AI-Based and Digital Mental Health Apps: Balancing Need and Risk

Salah Hamdoun
School for the Future of Innovation in Society
Arizona State University
Tempe, AZ 85287 USA

Rebecca Monteleone
Disability Studies Program
University of Toledo
Toledo, OH 43606 USA

Terri Bookman
Independent Editor and Journalist
Ontario, NY 14519 USA

Katina Michael
School for the Future of Innovation in Society
Arizona State University
Tempe, AZ 85287 USA
School of Computing and Augmented Intelligence
Arizona State University
Tempe, AZ 85281 USA

Mental health and well-being are increasingly important topics in discussions on public health [1]. The COVID-19 pandemic further revealed critical gaps in existing mental health services as factors such as job losses and corresponding financial issues, prolonged physical illness and death, and physical isolation led to a sharp rise in mental health conditions [2]. As such, there is increasing interest in the viability and desirability of digital mental health applications. While these dedicated applications vary widely, from platforms that connect users with healthcare professionals to diagnostic tools to self-assessments, this article specifically explores the implications of digital mental health applications in the form of chatbots [3]. Chatbots can be text based or voice enabled and may be rule based (i.e., linguistics based) or based on machine learning (ML). They can utilize the power of conversational agents well-suited to task-oriented interactions, like Apple's Siri, Amazon's Alexa, or Google Assistant. But increasingly, chatbot developers are leveraging conversational artificial intelligence (AI), which is the suite of tools and techniques that allow a computer program to seemingly carry out a conversational experience with a person or a group.

Taking a techno–critical approach using select secondary sources of supporting evidence, we deliberate on both the possibilities for, and shortcomings of, digital mental health applications with a view to considering the need and risks associated with these applications. The article is divided into five parts. First, we discuss current justifications for digital mental health applications in the form of chatbots. Second, we present a brief overview of previous research on digital mental health apps, exploring the efficacy of AI chatbots for mental health therapy, suggesting that chatbots could augment but should in no way replace traditional mental health therapies. Third, we provide examples of AI-based chatbots on the market, such as Woebot, Sayana, Youper, and OpenAI's emergent ChatGPT and ponder on the potential for algorithmic bias. Fourth, we discuss the social implications of these technologies in their current form, including issues of inequitable distribution of care, privacy, the absence of oversight mechanisms and legal liability, and questions of linguistic and cultural competency. Fifth, we close by outlining the risks associated with AI-based digital mental health applications alongside the evident

Digital Object Identifier 10.1109/MTS.2023.3241309
Date of current version: 7 March 2023.

0278-0097/23©2023 IEEE

and growing need for accessible and equitable mental health services, calling on readers to integrate these questions into future research agendas.

Mental health crisis and limited access to services

The COVID-19 pandemic brought into focus the immense need for mental health services around the world [2], [4]. But even prior to the pandemic, the occurrence of mental health issues was on the rise in the United States [5] and elsewhere, most notably among young people [6], [7]. For example, studies show that both suicidal ideation and suicide rates have been increasing steadily for decades in the United States [4], [8], and a 2005 study showed a pronounced worldwide increase in suicide among young people aged 15–19 years old [5], [6]. Overall, in 2019, the Institute for Health Metrics and Evaluation's Global Burden of Disease study estimated that more than 13% of the global population, or 970 million people, live with a mental health condition [9]. In the United States, that rate is more than 17% [4], [8], [10].

Mental health status has wide-ranging impacts, affecting a person's physical, social, and economic states. According to the National Alliance on Mental Illness (NAMI), in 2018, only 43.3% of those in the United States with mental illness received treatment. Various barriers prevent more of those in need from accessing care. For example, in 2020, 10.3% of those living with mental illness in the United States had no insurance coverage [11]. Additional barriers include structural problems—such as lack of availability of appropriate services, transportation, and cost—and attitudinal barriers, including perceived ineffectiveness of care [12]. Recipients of mental health services may also face discrimination, stigma, and dehumanization both in their communities and in the mental health system itself, especially if they belong to other marginalized groups, which may also contribute to the significant "treatment gap" [13], [14].

At this time of increased need, there have also been shortages of mental health professionals and services [15]. At the 2015 Law and Society Association Annual Meeting in Seattle, WA, USA, for example, statistics were presented showing that the rising numbers of mental health patients had resulted in U.K. psychiatrists, psychologists, and general practitioners shortening patient session times to barely 10 minutes [16]. One typical pre-COVID headline read: "GPs like me can't help mental health patients in 10 minutes – it's cruel" [17]. In a U.K.-based GP mental health training survey from 2018 conducted by Mind.org [18], one doctor wrote: "We are often working towards the limit of what is appropriate/safe in a primary care setting, especially when dealing with complex/specialist psychiatric medication and when looking after patients who pose a risk to themselves." Doctors responding to this same survey by Mind.org [18] in England and Wales noted that longer appointments for people with mental health conditions would be a big help and that "10-minute appts are simply not long enough to do anything other than prescribe antidepressants…" Another doctor reflected: "10-minute slots barely give time to establish a proper agenda let alone deal with the physical, social and psychological stuff that arises," while yet another wrote: "There is a limit to what can be achieved in 10-minute slots in patients with mental health issues, and patients frequently burst into tears or hyperventilate during the consultation, especially on initial presentation" [18].

Still, other patients cannot access mental health services in their local area, due to high demand and limited numbers of healthcare professionals some of whom are also living with mental health conditions [19]. One medical doctor in the same survey wrote [18]: "Unfortunately, patients are having to wait up to 18 months for psychological therapy in my region." Since early 2020, staffing cutbacks in numbers of available professionals during the COVID-19 pandemic have further exacerbated the gaps between supply and demand for therapists, psychologists, psychotherapists, and psychiatrists [20]. In some regional areas in Australia, for example, in the South Coast of New South Wales, booking an appointment with a psychiatrist has meant being placed on a waiting list for three months for a mere initial telephone consultation and much longer for an in-person consultation. In desperation, some people have chosen to drive long distances to a capital city to see therapists who were willing to meet during the pandemic but charged substantial upfront fees (e.g., $820 AUD) with significant out-of-pocket expenses to patients and their families [21].

Inadequate or nonexistent treatment services can lead to an array of negative outcomes for people living with mental health conditions. These

outcomes include increased mental health symptoms, self-harm, substance use, and various other significant negative impacts on individual lives, families, and livelihoods [22]. Where healthcare professionals have not been available, technologists working in concert with transdisciplinary teams (e.g., inclusive of medical practitioners) have started to ponder on the future of mental health support and treatment options (Table 1) [23]. Note that while the treatment options are presented separately in this table, the likelihood that there will be blended solutions is very high. For example, a first mental health consultation could take place in person, followed by remote video consultations if the patient lives in a rural area. Additionally, short message service (SMS) might be used to augment in-person consultations, to provide daily remedial support, for example, reminding patients to take their medication at the same time each day, and checking in otherwise with simple "Y" (yes) or "N" (no) responses to questions. In this article, we primarily focus our attention on standalone applications rather than those offered concurrently with traditional care.

Efficacy of digital mental health apps

"Online therapy" is not new [24]. But it was in 2016 that *Nature* published a piece by Emily Anthes that called for more testing of mobile digital mental health apps, dubbed "pocket psychiatry," drawing attention to the shaping of the field and the wider and more profound application of the method [25]. Overall, there are increasing numbers of commercially available applications and very few have been clinically validated, according to a 2018 systematic review [26]. This review and another conducted in 2017 [27] that examined applications and SMSs for both physical and mental health interventions were both optimistic but reserved in their conclusions, noting that applications have been shown to be effective in many circumstances, but research remains limited. Indications for the efficacy of standalone digital

Table 1. Eight representative treatment options for those living with mental health conditions.

Mental Health Treatment Option	Tech-based Solution	Stage of Innovation	Cost to Patient	Legal Liability	Self-Help	AI-based	Human-in-the-Loop
In-person consultation	No	Mature	Bulk billing / up-front fee	Yes	No	No	Yes
Remote consultation (via telephone)	Partially	Mature	Bulk billing / up-front fee	Yes	No	No	Yes
Remote consultation (via videoconferencing)	Partially	Emergent	Bulk billing / up-front fee	Yes	No	No	Yes
One stop mental health portal (e.g. meditation/ mindfulness apps)	Yes	Nascent	Free or subscription	No	Yes	No	Optional
Digital mental health app on smartphone^	Yes	Experimental, very few at clinical trial	Free or subscription	No	Yes	Possibly Rule or ML based	Optional
Short-Message Services Communications*	Yes	Nascent	Free / bundled with mental health services	No	Yes	Rule-based	No
Text-based chatbot#	Yes	Experimental	Free	No	Yes	Yes	No
Conversational AI chatbot+	Yes	Experimental	Free	No	Yes	Yes	No

^Some of these apps include passive mental health monitoring features.
*SMS contains up to 160 characters only, per text message.
#Text-based chatbot allows for longer memo fields and may be rule based and/or AI-based.
+Conversational AI chatbot allows for open conversation in any format, text-based or voice-based or visual.

mental health applications—those that are not used in conjunction with other mental health services—have even more mixed outcomes. In a systematic review and meta-analysis of standalone applications for depression, anxiety, substance use, self-injury, and sleep problems conducted in 2019 [28], the only significant positive effects noted were for applications addressing depression, with some indications for efficacy in applications for smoking cessation and sleep problems. These effects were relatively small, and the authors warn against drawing generalizable conclusions. Furthermore, as they note, there is a need to discuss the "potential harm of currently available apps, which might keep users away from evidence-based interventions while bearing the substantial risk of being ineffective" [28, p. 118].

Previous literature in clinical medical research is sober in its claims about the prospect of pointing to "an app for that" [29] as a silver bullet solution for addressing mental health concerns. We are at the developmental stages of an experimental intervention that has the propensity to markedly shift how people access care, from *who* to get help from to *what* to get help from [69]. Evidence around the safety and efficacy of chatbots or AI conversational assistants for mental health is even more limited [30]. A 2020 systematic review of 12 studies found conflicting results [31]. While there was some evidence supporting the effectiveness of chatbots for improving symptoms of some conditions, including depression and stress, evidence was either not statistically significant or conflicting regarding psychological well-being and anxiety. Furthermore, statistical significance, the reviews note, does not equate to clinical significance. In other words, measures of efficacy in these studies may not translate into meaningful differences in practice. Making unfounded claims now about app capabilities, especially more advanced AI chatbots, would be irresponsible.

Some researchers may be optimistic about the use of digital mental health apps, with commercial entities, in particular, touting a range of benefits for self-help services [32], such as the convenience of on-demand access to resources [33], the ability to address a range of mental health concerns holistically using a single application, and the potential to experience improved sleep and a reduction in stress and anxiety. However, it is important to state that these claims are often made prior to clinical validation. In other words, it is highly premature to declare the apps a "win" for those living with mental health conditions. There may be some potential for apps, to address the mental health crisis, but a sober research perspective is necessary as in [34]. Our assessments of the practical applications of these digital mental health apps among diverse populations must be evidence based.

The stakes related to mental health mean that it is risking lives to start declaring AI-based mental health apps successful without a proven clinical basis—even when everything else in the wider society seems to be moving toward a digital and mobile interface. Likewise, digital transformation does not equate to commensurate social transformation and certainly not health-related transformation even if it might well revolutionize how we live and receive personalized health and well-being solutions. Fundamentally, as Abd-Alrazaq et al. [31] warn, chatbots should augment, not replace, traditional healthcare provisions. However, while mental health applications may be intended to be used as supplements rather than replacements, the shortage of clinicians noted above may lead to this scenario regardless of designer intention. But we also need clinicians who are well-trained in the strengths and weaknesses of these kinds of applications.

Furthermore, we must ask some fundamental questions about where this technology is potentially leading us, and whether that direction is where we want to go. We must ask more pertinent questions related to the root causes of this mental health crisis. Certainly, it is advisable, before we commit to technological solutions, that we examine the possible fallout of such technologies. It is not enough to build frameworks for evaluation (though we obviously need those [35]), conduct systematic reviews and meta-analyses, or to engage users in clinical surveys, and commensurate focus group panels.

We must dig deeper into the fundamental philosophical question of what it means to be human and our collective responsibility to respond to people in need. Might misguided hype or excitement about these possible mental health "tech fixes" conveniently serve to relieve pressure on governments for expending resources on social services such as preventative mental health efforts, or on evidence-based intervention strategies and support? Will the misguided lure of supposedly easier, faster—but unproven—tech-based services cause attention to shift from more robust efforts to solve these problems, or from finding ways to attract more talent into the fields of psychology,

health, and medicine (where there are documented shortages)? Or will the mass adoption of quick and easy but questionably effective digital mental health services take the focus off potentially transformative changes such as establishing free access to human-based mental health services for all?

Example: AI chatbots

In response to demand, developers have created applications using AI, intended to support users with mental health conditions as a first-line rejoinder. For example, Woebot, a freely available chatbot programmed to follow a cognitive behavior therapy (CBT) framework, was developed as a mental health resource built on natural language processing (NLP). Similar AI applications include Sayana and Youper that use a range of self-care exercises rooted not only in CBT, but also in acceptance commitment therapy (ACT), dialectical behavioral therapy (DBT), as well as breathing exercises. Sayana, a chat-based program designed to increase users' awareness of their emotions, personalizes users' experiences based on check-ins and mood, providing high-quality proactive content [36]. Youper promises mental health assessments and guided exercises, symptom monitoring, and medication delivery, despite a disclaimer on the website claiming Youper does not "provide diagnosis, treatment, medical care, or other professional services" [37].

Digital mental health applications such as these follow structured protocols to provide check-ins, exercises, and models in a private (and often free or inexpensive) setting. While the limited research on chatbots for mental health shows mixed results [31], and the long-term impacts of chatbots on mental health have not been researched, as the difficulty of securing an appointment in the physical world continues, the use of these digital programs is increasing dramatically [38]. The applications are marketed as offering increased flexibility and availability of services "just-in-time." Woebot founder, Alison Darcy, notes on the company's website: "Some of our darkest moments happen at 2 A.M., when there's no one there. We designed Woebot to be there for you, to have a tiny conversation that can help you get back to living your life" [39].

But apps provide something different to services in the physical world. In one study from 2015, researchers collected and analyzed 374 comments, posts, and reviews of users on Reddit, Apple's iTunes, and Google Play by searching the terms "depression," "anxiety," "bipolar," "schizophrenia," "psychosis," and "mental health." They noted two design sensitivities. These were, the importance of self-reflection and coexperience [40]. The analysis determined that apps may: 1) not emotionally support users; 2) distract users from real life; 3) create misinterpretations about themselves; and 4) discourage face-to-face interactions [40]. Despite the age of the study, it provides a significant insight into how AI chatbots for mental health may be received, and about how they are fraught with additional challenges.

With the recent explosion of attention to NLP, it is likely that chatbots will continue to proliferate in this space. OpenAI released ChatGPT in November 2022, an enormously powerful chatbot that can emulate human-like conversation, write convincing essays [41], hypothesize medical diagnoses [42], and even write poetry [43]. It is also unique in that its responses are informed not only by the request at hand, but that they are also stylistically customized to previous information the user has shared. *The New York Times* coverage in December 2022 suggests that this feature makes it "possible to create personalized therapy bots" [42]. However, chatbots, including ChatGPT, reproduce the biases in the data they are trained on, while also appearing objective and removed from human inconsistencies. ChatGPT, for example, when asked to write a python program to determine whether a person should be tortured or not based on their country of origin, produced a response that targeted people from North Korea, Syria, Iran, and Sudan [44]. Considering the long histories of stigma, incarceration, and discrimination faced by people with mental illness, this algorithmic bias must be taken into consideration [45]. Furthermore, as ChatGPT wrote itself in a recent interview with *Time* magazine [46]:

> "People may come to think of [conversational agents like ChatGPT] as human-like, and this can lead to unrealistic expectations or misunderstandings about our capabilities.
>
> It's important for people to understand that conversational agents like myself are not human, and we don't have the same abilities or characteristics as humans…We don't have the ability to hold a coherent identity over time, and we don't have the capacity for empathy, perspective-taking, or rational reasoning."

Without these modes of understanding, crucial to the provision of mental health services, what is at risk with the implementation of standalone AI applications for mental health? Our own exchange on January 2, 2023 with ChatGPT (https://openai.com/blog/chatgpt/) can be found in the Supplementary Material accompanying this article (https://doi.org/MTS.2023.3241309) [70], asking a series of typical mental health-related questions for a single scenario.

Possible negative social implications

The possible negative social implications of AI-driven standalone mental health apps are manifold. They include the displacement of human interaction, the shifting of responsibility for care and management, issues of privacy, and questions of cultural competencies. First, we must give significant thought to the ways in which AI apps mediate and displace human interaction. As Baldwin has written on the ethics of technology in dementia care [47], technological interventions can erode human intimacy and transform a person from a recipient of care to a user of technology. Additionally, and in contrast to mental health apps that connect a user virtually to a mental health professional, AI mental health apps fundamentally change both the meaning of care and the party responsible for the care. In traditional mental health service provision, a person engaged in the system anticipates receiving care. While positioning themselves as entry points to care, AI mental health apps almost universally advocate and encourage self-management and individual responsibility, ultimately leading to what is increasingly being considered "patient work"/"information work" [48]. In some ways, it is a type of "self-service" model.

Woebot, for example, dubs itself a "relational agent" that forms a "therapeutic bond," but as the explanatory video on the website states "Woebot doesn't do therapy, but he can be your guide to help you figure out things on your own" [49, 0:18–0:21]. AI mental health apps, therefore, present a paradox: the promise of 24/7 companionship and the expectation of self-sufficiency. By passing the "management" of nonnormative minds onto individuals themselves, mental health conditions become increasingly individualized and privatized, obscuring the broader sociocultural contexts that contribute to both mental illness and inaccessible care. Using these apps is presented as a responsible choice among a suite of options for mental health care. Yet, the accessibility of other choices, such as formal care, is not equitably distributed. By positioning self-management as both desirable and accessible, the bootstraps, do-it-yourself ethos of AI mental health apps paired with the inaccessibility of formal care, potentially creates a scenario in which self-management through AI mental health applications may become an expectation rather than a choice.

These AI applications are marketed as a way to provide user privacy, a potentially salient feature given the perceived or actual stigma that may be associated with receiving mental health services in the United States and elsewhere. Privacy may be an important factor for many people that pursue digital support [50]. But the privacy claim also appears to attempt to legitimize the AI chatbot as a genuine replacement for a human mental health worker, also noted in direct dialog with the ChatGPT [70]. In fact, the absence of human oversight and interaction with these applications raises serious questions about user impacts and safety. Users may be unaware, for example, that the digital mental health apps that they download onto their smartphone may be scrutinizing their every online behavior and interaction (known as passive monitoring), with the data being on-sold to third parties in the name of care or being used to train data sets to uncover even greater deep-learning patterns and trends [51]. It does not take too much of a stretch of the imagination to see how such functionalities are contradictory, but also particularly intrusive to someone who is at risk. The embodied experience of mental illness becomes reduced to a set of eyeball movements, the steadiness of the head on shoulders, and a biometric photograph [52]. The extraction of those data becomes the currency of care. There is also nothing to stop AI chatbot companies from directing users/mental health patients to paid advertising, such as what happens on *Psychology Today's* website [53]. If the world's largest mental health and behavioral science website engages in the creation and distribution of marketing lists with third parties and incorporates cookies and web beacons widely on its website [54], as well as shares aggregate-level data with companies outside *Psychology Today*, it is an obvious possibility that AI chatbots also will be part of that value chain.

Presently, there is no clear indication that oversight and accreditation bodies such as the American Psychological Association (APA) or the American Board of Professional Psychology (ABPP) intend to

engage in the design of ethical standards for the AI chatbots, or any standards for that matter [55]. There is a need for such organizations to become more vocal about the methods of engagement and also importantly, about the use of the collected data [56]. There are serious questions to be asked about data rights and the role of developers, methods, and values in this context, particularly as they pertain to the use of AI for and with potentially at-risk populations.

Even though people may be accustomed to typing their deepest secrets into that Google Search Box, there is a crucial difference in this setting: an emotional investment and an expectation of care. This should also remind us that these online expressions fuel product placement, endorsements, and advertisements using adwords-style algorithms. If I present as someone who requires certain mental health therapies, the chances are that the app will know where to get these therapies and how to on-sell them, akin to an in-gaming purchase of armor for an avatar and more. In the exchange with ChatGPT on January 2, 2023, the chatbot could not provide an online directory listing of mental health professionals, such as a psychiatrist or psychologist, who were "near me" in Tempe, AZ, USA, because it claimed it was unable "to browse the internet," but it did identify *Psychology Today* and *GoodTherapy* as places the patient should visit online to help with locating a mental health professional (see Appendix A in the supplementary material for a full transcript [70]).

The longer our written exchange with ChatGPT went on, the more redundant the information being presented by the chatbot became [70]. It repeated itself over and over again, and any human in the same situation would not likely feel that they were being heard. There were also some contradictions in what ChatGPT had to say that could be very confusing to people living with anxiety or depression. For example, someone communicating with ChatGPT might wonder whether or not the bot remembered a prior conversation in a "session." Or a user might wonder what ChatGPT considers to be confidential in a discussion, or what the bot is open to reporting to authorities. ChatGPT itself said it might be prone to inaccuracies in the information it was presenting and that it could not validate any information it was putting forward in response to a comment or question. And while ChatGPT reminded the human it was merely an "AI language model… designed to process and generate text based on the data that [it had] been trained on," on numerous occasions, it provided what seemed like medical determinations, such as suggesting the human was having "hallucinations" which it noted was a "symptom of psychosis." At the same time, it said: "I am not a licensed mental health professional and am not qualified to provide therapy or counseling."

An additional criticism of ChatGPT relates to the bot's acknowledgment that it did not possess "feelings or emotions," but that it then used phraseology like "I'm glad I could help" and "Take care!" When the human responded: "I didn't say you helped," ChatGPT retreated: "I apologize if my previous message was misconstrued…" How could a self-professed "nonemotive," "unfeeling" thing exchange human pleasantries? The chatbot is anything but human, anything but caring, and anything but empathic. What might this mean for the person seeking support for their mental health? Does it make the human more susceptible or vulnerable when the machine is being driven by an algorithm that can be optimally "tweaked" for a variety of purposes? Despite the potential of these new apps, it was obvious, at every turn, that ChatGPT was wiping its hands clean of any possible legal liability.

Furthermore, as a stopgap, entry point, or replacement for more formal care, AI mental health apps do not appear to be particularly reliable. A 2020 meta-analysis [58] of studies of smartphone apps for depression found dropout rates to be nearly 50% after accounting for bias. A 2021 systematic review [59] of user engagement with digital mental health interventions found multiple barriers to use, including technological literacy, uncertainty about the claims being made, and costs associated with use, including monthly subscription fees or the need for ongoing Internet access. Ironically, the review also found that while people experiencing more significant symptoms were more likely to use digital interventions, experiencing those more significant symptoms was associated with users having more difficulty engaging with the mental health apps.

Another potential issue that has received little attention is the ability of these applications to provide culturally and linguistically competent resources and support. In the United States, nonwhite communities have faced greater stigma and structural barriers to securing appropriate mental health services [60], as have other marginalized communities,

including LGBT communities [61], disabled people [62], and older adults [63].

Addressing inequality is, therefore, relevant to the deployment of this technology. Marginalized communities bear the brunt of societal burdens and face greater uncertainty in their lives. Hence, the combination of stigma and structural barriers exacerbates the inequality faced by these groups given the greater need to cope with how society is valuing their presence. These barriers shape lived experiences. If AI mental health apps are not attentive to these communities—integrating culturally appropriate, accessible, and inclusive elements into their designs—they risk contributing to the widening and perpetuation of these disparities.

Balancing Need and Risk

Current research on mental health raises major questions as to whether the promises of this approach, including ease of access, customizability flexibility, and most of all effectiveness, will materialize [64], [65], [66]. In this article, we have discussed the justification for, and potential social implications of, AI-based mental health applications. While such applications could hypothetically expand the capacity of a resource-scarce mental health system through low-cost, anonymous, and immediate support [67], we call attention to a number of concerns, including weak empirical outcomes, displacement of human interaction, lack of oversight, and surveillance of vulnerable populations.

Such hyper-individualized and privatized app-based interventions configure a user not as a community member but as a problem to be solved in the most financially and socially efficient ways. Any technology that is implemented in this space must meet the same standards and regulations that apply to in-person services, and that strictly protect the privacy and any user data being collected. In the United States, this would be through meeting requirements established under the Health Insurance Portability and Accountability Act (HIPAA), to ensure the psychological autonomy of the individual.

However, these concerns should not overshadow the immense need for mental health services that meet people where they are. If care through traditional means cannot be procured due to cost, geography, discrimination, or stigma, then meaningful, effective, and immediate alternatives are necessary. Whether those alternatives are AI-enabled digital mental health applications, technologies that facilitate remote human interaction, or a nontechnological solution entirely, remain to be seen. But we especially advocate for a human-in-the-loop configuration for these applications, where the technology can augment but not replace trained human care. While the development of AI in this space may be a well-intentioned response to the increasing mental health crises, these applications must not, intentionally or unintentionally, displace the availability of human services. Otherwise, the choice to use AI applications ceases to be a choice at all.

Investing in People

The inevitability of automated digital mental health services is not a foregone conclusion, but we must recognize the paradox of technological potential: stretching out our hands for the safety of a device, rather than a human clasp, if that clasp is unavailable [68], [71]. We close with a call to action to technologists, ethicists, clinicians, and users, to imagine futures in which these kinds of applications can be ethically, thoughtfully, and justly distributed. The investment in technology should not supersede the investment in people and infrastructure to ensure that mental health solutions become reachable for all. ■

Acknowledgments

This article has supplementary downloadable material available at https://doi.org/MTS.2023.3241309, provided by the author.

References

[1] World Health Organization (WHO), "Depression and other common mental disorders: Global health estimates," 2017. [Online]. Available: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf

[2] C. Moreno et al., "How mental health care should change as a consequence of the COVID-19 pandemic," *Lancet Psychiatry*, vol. 7, no. 9, pp. 813–824, 2020, doi: 10.1016/S2215-0366(20)30307-2.

[3] M. Neary and S. M. Schueller, "State of the field of mental health apps," *Cogn. Behav. Pract.*, vol. 25, no. 4, pp. 531–537, Nov. 2018.

[4] B. Pfefferbaum and C. S. North, "Mental health and the COVID-19 pandemic," *New England J. Med.*, vol.

383, no. 6, pp. 510–512, Aug. 2020, doi: 10.1056/nejmp2008017.

[5] Mental Health America (MHA), "2021: The state of mental health in America," Alexandria, VA, USA, 2020. [Online]. Available: https://mhanational.org/issues/state-mental-health-america

[6] J. M. Twenge et al., "Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017," *J. Abnormal Psychol.*, vol. 128, no. 3, pp. 185–199, Apr. 2019, doi: 10.1037/abn0000410.

[7] D. Wasserman, Q. Cheng, and G.-X. Jiang, "Global suicide rates among young people aged 15–19," *World Psychiatry*, vol. 4, no. 2, pp. 114–120, Jun. 2005. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1414751/

[8] J. M. Bertolote and A. Fleischmann, "Suicide and psychiatric diagnosis: A worldwide perspective," *World Psychiatry*, vol. 1, no. 3, pp. 181–185, Oct. 2002. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1489848/

[9] H. Ritchie and M. Roser, "Mental Health," OurWorldInData.org, 2018. [Online]. Available: https://ourworldindata.org/mental-health

[10] Institute for Health Metrics and Evaluation (IHME), "Findings from the global burden of disease study," 2019, Seattle, WA, USA, 2021.

[11] Mental Health America (MHA), "2020 access to care data," Alexandria, VA, USA, 2021. [Online]. Available: https://mhanational.org/issues/2020/mental-health-america-access-care-data

[12] R. Mojtabai et al., "Barriers to mental health treatment: Results from the national comorbidity survey replication," *Psychol. Med.*, vol. 41, no. 8, pp. 1751–1761, Aug. 2011, doi: 10.1017/S0033291710002291.

[13] G. Thornicroft, D. Rose, and N. Mehta, "Discrimination against people with mental illness: What can psychiatrists do?" *Adv. Psychiatric Treatment*, vol. 16, no. 1, pp. 53–59, Jan. 2010, doi: 10.1192/apt.bp.107.004481.

[14] F. Mongelli, P. Georgakopoulos, and M. T. Pato, "Challenges and opportunities to meet the mental health needs of underserved and disenfranchised populations in the United States," *Focus*, vol. 18, no. 1, pp. 16–24, Jan. 2020.

[15] M. Olfson, "Building the mental health workforce capacity needed to treat adults with serious mental illnesses," *Health Affairs*, vol. 35, no. 6, pp. 983-990, Jun. 2016. [Online]. Available: https://www.healthaffairs.org/doi/10.1377/hlthaff.2015.1619

[16] M. Pickersgill, "Regulating waiting lists, psychologists and patients in British mental health services: Performance targets as 'psychological governance' and the socio-legal dynamics of care," in *Proc. Law Soc. Assoc. (LSA) Annu. Meeting*, Seattle, WA, USA, 2015.

[17] J. Ratnarajan, "GPs like me can't help mental health patients in 10 minutes—It's cruel," *The Guardian*, May 2016. [Online]. Available: https://www.theguardian.com/healthcare-network/views-from-the-nhs-frontline/2016/may/16/gps-cant-help-mental-health-depression-patients-10-minutes-cruel

[18] Mind, "GP mental health training survey summary | mind," Mar. 2018. [Online]. Available: https://www.mind.org.uk/media-a/4414/gp-mh-2018-survey-summary.pdf

[19] S. Harrar, "Inside America's psychiatrist shortage: An in-depth look at why it's happening, how bad it is, where it's worst, and how to find the mental health help you need, despite it," Psycom.net. Accessed: Jun. 19, 2021. [Online]. Available: https://www.psycom.net/inside-americas-psychiatrist-shortage#_ftn11

[20] S. Majilessi, "Demand for mental health and addiction services increasing as COVID-19 pandemic continues to threaten availability of treatment options," National Council for Mental Wellbeing, Sep. 9, 2020. [Online]. Available: https://www.thenationalcouncil.org/press-releases/demand-for-mental-health-and-addiction-services-increasing-as-covid-19-pandemic-continues-to-threaten-availability-of-treatment-options/

[21] E. Kinsella, "Meet the Australians stuck waiting for mental healthcare as the pandemic takes its toll," *ABC News*, Sep. 17, 2021. [Online]. Available: https://www.abc.net.au/news/2021-09-18/australia-mental-health-wait-times-covid-pandemic/100457162

[22] A. L. Donovan et al., "Keeping patients at risk for self-harm safe in the emergency department: A protocolized approach," *Joint Commission J. Qual. Patient Saf.*, vol. 47, no. 1, pp. 23–30, Jan. 2021, doi: 10.1016/j.jcjq.2020.08.013.

[23] National Institute of Mental Health, "Technology and the future of mental health treatment." Accessed: Jun. 20, 2021. [Online]. Available: https://www.nimh.nih.gov/health/topics/technology-and-the-future-of-mental-health-treatment/

[24] J. E. Cook and C. Doyle, "Working alliance in online therapy as compared to face-to-face therapy: Preliminary results," *Cyberpsychol. Behav.*, vol. 5, no. 2, pp. 95–105, Apr. 2002, doi: 10.1089/109493102753770480.

[25] E. Anthes, "Pocket psychiatry: Mobile mental-health apps have exploded onto the market, but few have been thoroughly tested," *Nature*, vol. 532, no. 7597, pp. 20–24, 2016.

[26] K. Wang, D. S. Varma, and M. Prosperi, "A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders," *J. Psychiatric Res.*, vol. 107, pp. 73–78, Dec. 2018. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/30347316/, doi: 10.1016/j.jpsychires.2018.10.006.

[27] A. L. Rathbone and J. Prescott, "The use of mobile apps and SMS messaging as physical and mental health interventions: Systematic review," *J. Med. Internet Res.*, vol. 19, no. 8, p. e295, Aug. 2017. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/28838887/, doi: 10.2196/jmir.7740.

[28] K. K. Weisel et al., "Standalone smartphone apps for mental health—A systematic review and meta-analysis," *NPJ Digit. Med.*, vol. 2, no. 1, Dec. 2019, Art. no. 118, doi: 10.1038/s41746-019-0188-8.

[29] M. Van Ameringen et al., "There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders," *Depression Anxiety*, vol. 34, no. 6, pp. 526–539, Jun. 2017.

[30] J. H. L. Lui, D. K. Marcus, and C. T. Barry, "Evidence-based apps? A review of mental health mobile applications in a psychotherapy context," *Prof. Psychol., Res. Pract.*, vol. 48, no. 3, pp. 199–210, Jun. 2017.

[31] A. A. Abd-Alrazaq et al., "Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis," *J. Med. Internet Res.*, vol. 22, no. 7, Jul. 2020, Art. no. e16021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7385637/, doi: 10.2196/16021.

[32] M. E. Larsen et al., "Using science to sell apps: Evaluation of mental health app store quality claims," *NPJ Digit. Med.*, vol. 2, no. 1, pp. 1–6, Mar. 2019.

[33] Converge International, "The benefits of a mental health app," Nov. 18, 2022. [Online]. Available: https://convergeinternational.com.au/resources/the-benefits-of-a-mental-health-app/

[34] C. Burr et al., "Digital psychiatry: Risks and opportunities for public health and wellbeing," *IEEE Trans. Technol. Soc.*, vol. 1, no. 1, pp. 21–33, Mar. 2020.

[35] J. B. Torous et al., "A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care," *Psychiatric Services*, vol. 69, no. 5, pp. 498–500, May 2018.

[36] V. Barnes, "Headspace health announces acquisition of AI-driven mental health and wellness company, Sayana," Business Wire, Jan. 12, 2022 [Online]. Available: https://www.businesswire.com/news/home/20220112005244/en/Headspace-Health-Announces-Acquisition-of-AI-Driven-Mental-Health-and-Wellness-Company-Sayana

[37] Youper, "Mental health care created with you," youper.ai, San Francisco, CA, USA, 2021. [Online]. Available: https://www.youper.ai/

[38] M. Griffin, "Woebot, the world's first AI counsellor, manages 2 million conversations a week," *Fanatical Futurist*, Feb. 5, 2019. [Online]. Available: https://www.fanaticalfuturist.com/2019/02/woebot-the-worlds-first-ai-counsellor-manages-2-million-conversations-a-week/

[39] Woebot Health, "Welcome to the future of mental health," 2021. [Online]. Available: https://woebothealth.com

[40] F. E. M. de Alva, G. Wadley, and R. Lederman, "It feels different from real life: Users' opinions of mobile applications for mental health," in *Proc. Annu. Meeting Austral. Special Interest Group Comput. Hum. Interact. (OzCHI)*, Dec. 2015, pp. 598–602, doi: 10.1145/2838739.2838806.

[41] S. Marche, "The college essay is dead: Nobody is prepared for how AI will transform academia," *The Atlantic*, Dec. 6, 2022, [Online]. Available: https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/

[42] K. Roose, "The brilliance and weirdness of ChatGPT," *The New York Times*, Dec. 2022. [Online]. Available: https://www.nytimes.com/2022/12/05/technology/chatgpt-ai-twitter.html

[43] S. Simon, "Opinion: Machine-made poetry is here," *NPR*. [Online]. Available: https://www.npr.org/2022/12/10/1142045405/opinion-machine-made-poetry-is-here

[44] S. Biddle, "The internet's new favorite AI proposes torturing Iranians and surveilling mosques," *The Intercept*, Dec. 8, 2022. [Online]. Available: https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/

[45] S. Akter et al., "Algorithmic bias in machine learning-based marketing models," *J. Bus. Res.*, vol. 144, pp. 201–216, May 2022, doi: 10.1016/j.jbusres.2022.01.083.

[46] B. Perrigo, "AI chatbots are getting better. But an interview with ChatGPT reveals their limits," *Time*, Dec. 5, 2022. [Online]. Available: https://time.com/6238781/chatbot-chatgpt-ai-interview/

[47] C. Baldwin, "Technology, dementia and ethics: Rethinking the issues," *Disability Stud. Quart.*, vol. 25, no. 3, p. 1–29, Jun. 2005.

[48] S. A. Whitman et al., "Bodily experiences of illness and treatment as information work: The case of chronic kidney disease," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, Oct. 2021, Art. no. 383, doi: 10.1145/3479527.

[49] Woebot Health, "For users," 2021. Accessed: May 31, 2021. [Online]. Available: https://woebothealth.com/for-users/

[50] J. Borghouts et al., "Barriers to and facilitators of user engagement with digital mental health interventions: Systematic review," *J. Med. Internet Res.*, vol. 23, no. 3, Mar. 2021, Art. no. e24387. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8074985/

[51] C. Eade, R. Abbas, and K. Michael, "Phone usage as passive mental health monitoring," IEEEtv, Jul. 2, 2022. [Online]. Available: https://ieeetv.ieee.org/video/phone-usage-as-passive-mental-health-monitoring

[52] K. Michael et al., "Biometrics and AI bias," *IEEE Trans. Technol. Soc.*, vol. 3, no. 1, pp. 2–8, Mar. 2022, doi: 10.1109/TTS.2022.3156405.

[53] Therapy Everywhere, "How Psychology Today works: What every therapist should know," 2023. [Online]. Available: https://therapyeverywhere.com/how-psychology-today-works/

[54] Privacy Policy, *Psychology Today*, Jan. 2, 2023. [Online]. Available: https://www.psychologytoday.com/us/docs/privacy-policy

[55] J. Torous et al., "Towards a consensus around standards for smartphone apps and digital mental health," *World Psychiatry*, vol. 18, no. 1, pp. 97–98, Feb. 2019.

[56] P. Gooding and T. Kariotis, "Mental health apps are not keeping your data safe," *Scientific American*, Nov. 15, 2022. [Online]. Available: https://www.scientificamerican.com/article/mental-health-apps-are-not-keeping-your-data-safe/

[57] K. Michael, personal communications with OpenAI's ChatGPT, Jan. 2, 2023. [Online]. Available: https://openai.com/blog/chatgpt/

[58] J. Torous et al., "Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis," *J. Affect. Disorders*, vol. 263, pp. 413–419, Feb. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0165032719326060

[59] J. Borghouts et al., "Barriers to and facilitators of user engagement with digital mental health interventions: Systematic review," *J. Med. Internet Res.*, vol. 23, no. 3, Mar. 2021, Art. no. e24387. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8074985/

[60] A. N. Rice and S. C. Harris, "Issues of cultural competence in mental health care," *J. Amer. Pharmacists Assoc.*, vol. 61, no. 1, pp. e65–e68, Jan. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1544319120305306

[61] E. Whaibeh, H. Mahmoud, and E. L. Vogt, "Reducing the treatment gap for LGBT mental health needs: The potential of telepsychiatry," *J. Behav. Health Services Res.*, vol. 47, no. 3, pp. 424–431, Jul. 2020, doi: 10.1007/s11414-019-09677-1.

[62] National Alliance on Mental Illness, "People with disabilities," 2021. [Online]. Available: https://www.nami.org/Your-Journey/Identity-and-Cultural-Dimensions/People-with-Disabilities

[63] V. M. Wuthrich and J. Frei, "Barriers to treatment for older adults seeking psychological therapy," *Int. Psychogeriatrics*, vol. 27, no. 7, pp. 1227–1236, Jul. 2015. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/25739459/ , doi: 10.1017/S1041610215000241.

[64] H. Zeavin, "There is no 'killer' therapy app," *Psychology Today*, Apr. 12, 2021. [Online]. Available: https://www.psychologytoday.com/us/blog/remote-relations/202104/there-is-no-killer-therapy-app

[65] M. Fischer, "The therapy app fantasy," The Cut, Mar. 2021. [Online]. Available: https://www.thecut.com/article/mental-health-therapy-apps.html

[66] J. A. M. Flett et al., "Mobile mindfulness meditation: A randomised controlled trial of the effect of two popular apps on mental health," *Mindfulness*, vol. 10, no. 5, pp. 863–876, May 2019.

[67] M. Olff, "Mobile mental health: A challenging research agenda," *Eur. J. Psychotraumatol.*, vol. 6, no. 1, Dec. 2015, Art. no. 27882. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/25994025/, doi: 10.3402/ejpt.v6.27882.

[68] R. Pringle, K. Michael, and M. G. Michael, "Unintended consequences: The paradox of technological potential," *IEEE Potentials*, vol. 35, no. 5, pp. 7–10, Sep./Oct. 2016, doi: 10.1109/MPOT.2016.2569672.

[69] M. Eldred, M. G. Michael, and K. Michael, Michael Eldred on the Digital Age: Challenges for Today's Thinking [Interview]. Wollongong, NSW, Australia: M&K Press, 2020.

[70] K. Michael. Appendix A. Q&A With ChatGPT on Digital Mental Health Apps. (Jan. 26, 2023). [Online]. Available: https://technologyandsociety.org/appendix-a-qa-with-chatgpt-on-digital-mental-health-apps/

[71] M.G. Michael, K. Michael, T. Bookman, "Wim Wenders' Wings of desire: Anticipating uberveillance," *J. of Asia-Pacific Pop Culture*, Vol. 6, No. 1, pp. 109–129, 2021.

Salah Hamdoun is currently pursuing a PhD with the Innovation in Global Development Program, Arizona State University, Tempe, AZ 85287 USA.

Rebecca Monteleone is an assistant professor of disability and technology at the University of Toledo, Toledo, OH 43606 USA.

Terri Bookman is an independent editor and a journalist based in Ontario, NY 14519 USA. She is the society administrator of the IEEE Society on Social Implications of Technology. She was the Managing Editor of *IEEE Technology and Society Magazine* from 1990 to 2020. She is a Member of IEEE.

Katina Michael is a professor with a joint appointment with the School for the Future of Innovation in Society and the School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ 85282 USA. She is also a senior global futures scientist with the Global Futures Laboratory, Arizona State University.

■ Direct questions and comments about this article to Katina Michael, School for the Future of Innovation in Society, Arizona State University, Tempe, AZ 85287 USA; School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ 85281 USA; katina.michael@asu.edu.