Generative AI Can Transform Mental Health: A Roadmap for Emerging Companies
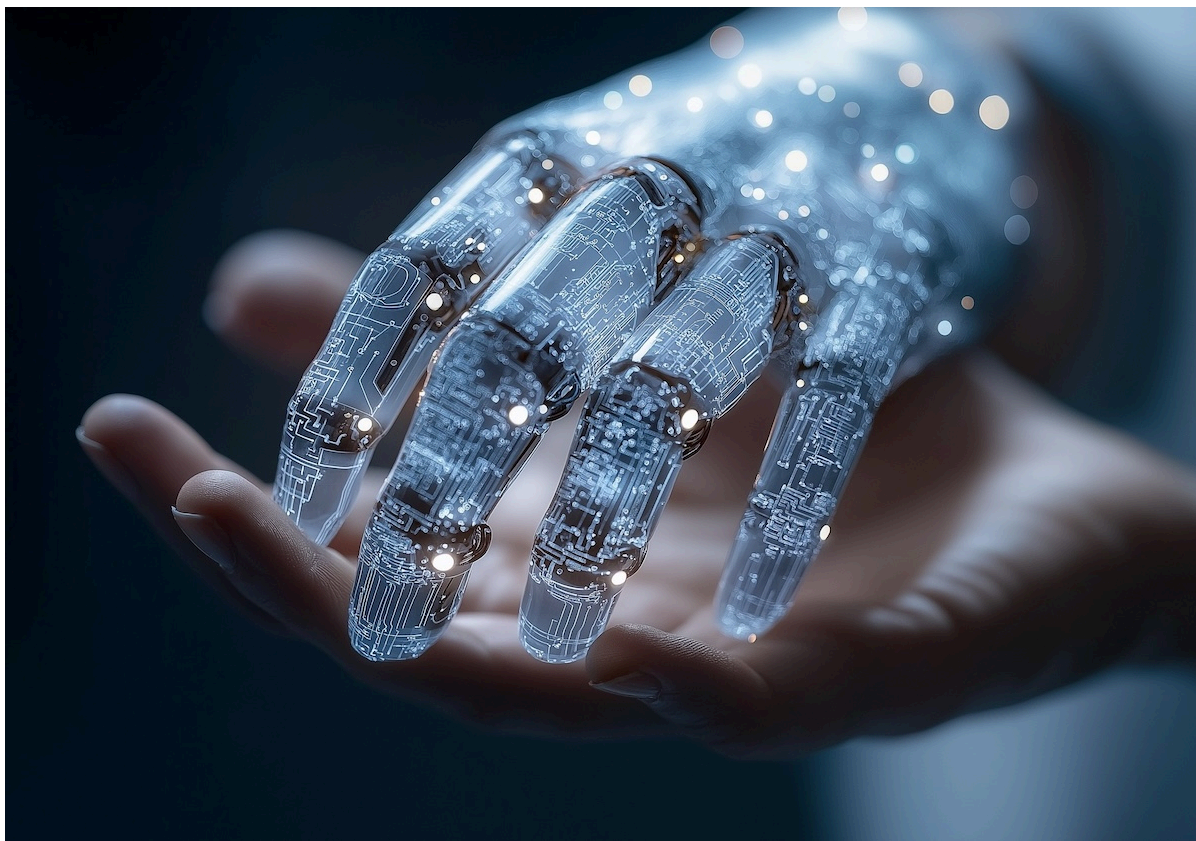
by Ajay Kumar and Jagpreet Chhatwal

Image Credit | Kate

Unlocking the Future of Mental Health using Generative AI

☑ INSIGHT | OPINION   07 May 2025

The rapid advancement of generative artificial intelligence (GAI) has generated new possibilities in healthcare, particularly mental health. Mental health issues can affect people at any stage of life, leading to considerable distress, dysfunction, and, in severe cases, mortality. A **recent CNN survey** found that an overwhelming majority of Americans think the United States (US) is experiencing a mental health crisis. Furthermore, two of the most common mental health conditions, depression and anxiety, are estimated to cost the global economy **US$ 1 trillion each year**. In the US, the healthcare industry is overburdened, especially after the COVID-19 pandemic, with fewer health workers and more patients than before the pandemic. In 2020, there were 350 individuals for each mental health provider in the US, with rates varying greatly by state. **Massachusetts**, for example, had the highest rate of mental health providers in the **US**, with 140 individuals for each mental health provider; meanwhile, Alabama had the lowest rate, with 850 individuals per provider. As a result, doctors often have long waitlists for new patients. **According to this AMA report**, in the US, a shortage of mental health specialists may result in new patients waiting time three months or more to get an appointment for their first consultation. A recent survey by the **Centers for Disease Control (CDC)** found that 57% of females in high school experienced persistent sadness and hopelessness, and 24% had engaged in suicide ideation. Moreover, this **report** revealed that over half (52%) of LGBQ+ students had recently experienced poor mental health; notably, 22% had attempted suicide in the past year. These data provide evidence for the increased prevalence of mental health problems and the imbalance between health providers and patients.

RELATED CMR ARTICLES

Junjie Zhou and Xing Wan, "**How to Build Network Effects on Online Platforms for Mental Health**," California Management Review, *64*/4 (2022): 20-46.

The Promise of GAI in Mental Health Prediction Applications

GAI can particularly improve two areas of mental health: early detection and diagnosis, and it can also serve as a therapeutic support assistant for the patients. In the diagnosis context, the assistance of GAI can help transform mental health care, primarily by increasing accessibility to mental health resources. By analyzing complex speech, text, and behavior patterns, GAI-based predictive models can potentially identify early signs of mental health issues that traditional AI methods may overlook. While traditional AI collects and analyzes historical textual datasets from social media posts, forums, or chat conversations to make predictions, GAI creates new data, multimodal content, such as facial expressions, tone, and body language, and ideas, based on its training data, in addition to the capabilities of traditional AI. This additional dimension allows for more personalized and innovative solutions and therefore can considerably enhance mental health care.

As a therapeutic support assistant, GAI can generates dynamic, empathetic **AI-based avatars** that can enhance traditional AI techniques that rely on static, prerecorded information. By training these avatars on datasets consisting of human interactions, which cover sentiments, speech patterns, and body language, these avatars can engage with patients in an empathetic and realistic manner. Establishing and maintaining a relationship and rapport between patient and therapist/care provider is crucial, but this is difficult to achieve when they are unable to communicate directly due to a language barrier. Further, as the world becomes increasingly globalized, such language barriers are more commonly encountered. GAI models can incorporate **advanced natural language processing techniques** to be able to understand and speak multiple languages, provide translations in real time, and convert spoken language into text (and vice versa), enabling effective communication between patients and mental health professionals, regardless of their respective native language. GAI-powered chatbots can engage users in conversations that mimic human interaction, providing immediate support and guidance, and are available 24/7.

GAI can significantly enhance traditional AI systems' ability to detect **subtle behavioral patterns** and identify mental health conditions at their early stages, especially in the context of **cognitive-behavioral therapy (CBT)**. GAI can also create customized therapeutic materials, such as mindfulness exercises, worksheets, or mood-boosting activities tailored to an individual's specific needs and preferences. GAI can assist mental health professionals by summarizing patient sessions, highlighting key emotional cues, and suggesting potential therapeutic interventions based on data analysis. GAI can also create realistic patient scenarios for training purposes, helping professionals develop skills in diagnosis and treatment planning.The technology remains in its nascent stage; however, companies should focus on leveraging GAI-based tools to develop applications in the context of mental health interventions and support.

Risks, Disadvantages, and Ethical Concerns of GAI in Mental Health Prediction Applications

While GAI offers several benefits, it also comes with significant disadvantages and risks. For example, a biased GAI-based application trained on data primarily from Western populations could lead to misdiagnosis the depression when used by a user from South Asia, as it may fail to accurately interpret culturally specific symptoms. In such a case, GAI may fail to recognize culturally specific distress symptoms, leading to misdiagnosis and reinforcing healthcare disparities. As a result, users may lose trust in GAI-based mental health solutions. If a user requires urgent care and is experiencing a life-threatening situation or severe suicidal thoughts, a GAI-based app may misinterpret their distress as mild anxiety and provide generic self-help advice instead of urging immediate medical intervention. While no predictive system can guarantee 100% accuracy, but in this case, a misinterpretation could result in the loss of a life. Therefore, over-reliance on GAI-based mental health apps may, in some cases, delay critical intervention, posing serious risks to users in need of urgent help.

Another example is when a GAI app company violates patient privacy by selling anonymized user data—such as moods, emotional states, sleep patterns, social interactions, daily activities, behavioual trends, diatery habits, and digital engagement —to an insurance agency. The insurance agency could then use this sensitive information to

deny coverage to users identified as at risk of depression or even to those simply registered on the mental health app. This could not only compromises user trust but also raises serious ethical and legal concerns regarding data exploitation and discrimination. This poses a significant privacy risk, especially when mental health predictions are shared with third parties, such as insurance companies, employers, advertisers or any profit-driven organizations. Such misuse could lead to discrimination, denial of services, and the unethical commercialization of sensitive user data and in some cases, GAI-based mental health app companies could exploit users rather than genuinely helping them.

A Strategic Policy Framework for developing GAI-based Apps in Mental Health

While GAI holds significant promise, leveraging its capabilities in mental health care presents unique challenges that must be carefully addressed. Organizations aiming to develop GAI-based mental health applications need to establish a comprehensive framework that reflects the unique capabilities and risks associated with GAI models. Key considerations include technological implications, ethical concerns, data governance, regulatory compliance, system design, technical integrity, and user impact.

To navigate these complexities, we propose the following strategic policy framework for organizations transitioning from traditional AI to GAI in mental health care:

1. Understanding Technological Shift and its Implications

GAI models hold great potential for unique applications in the mental health sector. However, organizations should understand the advantages and disadvantages of the model. Although they can be used to process and analyze large amounts of data for predictive purposes, the results can sometimes be out of context or shallow. Thus, it is vital to establish how such models will be utilized in predicting mental health conditions and their wider consequences. Stakeholders must be informed regarding these abilities with a clear explanation of how predictions are made, the possible limitations, and the significance of human intervention. In this regard, we recommend that all GAI

organizations train healthcare workers about the functioning of GAI. This training strike a balance between avoiding the replacement of automated professional pickers and helping staff understand AI's predictions.

2. Data Governance, Security Measures, and Privacy

Mental health data are very private and sensitive; therefore, companies should put more effort into securing such data by using GAI to develop models. Data breaches or misuse of personal information can have significant impacts because mental health data contain critical information. Thus, organizations must review policies to counteract the challenges transmitted by using GAI. In this context, we recommend that GAI organizations secure data in more sophisticated ways, such as storing and encrypting information. They should ensure that the consent form is properly updated so that patients know that they are using the GAI and regularly audit the systems for vulnerabilities. This approach can help prevent unauthorized access to this information, thus retaining the patient's trust.

3. Ethical Considerations and Bias Mitigation

GAI systems can either create or intensify biased values about culture, gender, and age. In the mental health department, this could lead to many serious problems, such as predictions or inappropriate treatment recommendations. Companies must regularly check and fix AI models to handle these biases and ensure ethical use of AI. In this regard, we recommend that GAI organizations regularly review their AI model outputs for any biases and make necessary corrections to ensure the effective functioning of models. Furthermore, organizations should create an ethics board to review the company's outputs and guarantee compliance with all relevant social responsibility, transparency, and sensitivity regulations. This ethics board will follow the guidelines established by **government** or **lawmakers**, such as the EU, and G7 that are trying to regulate the GAI and to focus on ensuring that the system adheres to ethical standards.

4. Regulatory Compliance and Standards

Every organization using GAI in healthcare, particularly in mental health, must follow current rules described and necessary by **HIPAA** and **GDPR** for all AI development. When new advancements are made in GAI technology, the rules must change, and we believe HIPPA and GDPR will release updated regulations for GAI in the coming months. It is essential to create GAI guidelines in mental health to maintain trust and be a good partner in the healthcare business. We recommend that GAI organizations make routine revisions and update the system to follow all compliances, laws, and regulations. Organizations can also collaborate with regulators to design transparent and trackable best practices and standards for GAI in healthcare development.

5. System Design and Risk Management

The issue of GAI models generating fake or incorrect information, which is usually referred to as "hallucinations," is a prominent threat to mental health predictions. Therefore, organizations must develop their GAI systems in a way that allows human professionals to make decisions regarding patient care, that is, when such specialists have reviewed GAI-generated outputs. In this regard, we recommend that GAI organizations ensure that humans make essential judgments in these applications. They can also include guidance on controlling GAI and other forms of risk in the design, as well as rapid post-launch moderating and error correction disciplines. They should also develop the ability of predictive systems to rapidly and effectively respond to unintended consequences.

6. Cross-functional Collaboration

When organizations develop a GAI predictive mental health application, they must collaborate with different fields and experts, including data scientists, mental health professionals, legal experts, and ethics consultants. Moreover, they must collaborate to ensure that the GAI system is entirely aligned with clinical, ethical, and legal standards. In this regard, we recommend that GAI organizations create diverse teams that regularly meet to discuss and assess the system. They can seek input from external sources such as patient advocacy groups, mental health organizations, and legal consultants to gain varied perspectives. This team-based approach can help ensure the system's safety, effectiveness, and alignment with every requirement.

7. Pilot Testing and Iteration

Pilot tests are essential before a GAI launches an update or application. These tests can help identify problems and areas that require improvement. Pilot tests can be conducted in real situations to make the system more accurate and reliable. In this regard, we recommend that GAI organizations conduct controlled pilot tests of their predictive systems in different clinical environments to receive feedback from mental health experts and patients alike. Use this input to change the system and refine the predictive models and policies to ensure the system works well. They must continue collecting data from cases of early use to make smart updates before implementing it.

8. Communication Strategy

Companies must introduce a GAI system for mental health care that everyone can understand, which means talking to all involved: doctors, patients, and the public. They must know what the system does, how it helps, and where it might fall short. GAI organizations should create two clear plans to talk about the system: one for staff to learn how it works and what changes to expect and the other for the patients; how well it can predict things.

9. Scalability and Future-Proofing

The GAI system should grow and change in light of future developments. As technology and healthcare systems improve, they should handle growth and changes without requiring complete reconstruction. Organizations must ensure that their GAI systems are designed to be flexible and easy to upgrade and expand. They should regularly review and revise the system and associated policies to integrate new technology and address essential healthcare requirements. This approach can help ensure that the system remains valuable and efficient as mental healthcare progresses.

10. Cultural Sensitivity and Localization

A GAI application designed for anxiety management adjusts its therapeutic approaches based on cultural norms. For instance, in some cultures, discussing mental health openly is stigmatized. The system adapts by offering more private coping strategies and using culturally appropriate language and metaphors.

11. Training and Capacity Building for Professionals

Providing training for mental health professionals on how to effectively use GAI tools enhances their ability to deliver care and increases acceptance of new technologies. Clinicians participate in workshops to learn how to interpret data from a GAI-driven mood monitoring app, enabling them to better understand patients' emotional states between appointments and adjust treatment plans accordingly.

12. Prevention of Over-Reliance on Technology

It is essential that clinicians use GAI tools to augment therapy sessions but continue to prioritize in-person or direct communication, ensuring that technology enhances rather than diminishes the human connection essential in mental health care.

Organizations can use the above policy framework and our strategic recommendations for transitioning to GAI, focusing ethics, bias mitigation, user safety, regulatory compliance, and continuous monitoring. It could balance innovation with responsibility and ensure that organizations can leverage the power of GAI while maximizing user trust and minimizing risks in their application development process.
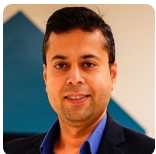
Ajay Kumar (Follow)

Ajay Kumar is an Associate Professor of Business Analytics at EMLYON Business School, France. Before joining EMLYON, Ajay held postdoctoral positions at the Massachusetts Institute of Technology, and Harvard University. Presently, he is also affiliated as a visiting senior fellow with the London School of Economics & Political Science, UK.

Jagpreet Chhatwal (Follow)

Jagpreet Chhatwal is the Director of the MGH Institute for Technology Assessment (ITA) and an Associate Professor at Harvard Medical School. He is also a faculty member of the Center for Health Decision Science at Harvard T.H. Chan School of Public Health and Meditation Research Program at Massachusetts General Hospital and Harvard Medical School. Dr. Chhatwal has co-authored over 100 original research articles and editorials in peer-reviewed journals. His work has been cited in leading media outlets, including CNN, Forbes, National Public Radio, New York Times, and Wall Street Journal.