The Winter of Cyber "Mind-Readers":

The Severe Challenges Faced by AI Psychology

EAP Class 21

12111524

Eliot Chen

Imaging that you are struggling with anxiety and depression. Being afraid of seeking help from a licensed psychology therapist, you turned to a popular AI-based mental health app. However, the app collected tons of your personal information and your privacy was exposed on the internet. As a result, your condition worsened. Before delving further into the topic, let us first define the meaning of AI psychology. AI refers to the simulation of human intelligence in machines that are programmed to think and learn like humans. Psychology, on the other hand, is the scientific study of human behavior and mental processes. AI psychology is the interdisciplinary field that create cyber intelligent systems that can analysis and intervene human psychological processes. Although threr are already some application of AI in the field of psychology, basing on the challenges and limitations of AI psychology, this paper will ultimately argue that AI could NOT provide extensive psychological support for humans in the future.

The extensive application of AI in the field of psychology is hindered by numerous formidable technical challenges. Despite significant breakthroughs in AI technologies in recent years such as General

Pre-Trained (GPT) Models, which offer opportunities for deeper observations of psychological phenomena (Schneider, 2023), the achievement of AI providing extensive psychological support for humans in the future faces significant technological gaps in scientific knowledge, AI model training, and human-AI interaction. As one of the most powerful Artificial General Intelligence models available, ChatGPT has been optimized primarily for text-based communication, limiting its ability to interact in the complex and diverse ways required for psychological applications. For instance, in conducting psychological diagnostics and analysis, experts need to perceive individuals' cognition and behaviors, continuously observe and analyze the interactions between their thoughts, emotions, and behaviors, and synthesize different types of information to analyze and predict their psychological states. In the process of psychological interventions and therapies, professionals engage in communication with individuals through techniques like breathing exercises and meditation, and acquire feedback using various specialized methods (Gusinow, 2022). Based on the aforementioned evidence, researchers must recognize the significant gaps in existing scientific knowledge and technology within the field of AI psychology applications. The pursuit of AI models tailored for psychological applications may necessitate the development of new and powerful models that can meet the diverse demands of interactive experiences. This would require substantial progress in foundational sciences, deepening the advancement of applied sciences, and countless efforts from researchers, making it a highly challenging endeavor. In conclusion, the technological gaps pose a fatal obstacle for AI in providing extensive psychological support to humans.

In addition to the technological challenges, one crucial reason why AI cannot provide extensive psychological support for humans in the future is due to the security limitations. Given that AI

psychology requires the digitalization of sensitive user information and that AI-based psychological diagnosis could threaten users' personal dignity and exceed existing technical ethical systems which would create unknown security risks (Quill, L. 2021), the security limitations pose a fatal barrier to the potential development and implementation of AI-powered psychological support. In 2020, the National Health Service (NHS) Apps Library website featured at least 20 mental health apps that underwent a thorough digital assessment, conducted by NHS Digital, which scrutinized the apps for clinical safety, data protection, security, usability, accessibility, interoperability, and technical stability before being included on the NHS site. However, despite this assessment, a lengthy disclaimer absolves the NHS of any liability associated with the purpose of the app. Therefore, inspection agencies and developers have shifted responsibility onto users, and the rules for auditing user data security are not comprehensive, which means that the AI psychology applications cannot guarantee the security of user data (Quill, L. 2021). The security limitations during AI-based psychological diagnosis on personal dignity and technical ethical framework pose a significant obstacle to the potential development and implementation of AI-powered psychological support. In conclusion, security limitations associated with AI psychology represent a significant barrier that difficult to be overcome to realize its potential.

The current state of development in AI psychology may lead some to believe that AI has made certain achievements. For instance, the experimental application of AI-based partner robot "PALRO" has shown promise in providing companionship and psychological assistance to the elderly, aiding them in better adapting to social environments. However, it is important to emphasize that this application is merely an experimental use of AI psychology for non-professional psychological support within a limited population. In order for AI to provide extensive psychological support for humans in the future,

there are two key considerations. Firstly, for professional applications, psychologists and AI developers need to conduct numerous experiments and research to carefully examine the performance measures of AI psychological models, thus further substantiating the effectiveness of AI psychology applications. Secondly, for wide-ranging applications, developers must design and extensively test AI psychological systems to ensure the security, beneficence, and justice of the information being delivered to the public. This ensures that AI psychology applications operate within a secure framework that aligns with technical ethics and human societal morals. In conclusion, the existing small-scale non-professional applications of AI psychology do not directly demonstrate AI's ability to provide extensive psychological support for humans in the future. The further development of AI psychology requires addressing both the professional and large-scale aspects. This implies that psychologists and AI engineers still face significant challenges in this field.

In general, the prospects for AI psychology to provide extensive psychological support for humans in the future are not very promising due to technological barriers and security limitations. Stakeholders from both within and outside the field should objectively acknowledge the significant gap between the current technological level and the technological requirements for widespread AI psychology applications, as well as the existing security risks associated with AI psychology. Generally speaking, it is crucial for us to maintain humility and caution regarding the development of science and technology, objectively assessing the current state and future prospects. This ensures that science and technology genuinely serve humanity's best interests.

Reference

Schneider, S., Lee, J.H. & Mathis, M.W. *Learnable latent embeddings for joint behavioural and neural analysis.* Nature 617, 360–368 (2023). https://doi.org/10.1038/s41586-023-06031-6.

GUSINOW, S. (2022). SECOND NATURE? Health monitoring, remote therapy and predictive Al have emerged to combat the growing mental health crisis - but too much tech is part of the problem. *Oregon Business Magazine,* 45(2), 28–32.

Quill, L. (2021). The software will see you now: The promise and peril of digital therapy. *Counselling & Psychotherapy Research,* 21(3), 498–501. https://doi.org/10.1002/capr.12404.

Masuta, Onishi, T., & Hun-ok Lim. (2012). Human-robot interaction based on frankl psychology for a partner robot. 2012 Proceedings of SICE Annual Conference (SICE), 79–84.

Handle with care: Assessing performance measures of medical AI for shared clinical decision-making. (2022). Bioethics, 36(2), 178–186. https://doi.org/10.1111/bioe.12930.