325(Accessed on 29 September 2023).

Van Breda, W., Pastor, J., Hoogendoorn, M., Ruwaard, J., Asselbergs, J., & Riper, H. (2016). Exploring and Comparing Machine Learning Approaches for Predicting Mood Over Time. Smart Innovation, Systems and Technologies, pp.37–47. https://doi.org/10.1007/978-3-319-39687-3_4.

Wadden, J. J. (2021). Defining the undefinable: the black box problem in healthcare artificial intelligence. Journal of Medical Ethics, 48(10), 764–768. https://doi.org/10.1136/medethics-2021-107529.

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. Communications of the ACM, 9(1), pp. 36–45. https://doi.org/10.1145/365153.365168.

Wisniewski, H., Gorrindo, T., Rauseo-Ricupero, N., Hilty, D. M., & Torous, J. (2020). The Role of Digital Navigators in Promoting Clinical Care and Technology Integration into Practice. Digital Biomarkers, 4(Suppl. 1), 119–135. https://doi.org/10.1159/000510144.
World Health Organization (WHO). (2017). Depression and Other Common Mental Disorders: Global Health Estimates. https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf(Accessed on 29 September 2023).
World Health Organization (WHO). (2019). Basic Documents: Fourty-ninth edition (including amendments up to 31 May 2019). https://apps.who.int/gb/bd/pdf_files/BD_49th-en.pdf(Accessed on 29 September 2023).
World Health Organization (WHO). (2022). International Classification of Diseases for Mortality and Morbidity Statistics -Eleventh Revision. https://icd.who.int/browse11/l-m/en#/?view=G0(Accessed on 29 September 2023).
World Health Organization (WHO). (2023). Depressive disorder (depression). https://www.who.int/news-room/fact-sheets/detail/depression(Accessed on 29 September 2023).
Wu, J., Yu, L., & Chang, P. (2012). Detecting causality from online psychiatric tests using inter-sentential language patterns. BMC Medical Informatics and Decision Making, 12(1). https://doi.org/10.1186/1472-6947-12-72.
Yan, D., Zhao, L., Song, X., Zang, X., & Yang, L. (2022). Automated detection of clinical depression based on convolution neural network model. Biomedizinische Technik. https://doi.org/10.1515/bmt-2021-0232.
Yang, S., Zhou, P., Duan, K., Hossain, M. S., & Alhamid, M. F. (2018). emHealth: Towards Emotion Health Through Depression Prediction and Intelligent Health Recommender System. Mobile Networks and Applications, 23(2), pp. 216–226. https://doi.org/10.1007/s11036-017-0929-3.
Ye, Z., Rae, C. L., Barcia, J. A., Ham, T. E., Rittman, T., Jones, P. B., Rodríguez, P., Coyle-Gilchrist, I., Regenthal, R., Altena, E., Housden, C. R., Maxwell, H. S., Sahakian, B. J., Barker, R. A., Robbins, T. W., & Rowe, J. B. (2016). Predicting beneficial effects of atomoxetine and citalopram on response inhibition in Parkinson’s disease with clinical and neuroimaging measures. Human Brain Mapping, 37(3), pp. 1026–1037.https://doi.org/10.1002/hbm.23087.
Zhang, X., Hu, B., Zhou, L., Moore, P. K., & Chen, J. M. (2012). An EEG Based Pervasive Depression Detection for Females. Springer eBooks, pp. 848–861. https://doi.org/10.1007/978-3-642-37015-1_74.