“My Therapist Sounds like a Robot” – The Peril and Promise of Chatbot Therapy

“My Therapist Sounds like a Robot”

The Peril and Promise of Chatbot Therapy

In recent years, healthcare organizations and providers are increasingly using various forms of artificial intelligence (AI), such as chatbot technology, to interface with patients or to provide personalized communication on a larger scale. In early 2020, the World Health Organization (WHO) created a Chatbot version of its Health Alert platform for Facebook Messenger and WhatsApp. WHO designed the platform to provide accurate and up-to-date COVID-19 information to the public in several languages. The system assisted roughly 12 million people, though the particular technology amounted to little more than an automated customer service system. Instead of dialing a customer service number and selecting answers using a numeric keypad, users clicked through options via the App or Facebook’s website. Limitations in natural language processing (NLP) ultimately limited the utility of this resource, leading many users to seek information elsewhere.

In other corners of the healthcare industry, providers are more aggressively testing the boundaries of medical communication enhanced by or delegated to artificial intelligence. Behavioral health, specifically, stands out for the number of providers trying to wed NLP technology with digital mental health: Woebot, which utilizes a chatbot for cognitive-behavioral therapy (CBT); Ginger, with a chatbot to provide emotional support and connect users with licensed therapists; Wysa, which offers users a way to vent or talk through negative thoughts and feelings; 7 Cups of Tea, which connects users to caring listeners for free emotional support; and Koko, a non-profit founded on the idea that AI could spot individuals at risk of self-harm.

Until recently, the biggest challenge for platform creators using chatbots has been that, while users want a personal, human feel, chatbots have been limited to responses that come across as scripted and automated. Of late, however, AI technology is getting better and better at sounding human. In December 2022, OpenAI launched ChatGPT, a natural language model trained to generate human-like text, and to converse proficiently on a vast range of topics. Last year, it became the second AI to pass the Turing test (following Google’s LaMBDA earlier in the year), which meant that it simulated human-like communication to a level that a judge was unable to tell the difference between the machine and the human control. Those who have experimented with ChatGPT’s free public version have discovered that its capacity to communicate is of a different caliber altogether. Not only can ChatGPT converse at length, but it can spontaneously compose stories, write college papers, translate language, summarize text, troubleshoot code, and much more. Its developers warn that it can also generate factual-sounding answers that are completely false. But for behavioral healthcare, GPT stands out for its advanced cognitive intelligence, or what we sometimes call “soft skills,” such as the human knack for active listening, humor, reflective communication, and the art of reading between the lines.

In December, Koko conducted an experiment using OpenAI’s GPT-3 technology. Over the course of several weeks, around 4,000 people received responses from Koko that were partially or entirely written by artificial intelligence. Unwitting users, many of whom were struggling with depression, PTSD, or anxiety, rated messages composed by the AI as being significantly *better* than those written solely by humans. However, when Koko revealed to users that the messages were composed by a machine, satisfaction plummeted. Koko’s co-founder, Robert Morris, noted that “simulated empathy feels weird, empty.” The company’s revelation led to criticism and accusations of unethical behavior, as users felt they had been tricked into participating in the experiment. But the experiment, poorly executed or not, suggests that AI involvement in behavioral healthcare is likely to remain. With a shortage of mental health professionals and an increase in demand for counseling and psychiatric assessments exacerbated by the pandemic, many no longer question *if* AI can help to fill the gap in mental healthcare, but when and how such changes will be implemented.

There is no shortage of legal quandaries in bringing GPT-like chat technology to mental and behavioral healthcare. Our current system is predicated on the licensing of human professionals, who are empowered by their licenses to

diagnose and treat patients. Is it even legal for machine-driven platforms to encroach on licensed practice? Could purveyors of AI technology be accused of practicing a health profession without a license? The protocols for human involvement or supervision of emerging AI technologies are still in their infancy. In other areas of practice (such as teleradiology), it has become common practice for professionals to review computer analyses (or "read" analyses performed by overseas physicians) to validate the accuracy and completeness of test results. But this layering on of professional oversight in diagnostic healthcare does not translate so easily into mental health. Like self-driving technology, chatbots in direct unscripted communication with patients is a new frontier: the place of human supervision is still being weighed practically, and life-or-death outcomes may hang in the balance.

Data privacy and security is another challenge. What are the limits of AI platforms collection, storage, and use of sensitive personal information (PMI)? Since AI providers cannot bill insurance (with limited exceptions), they are mostly not subject to HIPAA. What is the status of the resulting mental health records reflected in chat conversations? Which confidentiality requirements apply? While human therapists can provide continuity of care by relying on memory and a few jotted notes, AI providers may end up preserving entire session logs so the AI can pick up where it left off for continuity of care from one interaction to the next. How long will it be until exposure or mishandling of sensitive conversations with AI chatbots leads to significant legal repercussions?

GPT technology opens a world of potential malpractice risk in scenarios where an AI provides inaccurate information or inadequate care. So far, ChatGPT can generate responses that sound convincingly human. But it is not sentient. It cannot, in any way, sympathize, empathize, or even understand a patient's unique needs. AI follows the algorithms of its creators, whether it is for conducting intake, screenings, or prognosis. The fact that patients are warned of the risks beforehand is unlikely to indemnify providers. Similarly, dataset biases can result in inferior outcomes for different populations or ethnic groups.

Finally, chatbot therapy raises ethical concerns. In 1968, MIT professor Joseph Weizenbaum conducted an experiment: a proto-chatbot called ELIZA, was designed to mimic Rogerian psychotherapy techniques. Using a simple "pattern matching" technique to trigger pre-written responses, the program asked trial participants about their lives and prompted them to examine their thoughts and emotions. Weizenbaum was shocked to discover that many participants were convinced that the psychotherapeutic chatbot genuinely empathized with them. With GPT, and other language models, which benefit from over five decades of advancements in AI technology, it is important to consider the potential for individuals struggling with mental health issues, such as paranoia, schizophrenia, substance abuse, or even loneliness, to mistake the illusion of sentience for the real thing. This was demonstrated quite poignantly last July, when Google had to dismiss a top LAMBDA engineer who publicly claimed that the search giant's AI was self-conscious and deserving of rights. We can look ahead to a new era of both promise and peril in chatbot therapy.

Authored By:

Harry Nelson, Managing Partner, Nelson Hardiman

Yehuda Hausman, Law Clerk, Nelson Hardiman

Nelson Hardiman LLP

Healthcare Law for Tomorrow

Nelson Hardiman regularly advises clients on new healthcare law and compliance. We offer legal services to businesses at every point in the commercial stream of medicine, healthcare, and the life sciences. For more information, please contact us.