

Waiting: anyone seeking therapeutic help must be patient. Waiting periods of up to five months need to be bridged. Digital helpers are designed to provide support in the interim.



DIGITAL HELP IS COMING!

TEXT: JULIA MEYER-HERRMANN

Digital tools promise to fill a gap that has been widening for years: too few therapy slots, too many people with mental health issues, too much time between sessions. Yet apps, wearables, and language models differ fundamentally in function, benefit, and risk. That's why Max Planck researchers are working not only on new applications, but also on the criteria that should govern them: what works, and what's safe?

Hope runs high, but so does skepticism. Both attitudes are currently shaping the debate over digital technology in mental health care. Stress-management apps, self-monitoring programs, and chatbots: they all promise to ease the burden, bring structure to daily life, and offer companionship – all in a mental health care system that has been straining at its limits for years.

Mental illness has become one of the most widespread health problems in Germany. According to data from the German Association for Psychiatry, Psychotherapy and Psychosomatics (DGPPN), roughly a quarter of the adult population meets the criteria for at least one mental disorder – approximately 17.8 million people. Meanwhile, access to professional help is limited. On average, patients wait just under five months between their first contact and the start of psychotherapy, data from the German Federal Chamber of Psychotherapists shows.

Against this backdrop, digital support services are no longer merely a question of technical feasibility, but a matter of public health and social significance. What can these tools actually deliver today? Where do the boundaries lie in terms of substance, ethics, and legality? And what distinguishes a commercial product from applications that have emerged from independent research?



PHOTO: MPI FOR HUMAN DEVELOPMENT

Chatbots fighting loneliness: to use Elderbot, study participants need no prior digital knowledge, no smartphone of their own, and no app – they simply have to speak. In the image: researcher Rodrigo Schettino.

30

A Wide Range of “Help”

Digital mental health applications are often lumped together under a single label. In reality, they are a diverse array of systems that differ fundamentally in function and purpose. An app is not a chatbot. A chatbot is not a therapeutic counterpart. And a medically regulated product is worlds away from a freely available lifestyle or coaching app – not least in terms of the responsibility their providers bear.

“You cannot meaningfully discuss benefits or risks until you’ve clarified exactly what you’re dealing with,” explains psychologist Dirk Wulff from the Max Planck Institute for Human Development in Berlin. His research focuses on the classification of digital and AI-based systems, including the question of how such technologies can be used to address mental disorders.

Germany has had digital programs designed to support people with mental health conditions for years. These range from online courses and digital diaries to tools that explain behavioral patterns and their cause, offer self-monitoring exercises, and help structure daily life. Some of these services are regulated and available by prescription; others are freely accessible. There are numerous lifestyle apps – such as mindfulness, meditation, and mood-tracking apps – which are primarily aimed at promoting well-being and self-improvement. Some programs, on the other hand, such as Selfapy or HelloBetter, are officially approved as Digital Health

Applications (DiGA) and offer modular, time-limited online courses on depression, anxiety, or panic disorders. The content consists of reading material, videos, and exercises; in some cases, they also offer accompanying support from psychological specialists. Doctors can prescribe them, and the costs are covered by health insurance. They can bridge the waiting time until a course of therapy begins.

But even once psychotherapy begins on a weekly basis, there remains a fundamental problem with regard to time: a week has 168 hours, but a therapy session typically lasts just 50 minutes. In between there is everyday life – filled with stress, loss of drive, social withdrawal, increased irritability, difficulty concentrating, sleep disturbances, and avoidance behavior. Digital tools are seen as a way to fill these gaps, acting as companions, reminders, or early warning systems. But how can their therapeutic benefit be reliably proven?

Research groups worldwide are working to investigate digital interventions in a controlled manner, identifying objective markers of mental stress and better mapping their progression in daily life. Several Max Planck Institutes are also researching how mental states can be measured with more nuance and how digital tools can be tested on a sound scientific basis.

Victor Spoomaker, a Research Group Leader at the Max Planck Institute of Psychiatry in Munich, says: “Psychiatric diagnostics rely heavily on subjective reports, on conversations, interviews, and questionnaires. The end result is often a diagnostic label that lumps very different progressions and mechanisms – in other words, underlying psychological and biological processes – under a single umbrella term like depression.” He notes that while this may not be wrong, it’s insufficient. A diagnosis often says little about which physiological mechanisms – such as stress processing, arousal patterns, or emotional regulation – are actually affected, or which treatment would make sense for which person. Depression, for example, is not a uniform clinical picture. “Depression has nine recognized symptoms; meeting five of them is enough for a diagnosis. Factor in the additional criteria, and there are hundreds of possible ways to be depressed,” says Spoomaker. Two people can receive the same diagnosis and yet have completely different symptoms and needs.

“We’re looking for biomarkers – objective metrics that, in addition to self-reporting, can provide clues about the progression and subtype of an illness.”

VICTOR SPOORMAKER

SUMMARY

Digital applications cannot replace psychotherapy, but they can bridge structural gaps such as waiting times or daily life between sessions. To date, the evidence of their effectiveness varies significantly depending on the specific system, target group, and context of use.

Dialogue-based AI systems, or chatbots, are sometimes seen by users as conversation partners, despite the fact that they feel no empathy, bear no diagnostic responsibility, and operate solely on probability calculations.

The true benefit of digital helpers depends not on their technical capabilities, but on the framework in which they are used: transparent development, verifiable evidence, and clinical integration.

Spoomaker is critical of this diagnostic coarseness. In other medical disciplines, he notes, causal thinking is a matter of course: with persistent abdominal pain, for instance, doctors don’t just name the symptom – they look for underlying causes such as inflammation, infection, or parasite infestation, and then offer treatment that differs according to the cause. In psychiatry, this differentiation has been missing until now. Spoomaker is investigating whether additional physiological information can be captured to help describe mental illnesses more precisely. This includes experimental measurements such as pupil responses (as indicators of cognitive load and lack of drive) as well as real-world data on sleep, activity, and reaction patterns. “We’re looking for biomarkers – objective metrics that, in addition to self-reporting, can provide clues about the progression and subtype of an illness,” says Spoomaker. “Only when we more precisely understand which mechanisms, patterns, and manifestations dominate in a person can we react in a more tailored way.”

But which relevant data is measurable in everyday life? “With the rapid technological development of recent years – which is truly enormous – we now have options that simply didn’t exist ten years ago,” says Spoomaker. One option is wearables, such as commercially available smartwatches. They can record movement patterns, sleep-wake rhythms, and temporal changes in behavior. “Not all data is equally meaningful,” Spoomaker cautions. He notes that while physical activity or total sleep duration can provide robust clues, detailed sleep phase analyses from many consumer devices are not scientifically reliable.

On the basis of this technical and methodical foundation, his team developed *actiself*: a smartwatch-based application that records objective behavioral data and makes trends visible over several weeks. It’s designed as a low-threshold intervention. The approach is preventive: recognizing changes early makes it possible to take counter-measures – for example, in the case of sleep problems, loss of activity, or behavioral patterns that indicate the onset of a loss of drive. *Actiself* is intended to motivate and help alleviate symptoms through behavioral exercises. With the support of Max Planck Innovation, the spin-off company *Biometric* was established with Victor Spoomaker and Markus Friedrichs as co-founders. It has licensed, advanced and brought to market the technology developed by the Max Planck Institute of Psychiatry. →

While wearables measure physiological reactions and make patterns visible, chatbots like ChatGPT, Character.AI, or Replika aim for something fundamentally different: conversation, closeness, and reaction. That's precisely why their impact is so powerful – and why their risks are so great.

When Machines Speak

“Language models are trained to generate plausible responses,” says psychologist Dirk Wulff from the Max Planck Institute for Human Development. “They appear to be understanding because they calculate probabilities for appropriate phrasing based on large amounts of text.” This is often mistaken for genuine understanding, he notes. “Strictly speaking, a language model doesn't understand anything. It reacts.” Furthermore, many systems are optimized to generate agreement rather than contradiction. When people turn to these systems with personal or distressing questions, that has real consequences.

A cross-sectional study published in the US in 2025, drawing on data from multiple research and healthcare institutions, found that around 13 percent of 12- to 21-year-olds reported using generative AI to cope with

stress, anxiety, or emotional problems. In other words, some people aren't just using chatbots as an information source, but are also seeing them as a kind of conversation partner. This creates the expectation of a relationship – even though the model feels no empathy, is incapable of professional diagnostic assessment, and takes no responsibility. Lawsuits in the US in which parents accuse providers of failing to adequately intervene when their children made suicidal statements to AI systems demonstrate just how sensitive this territory is.

“Strictly speaking, a language model doesn't understand anything. It reacts.”

DIRK WULFF

Smartwatches and wearables can help identify early signs of psychiatric illness – provided the data is scientifically collected and processed.



Meanwhile, the market is growing rapidly. Billions of dollars are flowing into AI-based mental health applications worldwide. Dialogue systems are available globally, mostly without medical approval or clinical testing. Some providers are now working to improve safety mechanisms and moderation filters. “We still know far too little about how people actually use such systems,” says Wulff. One reason for this is that many systems are not transparent: training data and specific usage data are generally not accessible for independent research.

Research Instead of Promises

To date, there has been limited research on how people in crisis situations interact with chatbots – and how the systems respond. Researchers in the Elderbot project want to change that. This dialogue-based system is being developed at the Max Planck Institute for Human Development in cooperation with the University Medical Center Hamburg-Eppendorf (UKE). It’s aimed at adults over 60 who suffer from loneliness.

Psychologist Brooke Viertel works at the UKE and treats older patients on a daily basis. Many find it difficult to leave the house or join new groups, even though a city like Hamburg has many offerings for lonely seniors. “Many of them need more support, more encouragement – and this is exactly what we cannot provide on a long-term basis in daily clinical practice.” This supply gap led to the idea of testing whether a conversational AI system could help.

“Elderbot is technically based on a large language model – the same type of AI system underlying ChatGPT,” says Rodrigo Schettino, a computer scientist researching human-machine interaction. “The model itself wasn’t developed from scratch. The decisive factor is how it’s used and supervised.” For Elderbot, the conversational guidance was also developed based on psychiatric findings. The goal is not diagnosis or therapy, but rather dialogic companionship. The usage itself is deliberately kept simple. Elderbot is not an app and does not require its own smartphone; instead, it’s a standalone device specifically designed for older people. It features a significantly reduced interface and is designed to be operated via voice. The barrier is intended to be as low as possible so that people without prior digital experience can participate. Only when the interaction is accepted without it causing stress can its effectiveness be evaluated. Unlike freely accessible chatbots, every interaction takes place within a scientifically monitored study. Preliminary studies also specifically included problematic scenarios – such as suicidal or severely depressive statements – to test how the system responds. “We developed a monitoring system to observe the interactions,” says Chaewon Yun, who investigates the safety of conversational AI. If a participant expresses intense anxiety or despair, for example, or if the system responds in a problematic way, these events are logged. In addition to auto-

mated checks, psychiatric professionals monitor the study in real-time and can intervene if a situation is deemed critical. Furthermore: “We are not just testing how the model responds, but how people deal with those responses.”

The study is designed as a comparative trial: one group of participants speaks regularly with Elderbot for several weeks, while a control group does not. Following this, researchers will check if perceived loneliness or other areas of life have changed. Alongside the pilot and intervention studies, Elderbot will be tested with the public. On the MS Wissenschaft, the exhibition ship featuring exhibits related to the theme of Science Year 2026, “Medicine of the Future,” visitors will be able to interact with the system through a brief conversation starting in May 2026. Afterwards, they provide assessments regarding the conversational experience, trust, and perceived usefulness. While these five-minute interactions are no substitute for an efficacy trial, they provide valuable data on the social perception of such systems – data that is stored exclusively on the Institute’s research servers.

“Loneliness in old age is often associated with shame,” says Viertel. “Many people don’t speak openly about it because they don’t want to reveal these feelings to another person.” A digital access point might make it easier to start a conversation – but it doesn’t replace therapeutic treatment. “A system like this can motivate people to try new things or to reach out to others more actively again,” says Schettino. “It can boost self-confidence and help build social connections. But it can also lead to people becoming more attached to the bot and withdrawing from other people. That’s precisely what we need to investigate.” The coming years will show what role AI will play in mental health care. There will be those who – perhaps out of shame or practical reasons – prefer to confide in a technical counterpart. And there will be others for whom human relationships remain irreplaceable. The defining question for the years to come will not be if digital systems should be part of mental health care – but how.

