Critical Insights into the Impact of Artificial Intelligence on Mental Health, Patient Rights, and Human Rights

Sangramjeet Chavan*[1], Vojjala Sahiti[2], Dr. Shashikala Gurpur[3] and Dr. Atmaram Shelke[4]

Abstract

Artificial Intelligence (AI) technologies are changing many aspects of contemporary life, including the emergence of mental healthcare as a field for innovation. This interdisciplinary research investigates the complex interactions between Artificial Intelligence, Mental Health Care, Patient Rights, and Human Rights to assess the advantages and disadvantages of incorporating AI into Mental Health Services. Globally, millions of people suffer from mental disorders, placing significant strains on health systems and societies. Mental health care has always been restricted due to stigma, structural barriers and lack of resources, often leading to an inability to provide accessible and effective treatments despite increasing patient demands.

AI technology has emerged as a possible solution for age-old issues like efficient early detection methods, treatment strategies, and personal follow-up. Conversely, AI application in this delicate area raises ethical and legal dilemmas about people's freedoms, which must be assessed critically, and predictable steps must be taken to protect patient rights.

Keywords: Artificial Intelligence, Mental Health, Patient Rights, Human Rights, Healthcare Ethics

[1] Ph.D. Scholar, Symbiosis Law School, Pune, Symbiosis International (Deemed University). sangramjeet.chavan@symlaw.ac.in, *corresponding author.

[2] Teaching Associate, Ph.D. Research Scholar, Symbiosis Law School, Pune, Symbiosis International (Deemed University) vojjala.sahiti@symlaw.ac.in

[3] Professor and Director, Symbiosis Law School, Pune, Head, Symbiosis Centre for Advanced Legal Studies and Research (SCALSAR), Symbiosis International (Deemed University) director@symlaw.ac.in

[4] Associate Professor, Symbiosis Law School, Pune, Symbiosis International (Deemed University). ashelke@symlaw.ac.in

Introduction

Artificial Intelligence (AI) technology is rapidly advancing and changing many facets of modern society, including mental health services and healthcare. Artificial Intelligence algorithms and systems are used frequently due to the potential to provide more individualised, effective, and readily available care. It has been used more for the diagnosis of mental health problems, planning the treatment, and monitoring the progress of the treatment and aftercare. The use of AI in mental health care also brings up important issues related to human rights, the law, and ethics that call for further research and preventative action (Dotsen, Ghazal, Gover & Larson, 2021). Millions of people worldwide are affected by mental health illnesses, which represent a substantial global burden. Upholding patients' rights and ensuring access to high-quality mental healthcare treatments have always been complex tasks, made more difficult by structural impediments, stigma, and resource limitations (Kun, L. G. 2018).

AI tools such as mental health apps serve as vital tools to overcome these issues, especially by being accessible to a broader population. Still, these AI technologies bring both potential and threats. Although AI-powered technologies have the potential to increase access to care, improve treatment outcomes, and improve early diagnosis, they also carry dangers pertaining to algorithmic bias, data privacy, and potential violations of patient autonomy and human rights standards, (Gerke et al., 2020), multidisciplinary strategy is necessary to deal with the complicated and multidimensional terrain at the nexus of AI, mental health, patient rights, and human rights. AI systems must be created and implemented responsibly and fairly, which requires careful consideration of ethical issues, including informed consent, privacy, and non-discrimination (Graham et al., 2020). Additionally, a critical analysis of the current legal and regulatory frameworks is required due to the possible effects of AI on fundamental human rights, such as the right to privacy, the prohibition of discrimination, and the availability of healthcare. (Luxton, 2014).

Human rights are mostly internationally accepted and can act as guidelines for formulating global regulations. However, as philosophical and ethical environments vary across nations, subsequent policies reflect varying conceptions and fulfilments of human rights. One of the biggest challenges in the future of AI in mental health would be the protection of human rights in a practical manner. To develop better AI, better healthcare and, thus, better AI in healthcare, one must adhere to the human rights that are granted by various declarations, covenants, legislation, and

constitutions. However, as we explore this notion of regulation of human rights, we find agreement, disagreement, and variability on a global scale. These variabilities may well hamper the ethical development of AI in health care internationally. This multidisciplinary study intends to investigate the complex interactions among AI tools, mental health, patient rights, and human rights. This study examines the possible advantages, hazards, and difficulties connected with integrating AI in mental healthcare by utilising insights from various domains, including computer science, bioethics, law, psychology, and human rights. It also aims to provide guidelines and suggestions for the responsible development and application of AI technologies in this field and to build a thorough knowledge of the ethical, legal, and practical ramifications.

This study aims to defend patient rights and maintain human rights principles while significantly contributing to the current conversation on AI ethics and governance through a thorough examination and analysis. The ultimate objective is to recommend a regulatory framework that enables the use of AI in mental healthcare that can be used to enhance lives while preserving each person's basic rights and dignity.

This study shall further analyse the existing legal framework, which includes international conventions and organisational policies in relation to AI, mental health, patient rights, and

Human rights. It shall further study if the existing AI applications in mental healthcare comply with existing regulations concerning patient rights, data protection, and human rights.

This study shall also examine case studies about AI integration in health care, especially focusing on chatbots providing mental healthcare services for the purpose of assessing the impact on patient rights and human rights to analyse both the benefits and challenges, thereby identifying best practices, lessons learned, and areas for improvement.

Literature Review:

The advent of AI in health care has been pivotal in making health care affordable and accessible to a large population. AI in the medical field is used in various areas, including physical and mental health. There are multiple applications of AI in the healthcare sector that cover patient engagement and adherence, diagnosis and treatment recommendations and administrative tasks, such as authorisation, maintenance of health records, etc. (Davenport & Kalakota, 2019).

With its numerous advantages and challenges, the introduction of artificial intelligence (AI) in healthcare has attracted substantial interest

among scholars, professionals, and decision-makers, especially in mental health.

Mental health, with its significant effects on people, communities, and economies, is one of the important worldwide issues. To address the issues of stigma, cost, and accessibility in mental health care, digital solutions for mental health, such as the use of AI chatbots, could transform digital mental health. (Almeida & Silva, 2021).

With millions of people suffering from mental illnesses globally and a lack of readily available care, utilising technologies like Artificial Intelligence (AI) could improve mental health via both prevention and therapy. AI-enabled solutions that identify higher-risk populations and facilitate faster intervention can stop the development of more serious mental diseases. Stress can be identified, evaluated, and predicted by AI, which is capable of processing natural language from electronic health records to identify early cognitive impairment or child abuse.Both of which can have long-term consequences for an individual's mental health. (Catherine K and et al., 2023).

Numerous studies have investigated how AI might transform mental healthcare delivery. Early diagnosis and risk assessment of mental health problems have shown potential for AI-powered diagnostic tools and predictive models. (Barocas, S., & Selbst 2016). Artificial intelligence (AI) in mental health treatment could significantly increase the accessibility and scalability of online interventions if AI is perceived to be just as effective as human specialists, but this empirical study comparatively analysing responses of AI and human beings found that the response given by AI often lacks authenticity or professionalism (Gagan Jain & et.al, 2024).

Artificial intelligence (AI) has brought forth several legal and ethical challenges for society, including those involving privacy and surveillance, bias and discrimination, and the role of human judgment. These challenges may also pose a philosophical difficulty. Introducing more recent digital technologies has given rise to worries that they could become a new source of errors and data breaches. (Nithesh Naik et al.,2022).

Although there is general agreement on the potential benefits of AI, several questions have been raised about the ethical implications, patient rights, and human rights. For this reason, it is essential to have a thorough awareness of the body of research already done in this area.

Gaps and Challenges:

Although the works now in publication offer insightful analyses of the distinct elements of AI, mental health, patient rights, and human rights, a more thorough and integrated strategy is required to address the intersections and interdependencies of these domains, particularly when seen from a multidisciplinary approach that takes legal, sociological, and cultural concerns into account, there are still gaps in our knowledge of the practical and ethical consequences of AI in mental healthcare. Furthermore, there are not enough thorough ethical and legal frameworks to direct the responsible development and application of AI in mental healthcare while respecting the rights of patients and the fundamentals of human rights.

By examining the effects of AI on mental health and thereby further impact on patient rights and human rights through a multidisciplinary lens, this study seeks to close these gaps. This research aims to add to the current discussion and provide guidance for creating moral standards and legislative frameworks for the proper integration of AI in mental healthcare by utilising a variety of viewpoints from different sectors.

Methodology – Tools and Techniques:

The doctrinal methodology adopted for this research shall consist of extensive theoretical analysis and a comprehensive literature review of relevant literature.

To comprehensively investigate the implications of artificial intelligence (AI) on mental health, patient rights, and human rights, this study will adopt a multi-pronged methodological approach, drawing upon various qualitative and analytical techniques.

This study takes an interdisciplinary angle by incorporating knowledge from computer science, ethics, law, and psychology, among other disciplines, to explore the complex dynamics between AI usage in mental health care with a focus on patients' rights and human privilege at large. The framework used to evaluate is that of creating a balanced view on the different sides of these disciplines about AI involvement in the mental healthcare field, generalising ethical implications with legal aspects that accompany it while adhering to practical issues regarding its operationalisation.

The research discusses increasing instances that demonstrate the transfusion of AI into mental health care. It includes diagnostic tools powered by AI, models that predict early risks, and therapies that are aided by it. However, while these tools show hopes for increasing access to treatment and improving health results, they also have weaknesses related to discrimination based on algorithms, the nature of data and misleading diagnoses or inappropriate treatments. Moreover, the usage of AI in psychiatry generates important ethical dilemmas about patients' independence from authorities of hospitals that prescribe medicine and manage medical records.

An analysis of perspectives on human rights reflects on how artificial intelligence influences basic freedoms related to including but not limited to private life behaviour and denial of free movement among individuals who need health services alike. One primary area for concern is whether using artificial intelligence within mental health services is shaping already existing inequalities experienced by people seeking help regarding their mental well-being since this tool can perpetuate them as well as worsen them through its application. In addition, this study looks into whether existing legal frameworks are adequate in considering human rights involving artificial intelligence in treating various psychological disorders by outlining some of these deficiencies and possible enhancements.

Concrete instances of the advantages or challenges faced by real-world implementations of AI in mental health services can be found in case study analyses. An integral part of the research is an ethical and legal analysis using established frameworks and methodologies to evaluate whether AI systems adhere to current statutes, regulations, and ethical precepts.

Research Objectives:

The primary objective of this research is to critically examine the potential benefits, risks, and challenges associated with integrating AI into mental healthcare, focusing on patient rights and human rights implications.

This research aims to explore the ethical, legal, and practical considerations that need to be addressed to ensure the responsible, rights-respecting development and deployment of AI technologies in mental healthcare and to develop a comprehensive understanding of the intersections between AI, mental health, patient rights, and human rights

from a multidisciplinary perspective, drawing insights from various fields such as computer science, bioethics, law, psychology, and human rights. It also aims to propose guidelines and recommendations for the ethical and rights-based governance of AI in mental healthcare, balancing technological innovation with the protection of patient rights and human rights principles.

AI and Mental Health

The usage of AI in healthcare majorly emerged during the COVID-19 period when there was an ever-rising demand for healthcare and there was a lack of skilled workforce in many nations. "AI in health care was used as a diagnostic tool and for treatment of COVID." (Chang et al., 2021). The usage of advances in AI has opened new opportunities to tackle the workforce shortage and provide healthcare services to a larger population. During the period of COVID-19, in addition to diagnostic tools, AI has also emerged in the field of mental health care, where chatbots were used as conversational agents to provide support in mental health care.
In an effort to deliver individualised and ongoing care, AI-assisted therapies and treatment monitoring systems have also been created. (Cearns, Hahn and Baune, 2022). On the other hand, issues including algorithmic bias, data quality, and the possibility of incorrect diagnosis or treatment recommendations have been brought to light (Council of Europe. 2020). Chatbots are prominent emerging tools utilised for mental health applications.

Chatbots: Chatbots have been defined as *'any software application that engages in a dialogue with a human using natural language'* (Dale R, 2016); other terms for chatbots include *dialogue agents*, *conversational agents, and virtual assistants.* Previous scholarly research has outlined the ethical benefits and drawbacks of chatbots and related technologies. Some studies have also provided ethical frameworks or standards that are pertinent to chatbots for mental health. (Simon Coghlan et al., 2023)
General application process of chatbots: Chatbots are majorly used in two kinds of applications:

i. a diagnostic tool for mental ailments;

ii. an application to assess mental and emotional well-being.

Effectiveness of chatbots:

Multiple studies concluded that awareness of AI involvement alters user perception of the interactions, with human responses generally perceived as more authentic and practical than AI. This highlights the irreplaceable value of human empathy in mental health support. Although AI is recognised for its professional and informative role, it is seen as a supplement rather than a replacement for human interaction in mental health services. The trust in AI was notably correlated. (Jain, Pareek & Carlbring 2024).

Potential Risks of Mental Health Apps:

More general 'mental health' apps that purport to assist with anxiety, depression, and other conditions have been used with varying levels of success. Considering the existing position of AI in mental health, a wide range of mental health apps are suffering from a frequent lack of an underlying evidence base, a lack of scientific credibility, and subsequent limited clinical effectiveness. Moreover, a comprehensive analysis of AI states that AI can provide primary care where the responses provided by it are not human equivalents. Comparative studies.

Analysing the response of AI and human beings, we found that in comparison with AI, the response provided by humans is more practical and empathetic.[18]

There are clear risks with hyping technology, especially for disadvantaged people and without a commensurate evidence base to justify the enthusiasm. These risks appear at both the individual and population levels, from shaping individual users' preferences and expectations, (Gerke et al., 2020).

Additionally, there is a chance that algorithmic prejudice and discrimination in AI systems will worsen already existing gaps in the availability and calibre of mental healthcare.

Human Rights Implications of AI:

An increasing corpus of research explores how AI technologies affect human rights as they become more prevalent. Particularly in the context of AI systems that depend on enormous volumes of personal data, the right to privacy and data protection has been a primary issue.[20]Concerns

regarding the right to equality and non-discrimination have also been highlighted by the possibility that AI algorithms would reinforce or magnify inequality and discrimination. (Gianfrancesco et al. 2018). The necessity of ensuring that AI technologies respect and advance human rights norms has also been underlined by the UN and other international organisations (United Nations, 2019; Council of Europe, 2020), (Obermeyer et al., 2019).

Patient Rights in the AI Era:

According to the WHO Patient Safety Rights Charter, the 10 fundamental patient safety rights outlined in the Charter are the right to:

1. Timely, effective, and appropriate care;
2. Safe health care processes and practices;
3. Qualified and competent health workers;
4. Safe medical products and their safe and rational use;
5. Safe and secure health care facilities;
6. Dignity, respect, non-discrimination, privacy, and confidentiality;
7. Information, education, and supported decision-making
8. Access to medical records;
9. To be heard and fair resolution;
10. Patient and family engagement." (World Health Organization, 2023)

The use of AI in health care has raised significant concerns regarding privacy rights and non-discrimination. The implications of using chatbots or mental health care-related apps shall be further analysed in detail.

Privacy Rights and AI Mental Health Tools:

Privacy rights are at the forefront of concerns regarding smart mental health tools like Kintsugi and Mind Strong Health. These technologies collect highly sensitive biometric and behavioural data, which are considered special categories of personal data under many data protection laws.

Examining the AI in mental health, "People who utilise chatbot apps to

improve their mental health share various healthcare-related personal information while using these apps such as burnout, mood swings, and details about mental disabilities, neurodiversity, or other diseases. As every user has a unique identity, the data that the chatbot collects is not necessarily anonymous. This private data of patients is kept on the cloud servers that the app developers give or on their own servers."(Gallese 2022) These cloud servers are usually under third-party management, which causes the possibility of data breaches; this practice jeopardises the right to privacy of the patients. There is no proper regulation about processing the personal data of the patient; there are high chances of breach of confidentiality and, thereby, violation of the patient's right to privacy. A more stringent regulatory framework is necessary because of the dire repercussions that could arise from a breach of mental health data.

Internationally, the right to privacy is enshrined in Article 12 of the Universal Declaration of Human Rights (UDHR) and Article 17 of the International Covenant on Civil and Political Rights (ICCPR). These provisions protect individuals from arbitrary or unlawful interference with their privacy. The UN Human Rights Committee's General Comment No. 16 on Article 17 of the ICCPR further elaborates that protecting personal data is a crucial aspect of the right to privacy in the digital age.

Data Protection in the Context of AI Mental Health Applications:

Internationally, the OECD Privacy Guidelines and the EU's GDPR set standards for data protection. In India, the Information Technology Act 2000 and IT Rules 2011 provide some safeguards, and the Digital Personal Data Protection Act 2023 aims to provide comprehensive data protection for rights. The need for a robust data protection regime was emphasised in the "*Puttaswamy judgment"* (AIR 2018 SC (SUPP) 1841). For AI mental health tools, data protection rights should include the right to access one's data, rectify inaccurate data, erase data, restrict processing, and ensure data portability.

Non-Discrimination and Access Barriers: Reduce AI Algorithmic Bias:

One of the significant downsides of AI is algorithmic bias, which could lead to racial, gender, and ethnical discrimination. To ensure the patient's rights to non-discrimination, these systems must be carefully designed

and monitored to prevent discrimination. Internationally, the right to non-discrimination is protected under numerous treaties, including Article 7 of the UDHR, Articles 2 and 26 of the ICCPR, and the International Convention on the Elimination of All Forms of Racial Discrimination. In India, Article 14 of the Constitution guarantees equality before the law, while Article 15 prohibits discrimination based on religion, race, caste, sex, or place of birth. The Rights of Persons with Disabilities Act of 2016 specifically prohibits discrimination against persons with disabilities, which is relevant in the context of mental health. Landmark cases like *Naz Foundation v. Govt. of NCT of Delhi (2009) 160 DLT 277* and *National Legal Services Authority v. Union of India (2014) 5 SCC 438* have emphasised the importance of non-discrimination based on sexual orientation and gender identity, principles that should be incorporated into AI mental health tools.

Ethical Considerations in the Development of AI: The development of AI should follow certain principles to address ethical concerns and uphold human rights for this purpose. The five key principles in this framework were identified. They are (a) non-maleficence, (b) beneficence, (c) respect for autonomy, (d) justice, and (e) explicability (Floridi & Cowls, 2019).

The High-Level Expert Group on Artificial Intelligence has a tiered level of principles. The expert group proposes a framework for trustworthy AI that consists of lawful AI (which they do not cover), ethical AI, and robust AI. This framework is based on four ethical principles: respect for human autonomy, harm prevention, fairness, and explicability. From these principles, they deduce seven key requirements for the realisation of trustworthy AI, namely:
1. human agency and oversight 2. technical robustness and safety 3. privacy and data governance 4. transparency 5. diversity, non-discrimination, and fairness 6. Social and Environmental well-being 7. Accountability. (Stahl, B. C, 2021) from there, they then develop assessment methods for trustworthy AI and policy recommendations. (Durstewitz, Koppe, & Meyer-Lindenberg, 2021).
Regarding patient autonomy and rights, using AI in mental healthcare presents serious ethical questions. Given the sensitivity of mental health data and the possibility of misuse or unauthorised access, informed consent, privacy, and data protection have become critical issues.
Unauthorised access to patient data may have serious repercussions, such as stigma, discrimination, or psychological distress for patients. Additionally, patient privacy may be further jeopardised because even

anonymised data can occasionally be re-identified, especially when paired with other datasets.
Obtaining informed consent is one of the challenging aspects of using AI. AI's intricate algorithms and decision-making procedures are opaque; patients may find it challenging to understand how their data will be used. They may agree to the use of AI tools without fully understanding the consequences, which would compromise the validity of their consent and, thereby, their right to privacy.

A balance needs to be established between the development of AI and the protection of human rights. This approach will be crucial in realising the potential of AI to transform mental healthcare while upholding the highest ethical and human rights standards. The balance between innovation and regulation must be carefully struck to avoid stifling progress while safeguarding against potential harm. (Simões de Almeida & Pereira da Silva, 2022).

Guidelines for Regulation of AI: A comprehensive legal framework should be introduced for effective regulation and rights-based governance of AI in mental healthcare. The following guidelines can be adapted to protect the patient's rights and human rights.

1. **Classification of Mental Health Data as Sensitive Personal Data:** Mental health data should be classified as sensitive personal data under the Data Protection and Digital Privacy (DPDP) Act, 2023. Such data must be handled with a high level of protection, ensuring that the mental health information of individuals is kept confidential and secure, with strict compliance with privacy laws.

2. **Examination and Certification by Mental Health Professionals:** To ensure effective human intervention (Human in the Loop) as WHO recommends, mental health apps and chatbots must undergo rigorous examination before launch. The standards should be certified by mental health professionals, including psychologists, sociologists, and other relevant experts. This process ensures that the AI tools align with ethical standards and clinical effectiveness before being made available to the public.

3. **Customisation to Regional and Cultural Contexts:** Mental health apps should be customised to reflect regional and cultural nuances. This ensures that the AI's prompts and suggestions are relevant and non-

discriminatory with respect to race, region, and ethnicity. Tailoring these tools appropriately will foster inclusivity and improve their efficacy in different demographic contexts.

4. **Periodic Evaluation by Experts**: The performance and effectiveness of AI-powered mental health apps must be periodically evaluated by a team of experts. Regular assessments will help identify and rectify any issues, ensuring the tools remain effective, up-to-date, and aligned with current medical and psychological standards.

5. **Integration with Psychology Departments of Hospitals to Ensure Effective Human Intervention:** AI-based mental health apps should be integrated with real-world psychology departments and hospitals. This integration is crucial in cases where the AI identifies high-risk diagnoses or situations. Connecting users with actual doctors for more accurate treatment helps to minimise the risk of misdiagnosis and ensures that users receive the necessary care.

These guidelines offer an effective legal framework for the responsible development and use of AI in mental health, ensuring both the protection of users and the ethical use of technology in sensitive healthcare areas.

Conclusion:

From the above analysis, it can be concluded that using AI in mental health care applications such as chatbots has been highly beneficial in assisting patients. Nevertheless, certain regulatory steps are required to address the challenges, such as data privacy, patient safety and the protection of patients' rights.

Incorporating AI and human rights perspectives into the current mental healthcare framework requires a comprehensive and multifaceted approach. By updating existing legislation, creating effective human interventions to monitor AI, introducing new laws, establishing regulatory bodies, and developing comprehensive policies, we can create a foundation for the responsible use of AI in mental healthcare. The focus on human rights safeguards, ethical considerations, and patient empowerment ensures that the benefits of AI are realised without compromising individual rights and autonomy.

Technical standards, accountability measures, and ongoing evaluation processes will help maintain the quality and reliability of AI systems in

mental healthcare. Meanwhile, research, workforce development, and international cooperation investments will drive continuous improvement and adaptation to emerging challenges.

A regulatory framework that allows for innovation in AI mental healthcare while prioritising patient well-being, privacy, and rights should be introduced. Legal frameworks must ensure that these chatbots provide accurate, evidence-based support while safeguarding user rights and ensuring accountability for harm.

Findings:

The findings and the conclusions drawn from this research indicate the ability of AI to aid mental health care delivery but with authors' inputs of emphasis on consideration of ethical principles and legal considerations, especially the possible human rights implications. The research highlights how the ethical principles of fairness, transparency, and accountability should be integrated into mental health services. One of the important findings is that the current governance structures are weak. Thus, the authors suggest a flexible regulatory framework that keeps up with the rapid advancements in AI.

As such, this study recommends ethical and rights-based governance of AI in mental healthcare. Emphasis is laid on the necessity of transparency, accountability, and fairness in developing and applying AI systems. This does not imply that they never do this, for instance, as it pertains to credible consent towards the protection of a patient's anonymity and confidentiality when it comes to data as well as removing possible bias in AI-driven systems of supporting people with mental problems with an aim that no one should be left out.

The article further suggests that integrating AI systems into mental healthcare should occur within parameters defined by human rights; therefore, protecting patients' rights must be given the utmost priority. It advocates for continuous collaboration across disciplines and engagements with stakeholders who play key roles in mental health, such as community-based organisations or civil society organisations, to tackle some ethical dilemmas concerning Artificial Intelligence (AI) today. It also provides suggestions for collective efforts aimed at combating cultural beliefs influencing patients' perceptions about mental health services and suggesting ways that these technologies could improve lives, particularly considering inequitable access due to poverty.

The study's limitations are acknowledged, including the rapidly evolving nature of AI technologies and the need for more extensive longitudinal research. Future research directions are proposed, emphasising the importance of expanded stakeholder perspectives in the long term.

Recommendations:

Upon analysing the existing concerns regarding AI in mental health applications, it can be recommended that ensuring explainability and effective human intervention can address the concerns of privacy, non-discrimination, and patient safety.

Need to provide right to Explainability: The Right to Explanation for AI-driven decisions is closely linked to the right to human intervention. This concept is gaining traction, particularly in the EU's GDPR (Article 22). While not yet explicitly recognised in Indian law, it can be inferred from principles of natural justice and Article 21 of the Constitution. For mental health AI tools, this could mean providing clear, understandable explanations of how the AI reached its conclusions, offering insights into key factors influencing decisions, and allowing for the contestability of AI-driven decisions. The Right to Health is fundamental in the context of AI mental health tools. Internationally, this right is recognised in Article 25 of the UDHR and Article 12 of the ICESCR. In India, while not explicitly stated in the Constitution, the Right to Health is considered part of the Right to Life under Article 21, as interpreted in cases like *Paschim Banga Khet Mazdoor Samity v. State of West Bengal (1996) 4 SCC 37* and *Bandhua Mukti Morcha v. Union of India (1984) 3 SCC 161*. The National Health Policy, 2017 and the Mental Healthcare Act, 2017 further reinforce this right. AI mental health tools have the potential to significantly advance the right to health by improving access to care and enhancing the quality of interventions.

Need to have Human Intervention: The concept of Human Intervention in AI-driven decisions is an emerging concept; it is also called ‘Human in the Loop’ (HITL), according

to the latest ICMR guidelines. This intervention is particularly important in healthcare. While not explicitly stated in traditional human rights documents, it aligns with the right to the highest attainable standard of health (Article 12 of ICESCR) and the right to benefit from scientific progress (Article 27 of UDHR). In the European context, Article 22 of the EU General Data Protection Regulation (GDPR) recognises the right not to be subject to automated individual decision-making. In India, while there is no specific legislation addressing this right, it can be interpreted

under the broad ambit of Article 21 of the Constitution. For tools like IBM's Watson Health and Cerebral, there should be clear protocols for human healthcare provider intervention, mechanisms for patients to request a human review of AI-generated decisions and transparency about the role of AI in the decision-making process.

The concept of 'Human in the Loop' (HITL) places human beings in a supervisory role and is more relevant for healthcare purposes. This will ensure individualised decision-making by the health professionals, keeping the patient's interest at the centre. Adopting the HITL principle throughout the development and deployment of AI for health also helps in the optimal sharing of accountability by the team involved in developing and deploying AI-based algorithms.
It is critical to ensure that the entity(ies) seeking such responsibility have proper legal and technical credentials in the area of AI technologies for health impact studies, as well as an ongoing examination of emerging AI technologies and their implications for mental healthcare and human rights.

References:

Almeida, R. S. de, & Silva, T. P. da. (2021) 'AI Chatbots in Mental Health', in Advances in Psychology, Mental Health, and Behavioral Studies. pp. https://doi.org/10.4018/978-1-7998-8634-1.ch011

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104, 671-732.
https://doi.org/10.15779/Z38BG31

Cearns, M., Hahn, T., & Baune, B. T. (2022). Artificial intelligence in mental health. Current Opinion in Psychiatry, 35(2), 116-122. https://doi.org/10.1097/YCO.0000000000000792

Chang, Z. et al. (2021) 'Application of artificial intelligence in COVID-19 medical area: a systematic review', J. THORAC. DIS., 13, p. 7034. https://doi.org/10.21037/jtd-21-747

Coghlan, S. et al. (2023) 'To chat or bot to chat: Ethical issues with using chatbots in mental health', Digital Health, 9
https://doi.org/10.1177/2055207623118354

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. FUTUR. HEALTHC. J., 6(2), 94. https://doi.org/10.7861/futurehosp.6-2-94

Dale R, (2016), The return of the chatbots, Natural Language Engineering , Volume 22 , Issue 5 , September 2016 , 811 – 817. https://doi.org/10.1017/S1351324916000243

Dotsen, E. K., Ghazal, L. V., Gover, P. E., & Larson, K. A. (2021). The future of artificial intelligence in mental health care. The Journal of Technology in Behavioral Science, 6(1), 1-14. https://doi.org/10.1007/s41347-020-00174-5

Durstewitz, D., Koppe, G., & Meyer-Lindenberg, A. (2021). Deep neural networks in psychiatry. Molecular Psychiatry, 26(11), 6487-6503. https://doi.org/10.1038/s41380-021-01174-z

Ettman, C. K. & Galea, S. (2023) 'The Potential Influence of AI on Population Mental Health', JMIR Mental Health, 10(1), e49936. https://doi.org/10.2196/49936

Floridi, L. & Cowls, J. (2019) 'A Unified Framework of Five Principles for AI in Society'. Available at: https://papers.ssrn.com/abstract=3831321 https://doi.org/10.2139/ssrn.3831321

Gallese. (2022). Legal Issues of the Use of Chatbot Apps for Mental Health Support, Communications in Computer and Information Science 258, 258-267. https://doi.org/10.1007/978-3-031-18697-4

Gerke, S., Shachar, C., Fergerson, N. T., & Cohen, I. G. (2020). The legal and ethical considerations of artificial intelligence in mental healthcare. Current Psychiatry Reports, 22(11), 1-9. https://doi.org/10.1007/s11920-020-01197-9

Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. JAMA Internal Medicine, 178(11), 1544-1547.

https://doi.org/10.1001/jamainternmed.2018.3763

Graham, A. K., Muruthi, M., Musich, S., Deer, L., Miller, L., Shea, E., & Kathol, K. (2020). Artificial intelligence in mental health care: Speculation to application. Harvard Review of Psychiatry, 28(6), 396-406. https://doi.org/10.1097/HRP.0000000000000276

Jain, G., Pareek, S., & Carlbring, P. (2024). Revealing the source: How awareness alters perceptions of AI and human-generated mental health responses, 36 INTERNET INTERV. 100745, https://doi.org/10.1016/j.invent.2024.100745.

Kun, L. G. (2018). Artificial intelligence and the future of mental health. The Journal of Technology in Behavioral Science, 3(2), 109-122. https://doi.org/10.1007/s41347-018-0064-1

Luxton, D. D. (2014). Artificial intelligence in psychological practice: Current and future applications and implications. Professional Psychology: Research and Practice, 45(5), 332-339.
https://doi.org/10.1037/a0034559

Naik, N., et al. (2022). Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility, 9 FRONT. SURG. 1
https://doi.org/10.3389/fsurg.2022.862322

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453.
https://doi.org/10.1126/science.aax2342

Simões de Almeida, R., & Pereira da Silva, T. (2022) 'AI Chatbots in Mental Health', in Advances in Psychology, Mental Health, and Behavioral Studies. pp.
https://doi.org/10.4018/978-1-7998-8634-1.ch011

Stahl, B. C. (2021) 'Concepts of Ethics and Their Application to AI' https://doi.org/10.1007/978-3-030-69978-9_3