

eScholarship@UMassChan

The Use of Artificial Intelligence in Suicide Prevention and Mental Health Support

Item Type	Other
Authors	Harris, Leah;Goodman, Erin;Lezine, DeQuincy;McQuade, T'Leah;Glynn, Jennifer;Do, Sara;Strait, Rick;Kelly, Lourah
Citation	Tavory T Regulating AI in Mental Health: Ethics of Care Perspective JMIR Ment Health 2024;11:e58493 URL: https://mental.jmir.org/2024/1/e58493 DOI: 10.2196/58493
Publisher	University of Massachusetts Chan Medical School
Rights	Attribution-NonCommercial 3.0 United States
Download date	2026-03-13 23:20:09
Item License	http://creativecommons.org/licenses/by-nc/3.0/us/
Link to Item	https://hdl.handle.net/20.500.14038/55217



These are questions and guidance for researchers developing or applying for funding related to artificial intelligence for use in mental health and suicide prevention. Persons with a range of lived experiences—from acute to chronic suicidal despair—should be meaningfully included in the design and development of AI tools, to ensure the content is grounded in real world perspectives. See also: Tavory T. (2024). Regulating AI in Mental Health: Ethics of Care Perspective. *JMIR Mental Health*, 11, e58493. <https://doi.org/10.2196/58493>

Questions to Ask Prior to and During AI Development

1. What are the ethics of offering AI tools in place of human interactions?
2. What are the ethics of offering AI if the alternative is for someone to receive no support?
3. Is it possible for therapists or other providers to see a summary or dashboard of interactions with an AI tool, to inform treatment or offer other support?
4. Are people training the AI tool by using it? Does the AI learn from users to inform the general model (what responses are given to all users) or the personal model (what happens in only that person's interactions)—and do users know this?
 - Can users train the tool to learn and suggest a wide variety of coping skills and coping thoughts?
 - Can the AI be programmed to identify more negative perspectives and encourage alternative ways of thinking or more balanced perspectives
 - What safeguards need to be in place to avoid the AI supporting suicidal behavior or worsening of mental health? Have these safeguards been fully tested?
5. People with advanced training and/or degrees with AI and natural language processing should collaborate with mental health experts and people with lived expertise on potential responses and resources

Potential Risks to Mitigate and Considerations Prior to Implementing AI for Suicide Prevention

1. There are risks of AI “hallucinations” and for the AI to respond in unexpected ways
2. What are the safety procedures for use of AI with people when they are in a mental health crisis? If used for mental health, it should be trained on evidence based mental health interventions and expertise, rather than searching the internet
 - How can the tool be blocked from providing support for suicidal or self-harm behavior? On the other hand, people should not have an ambulance or police sent if an AI considers them to be at risk because an AI cannot accurately assess acute suicide risk
3. AI is going to reflect the perspectives of the person using it—and so might encourage unhealthy behaviors or thinking patterns, including reinforcing paranoia
4. AI can potentially be used to identify patterns and intervene with support, but may rely on people being willing to have their phone physically with them and to have their data mined
5. AI tends to have short “memory” and only use information from the same day or previous few hours, but can seem like it does remember a long history
6. People may be looking for reassurance from an AI chat bot, and may benefit instead from advice on how to manage the anxiety or strong feelings that stem from not knowing or being able to interpret social interactions
7. AI may provide too much validation and tell a person they are correct, when a more balanced view might be healthier or realistic
8. AI may misinterpret the language especially of certain cultural or age groups

Potential Safeguards against Harm from AI to Consider

1. Do not name the AI—this makes the interactions seem like they occur with a real person, even if the user was originally told that the tool is AI
2. Do not set daily reminders to engage with the AI, which can fuel reliance on the tool over interactions in real life
3. Build in breaks, since longer daily conversations with AI have been linked to more negative outcomes
4. AI should be programmed to offer support and resources for persons who ask questions about suicide methods or share suicidal thoughts, and program safeguards to prevent the AI from reinforcing these thoughts and plans
 - Consider adding a pop-up, clickable banner, or link in the footer with mental health resources (without users needing to ask about suicide or mental health)

Suggested Citation: CAPES Lived Experience Advisory Board, Harris, L., Goodman, E., Lezine, D., McQuade, T., Glynn, J., Do, S., Strait, R., Kelly, L. M. (2026). *The Use of Artificial Intelligence in Suicide Prevention and Mental Health Support*. Worcester, MA: UMass Chan Medical School, Department of Emergency Medicine, Center for Advancing Practices to End Suicide (CAPES).

This tip sheet was supported by NIMH grant P50MH129701. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

