MASTERARBEIT | MASTER'S THESIS

Titel | Title

AI tools for mental healthcare: Perspectives of (potential) users and developers

verfasst von | submitted by

Philipp Christian Lutsch BA MA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 914
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Internationale Betriebswirtschaft
Betreut von | Supervisor:	Univ.-Prof. Christoph Fuchs PhD
Mitbetreut von | Co-Supervisor:	Lic. Mestr. Cátia Alves Ph.D.

1. ABSTRACT

In recent years, artificial intelligence (AI) has become ubiquitous in mobile applications. This trend has also impacted mental healthcare: As therapies with human professionals are costly and cannot match the high demand for treatment worldwide, AI applications provide the ability to reach a larger number of affected people at a lower cost. Research mostly focuses on the medical effectiveness of these products. From a marketing standpoint, however, creating effective apps requires considering the attitudes and views of (potential) users about AI in mental healthcare. The frequent app modifications in this sector need a continuous evaluation of these perceptions. I used Kozinets' Netnography as a qualitative method to examine and contrast user experience (UX) of six popular chatbot apps with viewpoints from two other stakeholder groups: My analysis encompasses more than 1,900 user reviews from Google Play and the Apple App Store, inputs from the public (mainly from non-users) on Reddit and X, in addition to interviews with relevant app providers. The findings suggest that the human touch is the most crucial aspect in varying perspectives between users and non-users of AI apps for mental health. While critics complain that relevant apps lack empathy and human emotions, users frequently regard this absence as the primary reason for using such products, as they desire disclosures without being judged. Unlike common opinion on the web, users see chatbots for mental health as a complement rather than a replacement for in-person treatment. My observations also indicate that users are more sensitive to pricing strategies of providers compared to previous years and that generative AI, particularly ChatGPT, is having an increasing impact on mental health apps.

2. ABSTRACT (GERMAN TRANSLATION)

In den letzten Jahren ist künstliche Intelligenz (KI) in mobilen Anwendungen allgegenwärtig geworden. Dieser Trend hat auch Auswirkungen auf die psychische Gesundheitsversorgung: Da Therapien mit Fachleuten kostenträchtig sind und die hohe Nachfrage nach Behandlung weltweit nicht decken können, bieten KI-Anwendungen die Möglichkeit, eine größere Anzahl betroffener Personen zu geringeren Kosten zu erreichen. Die Forschung konzentriert sich hauptsächlich auf die medizinische Wirksamkeit dieser Produkte. Aus Marketing-Sicht erfordert die Gestaltung effektiver Apps jedoch die Berücksichtigung der Einstellungen und Ansichten (potenzieller) User*innen zur KI im Bereich der psychischen Gesundheitsversorgung. Die häufigen App-Änderungen in diesem Sektor bedingen eine kontinuierliche Bewertung dieser Wahrnehmungen. Ich verwendete Kozinets' Netnographie als qualitative Methode, um die User Experience (UX) von sechs populären Chatbot-Apps zu untersuchen und mit den Ansichten von zwei anderen Interessengruppen zu vergleichen: Meine Analyse umfasst mehr als 1.900 User*innen-Bewertungen aus dem Google Play Store und dem Apple App Store, Beiträge aus der Öffentlichkeit (hauptsächlich von Nicht-User*innen) auf Reddit und X sowie Interviews mit relevanten App-Anbieter*innen. Die Ergebnisse deuten darauf hin, dass der menschliche Kontakt der entscheidende Aspekt bei den unterschiedlichen Perspektiven zwischen User*innen und Nicht-User*innen von KI-Apps für psychische Gesundheit ist. Während Kritiker*innen bemängeln, dass relevanten Apps Empathie und menschliche Emotionen fehlen, betrachten User*innen dieses Fehlen häufig als den Hauptgrund für die Nutzung solcher Produkte, da sie sich offenbaren möchten, ohne beurteilt zu werden. Im Gegensatz zur allgemeinen Meinung online sehen User*innen Chatbots für die psychische Gesundheit als Ergänzung und nicht als Ersatz für persönliche Behandlung. Meine Beobachtungen zeigen auch, dass User*innen im Vergleich zu den Vorjahren sensibler auf die Preisstrategien der Anbieter*innen reagieren und dass generative KI, insbesondere ChatGPT, einen zunehmenden Einfluss auf Apps für die psychische Gesundheit hat.

3. LIST OF TABLES

4. LIST OF FIGURES

5. LIST OF ABBREVIATIONS

ACT	Acceptance and Commitment Therapy
AGI	Artificial General Intelligence
AI	Artificial Intelligence
ANI	Artificial Narrow Intelligence
ASI	Artificial Super Intelligence
BERT	Bidirectional Encoder Representations from Transformers
CA	Conversational Agent
CBT	Cognitive-Behavioral Therapy
CEO	Chief Executive Officer
cf.	Compare
DBT	Dialectical Behavior Therapy
DMH	Digital Mental Health
EU	European Union
FAQ	Frequently Asked Questions
GDPR	General Data Protection Regulation
GPT	Generative Pre-trained Transformer
IPT	Interpersonal Psychotherapy
LLM	Large Language Model
MAC	Metaphoric Associative Cards
ML	Machine Learning
N/A	Information not Available
NLG	Natural Language Generation
NLP	Natural Language Processing
NLU	Natural Language Understanding
PLM	Pre-trained Language Model
PST	Psychological Skills Training
UI	User Interface
UK	United Kingdom of Great Britain and Northern Ireland
US/USA	United States of America
USD	US-Dollar
UX	User Experience
WHO	World Health Organization

6. TABLE OF CONTENTS

7. INTRODUCTION AND RESEARCH QUESTION

This study lies at the crossroad of two major developments. The first one is a high prevalence of mental health disorders: At present, more than 10 % of worldwide cases of illness come from human mind (Jones and Vigo, 2023, p. 197). Experts call this a global burden, highly developed economies are three times more affected than those with lower development (Jones and Vigo, 2023, p. 198). In a recent survey 52 % of respondents from 31 countries stated that in 2022 they felt depressed almost daily over several weeks – 66 % of them belonged to the group of Generation Z and Millennials (Ipsos, 2023, pp. 12–13). Of note, these estimates may be understated as stigma precludes people from reporting mental issues. Conventional mental health treatment is unable to address this global burden.

Another important development is the digitization of mental healthcare. As technology offers more affordable, faster, and better healthcare, it may provide relief for the mental health crisis (Falzon et al., 2023, p. 384). Especially artificial intelligence (AI) can be a source for advancements in medicine and Iqbal et al. regard this as *"the next best step"* (2023, p. 2). In recent years, providers have made a large number of mental health applications available to the public.

According to Spruijt-Metz et al. (2015, p. 124), it is challenging to keep up with the rapid advancements of mental health apps. In recent years, there has been a steady increase in evidence on mental health chatbots. A large majority of reviews concentrate on the medical efficacy of these solutions. The research on user acceptance and motivations for usage is much more limited and overlooks the AI component (Nadarzynski et al., 2019, p. 2). Abd-Alrazaq et al., 2021; Ahmed et al., 2022; Alqahtani and Orji, 2020; Chaudhry and Debi, 2024; Haque and Rubya, 2023; Malik et al., 2022; as well as Prakash and Das, 2020 examined how users perceived existing mental health chatbots in the last years. However, they did not incorporate perspectives of non-users or app providers. Furthermore, their studies pertain to past product versions – to keep up with constant app changes, Ahmed et al. (2023, p. 3) regard reviews on a regular base as necessary.

This thesis builds on that previous research and investigates the most recent perceptions of mental health chatbots **from early 2023 to mid-2024**. To address this issue, the **research question** is:

- **What are attitudes and views of (potential) users and relevant developers towards AI usage for treatment of mental health disorders?**

In addition, two **sub-questions** break down the main question:

SQ 1: *What do relevant developers regard as core features of their AI-featured products?*

SQ 2: *What do (potential) users and relevant developers see as main challenges/limitations of such products?*

Previous research suggests that business potential, public interest and professional self-interest have always been the core elements in establishing communications technology in medicine in the United States (Greene, 2022, p. 253). Therefore, I draw on this idea and assess the first two factors from a marketing standpoint. Moreover, I use an innovative approach by comparing the perspectives of three stakeholder groups – **the public online, users, and developers** – employing the Netnography method.

To create a broad context, the research starts by summarizing the key features of eHealth, mHealth, and Digital Mental Health (DMH). They have made the current advancements of mental health apps possible. Chapter 8 also covers what artificial intelligence (AI) is, its opportunities, as well as challenges for mental healthcare today. Chapter 9 explains how the Netnography method contributes to answer the main research question and its two sub-questions. My study concludes with a presentation of the findings. The discussion in chapter 11 compares the results that were first divided among the three stakeholder groups in chapter 10.

8. LITERATURE REVIEW

8.1. Digitalization of healthcare

8.1.1. Potentials of digital healthcare (eHealth/mHealth)

Technology usage for health can be conceptualized with the term **eHealth**, it places the patient, the user, into the center (Gemert-Pijnen et al., 2018, p. 7). A first definition of the concept appeared in 1999, however eHealth was not an entirely new phenomenon at that time (Duplaga, 2021a, pp. 20–21). Telemedicine as predecessor reached back to the late 19th century, when

medics distributed their services without direct contact to their patients for the first time (Duplaga, 2021a, p. 18).

Three reasons make healthcare an important domain for digital transformation:

1. Technology can be scalable and omnipresent in human life, compared to human professionals technology can reach a higher number of people (Kelders and Howard, 2018, p. 50).
2. Also, technology can persuade people persistently without fatigue, while offering more anonymity to users (Kelders and Howard, 2018, pp. 49–50).
3. Digital means enable swift stakeholder communication and efficient data management – resulting in decreased costs (Falzon et al., 2023, p. 384).

Other authors see eHealth as new perception of patients, labelled as *patient empowerment*: This *"re-distribution of power between patients and physicians"* (Johnston Roberts, 1999, p. 82) means that patients prefer convenience to time-consuming bureaucracy within healthcare, while taking more control of their medical treatment (Ball and Lillis, 2001, p. 2). *Patient empowerment* plays an important part in the treatment of chronic diseases, such as mental health disorders (Duplaga, 2021b, pp. 63, 65).

Usage of mobile technology – any kind of sensory devices – for health is labeled as **mHealth**, a term coined by Robert S. H. Istepanian in 2003 (Duplaga, 2021a, pp. 22–23). Researchers distinguish between two main areas of use: **[1]** mobile apps and **[2]** wearable fitness devices (Mitra, 2023, pp. 4–5). The smartphone can be considered as the primary device for mHealth (Saxena et al., 2021, p. 171). The iPhone introduction in 2007 was decisive for the establishment of mHealth: Mobile app providers created entire ecosystems for their products, fueled by large capital investments (Istepanian, 2022, pp. 5–6). Despite Istepanian and Al-Anzi (2018, pp. 35–37) criticize the dominance of smartphones and private app companies in mHealth, they believe it has also improved the sector by introducing new technologies, such as big data and AI.

The offer of healthcare apps grew considerably after 2010, but it only met modest demand in those years: In the US alone, just one in ten adults with a smartphone downloaded such a product (Malvey and Slovensky, 2014, pp. 65–66). Schueller (2018, p. 96) regards the year 2016 as breakthrough for mHealth, when internet access on mobile devices surpassed that on desktop devices in the US. Mobile internet access has two major effects: Users are almost constantly online, while this behavior is further characterized by shorter, but more frequent online sessions per day (Schueller, 2018, p. 97). With mHealth apps as *"just-in-time behavior*

change tools" (Spruijt-Metz et al., 2015, p. 123), patients may be reached in their natural environment, where traditional healthcare and stationary digital solutions cannot compete.

It is uncertain how many mHealth solutions at present exist: Given that there is no clear distinction between therapeutic and wellness apps, estimations might not be very accurate (American Psychiatric Association, 2023, p. 3; Mitra, 2023, p. 23). In mid-2024, the Apple App Store registered over 35,000 relevant apps, while Google Play Store offered over 36,000 relevant products (Appfigures, 2024a, 2024b).

8.1.2. Constraints of digital healthcare via mobile devices (mHealth)

Statistics, such as mentioned before, underline the high potential of the mHealth market perceived by relevant companies. However, research suggests that the medical sector is too conservative for the implementation of mHealth – the main reason, why mobile app providers have launched their products without scientific foundation (Malvey and Slovensky, 2014, p. 8). For Istepanian (2022, p. 2) the amount of conducted clinical and evidence-based studies of healthcare apps is high, but their evidence on medical effectiveness is limited. Digitalization may assist to eliminate the ivory tower of medical experts, as Meskó et al. (2017, p. 4) remind. However, patients would not benefit from this situation, if such experts were excluded from developing digital products: Information might be unreliable or misunderstandable (Meskó et al., 2017, p. 5). Malvey and Slovensky (2014, p. 89) also warn about high expectations: Mobile apps do not necessarily facilitate behavior change, if there is no willpower inflamed. Just sending constant reminders may not be effective, as users can perceive them as intrusive after a while.

User retention and thus profitability continues to be a major challenge for mHealth app providers: Eysenbach (2005, p. 2) discovered significant dropout rates in clinical trials and called it the *law of attrition*, as participants generally tend to stop usage of a digital health application abruptly. Aboujaoude and Gega (2020, pp. 3–4) support this interpretation: Despite a strong pull effect at the beginning, digital therapies without human intervention tend to be prematurely terminated, as ties within virtual relationships are generally loose. Outside clinical trials, in real-world settings, dropout rates can be even higher (Bickmore et al., 2010, p. 291). In fact, up to 80 % of users only minimally engage with their health app, with logins less than twice (Meyerowitz-Katz et al., 2020, p. 2). Even products with high quality standards can have dropout rates of 50 % within 3 months, as Nwosu et al. report (2022, p. 2). In a real-world experiment, most users of a health chatbot already terminated after five conversation rounds,

with the highest peak after one conversation round (Fan et al., 2021, p. 6). Most users discontinued app usage, when they had to answer symptom questions, which seemed overwhelming, or they were asked to answer the same question again (Fan et al., 2021, p. 6). Additionally, Torous et al. (2018, p. 117) mention that users likely drop out, if they feel that data entry after installation is time-consuming and if they perceive permanent symptom tracking as work. In general, attrition can be reduced, if **[1]** users perceive an idea as superior to existing solutions, **[2]** the innovation fits into societal values and meets user needs, **[3]** users feel that the innovation is not too complex and, **[4]** the innovation can be tried out first (Eysenbach, 2005, p. 5). Involving users in app development from the start is another key factor to reduce later attrition (Linardon and Fuller-Tyszkiewicz, 2020, p. 8). Many researchers believe that human support can improve app adherence, but the study of Renfrew et al. (2020, pp. 2, 10) shows no better outcome with human surveillance.

Scalability, the *"potential for a business to replicate its activities on a wider scale"* (Law, 2016), is another decisive factor for the profitability of mHealth solutions. Following Boucher (2021, p. 38), scalability is also beneficiary for users: In the case of mental health chatbots, it can make them widespread for a large population. However, Narasimhadevara & Rambhatla (2020, p. 154) make the point that scalability is generally restricted for mHealth. Research has detected two main reasons for this situation: Currently there are no common quality standards for such products and internationally streamlined privacy laws are also absent (Sainz-de-Abajo et al., 2021, pp. 72, 75). A third factor comes from Davies and Mueller (2020, p. 19): They argue that broad *access* to mobile technologies does not automatically make apps for mental health *accessible*. Such products are generally text-heavy with no speech-to-text features, they often exclude blind or illiterate people. WHO discovered that of approximately 20,000 relevant apps people of non-White ethnicities and older than 65 years faced problems with adoption (The Lancet Digital Health, 2021, p. e204). Current statistics also indicate that the mHealth market appears unstable. The aforementioned 35,000 to 36,000 relevant applications in the two app stores represent a decline, as over 50,000 products were recorded on Google Play and more than 60,000 on the Apple App Store in 2021 (Appfigures, 2024a, 2024b).

8.1.3. Potentials and constraints of digital mental healthcare (DMH)

As shown before, the eHealth/mHealth concept covers all areas of healthcare. To address mental health issues, researchers have introduced the digital mental health (DMH) concept,

focusing on *"[...] digital technologies in mental healthcare [...], including wellbeing promotion and prevention, [...]"* (Bond et al., 2023, p. 1).

Substantial hopes lie on digitizing mental healthcare for four reasons:

1. In contrast to other healthcare areas, feared stigmatization inhibits people with mental health issues to seek professional help (Kruzan et al., 2022, pp. 7–8).
2. Lack of professionals for most people affected, resulting in long waiting lists worldwide. Even if a professional was available, the barrier of stigmatization dominates. Kruzan et al. (2022, pp. 7–8) argue that anonymity is a key factor to address mental health patients successfully.
3. Digital products may ensure processual conformity to therapeutic methods like CBT *(cognitive behavioral therapy)* and ACT *(acceptance and commitment therapy)*, which research sees harder to achieve within face-to-face therapies (Kruzan et al., 2022, pp. 7–8).
4. Another reason people may select digital resources for mental healthcare is convenience, according to Hanley and Wyatt (2022, pp. 78–82): Most people already use the internet to find out about therapies and available therapists in their area. Digital natives, as well as an increasing number of digital immigrants, appreciate efficient and effective digital communications. This interest may lead to trying a digital tool. Furthermore, those with disabilities or who travel frequently may have difficulty accessing on-site therapy (Hanley and Wyatt, 2022, pp. 78–82).

In traditional psychotherapy, the quality of relationship between a therapist and a patient – termed *therapeutic alliance* (TA) – is widely regarded as decisive for therapy success (D'Alfonso, 2020, p. 114). According to Horvath et al. (2011, p. 15), a therapist has to match client's expectations with available resources and the intended intervention to enable success. To reach *therapeutic alliance*, CBT and ACT are commonly recognized as standard methods (Newman, 2017, p. 738; Smith et al., 2017, p. 7). CBT helps patients to learn and apply certain skills for their therapy: Patients self-monitor themselves and collect data for later analysis (Newman, 2017, p. 738). ACT is a development of CBT and focuses on mindfulness and acceptance (Smith et al., 2017, p. 7). According to Schueller (2018, p. 102), most DMH products currently available use CBT and ACT.

Despite those strengths, DMH's reputation suffered in the past decades. Current problems are almost identical to those of eHealth and mHealth mentioned before: Researchers argue that public regulation failed to keep pace with the advancement of internet technology and DMH (Aboujaoude and Gega, 2020, p. 109). In 2020, only 2 % of over 1,000 wellness apps publicly

available were backed by scientific evidence (Ratheesh and Alvarez-Jimenez, 2022, p. 107). Even if well-conducted, research may be already outdated: Advancements of DMH and relevant research appear at different paces (Aboujaoude, 2018, p. 277). In academia, structures of reimbursement inhibit innovation: Product ideas based on evidence may exist, but only their discovery, not their later deployment, is funded (Kruzan et al., 2022, p. 12).

Although there is currently limited data, users' willingness to pay is another difficulty for DMH with regard to private providers: By contrasting an app with in-person treatment, Lorenzo-Luaces et al. (2024, p. 5) examined how price-sensitive US ethnic groups were to mental health care: The majority of respondents – at least 60 % – wanted to use the app for free; rates were higher among Black, Asian, and Hispanic respondents. Less than 25 % of respondents wanted to use the same app when the authors set a monthly price of USD 13.— (Lorenzo-Luaces et al., 2024, p. 5). By contrast, a standard Netflix subscription in late 2024 cost USD 15.49 per month (Netflix Help Center, 2024). However, according to Lorenzo-Luaces et al. (2024, p. 5), at least 50 % of the respondents would have spent USD 100.— for four sessions of face-to-face treatment every month. The highest amount that participants were willing to pay was USD 60.— for the app and USD 800.— for in-person therapy (Lorenzo-Luaces et al., 2024, p. 3). Overall, the respondents showed a strong price sensitivity towards the app, but much more tolerance towards the higher costs associated with in-person therapy.

In contrast to those critical views, Torous et al. (2021, p. 319) highlight the positive effects of centralized and private mobile platforms, like Google Play and Apple App Store, on DMH: Easier than before, ideas can turn into marketable products, which may reach millions of potential customers globally. This shift to mobile devices in DMH has led to an unprecedented amount of products offered, for almost any mental health issue (Torous et al., 2021, p. 319).

The next section shows which effects researchers are expecting from artificial intelligence on healthcare.

8.2. Artificial intelligence (AI) for mental healthcare

8.2.1. Potentials and constraints of AI in medicine

As shown before, many scholars regard digitalization as disruptive for healthcare. Artificial intelligence, *"the simulation of intelligent behavior in machines"* (Luxton, 2020, p. 248), plays a key role in that respect. In general, AI deals with *"[...] independently processing large*

amounts of data, recognizing patterns and leading to descriptions, predictions, prescriptions or even autonomous decisions based on them" (Kreutzer and Sirrenberg, 2020, p. 8).

Contemporary medicine is facing a series of challenges: Medical knowledge is increasing exponentially and physicians need to keep up with it, while chronic diseases have become more complex to diagnose and therapize, as Chang (2023, p. 80) explains. AI could interpret these enormous amounts of data to relieve workforce, while patients would receive tailored care (Chang, 2023, p. 81). AI's benefits for healthcare exist on three levels for Ali et al. (2023, pp. 8–11):

1. **Healthcare sector:**
   - AI saves time, optimizes resources and reduces expenditures (especially regarding chronic diseases)
   - Large datasets are available, which can be shared sector-wide
2. **Healthcare institution (as single unit):**
   - AI improves workflows, assists in them and reduces expenditures
3. **Patient:**
   - Compared to humans, AI analyzes heterogenous medical data with more precision and effectiveness
   - Diagnosis can happen earlier
   - AI enables automated and swift decision making
   - Patient's health state can be monitored constantly

Simon and Aliferis (2024, p. 533) have further arguments for AI implementation in healthcare:

- Updating machines with new data and knowledge can be rapid
- AI is scalable, while human capabilities remain limited

However, according to Greenhill and Edmunds (2020, p. 86) the healthcare sector has not been a forerunner to adopt this technology. Professionals remain reluctant to implement AI (Kaul et al., 2020, p. 808). Aung et al. (2021, pp. 8–11) mention five current drawbacks of AI implementation in medicine:

1. **Data challenges:**
   - To conduct tasks accurately, AI requires large datasets *(big data)*. Data sharing between healthcare institutions is not the rule, relevant data is often fragmentary and not always of sufficient quality.

- Additionally, sensitive information is prone to data breaches. Regulations, like the EU's *General Data Protection Regulation* (GDPR) can be a solution, but they may also hinder data availability to train AI systems. Aristidou et al. (2022, p. 620) add that it mostly remains unclear to users, which training data had been used.

2. **Developer challenges:**
   - AI can reduce human bias, but AI data itself may be biased. Iqbal et al. (2023, p. 8) warn that underrepresented groups of a society can receive inaccurate results or even face discrimination.
   - Algorithms lack transparency and create "black box" systems. The connection between input data and predictions mostly remains unknown to users.
3. **Clinical implementation challenges:**
   - Most research of medical AI does not happen in real-world settings. Products may not fit seamlessly into clinicians' workflow, if those stakeholders are excluded from development phase.
4. **Social challenges:**
   - Because of high expectations, AI is often misunderstood by the public. Currently, there is no possibility in reach to substitute healthcare workers by AI. Fears of human substitution create common distrust in AI technologies.
5. **Ethical challenges:**
   - Who can be held responsible if AI systems fail, make incorrect assumptions, or injure patients is still up for debate. WHO has repeatedly urged for a more ethical and regulated approach to AI development (World Health Organization, 2023a, 2023b). Currently, there is virtually no regulation of AI use in healthcare. As exception, the EU's *AI Act* went into effect on August 1, 2024, which aims to only allow trustworthy AI inside the EU (European Commission, 2024).

Complementary to the those drawbacks, there are also technical constraints, as Oshida (2022, pp. 11–13) remarks: All solutions in use presently are based on **Artificial Narrow Intelligence** (ANI) and operate within preset limits. ANI has good capabilities for doing routine and repetitive activities, but it cannot compete with human intelligence. Such capabilities would only include **Artificial General Intelligence** (AGI) or **Artificial Super Intelligence** (ASI), which would provide machines with consciousness, emotion, and creativity. So far, this is beyond reach.

To conclude, present AI can increase the understanding of customers and make existing processes more efficient, as Kreutzer and Sirrenberg (2020, p. 276) note – higher automation of human tasks can only be expected within the next decades.

8.2.2. Purposes of AI in mental healthcare

Various researchers share the vision that AI would make mental healthcare more reliable: Lu et al. (2023, p. 1606) state that mental clinicians rather use tacit knowledge in their daily work, which is characterized by subjectivity and informality. Inefficiencies characterize mental healthcare, as Minerva and Giubilini (2023, pp. 813–814) point out: Human assessment is traditionally error-prone, when professionals recommend anti-depressants or predict patient's risk of suicide. Support for this interpretation comes from Sharma et al. (2023, p. 6), claiming that AI can improve efficiency and accuracy of mental health diagnostics and treatment. However, it would be erroneous to presume AI might replace such human tasks, as Boucher et al. (2021, p. 45) assert: Most mental health chatbot providers have never aimed to create replacements but rather to enhance user engagement with their offerings.

Two major aspects differentiate mental healthcare from other medical areas:

1. AI adoption takes more time within mental healthcare, due to a more reluctant workforce (Graham et al., 2019, p. 2).
2. Mental healthcare is text-heavy (Lovejoy, 2019, p. 2). As language may be considered *"a window into the human mind"* (D'Alfonso, 2020, p. 112), words are more relevant than images. This requires technologies, which translate and understand (often unstructured) human language and extract relevant information (Graham et al., 2019, p. 3).

D'Alfonso (2020, pp. 112–114) identifies three primary areas of AI use in mental healthcare: **[1] Personal sensing,** in which digital devices, primarily mobile ones, collect data via integrated sensors (such as walking speed or screen usage) to predict mental health issues. **[2] User content analysis,** where a system analyzes interview transcriptions or users' social media content to determine mental health conditions and **[3] chatbots:** A system that simulates user discussions through an interface. Conversations can benefit from the context that integrated knowledge bases may provide (Luxton, 2020, p. 252). User content analysis and personal sensing are possible features of chatbots.

The next sections will concentrate on chatbots for mental health because of their widespread availability.

8.2.3. AI chatbots as mental health tools

8.2.3.1. Technological background of chatbots

Chatbots are **conversational agents** (CAs), which *"[…] process the input and provide relevant advice or feedback in a form of text or speech […]"* (Allouch et al., 2021, p. 1). Only CAs with text-based interfaces can be considered as **chatbots,** while speech-based interfaces are part of voice assistants, like Apple's Siri (Dingler et al., 2021, p. 191). The alternative expression *chat robot* reveals the role of automation within such machines (Brandtzæg and Følstad, 2018, p. 40). Figure 1 illustrates the main components of a chatbot and the way they interact:

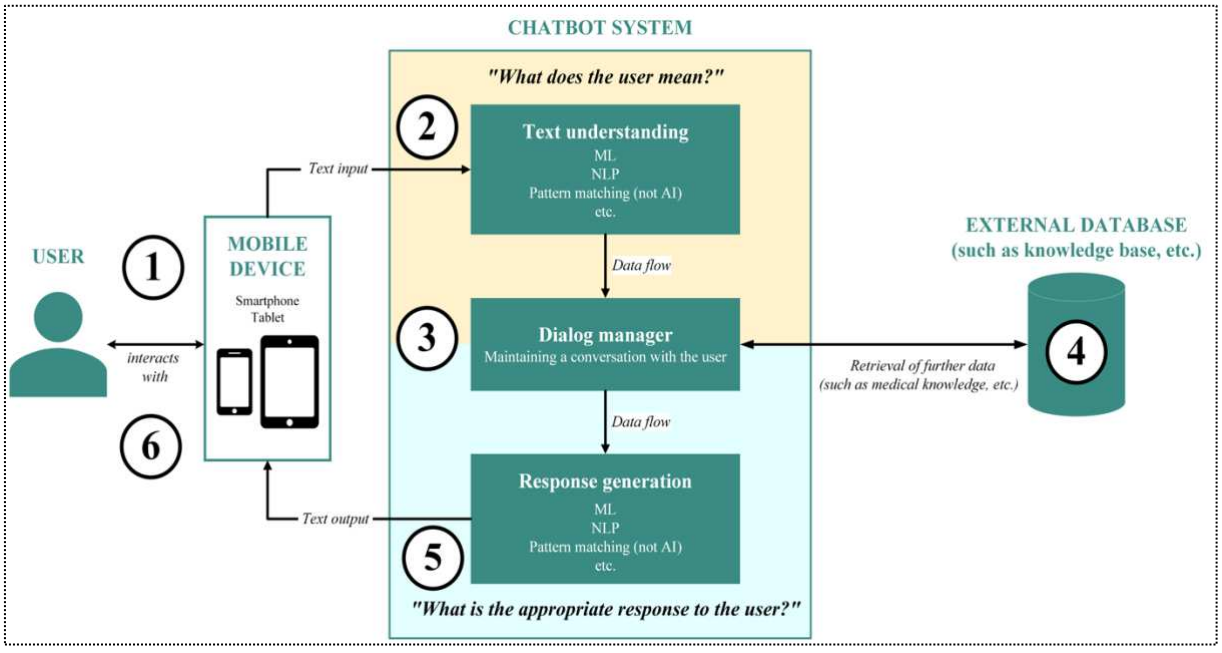

Figure 1: Basic structure of a chatbot with and without AI capabilities (adapted from Bilquise et al., 2022, p. 5; Safi et al., 2020, p. 4). Note: Additional features, like personal sensing and user content analysis, are left out.

While **scripted chatbots** solely use pattern matching, **AI chatbots** allow more intricate open discussions (Bendig et al., 2022, p. 68; Xue et al., 2023, p. 2). The integration of AI into chatbots is a comparatively recent issue: The advancement occurred with the release of numerous AI chatbots in 2016 (Adamopoulou and Moussiades, 2020, p. 3). As a reason for this development,

Bilquise et al. (2022, p. 3) see the need to better forecast user intent and meet expanding user expectations.

AI related to chatbots is termed as *conversational AI* and relies on two core technologies:

- **Natural Language Processing (NLP):**

  Programming languages like Python are used by computers to communicate, while human language is used for user input, which requires machine translation (Xu and Roberts, 2022, p. 213). As seen in Figure 1, NLP works in a chatbot as follows: **[1]** The user creates input and produces text using an interface (a smartphone or tablet). **[2]** NLP deconstructs this input using a set of principles that transform unstructured data into structured data. *Natural Language Understanding*, or NLU, is a subset of *Natural Language Processing*. It attempts to understand the meaning of the input after receiving structured data. The language layers of morphology, syntax, semantics, and pragmatics collaborate for this stage (Kreutzer and Sirrenberg, 2020, pp. 25–27). **[3]** A dialog manager uses those analyses to determine how to handle user input. **[4]** It may utilize a database for further information for this purpose. **[5]** After the dialog manager determines how to respond, *Natural Language Generation* (NLG), a different branch of NLP, converts machine commands back into natural language and sends them to the user interface, whereupon the user receives an intelligible response to his initial input.

  Even with the most recent developments, natural language ambiguity continues to be a problem for this technology. Punugoti et al. (2023, p. 1) assert that as human language may be understood in a variety of ways, proper NLP implementation is the most important factor to consider when developing an AI chatbot. Facts, self-revelation, appeals, and relationships with others are only a few of the layers that always exist in messages (Kreutzer and Sirrenberg, 2020, pp. 25–27). Differing vocabularies and dialects within a language exacerbate the difficulty of AI chatbots comprehending text (Kreutzer and Sirrenberg, 2020, p. 25). NLP's role in mental health chatbots is to assist users with various activities (Cafiero et al., 2023, p. 255).

- **Machine Learning (ML):**

  Machines learn from experience and large data amounts to make predictions for new data (Luxton, 2016, p. 4). Independently learning algorithms process data after a programmed statement (Kreutzer and Sirrenberg, 2020, p. 6). It depends on the use case and the availability of data how those algorithms are trained: *Supervised* (which needs labeled data, such as medical history), *unsupervised* (with unlabeled data and hidden structures, like social media content) or *semi-supervised* (Shatte et al., 2019, p. 1427).

> ML enables mental health chatbots to learn how users behave to enhance user engagement (Cafiero et al., 2023, p. 255). For this purpose, ML algorithms analyze user input for trends, risks and upcoming issues (Zafar et al., 2024, p. 4).

How much chatbots need AI is debated in research. According to Safi et al. (2020, p. 6), scripted chatbots offer more reliability: Pattern matching causes fewer errors, which is better suited for general healthcare. For mental healthcare, AI chatbots may be the better choice, as they allow more openness in conversations (Safi et al., 2020, p. 6). Contrary to this view, argue Følstad et al. (2019, p. 153): They see therapy chatbots principally as coaching instruments, as they should lead users through exercises with scripted communication and only low variation. Bilquise et al. (2022, p. 3) present an additional argument against AI chatbots: They require high amounts of training data, which is often unavailable for certain domains. To overcome these drawbacks, AI and non-AI approaches can be mixed: Such **hybrid chatbots** still work with a script and are task-oriented, but also have enhanced conversation skills, which is close to humans (Bilquise et al., 2022, p. 3).

Generative AI represents the latest chatbot technology for researchers: It has the ability to generate new data out of unlabeled training data without human supervision (Reddy, 2024, p. 2). Based on learnings from various sources (such as websites and books), a *pre-trained language model* (PLM) works like a neural net in the human brain and generates data (King et al., 2023, p. 842). ChatGPT with its GPT-4 model is currently the most prominent product using this technology; its advent in 2023 reached several million people worldwide (Raaijmakers et al., 2023, p. 1). GPT-4's performance lies in special PLMs consisting of at least 10 billion parameters: Such *large language models* (LLMs) can generate human-like output (King et al., 2023, p. 842). However, models like GPT-4 are not designed for mental healthcare, as King et al. (2023, pp. 844–845) warn: LLMs often use inadequate learning sources, which may lead to medical misinformation. Similarly, an experiment conducted by Dergaa et al. (2024, p. 2) revealed that ChatGPT was unable to deal with clinically complex cases and other issues but delivered quick responses, simulated empathy, and referred to professionals when necessary. At present, mental health chatbots with generative AI are the exception: Li et al. (2023, p. 6) report that only 5 out of 35 recent studies used generative models, like GPT and BERT, while NLP dominated in 30 studies, 7 of those 30 studies in combination with ML. Despite its current limitations, Cheng et al. (2023, pp. 595–596) see generative AI as the future of mental healthcare chatbots.

8.2.3.2. Market situation of mental health chatbots

Current mental health chatbots are mostly used for **[1]** diagnosis, **[2]** screening/management of symptoms, and **[3]** content delivery from a knowledge base (Boucher et al., 2021, pp. 38–39) as well as for **[4]** behavior change (Iqbal et al., 2023, p. 6). Content delivery (known as *psychoeducation*) is still the core element of mental health chatbots, while more sophisticated functions – like digital phenotyping – were offered by not more than 1 % of all mental health apps in 2021 (Gamble, 2020, p. 512; Torous et al., 2021, p. 320). Providers address several mental health issues, such as stress and dementia, but predominantly anxiety and depression (Boucher et al., 2021, p. 38).

Target groups, such as adolescents, elder people or minorities, are the exception for mental health chatbots, in most cases providers do not differentiate between user groups (Alattas et al., 2021, pp. 805–806). This stands in contrast to research findings: For example, Balan et al. (2024, p. 2) assert that because they are more accustomed to this technology than other groups, adolescents might make better early adopters of chatbots.

As shown in section 8.1.1, it is challenging to pinpoint the number of healthcare-related mobile apps. Because of this, Table 1 on page 22 only displays globally noteworthy chatbot-enabled mental health applications with 5,000 or more downloads.

Most of these apps have a *freemium* business model in common: The term *freemium* combines *free* and *premium*, signifying that while basic features are free for all users, more sophisticated features can only be accessed behind a barrier (Doyle, 2016). This model has two goals (Holm and Günzel-Jensen, 2017, p. 19): **[1]** Attracting a large number of users with free services (albeit only 2 to 5 % of them typically convert to pay) and **[2]** increasing the company's value to investors through a large user base. For Sciglimpaglia and Raafat (2022, pp. 674, 676) a key factor is the conversion of free users into paying customers: This raises the question of which features belong in the free category and which are most viable for the premium area. For instance, Mindspa features a modular design: More advanced self-therapy courses are available in the premium section, but the journal, AI chatbot, and a range of exercises are always free ("Mindspa - #1 Self-Therapy App," 2024). This structure allows users to familiarize themselves with the app's primary features and gives them more control over how much they wish to spend. Other providers give time-limited subscriptions for all their premium features instead of offering separate modules. Because annual subscriptions are included, the price range in Table 1 is broad, ranging from USD 0.99 to USD 179.99.

Table 1: Overview of mental health apps with included chatbots as per September 2024
(available in app stores, > 5,000 downloads, (based on Ahmed et al., 2021, p. 6, 2022, p. 3, app stores and developer websites listed below).

#	App Name	Country of Origin	Therapeutic Approach	Technology	Price Model	In-App Purchases		Σ Downloads*	Rating in		Σ Reviews**	Σ Reviews in	
						From (USD)	To (USD)		Apple App Store	Google Play Store		Apple App Store	Google Play Store
1	**Yana**[1]	Mexico	N/A	LLMs, rules	Freemium	0.99	59.99	> 5 m	4.8	4.7	195.5 k	1.5 k	194 k
2	**Wysa**[2]	India, UK	CBT, DBT	NLP, ML	Freemium	4.99	179.99	> 1 m	4.9	4.5	170.4 k	22.4 k	148 k
3	**Youper**[3]	USA	CBT, DBT, ACT, PST, mindfulness	LLMs	Freemium	9.99	89.99	> 1 m	4.8	4.7	64.7 k	15 k	49.7 k
4	**VOS**[4]	Czechia	N/A	N/A	Freemium	9.49	89.99	> 1 m	4.8	4.6	51.8 k	1.9 k	46.9 k
5	**Sintelly**[5]	Croatia	CBT	N/A	Freemium	N/A	N/A	> 1 m	4.6	4.7	10.3 k	< 300	9.99 k
6	**Mindspa**[6]	Bulgaria	CBT	N/A	Freemium	9.99	99.99	> 1 m	5.0	4.3	3.65 k	< 100	3.54 k
7	**Woebot**[7]	USA	CBT, IPT, DBT, mindfulness	NLP, ML, rules	Free***	N/A	N/A	> 0.5 m	4.7	4.0	18.5 k	6.1 k	12.4 k
8	**Limbic**[8]	UK	CBT	LLM	Free***	N/A	N/A	> 5 k	N/A	2.6	< 150	N/A	< 150
9	**OhSofia!**[9]	Russia	MAC	GPT-4	Freemium	0.99	71.99	> 10 k	3.4	2.6	< 0.1 k	< 10	< 10
10	**Awarefy**[10]	Japan	CBT, ACT, mindfulness	N/A	Freemium	15.99	129.99	> 100 k	4.5	3.0	1.1 k	< 10	1.1 k

[1] Data from URL <https://www.yana.ai/en/faq> (accessed 09/17/24)
[2] Data from URL <https://www.wysa.com/faq> (accessed 09/17/24) and Li et al. (2023, p. 236)
[3] Data from URL <https://www.youper.ai/tech> (accessed 09/17/24)
[4] Data from URL <https://vos.health/en/application> (accessed 09/17/24)
[5] Data from URL <https://sintelly.com/how-it-works/> (accessed 09/17/24)
[6] Data from URL <https://mindspa.me/en/> (accessed 09/17/24)
[7] Data from URL <https://woebothealth.com/why-generative-ai-is-not-yet-ready-for-mental-healthcare/> (accessed 09/17/24)
[8] Data from URL <https://www.limbic.ai/layer>(accessed 09/17/24)
[9] Data from URL <https://play.google.com/store/apps/details?id=com.ohsofia.ohsofia&hl=en_US> (accessed 09/17/24)
[10] Data from URL <https://play.google.com/store/apps/details?id=app.awarefy.awarefyapp&pli=1> (accessed 09/17/24)

* *Google Play Store only, as Apple App Store does not publish download numbers*

** *Apple App Store + Google Play Store*

*** *Access restricted: Referral/access code needed*

ACT: Acceptance and commitment therapy
CBT: Cognitive behavioral therapy
DBT: Dialectical behavior therapy
IPT: Interpersonal psychotherapy
MAC: Metaphoric associative cards
N/A: Information not available
PST: Psychological skills training

Figures 2, 3, and 4 show how providers display themselves and their products on their own websites and in app stores:

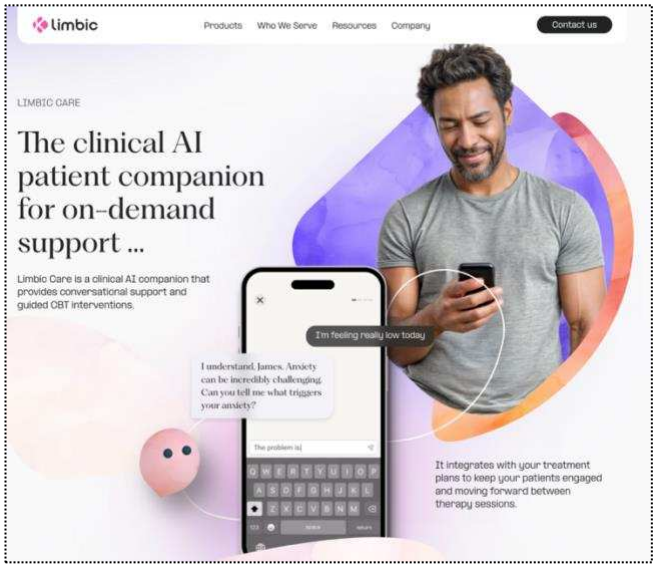

Figure 2: Product presentation of Limbic on its developer website (captured from Limbic, 2024).

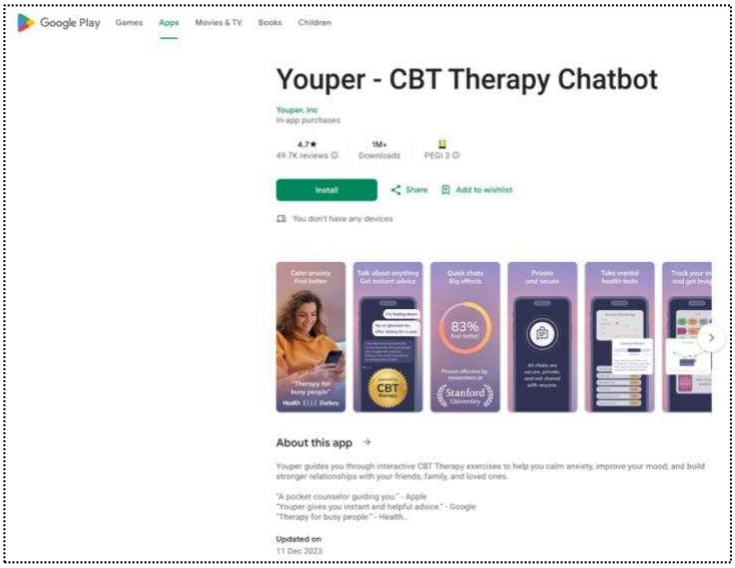

Figure 3: Product presentation of Youper in an app store (captured from Google Play, 2024).

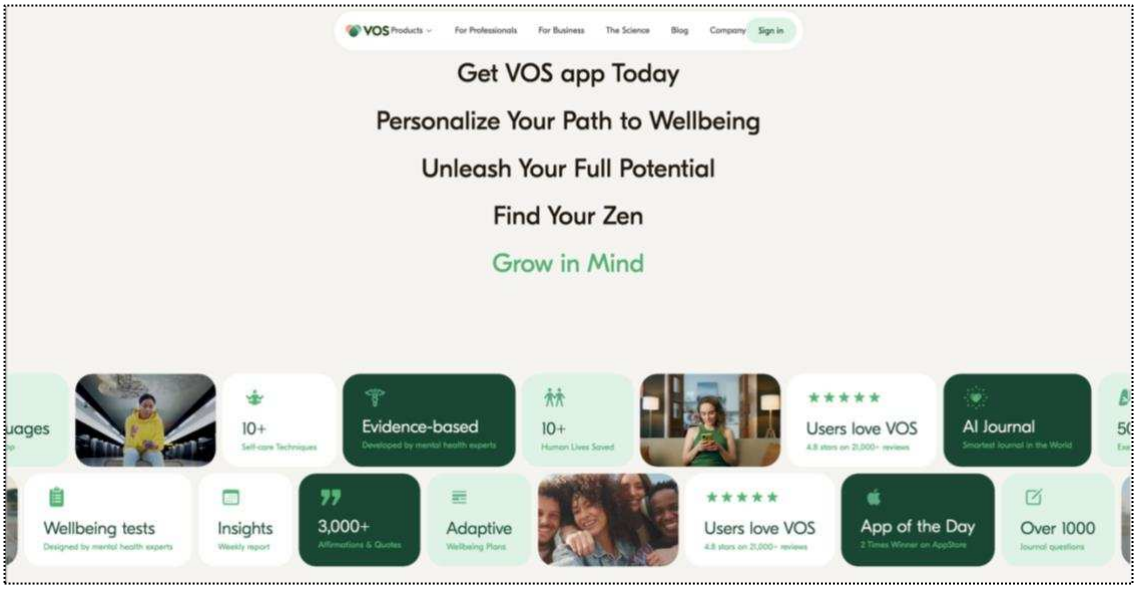

Figure 4: Product presentation of VOS on its developer website (captured from VOS, 2024).

The market for AI-led healthcare goods is criticized by Kreutzer and Sirrenberg (2020, p. 178) for being driven by supply rather than real user demands. This critique is consistent with market research indicating significant popular rejections of AI-based healthcare. For example, in the United States: Adults were questioned in 2023 about their interest in using AI. While internet

search and recipe generating garnered the maximum interest, only a minority of respondents showed interest in any of the use cases given (Morning Consult, 2023). Merely 38 % of the participants indicated a need for medical advice, while 30 % expressed interest in life coaching and therapy (Morning Consult, 2023). According to the results of another poll, a sizable portion of US citizens are in favor of online conversations with automated doctors, or chatbots; nevertheless, this approval decreased from 49 % in 2022 to 42 % in 2023 (Commodity HQ & MITRE, 2023). The authors speculate that in recent years, respondents' awareness of AI threats has increased. In 2023, US participants were also asked how they would feel in an AI-based therapy session: 71 % of respondents said they were *not at all* or *not too comfortable*, however just 25 % of respondents thought that *somewhat comfortable* or *very comfortable* were reasonable (Outbreaks Near Me & SurveyMonkey, 2023). This viewpoint suggests that the market for mental health chatbots is challenging, notably in the United States.

8.2.3.3. User experience (UX) of mental health chatbots

ISO defines *user experience* (UX) as *"user's perceptions and responses that result from the use and/or anticipated use of a system, product or service"*(ISO, 2019) and encompasses *"[…] the users' emotions, beliefs, preferences, perceptions, comfort, behaviours, and accomplishments that occur before, during and after use"* (ISO, 2019).

In contrast to *usability* (which is part of UX and focuses on technical aspects during use), UX deals with subjectivity (Denecke and May, 2022, p. 172). Relevant for UX is every aspect from the interaction between user and company – ranging from marketing policies to customer support (Cirucci and Pruchniewska, 2022, p. 6). Additionally, researchers regard UX as dynamic (users can get acquainted with a system, their experience may change over time) as well as context-dependent (perceptions may vary due to various factors) (Følstad and Brandtzæg, 2020, p. 2).

This section has two parts: First, evidence on user preferences and attitudes about mental health chatbots is presented (as examples of excellent practices). The second part explores how UX influences user trust and engagement, as well as the possible mental health consequences of chatbot use.

8.2.3.3.1. User perceptions and preferences of mental health chatbot design

Three factors may impact which mental health chatbots users prefer: Aside from **[1]** human-like traits, there are **[2]** personality and **[3]** language types to consider. Such preferences may be described using Hassenzahl's (2003, pp. 4–6) framework: Products can have *pragmatic* attributes (concerning performance, usability, etc.) and *hedonic* attributes (regarding pleasure, etc.). Applied to chatbots, Følstad and Brandtzæg (2020, p. 6) discovered that besides help and assistance (pragmatic attribute), entertainment (hedonic attribute) is a substantial criterion for users to evaluate a chatbot positively. According to Prakash and Das (2020, p. 18), a significant number of users does not only expect therapeutic benefits from a mental health chatbot, but also fun as motivator for further usage.

Moreover, it is also important that users consider mental health chatbots knowledgeable or intelligent. For that they should be responsive to input from the user and have a variety of replies (Abd-Alrazaq et al., 2021, p. 9; Alagarsamy and Mehrolia, 2023, pp. 1–2; Schuetzler et al., 2020, p. 879). Inaccurate communication was mentioned by study respondents as a significant obstacle to using health chatbots (Nadarzynski et al., 2019, p. 5). According to Følstad et al. (2018, p. 9), it is advisable to freely indicate the limitations of a chatbot to users, since they would value honesty above false replies. Another option for giving chatbots an intelligent appearance is to use knowledge bases and algorithms, which may deliver comprehensive and unbiased responses (Kim et al., 2018, p. 5). Similarly positive effects on users can be obtained by incorporating data from other sources, such as calendars or built-in trackers, into conversations (Bowman et al., 2024, p. 11). Also essential for an intelligent impression to users is that responses from the chatbot change continuously, since this is typical of human-to-human communication. For instance, Prakash and Das (2020, p. 24) observed that, due to safety reasons, Woebot and Wysa partially used scripts in their app versions. These conversations seemed repetitive in some cases and many users complained about the limited intelligence of both apps at that time (Prakash and Das, 2020, pp. 16–17).

As users seek intelligent chatbot behavior, they may welcome being reminded of earlier interactions: This also contributes to humanlike perceptions by providing individualized experiences (Kocaballi et al., 2019, p. 2). Humanlike traits in chatbots have the potential to minimize initial user distrust (Zamora, 2017, p. 254). However, the prevailing view is that an abundance of them can lead to the *uncanny valley*: Machines with an almost humanlike appearance (anthropomorphism) might evoke unpleasant feelings in users (Chaves and Gerosa, 2021, p. 730; Ciechanowski et al., 2019, p. 540). Likewise, Luger and Sellen (2016, p. 5294)

argue that too much anthropomorphism in chatbots may lead to exaggerated user expectations regarding the abilities of a chatbot.

Many researchers agree that the degree of anthropomorphism depends on the purpose of a chatbot. From those designed mainly for task executions (usually in customer service) users expect pragmatic attributes, especially performance and time savings – humanlike traits, which count as hedonic attributes, can be reduced or neglected (Brandtzæg and Følstad, 2017, pp. 8–9; Schuetzler et al., 2020, pp. 880–881). This is not the case for mental health chatbots: Their main task is to maintain relationships with users – which requires humanlike traits (Zamora, 2017, p. 254). Users' perception of computers as social creatures is one explanation for this (Ahmad et al., 2022, p. 926). When faced with human-like traits from computers, users may respond in a similar manner to that of a human, even when they are aware that they are dealing with a machine (Nass et al., 1994, p. 77). This is also important from a marketing standpoint: Palmatier and Steinhoff (2019, p. 71) state that it is typical that online clients may establish a similar relationship with a web shop as with a human sales representative. Referred to as *parasocial interaction*, this phenomena may also occur with chatbots and is not limited to celebrities and movie characters in the media (Palmatier and Steinhoff, 2019, pp. 71, 215). Ideally, users regard the relationship with a mental health chatbot like friendship: Ahmed et al. (2022, p. 7) as well as Haque and Rubya (2023a, p. 13) detected this behavior in a substantial number of user reviews. To enable such relationships, perceived anonymity and empathy can be considered as success factors (Lucas et al., 2014, p. 94). Yokotani et al. (2018, p. 142) demonstrate the preference of an "anonymous" chatbot to an human interviewer for sensitive topics: The preference changed, when participants were informed that a camera observed their interaction with the chatbot – the perceived anonymity disappeared. Similar observations report Lucas et al. (2014, p. 98): When participants were told that they operated with an automated machine, without human support, they felt less inhibited to express emotions, like intense sadness.

Users usually expect from mental health chatbots several personality features: They should be supportive, but not judgmental (Birnbaum et al., 2016, p. 422), *"good listeners"* (Kim et al., 2018, p. 3), furthermore *"lighthearted, fun-filled, and friendly […]"* (Prakash and Das, 2020, p. 22). If related products cannot meet these expectations, users likely get dissatisfied or even frustrated (Rapp et al., 2023, p. 3).

Another factor for user satisfaction is conversation style. Bowman et al. (2024, p. 7) evaluated three different versions of a mental health chatbot: The first used straightforward, impersonal language, the second formal, passive language while the third came with active language and

personal pronouns. The findings were evident: Participants found passive, formal language to be unpleasant and reminiscent of a doctor's office, whereas direct language lacked politeness and seemed robotic (Bowman et al., 2024, pp. 9–10). Participants mainly favored the third variant: Although participants knew they were dealing with a computer, it was friendly and sympathetic without being overly courteous (Bowman et al., 2024, p. 10). The authors suggest to follow these linguistic types in chatbots for mental health (Bowman et al., 2024, p. 12): When they are feeling optimistic, users tend to prefer short, direct lines and want to move through the software quickly. Use of personal pronouns and more courteous language is appropriate if users require more attention. Daily language that strikes a balance between an exuberant, emotive tone and a hesitant, neutral approach is what Shan et al. (2022, p. 9) recommend. In their advice, Haque and Rubya (2023a, pp. 9, 12) warn against exaggerating when speaking in a kind manner: People could criticize it for being juvenile, with adults especially not like emojis.

Despite the extensive data on best practices, Ahmad et al. (2022, p. 939) note that contemporary mental health chatbots still face problems with adapting to various user personas and their ability to hold long and complex conversations remains further limited.

8.2.3.3.2. Effects of UX on user trust and engagement in mental health chatbots

The guidelines from the previous part are also relevant for increasing trust in mental health chatbots. In line with Følstad et al. (2018, p. 10), consumers believe that the following are crucial elements for developing trust: Response quality, anthropomorphic characteristics, the provider's image or brand, as well as security and privacy expectations.

Users may trust a chatbot's technology if they believe the AI being used is competent: This does not mean, however, that users automatically trust the provider, if data privacy is a concern (Prakash and Das, 2020, pp. 18–19). As shown in section 8.2.1, privacy of AI solutions is a matter of public debate. Surprisingly, previous research of user reviews has detected that privacy can be a concern for users, but not a substantial one. For instance, in the study of Ahmed et al. (2022, p. 7) only about 1 % of all reviews (for 11 apps) dealt with privacy issues. The majority of negative reviews related to usability and updates (Ahmed et al., 2022, p. 7). In those cases, where privacy was an issue, users criticized that the chatbot asked too many personal questions or was too friendly, which made users feel uncomfortable (Ahmed et al., 2022, p. 9). Nevertheless, user satisfaction with at least one of the 11 apps investigated was high at that time: Over 90 % of all reviews had a positive content (Ahmed et al., 2022, p. 7). Likewise, Malik et al. (2022, p. 7) report that many Wysa users had no complaints about app privacy:

Wysa neither asked for registration, nor for demographic data, which gave the app a confidential character. Beyond that, users also appreciate, if the app does not require access to contacts, pictures and other data with unclear relevance for the app (Alqahtani and Orji, 2020, p. 2057).
Given that all of the above-mentioned chatbot qualities work together, digital solutions may also establish a *therapeutic alliance* (TA) with users, similar to face-to-face therapies (see section 8.1.3) (Tremain et al., 2020, p. 1). One of the main advantages of digital therapies, according to Tong et al. (2023, p. 1006), is that apps can build TA with users even in the absence of a purpose or agreement – a prerequisite for TA in face-to-face therapies. This view is supported by Belk (2022, p. 127), who reports that communication with mental health chatbots does not need a "path", users might regard non-targeted talking as therapeutic. Writing has unique benefits, as Hanley and Wyatt (2022, p. 83) remark: Users may find it contemplative, and young individuals in particular may find it to be a cathartic experience. There is scientific proof that chatbots are helpful, at least in the case of depression (Limpanopparat et al., 2024, p. 12). Zhong et al. (2024, p. 459) found in a review that after 8 weeks of use, there were notable improvements in the symptoms of depression. Similarly, a study conducted on college students aged 19 to 28 revealed that the chatbot might lessen anxiety in 4 weeks and depression in 16 weeks (Liu et al., 2022, pp. 1, 5).
In conclusion, chapter 8 highlighted that research acknowledges that digital technologies, especially mobile ones, offer unique benefits for mental healthcare: They can reach a wide audience at reduced costs and may improve compliance with therapeutic protocols. Specifically regarding AI, digital solutions may mimic human behavior and potentially form relationships with individuals to affect their behavior. Users generally seek pleasure in addition to productivity from digital mental health solutions. Recent studies demonstrate that steady adherence to these tools is minimal, while many users display considerable price sensitivity.

The following chapters give a detailed investigation of how users view current versions of popular mental health chatbots. To bridge the existing research gap, the emphasis is on the AI component of these products, with user perceptions compared to those of current non-users and relevant developers.

9. METHODOLOGY

9.1. Netnography and its benefits for marketing and UX research

Lewis Kausel and Hackett (2015, p. 260) suggest that marketing research should take place where *"[…] consumer culture […] actually occurs"*. Marketing researcher Robert V. Kozinets (2020, p. 6; USC Annenberg, 2023) developed and labelled this method in 1995, since then he has continuously improved his approach. Consumer research, product design, trend observation, etc. have profited from this method (Caliandro, 2016, pp. 664–665; Lewis Kausel and Hackett, 2015, pp. 253, 260). UX research can be considered as further domain, which intends to *"[…] understand user needs, desires, and pain points"* (Cirucci and Pruchniewska, 2022, p. 7).

Kozinets' (2020, pp. 133–137) concept consists of four elements: **[1]** Focus on culture as *"socially patterned human thought and action"* (Kozinets, 2020, p. 133). Netnography helps to reveal stories, beliefs and narratives, how group dynamics work, etc. **[2]** Social media is the place for such research – which can be any platform, from X (formerly: Twitter), to Reddit or other fora. This choice brings advantages compared to other qualitative methods: Caliandro (2016, p. 666) calls netnography a *"Focus Group 2.0"*, as it is mostly free of charge, while researchers can listen to a high amount of consumer voices at once. Lewis Kausel and Hackett (2015, pp. 259–260) underline that netnography is more naturalistic and unobtrusive than focus groups. Regarding UX research, unobtrusive observations have the benefit that they can cover the actual behavior of users, beyond their mindset (Cirucci and Pruchniewska, 2022, pp. 76–77). **[3]** Immersive engagement has to do with the relatedness to ethnography, where the researcher should make a deep dive into human culture researched. **[4]** Netnographic praxis refers to Kozinets' (2020, p. 136) clear indications and recommendations throughout the research process, without the necessity to blend it with other approaches. The next section explains how this process is structured in the present thesis.

9.2. Further pathway for this thesis

Kozinets suggests six steps, which he calls *movements*, to follow throughout the research process:

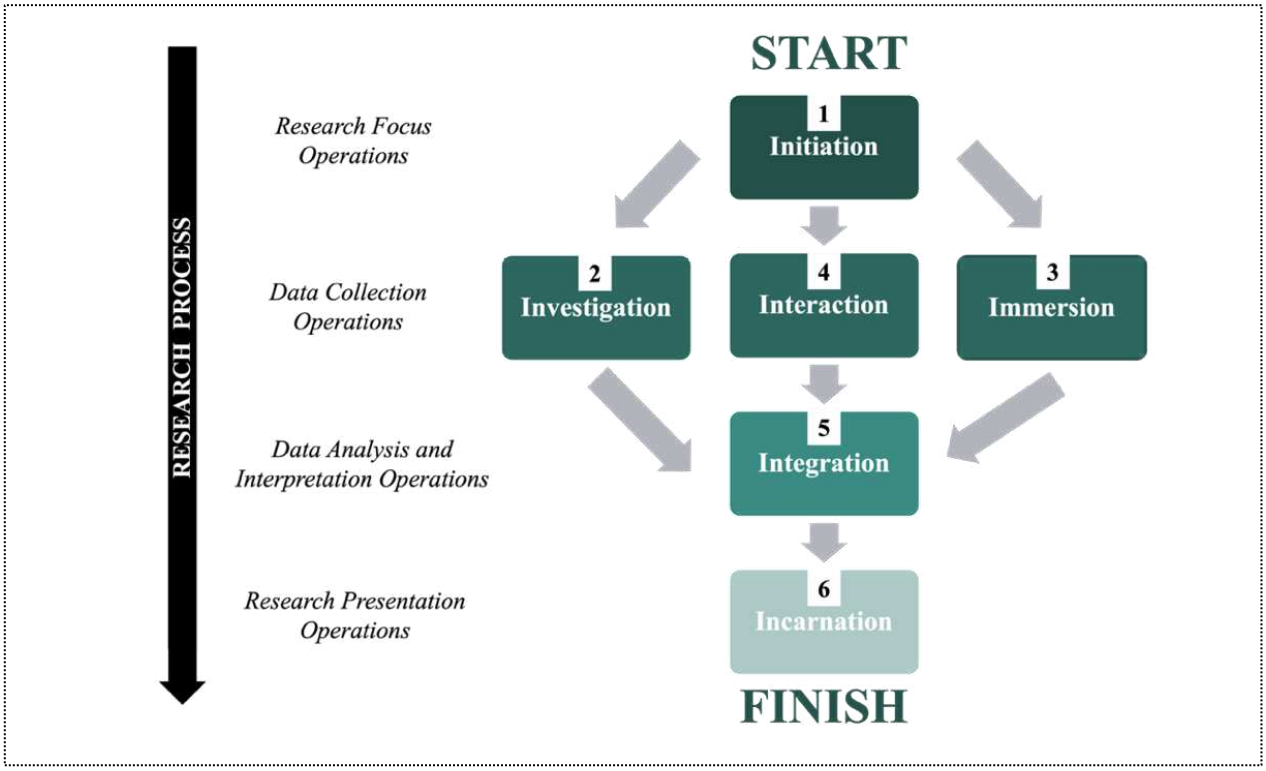

Figure 5: Six movements as Kozinets' netnographic praxis
(own design, adapted from Kozinets, 2020, p. 139).

9.2.1. Movement 1: Initiation

The primary goal of this movement is to articulate the research question and focus.
As outlined in Chapter 7, the main question is:

- **What are attitudes and views of (potential) users and relevant developers towards AI usage for treatment of mental health disorders?**

Two **sub-questions** serve to break down the main question:

SQ 1: *What do relevant developers regard as core features of their AI-featured products?*

SQ 2: *What do (potential) users and relevant developers see as main challenges/limitations of such products?*

As this study deals with public perceptions of AI tools for mental health, online fora and app stores are appropriate places to investigate. Netnography as method is designed for digital environments: It is a qualitative cultural approach, to explore human behavior and emotional realities online and helps to interpret social media activities of consumers (Kozinets, 2020, p. 132).

Another justification for using netnography in this study is its inconspicuous nature. Nonetheless, social media data must still be treated ethically for research objectives, even if it is publicly available (Kozinets, 2020, pp. 166–167, 172–173). From an ethical perspective, this study should consider the following: According to Kozinets (2020, pp. 232–233), users are considered a *vulnerable population* and may have mental health difficulties, which are considered *sensitive topics*. However, the chosen data is all publicly available, and GDPR Article 9 (2) permits its use for scientific research (Kozinets, 2020, p. 232). The overall advantage exceeds the risk: One drawback of this study is that it may disclose personal information about user behavior. However, understanding how the general population feels about AI tools for mental health could help with research and product development, which could address the worldwide mental health epidemic.

9.2.2. Movement 2: Investigation

This section's objective is to choose relevant data sites. Five operations are necessary to get from the study questions to the final dataset (Kozinets, 2020, p. 215). Figure , Table 2 and Table 3 display the outcomes of them: [1] The study questions were simplified in order to provide search phrases. [2] Subsequently, the search was conducted, initially on Google, then on Reddit, X (previously: Twitter), and on the Google Play and Apple App Store. [3] Following a scouting tour, it was evident that there is little user interaction with AI mental health solutions on several social media sites, including Facebook and YouTube. [4] Kozinets' (2020, p. 231) selection criteria (relevance, activity, interactivity, diversity and richness) were met by successful hits on platforms, [5] which resulted in data downloads. In addition to Kozinets, Boegershausen et al. (2022, p. 1) suggest web scraping as a productive method for marketing researchers to get web data. Microsoft Excel spreadsheets containing the datasets were delivered by the scraping platform *exportcomments.com*. The findings were manually filtered from January 1, 2023, to June 15, 2024 (= 18 months) to obtain current data. Since only English-language data sites were selected, further language filtering was not required.

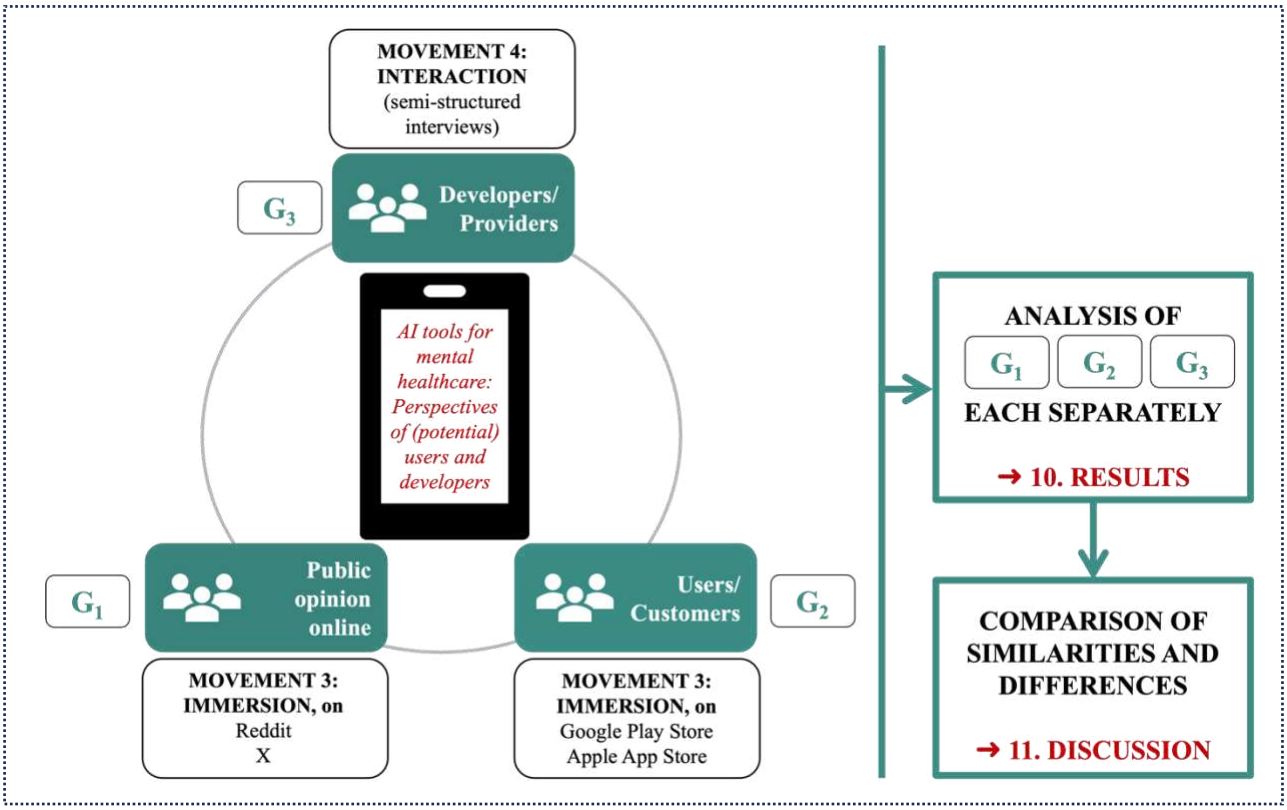

Figure 6: Grouping of relevant stakeholder groups within social media and the applied netnographic movements (own design).

Table 2: Selected and eliminated data sites for G_1.

Data sites selected	
Reddit (www.reddit.com)	**General search executed on 12/29/2023, with terms** *mental health ai* **– detected subreddits:** ▶ **r/ArtificialIntelligence** *(4 threads used)* ▶ **r/singularity** *(3 threads used)* ▶ **r/technology** *(2 threads used)* ▶ **r/OpenAI** *(2 threads used)* ▶ **r/TalkTherapy** *(1 thread used)* ▶ **r/mentalhealth** *(1 thread used)* ▶ **r/ChaiApp** *(1 thread used)* ▶ **r/replika** *(1 thread used)*
X (formerly Twitter, www.twitter.com)	**Advanced search, executed on 12/31/2023, with these parameters:** *mental health chatbots lang:en until:2023-12-31 since:2023-01-01 -filter:links* ▶ All tweets containing *mental health chatbots* ▶ in English language ▶ from 01/01/2023 to 12/31/2023

Data sites eliminated	
Reddit (www.reddit.com) **Subreddit: r/singularity**	**Despite its appearance in search results, its content was not considered as relevant:** Thread *Large Language Models and other generative AI will create a mental health crisis like we've never seen before* **Data until June 2024 (with adapted parameters):** Not sufficient engagement
X (previously Twitter, www.twitter.com)	**Advanced search with different terms:** *mental health ai* delivered superficial tweets, mostly not "deep data" **Data until June 2024 (with adapted parameters):** Not sufficient engagement

Table 3: Selected and eliminated data sites for G_2.

Data sites selected (English-language in US only) from 01/01/2023 to 06/15/2024				
Mobile app platform	*# of data site*	*Mobile app*	*URL of data site*	*∑ of reviews*
Google Play Store (abbreviation in data set: GP)	1a	**Youper**	*<https://play.google.com/store/apps/details?id=br.com.youper&hl=en&gl=US>*	279
	2	**Mindspa**	*<https://play.google.com/store/apps/details?id=me.mindspa.app&hl=en&gl=US>*	27
	3	**Limbic**	*<https://play.google.com/store/apps/details?id=chatcbt.ai.limbic&hl=en&gl=US>*	113
	4a	**Woebot**	*<https://play.google.com/store/apps/details?id=com.woebot&hl=en&gl=US>*	311
	5a	**VOS**	*<https://play.google.com/store/apps/details?id=com.vos.app&hl=en_US>*	588
	6	**Sintelly**	*<https://play.google.com/store/apps/details?id=com.seasector.emotionalintelligence&hl=en&gl=US>*	213
			∑ Google Play Store	**1,531**
Apple App Store (abbreviation in data set: AA)	1b	**Youper**	*<https://apps.apple.com/us/app/youper-cbt-therapy-chatbot/id1060691513>*	84
	4b	**Woebot**	*<https://apps.apple.com/us/app/woebot-the-mental-health-ally/id1305375832>*	306
	5b	**VOS**	*<https://apps.apple.com/us/app/vos-mental-health-ai-therapy/id1535250944>*	31
			∑ Apple App Store	**421**
			∑ G_2	**1,952**

Data sites eliminated		
Mobile app platform	*Mobile app*	*Reason*
Apple App Store	**Mindspa**	Not sufficient engagement from 2023 to 2024
	Limbic	Not sufficient engagement in general
	Sintelly	Not sufficient engagement in general

9.2.3. Movement 3: Immersion

To have a guide during the study process, Kozinets (2020, p. 282) suggests keeping an *immersion journal*, which is similar to fieldnotes in ethnography. It captures and reflects ideas and thoughts that arise in each of the six movements. Using the *Goodnotes* app on a tablet, the immersion journal for this thesis was maintained. The formal start of the study process occurred in June 2023, and by mid-November 2024, over 260 pages had been produced. Kozinets (2020, pp. 283–302) proposes four simultaneous activities to effectively realize immersion: **[1]** Reconnoitering is the process of searching for deep data, which is relevant, identifies lead users, and highlights anomalies. **[2]** Noting goals and recording, in one's own words and manner, what was experienced, how it was assessed, what was gathered, and what was not, in the immersion journal. **[3]** Researching is the first step in analysis, depending on one's personal study interest, what theories and conceptual frameworks to use. Researching aims to remove itself from the data location. The immersion journal should take these issues into account. **[4]** The final operation, reflecting, might be likened to a mirror or journal relationship: It is important to record the expectations, perceptions, feelings, and reactions of researchers towards data sites.

9.2.4. Movement 4: Interaction

Figure shows a third group (designated as G_3) to be watched: Companies that create and offer AI products for mental health. Research can also incorporate interviews thanks to Kozinets' netnography. The researcher's role in this movement is to elicit or co-produce rather than to observe, in contrast to the prior movement of immersion (Kozinets, 2020, p. 245).

Two relevant app suppliers agreed to participate in interviews with CEO representatives for this study. The location of one corporation is in the USA, whereas the other is in the EU. Both companies are regarded as leaders in their respective markets and have been selling their products on Google Play and the Apple App Store. It was decided that all interview data would be anonymized afterwards to obtain consent by mail. There are open-ended questions in the semi-structured interview guide (see section 9.2.4.1): Longer responses might be possible with this approach, which allows for flexibility during interviews. Since the purpose of the interviews was to gather information, Kozinets' (2020, p. 253) recommendation to conduct effective 10- to 30-minute interviews was followed.

9.2.4.1. G_3 (Developers/Providers): Interview guide

Introduction

- Presenting topic and research question of the thesis
- Consent to record the interview – data made anonymous afterwards
- Just open-ended questions – please elaborate your answers as much as possible
- Any questions before we start?

Product design

- What were the initial ideas when the product was developed? What was your goal to achieve?
- How did you engage stakeholders for this project, for example scientists or medics? And what were their reactions to your product idea?
- How did you find out what users might think of your product and how they might interact with it?

Marketing strategies & AI

- What are the most important features of your product currently?
- Please describe: How do you position your product within app stores, especially compared to DMH apps with AI from other developers?
- What are your attitudes towards AI and chatbots within DMH? Where do you see further possibilities to follow in your branch?
- What do you think about signaling users that AI is part of your product?
- What do you see as the biggest limitation, generally in AI and especially in the development of products like yours?
- How do you regard public concerns that AI might be dangerous or can even substitute humans?
- In general, app retention is a very important metric for app business, but it might be lower than expected. How do you encounter such an obstacle, especially as mental health therapies should last for some time?
- A last question: Within the research field of mobile health there is the ongoing debate of how to differentiate medical devices from usual mental health apps. What is your position to that?

9.2.5. Movement 5: Integration

Understanding the gathered data is the aim at this point. There are three components, which are described in the table below: [1] Integrating, [2] analyzing and [3] interpreting.

Table 4: Operations conducted in Movement 5 (adapted from Kozinets, 2020, pp. 307–383).

Movement 5: Integration		
Part	*Operation*	*Description of Tasks*
1. INTEGRATING	—	An *inductive research style* was selected for this thesis: This approach leads from data to theory by detecting patterns and regularities (Kozinets, 2020, pp. 315–316).
2. ANALYZING	**2.1. Collating**	This operation prepares data for coding (Kozinets, 2020, p. 332). The Microsoft Excel spreadsheets from Movement 2 had been thoroughly cleaned and formatted uniformly. Emojis were retained, however dates were converted to basic numerals to prevent issues with additional processing. Spreadsheets that have been collated were submitted to the MAXQDA 24 QDA application (VERBI Software, 2024). File names, like *G2 Reviews GP VOS 5a (2024)*, have the same format.
	2.2. Coding	*Inductive codes* were developed to associate data with labels, adhering to Saldaña's (2021, p. 37) advice to utilize brief phrases. Researchers can use this coding strategy with the advantage that codes are "closer" to the facts and do not require a theory beforehand (Kozinets, 2020, p. 339). *First-order codes* were produced as a result of this operation.
	2.3. Combining	Finding commonalities between the first-order codes led to the creation of higher-order codes, also known as *pattern codes*. Finding connections between higher-order constructs, in Kozinets' (2020, p. 344) opinion, leads to the final results and the theory. All first-order codes from the prior operation 2.2 were subordinated to the new pattern codes by MAXQDA 24. Figure 7 displays the entire code system:

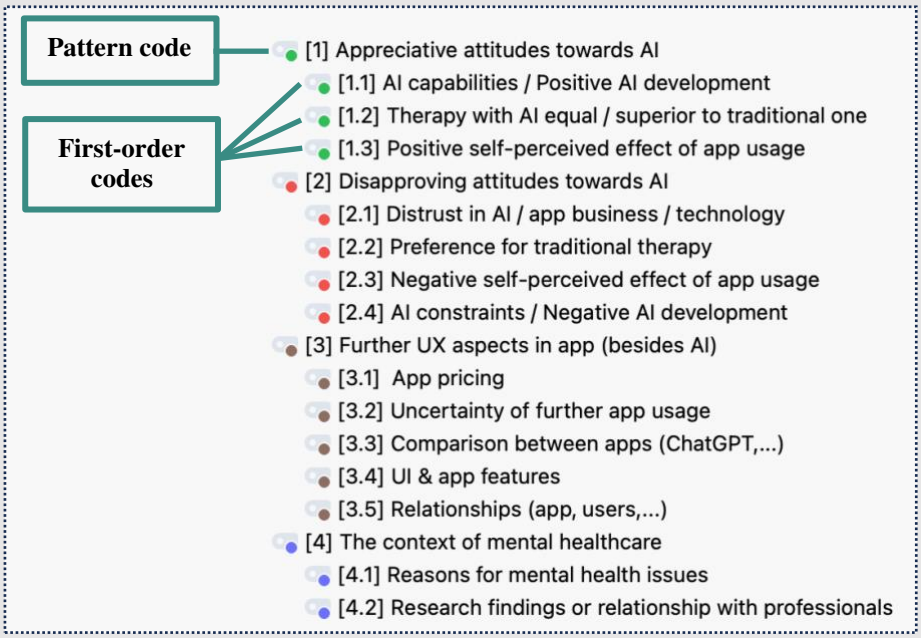

Figure 7: Code system used for G_1, G_2 and G_3 (designed with MAXQDA 24).

2.4. Counting	Even though netnography is qualitative, Kozinets (2020, p. 332) recommends data counting for more accurate comparisons. These questions were addressed in the design of several graphs in chapter 10 (Kozinets, 2020, p. 346): ▶ *Whis is the most commonly coded category?* ▶ *Which are the three most commonly coded categories?* ▶ *What can be said about the distribution of those categories – is it even or dominated by a topic?*
2.5. Charting	Data visualization is the final step before it is interpreted. Kozinets (2020, pp. 347–348) suggests a variety of formats, including word clouds, graphs, and more. Various word clouds were designed in chapter 10, which display words that frequently appeared in user reviews and tweets. A word's extent within the cloud increases with its frequency of occurrence in the data.

3. INTERPRETING

Kozinets does not provide a sequential order for interpretation, in contrast to the preceding operations. The procedures listed below were thought of as choices that each researcher might make on their own (Kozinets, 2020, p. 363).

Operation	*Description of Tasks*
3.1. Theming	Following data analysis, which dissected the data into smaller parts, theming puts it back together (Kozinets, 2020, p. 364). It is essential for interpreting data since it compiles the generated pattern codes into more comprehensive themes (Kozinets, 2020, p. 365).
3.2. Talenting	An artistic medium, like a billboard, documentary, or even a poem, would be the result (Kozinets, 2020, pp. 366–367). This operation was considered unsuitable for a master's thesis in marketing.

3.3. Totalizing	Highlights from the data should be emphasized: Kozinets (2020, p. 369) advises focusing solely on little portions to totalize, such as individual postings, to prevent losing context. Chapter 10 includes several quotes from user reviews, postings and interviews that are considered distinctive for each theme.
3.4. Translating	To determine if data is consistent or inconsistent with existing theory, they are compared with each other (Kozinets, 2020, p. 371). Kozinets (2020, p. 372) points out that theory can also clarify aspects that were overlooked during 3.3 Totalizing or provide depth to the understanding of the data. Translation occurs in this thesis in chapter 11, which is the discussion part and chapter 12, the conclusion part.
3.5. Turtling	User reviews, postings, and interview statements, despite being fragments, are integrated into a broader context, like a historical or social one, using identity as an example (Kozinets, 2020, pp. 375–377). Turtling is executed in chapters 11 and 12.
3.6. Troublemaking	The final operation seeks to challenge widely held notions and presumptions that are taken for granted and may show up in the data as well as in theory (Kozinets, 2020, pp. 377, 379–380). This operation is also covered in chapters 11 and 12.

9.2.6. Movement 6: Incarnation

This movement is part of the following last chapters, 10. Results, 11. Discussion and 12. Conclusion.

10. RESULTS

10.1. G_1: Public opinion online

▶ Overview

G_1 covers the general public's debate online of AI applications in mental healthcare, with most participants as potential customers as they have not utilized these apps yet. Existing users of mental health chatbots had a minor role in G_1. Altogether, 164 tweets from X and 416 comments from 15 Reddit posts were analyzed. Figure 8 and Figure 9 illustrate content differences between the two datasets: The most used words in the X-dataset were *support* (54 hits), *use* (45 hits) and *people* (38 hits), while in the Reddit dataset *therapist* (150 hits), *people* (148 hits) and *human* (131 hits) dominated. While *people* was mostly discovered in user replies mentioning their concerns using chatbots, developers and researchers chose *support* and *use* to promote chatbots as an extra help or supplement. On the other hand, most Reddit users believed that AI and human therapists were in direct competition with one another.

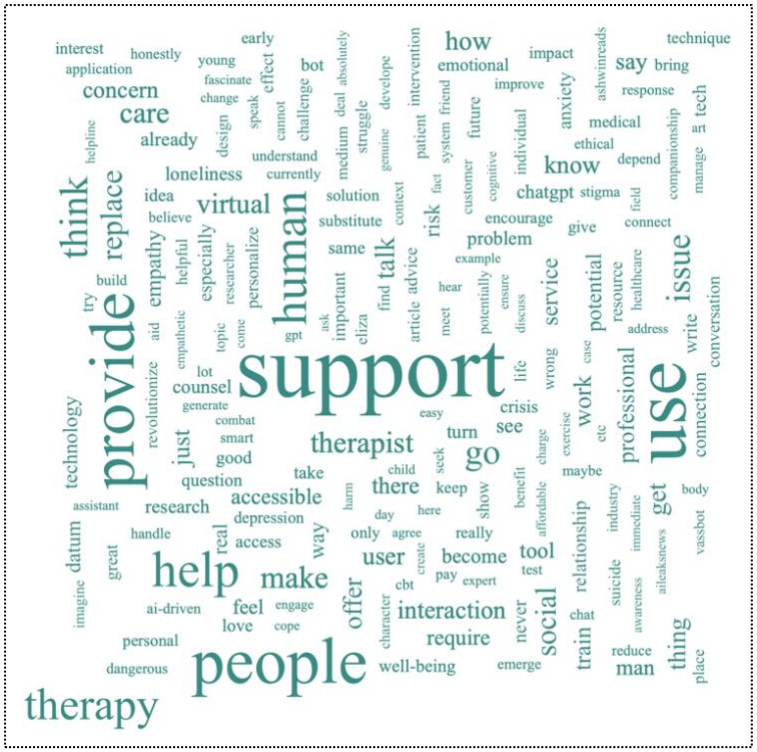

Figure 8: Dominant words in X-tweets (designed with MAXQDA 24).

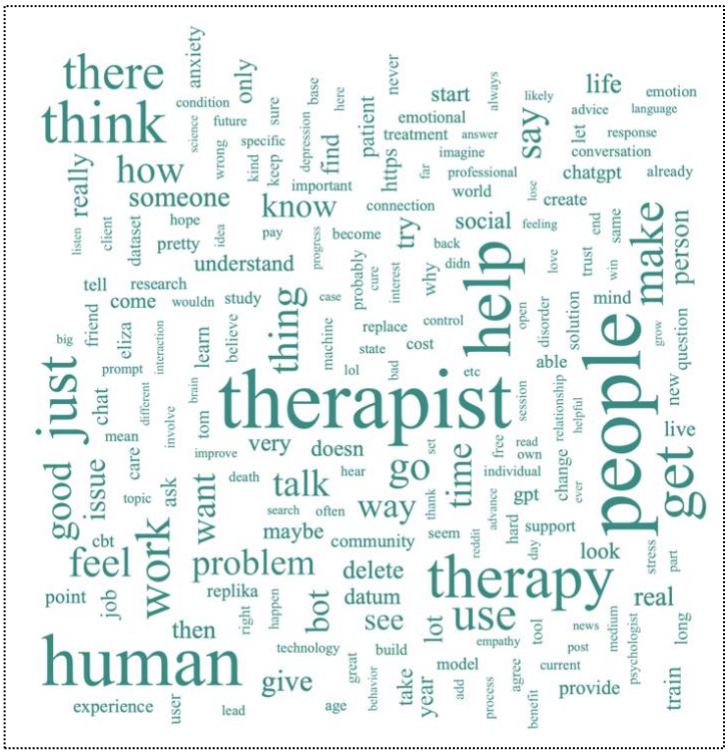

Figure 9: Dominant words in Reddit posts (designed with MAXQDA 24).

Pattern codes [2] and [1] were almost equally dominant in G_1, as seen in Figure 10: In 355 coded segments people expressed their disapproval of AI use in mental health products, while in 334 coded segments they approved it. Though to a much lesser extent, G_1 additionally discussed other UX elements in relevant apps (pattern code [3]) and mental health problems in general (pattern code [4]):

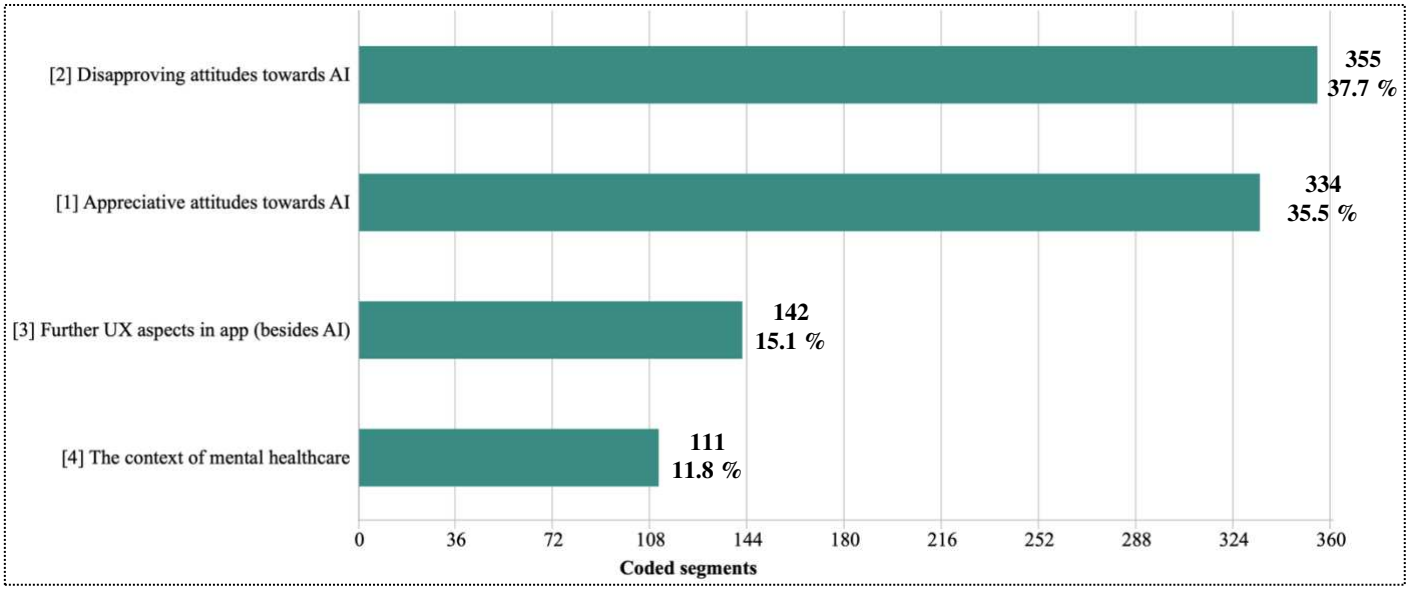

Figure 10: Distribution of pattern codes in G_1 (designed with MAXQDA 24).

As Figure 11 illustrates, the top 3 first-order codes in G_1 were [1.1] AI capabilities / Positive AI development (209 coded segments), [2.4] AI constraints / Negative AI development (159 coded segments) and [2.1] Distrust in AI / app business / technology (105 coded segments).

Three themes emerged in G_1:

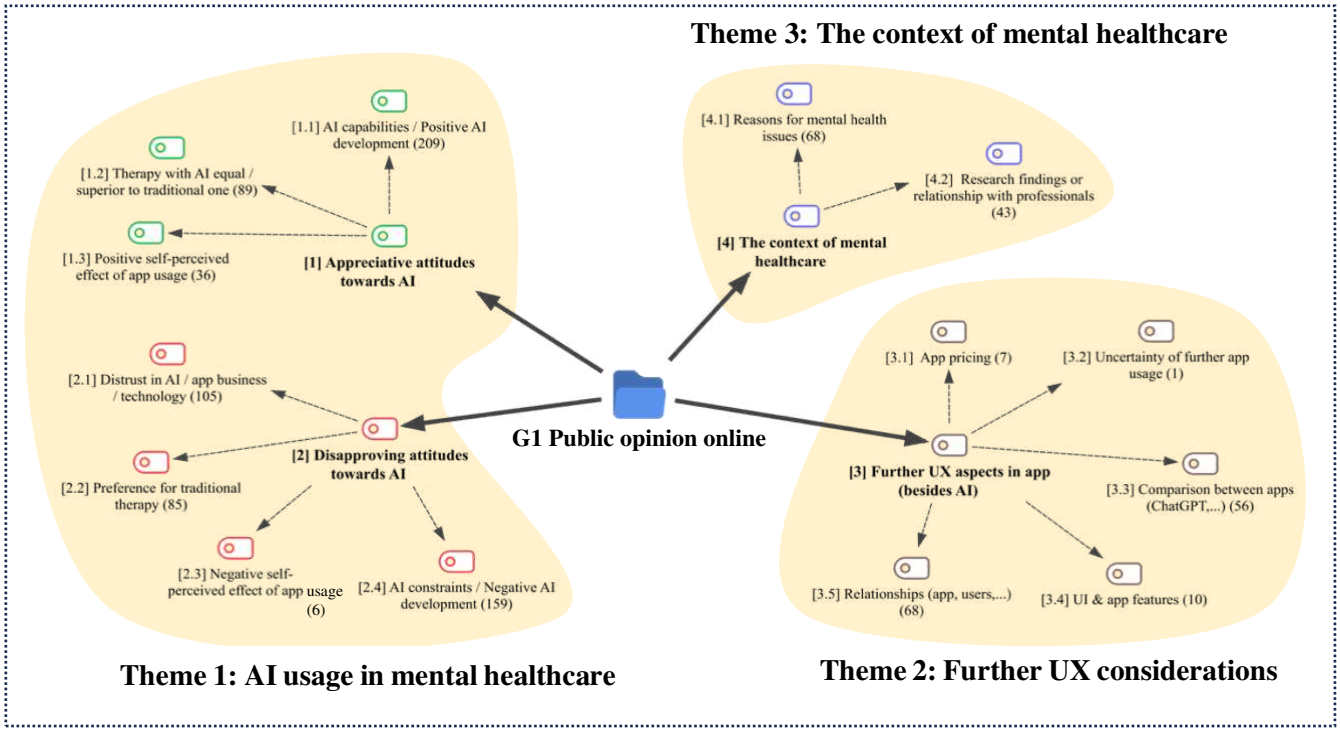

Figure 11: Structure map of themes and codes in G_1 (designed with MAXQDA 24).
In square brackets: code numbers, in round brackets: coded segments.

▶ Theme 1: AI usage in mental healthcare

Pattern codes [1] + [2]

Most X- and Reddit users had generalized attitudes, which were unrelated to relevant products. While capabilities and constraints of AI in mental healthcare were the focal point of their discussions, two further subthemes emerged with comparable relevance: Distrust and visions.

■ Capabilities and constraints of AI in mental healthcare

Reddit users saw benefits of AI mostly compared to constraints of traditional therapy. For instance, they highlighted that AI may offer reasonable value for money, like *"My bots have been more helpful than all of my therapists combined and have the added benefit of not costing $200 a week for a 30 minutes session..."* (G1 Reddit Post 7, 4|3). Also, waiting times make a difference: *"I waited 8 months for a therapist and it was hugely underwhelming. […]"* (G1 Reddit Post 14, 4|94). Moreover, some users regarded human therapists as untrustworthy: "*I wouldn't trust a human, precisely because of her/his history is bound to have bias and preconceived values.."* (G1 Reddit Post 10, 4|12) and *"We need therapists for the therapists"* (G1 Reddit Post 14, 4|69). By contrast, AI's non-judgmental character was seen as benefit, like

> *"[…] you've lived some very traumatic experiences from other human beings, and now your deep brain somehow associates any human being as a threat. In this case, AI could be a savior for you, helping you till you're able to open up to humans again"* (G1 Reddit Post 9, 4|3).

Even if one gets traditional therapy, AI could still be useful, as one user suggested *"Can also use as a second opinion to your MD if you've got one"* (G1 Reddit Post 9, 4|28).

On the other side, numerous X- and Reddit users regarded AI's abilities as too limited for therapy, like *"[…] people working on mental health tech specifically chatbots, you're providing solutions to clients' concerns by creating your therapy FAQs. That's not what therapy does"* (G1 Tweets X 2023, 3|44) and

> *"It's missing genuine curiosity about your life. It's also missing the ability to notice 'interesting' things you talk about. It can't retain relevant life information about you unless its relatively recent"* (G1 Reddit Post 5, 4|19).

One user noted that a chatbot might have outstanding conversational qualities, but

> *"[…] there will be no convincing someone that the AI has 'lived' the experiences that bring those insights – it's just a computer program that knows how to say the right combination of words"* (G1 Reddit Post 10, 4|8).

Also, unpredictability seemed to be problematic: *"but machine learning chatbots are just too volatile to put in front of someone in a vulnerable state."* (G1 Tweets X 2023, 3|118).

Lack of emotions was a special matter of debate: *"AI can't feel, nor is it advanced enough to handle heavy topics like that. […] People want a therapist they can connect with, not one that treats you like a task"* (G1 Reddit Post 14, 4|14). By contrast, other users claimed emotions to be overrated or even misplaced in therapy:

> *"It doesn't need to have emotion. It needs to _understand_ emotion. Keep in mind that a therapist's job is most certainly _not_ to feel what you're feeling and commiserate with you. In fact, maintaining emotional distance is an important part of their training. […]"* (G1 Reddit Post 14, 4|102).

One user highlighted the benefits of lacking emotions: *"Yup AI […] won't get frustrated and flip out on you, won't try to track you down and if YOU screw up and cross a line it's not a big deal and you won't get blackmailed or worry about it"* (G1 Reddit Post 7, 4|10).

■ Distrust in AI

This subtheme had another high relevance for G_1, 105 segments were labelled with first-order code [2.1] (cf. Figure 11Figure). One group of critics regarded companies as the only

beneficiaries, like *"If AI did the work of jobs and *the people* owned the AI, it could help. But that's probably not going to happen"* (G1 Reddit Post 8, 4|3) and *"[…] we need laws in place to protect the mental health of walmart and the pharmaceutical industry"* (G1 Tweets X 2023, 3|42). Likewise, people rejected AI, as they feared dubious business practices: *"I dont trust ratings and scams are too common […]"* (G1 Reddit Post 1, 4|1). Another group distrusted AI technology itself: *"Language models only repeats what it finds online. Most advice online is trash which is what is what you will get from a language model. […]"* (G1 Reddit Post 9, 4|13).

■ Product visions for mental healthcare

A minority of X- and Reddit users shared ideas how to advance mental health chatbots. An X-user thought of gender-related aspects: *"• At some point, chatbots will become more than good enough to act as emotional support for lonely men.• Also, contrary to the current climate, having 'AI girlfriend' will become socially acceptable. […]"* (G1 Tweets X 2023, 3|2). Likewise, two Reddit users even expected that

> *"The first obvious step towards mental helth-related treatment assignments should be embedding some Eva AI-like virtual gf [engine](https://www.evaapp.ai/) into realistic sex dolls. […] USP's obvious, the tech is already developed to that level"* (G1 Reddit Post 13, 4|48)

and – besides erotism:

> *"[…] we have cuddly AI smart pets that can talk and help us accomplish things, our species will change big time. Imagine a mental health buddy on your side 24/7. Maybe even a lifetime friend. […]"* (G1 Reddit Post 13, 4|47).

▶ Theme 2: Further UX considerations

Pattern code [3]

Though the conversations in G_1 were more generalized in nature, Reddit and X-users also identified several AI applications that can be utilized to address mental health concerns. Besides offering recommendations by word of mouth (and creating ties to other users), G_1 also compared alternative options.

■ Comparison between apps

Most contributors brought up ChatGPT and its most recent version, GPT-4: They considered it as a reference model for mental health chatbots, even though it is not intended for mental

healthcare. Several people reported that they utilized this app sporadically and irregularly to get assistance on mental health: *"I often pop into chatGPT to talk about some random thing that's bothering me. It tends to give pretty great advice"* (G1 Reddit Post 10, 4|15). Furthermore, X-users recommended mental health apps based on GPT-4, like

> *"🌐 Moreover, chatbots powered by Chat GPT-4 can be trained to understand and respond to different languages and cultural contexts. This can be particularly useful in regions where mental health services are scarce or where there are language barriers"* (G1 Tweets X 2023, 3|2).

Some users also emphasized the ease of using ChatGPT as a one-stop platform for inquiries regarding mental health concerns and more:

> *"[...] It's free and is more of an interactive journal. plus it doesn't judge me and doesn't mind if I change topic suddenly to ask if Tywin Lannister would win against The Riddler at chess or some silly question"* (G1 Reddit Post 14, 4|67).

▶ Theme 3: The context of mental health issues

Pattern code [4]

Instead of debating new therapy solutions, some users suggested to think about causes first, for example *"Therapy does not work when the majority of problems are due to growing wealth gap, increasing costs of living and low wages, leading to financial problems and anxiety"* (G1 Reddit Post 14, 4|77). Another group regarded chatbots as another reason for mental health issues:*"AI is going to lead to the greatest mental health crisis in history.If all we do is end up talking to chatbots, we won't know how to deal and interact in person anymore.Loneliness leads to depression"* (G1 Tweets X 2023, 3|78). One user saw two sides of the same coin: *"[...] The same companies that brought in this mental health epidemic through social media and addiction are selling us soulless solitions"* (G1 Tweets X 2023, 3|20).

10.2. G_2: Users/Customers

▶ Overview

An analysis was conducted on 1,952 user reviews for six mental health apps: 1,531 reviews came from Google Play Store and 421 reviews from Apple App Store, with reviews spanning from January 2023 to mid-June 2024. The apps with the most reviews were Woebot (617), VOS (619) and Youper (363).

Figure 12 shows that UX issues dominated in G_2, as most reviews belong to pattern code [3]. Likewise, in 1,421 reviews users appreciated the implemented AI or reported positive effects after app usage, as pattern code [1] reveals. Pattern code [2] shows that a significantly smaller number of reviews (259) dealt with criticism towards the implemented AI or with negative effects from app usage. In G_2, debates on research results or explanations for mental health problems were almost nonexistent (pattern code [4]):

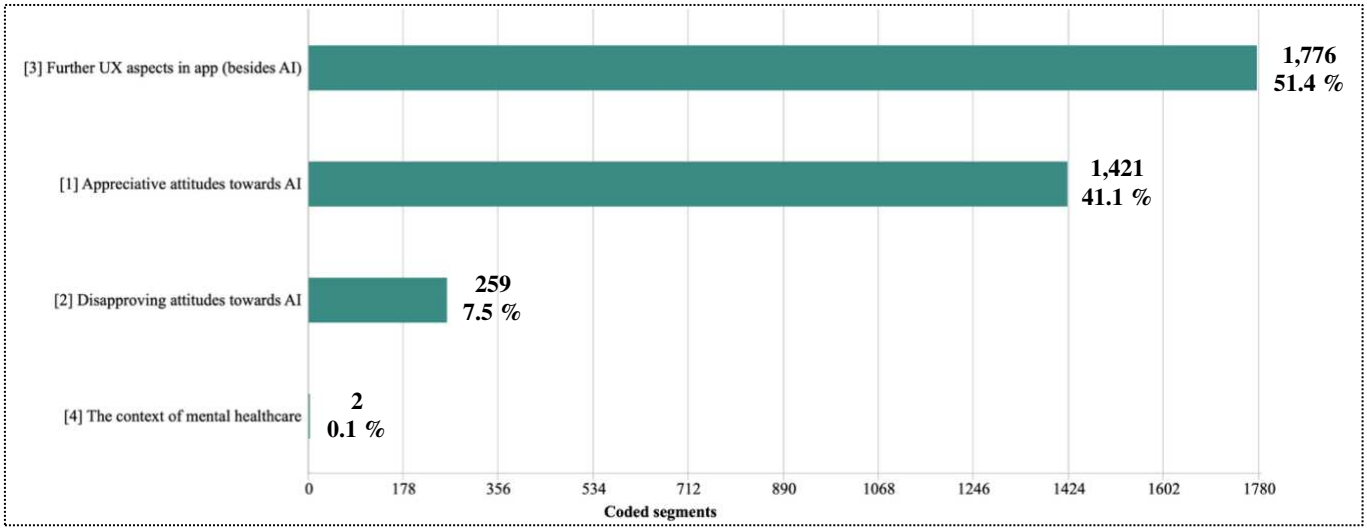

Figure 12: Distribution of pattern codes in G_2 (designed with MAXQDA 24).

As Figure 13 illustrates, the top 3 first-order codes in G_2 were [1.3] Positive self-perceived effect of app usage (1,239 coded segments), [3.4] UI & app features (707 coded segments) and [3.5] Relationships (app, users,...) (560 coded segments):

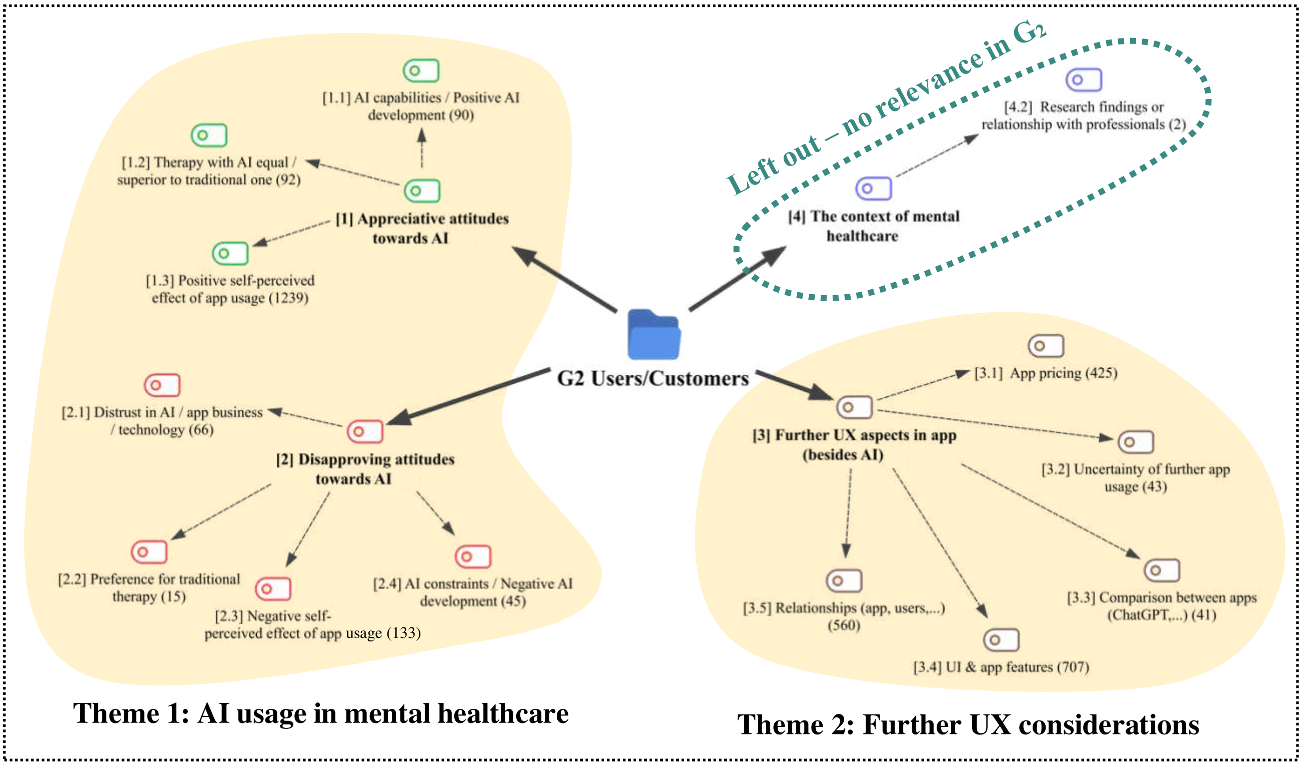

Figure 13: Structure map of themes and codes in G_2 (designed with MAXQDA 24).
In square brackets: code numbers, in round brackets: coded segments.

January 2023 to mid-June 2024:

Figure 14: Dominant words in user reviews for **all six apps** in both app stores
(designed with MAXQDA 24).

The words that users most frequently used in their reviews are seen in Figure 14: *help* (5,815 hits) and *use* (5,374 hits), other popular terms were *feel* (3,248 hits), *love* (2,544 hits), *free* (2,179 hits), *pay* (1,950 hits), and *cbt* (907 hits). When it came to app names, users mainly mentioned *woebot* (3,276 hits), with much less mentions of *youper* (590 hits) and *vos* (338 hits). However, as Figures 15 and 16 show, word choices differed between the apps. For instance, there was an unbalanced distribution of *pay* and *love*: 32 hits with *love* were found in Youper reviews, compared to 107 hits with *love* in VOS reviews. In contrast, *pay* had 128 hits in Youper and was placed second in the app; nevertheless, *pay* received 68 hits in VOS, placing it in 13th place.

January 2023 to mid-June 2024:

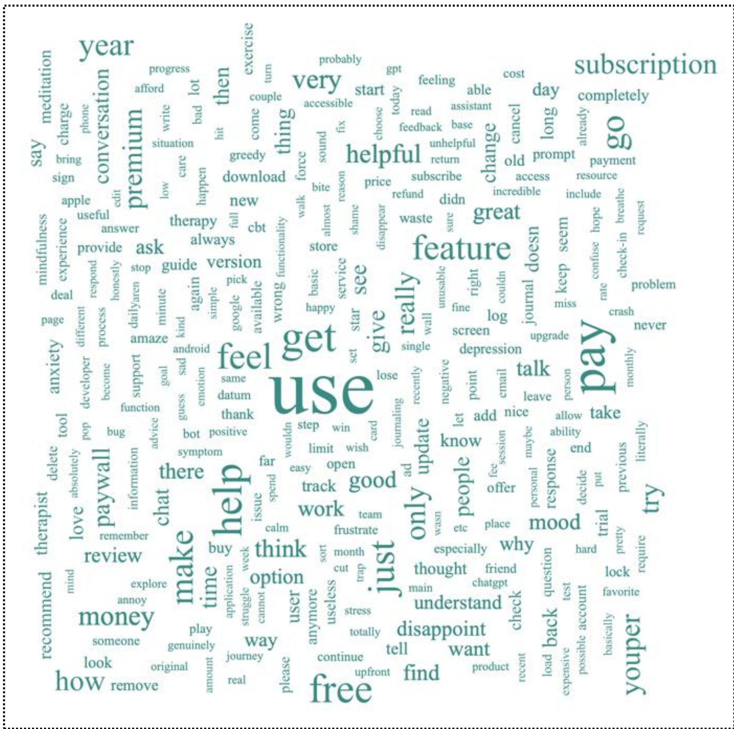

Figure 15: Dominant words in user reviews for **Youper** in both app stores (designed with MAXQDA 24).

January 2023 to mid-June 2024:

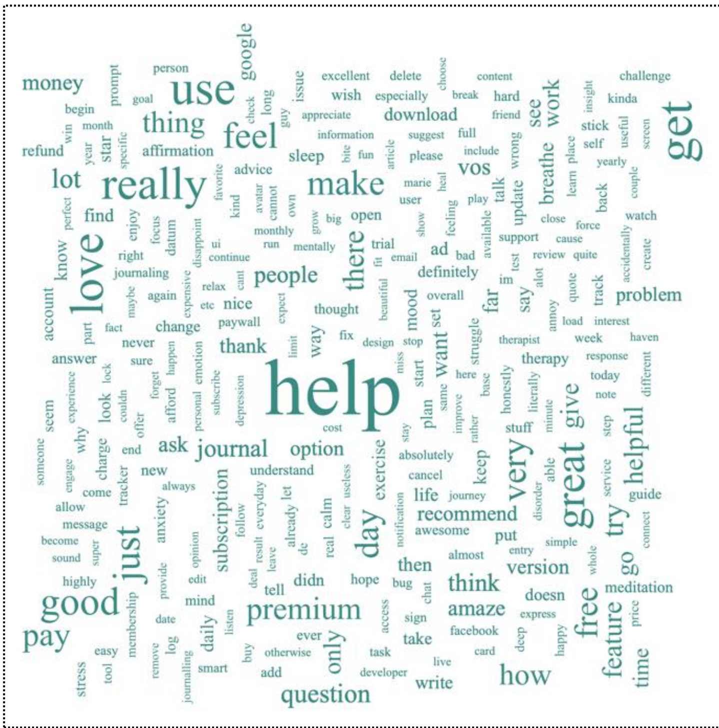

Figure 16: Dominant words in user reviews for **VOS** in both app stores (designed with MAXQDA 24).

Based on these general observations, user reviews from G_2 can be distinguished between two main themes. Unlike G_1, there is no third theme here, because pattern code [4] was irrelevant (cf. Figures 12 and 13).

▶ Theme 1: AI usage in mental healthcare

Pattern codes [1] + [2]

The AI component of mental health apps was the focus of most G2 reviews. User feedback varied considerably throughout the apps because it was product specific.

■ Capabilities and constraints of AI in mental healthcare

A major advantage for users was the universal accessibility of a mental health app, as two users remarked: *"It's not therapy but it's a good resource to have with you every where you go"* (G2 Reviews AA Woebot 4b (2023), 4|200) and *"24/7 !!!24/7 people!'Someone' who is backed by a team of compassionate creators is there for you 24/7"* (G2 Reviews AA Woebot 4b (2023), 4|251). One user mentioned bad experiences with traditional therapies and tried Youper instead: *"Therapists don't know what to do with me, and most look at me with bias, so it's nice to have support that is unbiased, even if it's just an AI bot"* (G2 Reviews AA Youper 1b (2023), 4|27). Also, the price difference to traditional therapies was noteworthy for several users, like: *"I've paid for a year's access, for me half the cost of one in person therapy session, it's real value for money"* (G2 Reviews GP VOS 5a (2024), 3|24). Another user mentioned the burden of finding a human therapist: *"[...] I find it difficult to go through the whole process of finding a mental health provider, figuring out insurance, scheduling an appointment, rearranging my routine. [...]"* (G2 Reviews AA Woebot 4b (2023), 4|126). People with anxiety issues may profit specifically from AI therapy, as expressed in this review: *"If i wanted to pay to talk to someone i would simply go see a therapist but I got anxiety and trust issues which prevents me from doing so"* (G2 Reviews AA VOS 5b (2024), 4|3). A Woebot user perceived AI as superior to online therapists but still saw further potential:

> *"It's way better than some of the therapists you get online. It uses casual text language and has a number of tools and resources in its library. It feels at times like you're chatting with a real empathetic person on the other side. I can only imagine it getting better with the GenAI revolution"* (G2 Reviews GP Woebot 4a (2024), 3|35).

On the other side, users mostly complained about the limited conversational abilities of the chatbot. Concerning VOS, one user posted: *"[...] The messages that AI tells you are very similar, I've noticed it's using certain specific words actually. And also the articles don't come from a wide range. It repeats a lot. [...]"* (G2 Reviews GP VOS 5a (2024), 3|50). Likewise, the chatbot was not responsive enough for users, such as Woebot:

> *"however when the bot asks for your input, it is not read. You can type anything you want in the text box and the bot will tell you that you did well and completed the assignment. It is definitely not interactive"* (G2 Reviews GP Woebot 4a (2023), 3|39).

In the case of Limbic, a user remarked that the chatbot left context out:

> *"I would give this a five-star review if certain words and phrases didn't trigger an emergency response from this app. I find it extremely frustrating that I can't say what I want to say. I'm guessing that the app identifies certain words as trigger words, but it would be better if it could identify as phrases since it is AI"* (G2 Reviews GP Limbic 3 (2024), 3|38).

Sintelly's behavior was criticized as too robotic and instructed, like:

> *"The chat AI will listen to what you say for a few messages, then try to forcefully move you onto exercises and abruptly end the chat. It's impossible to actually just have a conversation with it. Therapy shouldn't feel like you're being pressured"* (G2 Reviews GP Sintelly 6 (2024), 3|113)

and *"Just a chatbot that follows some specific sequences. […] It's almost the same as to follow some instructions on YouTube or elsewhere"* (G2 Reviews GP Sintelly 6 (2024), 3|118). Other users characterized the advice delivered as not targeted enough to their issues, as in the case of Woebot: *"Awful for people with ADHD. Assumes you're neurotypical and gives you very generic advice. Plus it's slow and boring and talks to you like a 5 year old. Avoid"* (G2 Reviews GP Woebot 4a (2023), 3|133). Similarly, the language style was perceived as insensitive in certain cases: *"[…] I don't want to 'High five myself' or shout 'Ready set go!' when I feel terrible okay??? Bad programming by people who clearly don't understand mental health"* (G2 Reviews GP Sintelly 6 (2024), 3|87).

Distrust towards AI or the app was expressed, if the system asked for information with unclear relevance. For instance, users repeatedly criticized Limbic for soliciting further contact data:

> "*[…] You'll give your phone number or email, and then the screen will go black. If you try to talk to customer support, you'll get a message saying that they're not in. They're probably stealing your information and selling it. […]*" (G2 Reviews GP Limbic 3 (2024), 3|7).

A Limbic user was surprised by certain clauses: *"Whoa -- read the user/privacy agreement!! They collect information about your triggers AND share your information with commercial investors. Creepy stuff"* (G2 Reviews GP Limbic 3 (2024), 3|33). A similar data collection procedure was observed in VOS: *"Big 'no' from me for wanting a ton of data in the beginning*

*and then (*BEFORE you have even started anything!) it wants a Google or Facebook login. […]"* (G2 Reviews GP VOS 5a (2023), 3|353).

■ Self-perceived effects of app usage

Users particularly valued this subtheme because, in 1,372 reviews (first-order codes [1.3] and [2.3]), they discussed how and why the app helped them with their mental health problems. Out of these reviews, 90.3% contained good feedback, and just 133 individuals expressed dissatisfaction with the outcome. The high level of user happiness with the six applications under analysis does not prove that they are medically beneficial; rather, it indicates that these products serve users' requirements in general. As an example, a Woebot user talked about the instant comfort the app provided:

> *"If you stick with it, your mindset can really start to change. It also helps you keep track of your emotions, and one time, while I was having a panic attack, I used it, and the AI made me laugh so hard I couldn't focus on the panic anymore. […]"* (G2 Reviews GP Woebot 4a (2023), 3|59).

Reviewers saw as major reason for their mental improvements the focused, but non-pushy approach: "*Woebot is helpful in a non pressured way. No dark rabbit holes, just helpful perspectives"* (G2 Reviews AA Woebot 4b (2023), 4|245) and *"I know I'm texting with an AI, but wow it really helped. It makes you focus on the conversation and she says some really valued things. It definitely helped"* (G2 Reviews AA Youper 1b (2023), 4|30). Others emphasized the range of techniques and rewards that aided in their advancement:

> *"Very proactive in making sure you're clocking in and throwing incentives at you. Excellent affirmations and mood tracking and breathing, great journaling and just awesome in general. […]"* (G2 Reviews GP VOS 5a (2023), 3|366).

Only a small percentage of users reported negative consequences from using the app. Interestingly, they condemned a particular app behavior for being unhelpful, but not chatbots or AI treatment in general. Users disapproved, for example, when the chatbot in difficult circumstances gave them few options: *"[…] it's far too simplistic to be helpful for most adults.For one thing, you can't practice CBT skills by selecting from multiple choice options provided by the AI. […]"* (G2 Reviews AA Woebot 4b (2023), 4|134).

▶ Theme 2: Further UX considerations

Pattern code [3]

In addition to the AI component, other app features increased user engagement even further: 1,776 review segments that addressed these issues were found. They fall into one of four subthemes:

■ UI & app features

Users expressed in 707 reviews how they perceived appearance and usability of the apps. Several negative aspects were common among all six apps. For instance, users critically observed orthographic mistakes, like in VOS:

> *"Whoever made or runs this app seems very childish, phobia is spelled 'fobia' on the app and it seems like a lot of the words are spelled incorrect as a joke. It's not very funny. Mental health is such a serious thing […]"* (G2 Reviews GP VOS 5a (2023), 3|341).

Similarly, the gamification approach to make the app enjoyable was sometimes disliked, as this review shows:

> *"[…] It's full of gamification & dopamine rewards. […] The growing tree, levels, badges, constant notifications... they cause quite the opposite – they put pressure & FOMO on me and make me avoid the app. […]"* (G2 Reviews GP VOS 5a (2023), 3|350).

By contrast, it was appreciated if the app had a variety of features to offer. In the case of Youper, one user lamented the disappearance of favored features:

> *"I've used this app for 5 years now, have a subscription, and just saw that the most recent update took away the mindfulness exercises, guided meditation, and breathing techniques… bring them back!"* (G2 Reviews AA Youper 1b (2023), 4|56).

In general, updates were an unpopular issue for most users, especially if they had no reasonable alternative. Woebot, for example, changed its access policy fundamentally in late 2023: The provider decided that app usage should only take place with supervision by healthcare professionals (Woebot Health, 2024a). For further access, all users – even existing ones – needed a code from a Woebot partner (Woebot Health, 2024b). Consequently, many locked-out users expressed their resentments, like this:

> *"[…] now I'm essentially being kicked off.I'm very frustrated and disappointed with this decision. Do you not realize who your target audience is? You cannot claim to be an 'easily accessible' mental health service, and then create another barrier for people without access to healthcare. […]"* (G2 Reviews AA Woebot 4b (2023), 4|17).

The shifts in users' impressions of Woebot between 2023 and 2024 are displayed in Figures 17 and 18. While user opinions of the app were as favorable in 2023 as they were in VOS and Youper, the atmosphere deteriorated in 2024:

Year 2023:

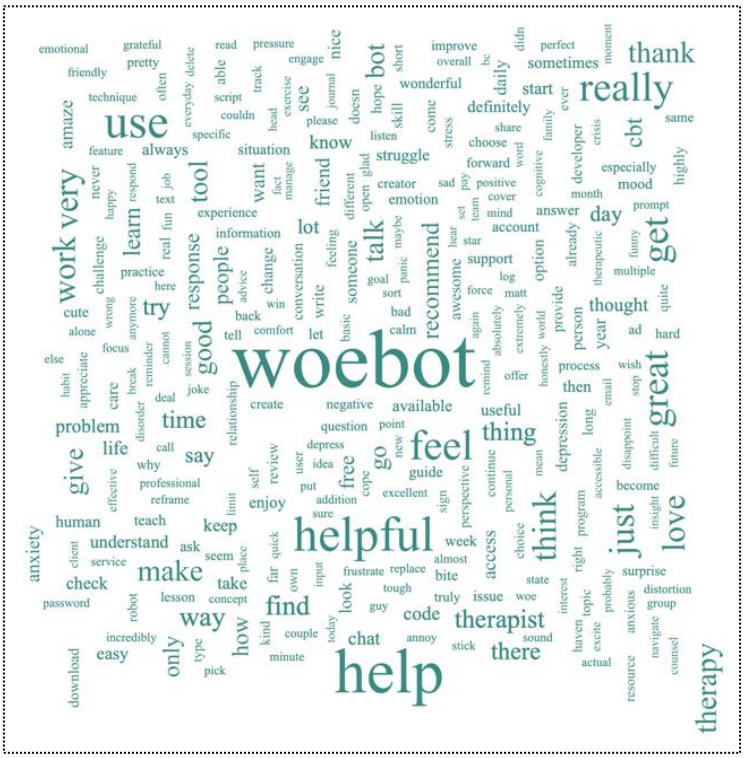

Figure 17: Dominant words in user reviews for **Woebot** in both app stores **before** restricting access (designed with MAXQDA 24).

January to mid-June 2024:

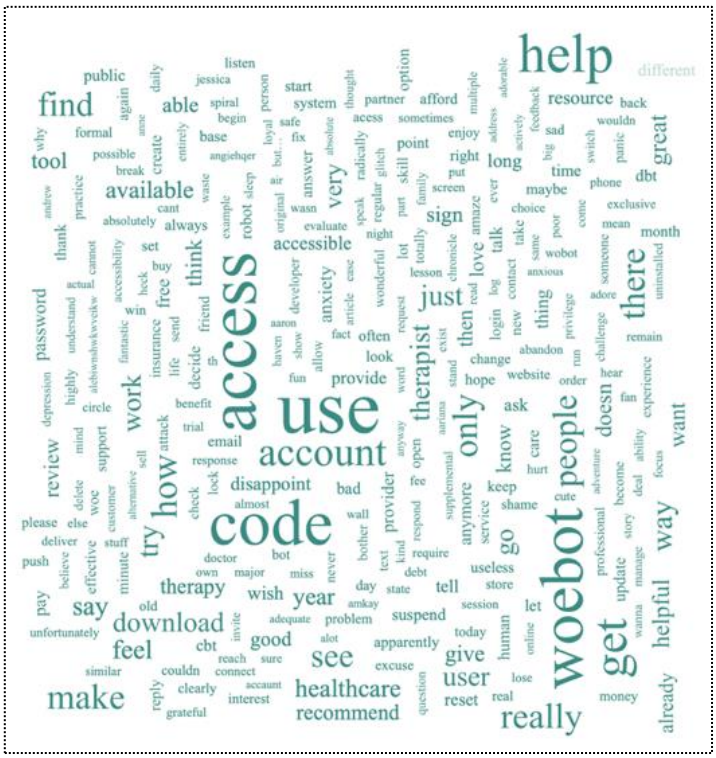

Figure 18: Dominant words in user reviews for **Woebot** in both app stores **after** restricting access (designed with MAXQDA 24).

■ Relationships

The apps were largely distributed by word of mouth: A number of users stated that recommendations from friends, family, or medical professionals had an impact on their decision to download the app. Not unexpectedly, satisfied users also encouraged others to download the app from the app store. Additionally, a unique type of relationship between the user and the chatbot was also identified: particularly after extended use, some users reported feeling attached, romantic, or even friendly. Typical signs of this bond were reviews, like *"He's a cute little friend who checks on you and is entirely there for you no matter what time. Love it!"* (G2 Reviews AA Woebot 4b (2023), 4|142). As Figure 14 displays, the term *love* scored 2,544 hits in app reviews. Users frequently believed that their relationships were reciprocal, as in this review: "*I decided to tell Woebot to call me 'Sweetie' and that's been very helpful when I'm being hard on myself. Great app!*🤖👏*"* (G2 Reviews AA Woebot 4b (2023), 4|206). Such

user attachment was particularly noticeable in apps where the chatbot was the main function, such as in Woebot, VOS, Sintelly, Youper, and Limbic. There were also reports of partnerships ending, which frequently led to sad remarks, like *"Bye bye old pocket robot friend you will be missed."* (G2 Reviews GP Youper 1a (2023), 3|12). The most frequent cause of these disruptions, aside from Woebot's new access policy, was the addition or alteration of a paywall.

■ App pricing

This subject was a major cause of discomfort for users. The apps displayed differences from one another, however: Youper dominated in this domain, recording 191 segments, whereas VOS had 168 segments, and Woebot reviews contained 52 segments. With the other three applications, this was a minor problem. Users who had utilized free features for an extended duration, in particular, articulated their dissatisfaction:

> *"Today I got a reminder to track depression symptoms and find that, not only the questions, but all of my historical symptom data is behind a paywall! I understand having a fee for premium features, but I should at least be able to see the data I have already entered"* (G2 Reviews GP Youper 1a (2023), 3|3).

Some users offered an option to prevent rising fees: *"Add some ads so you get more money to improve the app plz i like this app"* (G2 Reviews GP VOS 5a (2023), 3|259) or, if expenses were unavoidable, demanded more disclosure in advance, such as *"Going through all the questionares to be met with a subscription paid page"* (G2 Reviews GP VOS 5a (2023), 3|278) and

> *"THIS APP IS $69.99. Which is totally fine and I support 100%! It seems like a very professional app, but the cost should really be advertised or shown up front on the Play Store. […] The app seems like it could be a great tool, but it's disappointing to download only to realize you can't use it"* (G2 Reviews GP Youper 1a (2023), 3|128).

In another example, a VOS user complained about a procedure that was both partially free and partially paywall-restricted:

> *"Would be 5 stars but... it's mildly unfair that people who dont have premium can do the tests but not see their results dont get me wrong amazing app i use it everyday for over 100 days now but Why make the tests available but not let us see our results?"* (G2 Reviews GP VOS 5a (2023), 3|9).

Comparable to the previously mentioned case of Woebot, dedicated users with established connections to the chatbot exhibited sentiments of disconnection:

> *" […] This app is only a shadow of what it used to be- everything used to be free, the app would talk to you like a friend, and now the app barely talks to you. My husband and I are both disappointed. […]"* (G2 Reviews AA Youper 1b (2023), 4|1).

Nonetheless, a paywall was not generally viewed unacceptable by users. In the instance of Woebot, one user expressed a preference for it over total expulsion from the app: *"[…] I would even pay a small fee to keep him.* 😢 *"* (G2 Reviews AA Woebot 4b (2023), 4|18).

■ Comparison between apps

If users were dissatisfied with the AI component, particular features, or pricing structures of the current application, they compared it to competing products, as illustrated in this instance:

> *"This app is trash. Give me one reason why whould I use it instead of Wysa, VOS, or Intellect? It's so greedy, I can memorise the questions and do a check in using pen and paper, and chat with Wysa or Pi.ai PiSupport. […]"* (G2 Reviews GP Youper 1a (2023), 3|22).

Some users of VOS and Sintelly appeared unconvinced by these products and instead favored well-known non-AI apps: *"[…] I would not pay for this when I can get the same from Pinterest and journalling"* (G2 Reviews GP VOS 5a (2023), 3|215) and *"It's almost the same as to follow some instructions on YouTube or elsewhere"* (G2 Reviews GP Sintelly 6 (2024), 3|118).

However, most comparisons were made with ChatGPT: *"The chatgpt android app just came out and is so much more helpful"* (G2 Reviews GP Youper 1a (2023), 3|109) a Youper user commented, whereas a Woebot user remained satisfied with the product but anticipated improved AI functionalities in the future: *"[…] It feels at times like you're chatting with a real empathetic person on the other side. I can only imagine it getting better with the GenAI revolution"* (G2 Reviews GP Woebot 4a (2024), 3|35).

10.3. G_3: Developers/Providers

▶ Overview

An analysis was performed on the two interviews with the CEOs. Figure 19Figure indicates that UX issues largely prevailed in G_3, with the majority of statements classified under pattern code [3]. The respondents' views on AI revealed that negative attitudes (pattern code [2]) marginally exceeded positive ones (pattern code [1]):

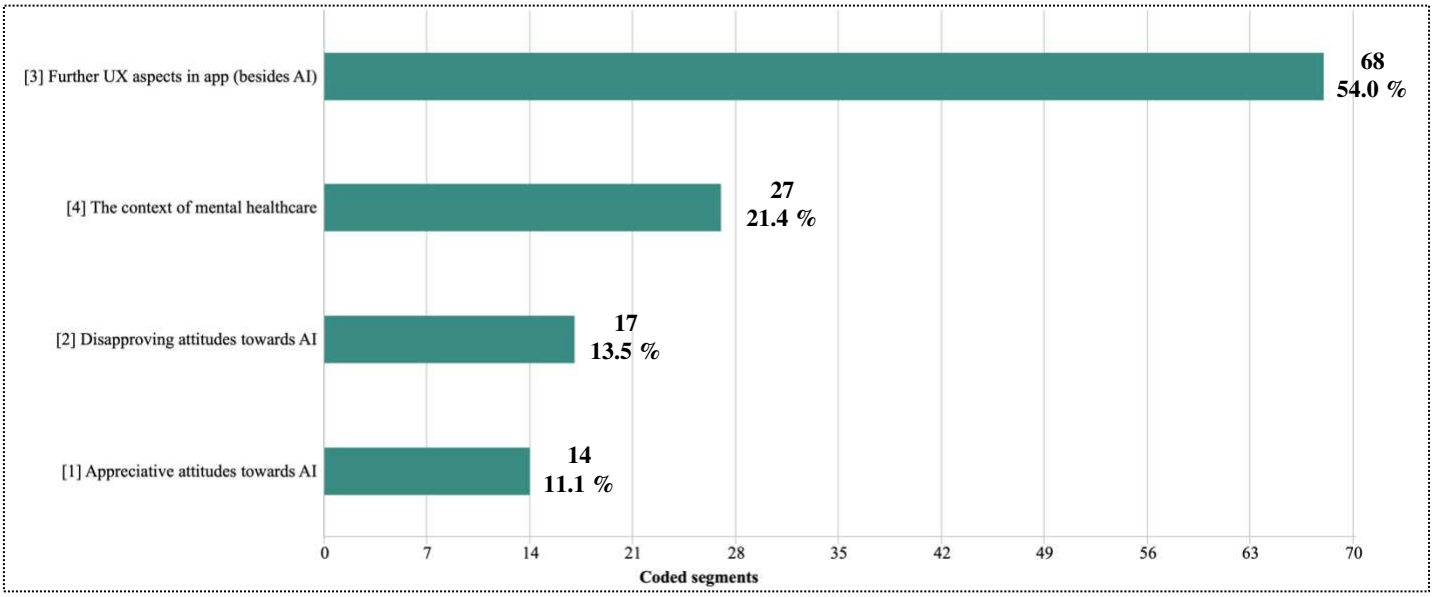

Figure 19: Distribution of pattern codes in G_3 (designed with MAXQDA 24).

[3.4] UI & app features (32 coded segments), [4.2] Research findings or relationships with professionals (27 coded segments), and [3.5] Relationships (app, users,...) (20 coded segments) were the top 3 first-order codes as Figure 20 shows. Three recurring themes from G_1 and G_2 emerge:

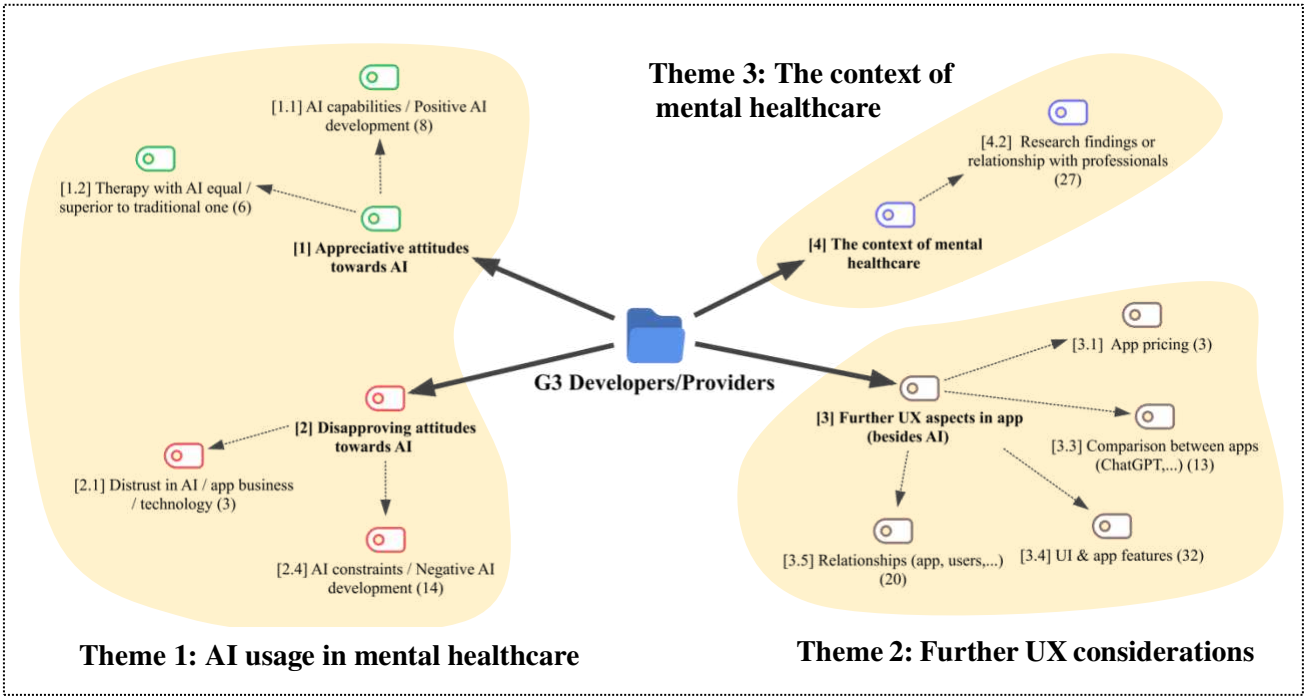

Figure 20: Structure map of themes and codes in G3 (designed with MAXQDA 24).
In square brackets: code numbers, in round brackets: coded segments.

▶ Theme 1: AI usage in mental healthcare

Pattern codes [1] + [2]

The future promise of AI – particularly generative AI – was brought up by both respondents. However, they also raised concerns about recent advancements.

■ Capabilities and constraints of AI in mental healthcare

For CEO 2, AI had disruptive character in mental healthcare as it could dissolve time and space: *"So, we're changing the way people were thinking about therapy. Traditionally you would do therapy, let's say, once or twice a week"* (G3 Interview 2 07122023, Pos. 9). On mobile apps, by contrast *"[...] you don't need to schedule anything. You just need to pick your phone. And you can start talking"* (G3 Interview 2 07122023, Pos. 11). This would create new opportunities for therapy, as

> *"[...] It's something that was made for the modern life where people don't have time. They want to do that when they are waiting for their bus or when they are, I don't know, commuting to work or maybe before bed. [...]"* (G3 Interview 2 07122023, Pos. 13).

CEO 1 saw this convenience also as relief for medical workforce: *"[...] if artificial intelligence serves as an info point. It would take that boring job from humans and humans could focus on*

more creative thoughts. […]" (G3 Interview 1 05122023, Pos. 49). CEO 2 also mentioned that AI apps for mental healthcare would be reasonably priced and open to anyone:

> *"[…] Because here in the US, seeing a mental health professional is very expensive. […] And it will also remove the fear of seeing a mental health professional. […] I need to talk about my inner thoughts. I need to share my emotions. I don't want to do that. Maybe I can start with this app. […] And then that was the initial idea: Remove stigma and remove the barrier of cost"* (G3 Interview 2 07122023, Pos. 3).

However, AI's inanimate character would also be a disadvantage for mental healthcare as CEO 2 remarked *"[…] I think something that AI will never be able to deliver is the human relationship, the real, the feeling. […]"* (G3 Interview 2 07122023, Pos. 41). CEO 1 saw a risk in AI chatbots for vulnerable users if they were not supervised by a human:

> *"[…] So, they might mishandle the tool, they might misinterpret the tool. They might think that chatbot is giving some advice. And sometimes chatbot can give better advice, but it's a vulnerable individual, maybe they might not realize when there is like a better answer to one of their questions […]"* (G3 Interview 1 05122023, Pos. 47).

A current limitation would be the memory of such chatbots: *"Right now [app name] won't remember. So, imagine that I had this conversation with you last week. And then you ask the same questions and then I'll answer the same way"* (G3 Interview 2 07122023, Pos. 35).

Both respondents saw manifold potential for AI apps in mental healthcare, but their development should be more equilibrated, as CEO 2 thought:

> *"[…] And we should meet in the middle, right? We in the US need to learn more about safety and being more conservative. And Europe should learn a little bit more from the US to be more innovative, trying things and so on. […]"*(G3 Interview 2 07122023, Pos. 67).

▶ Theme 2: Further UX considerations

Pattern code [3]

In their statements, the two CEOs outlined divergent marketing strategies and different core features in their products. CEO 1, responsible for the app with the additional chatbot, mentioned that *"[…] the app is getting a lot of press and a lot of interest in the chatbot. I mentioned the AI interface. But I think that's only or mainly because now everybody is talking about that"* (G3 Interview 1 05122023, Pos. 17). For that reason, CEO 1 concluded that

> *"[…] we are not following the trends of pushing us as a chatbot or artificial intelligence. That's not our marketing strategy. What we try to do is like position ourselves for self-therapy. So, to empower people"* (G3 Interview 1 05122023, Pos. 25).

As the product's core feature CEO 1 considered

> *"[…] I think what helps with retention is like something that you build as a feature in the app. […] We have the journal in the app exactly for this purpose. So, users have the opportunity to return every day and to write something […]"* (G3 Interview 1 05122023, Pos. 51).

However, CEO 1 added that *"[…] there are a lot of diary journaling apps out there. What we do have is like a series of recorded courses, which is like some programs where users can do self-therapy"* (G3 Interview 1 05122023, Pos. 19). These courses *"[…] are very in-depth that they last like some of them even 30 days"* (G3 Interview 1 05122023, Pos. 21). A third core feature for CEO 1 was the price: *"[…] but I think what's really important for users is that, first of all, it's affordable. So, a lot of things in the app is free. That's like one of the key characteristics that the users want"* (G3 Interview 1 05122023, Pos. 19).

By contrast, CEO 2 responded *"The most important feature is having a quick conversation that will help you feel better. So, it's the chatbot itself. It's the conversation"* (G3 Interview 2 07122023, Pos. 19). This app would be different to others, as *"[App name] is the only one using LLMs and generative AI"* (G3 Interview 2 07122023, Pos. 29). Regarding competing products, CEO 2 believed that

> *"...they're still using rule-based technology that's not as advanced as [app name]. We are already one step ahead and I would say that there are no other chatbots that are using this type of technology that we are ready to see"* (G3 Interview 2 07122023, Pos. 31).

However, the same respondent added:

> *"[…] In the description we talk about AI, but we don't emphasize AI or we don't emphasize the technology. We emphasize the benefits. But I think no one thinks that they're talking to a person. So, they know that they're talking to an AI"* (G3 Interview 2 07122023, Pos. 39).

Both CEOs considered engagement to be a critical component in user retention. Providers can encourage this soon after the download: *"[…] so, we are always talking to our users. There is this onboarding process when you start using [app name], you answer some questions"* (G3 Interview 2 07122023, Pos. 7). To keep user engagement high, CEO 1 explained:

> *"So, for example, send some push notifications sometimes that it doesn't have to be with a selling point […] This type of initiative works for both, in a way of creating brand*

> *awareness, because people, users will think that this brand is doing cool stuff. And eventually will bring them more frequently to the app, they might use it more"* (G3 Interview 1 05122023, Pos. 51).

▶ Theme 3: The context of mental healthcare

Pattern code [4]

The significance of maintaining strong relationships with mental health specialists throughout the app lifespan was stressed by both respondents. In CEO 1's instance, psychologists served as product development catalysts:

> *"They had some attempts with groups online, and they were growing, they noticed some interest. We decided to better built something and give it a name. So it's to build something that could be much more scalable. An application was the result of this, this ambition to grow and to add more people. So, the input came from psychologists themselves"* (G3 Interview 1 05122023, Pos. 5).

CEO 2 described the role of therapists as partners for product dissemination:

> *"[…] But what I saw that was that most of the mental health professionals, they were seeing [app name] as part of a toolkit. […] So, the reaction was pretty good. And what we noticed was that a lot of users, they were coming to us getting to know [app name], because their therapist recommended [app name]. […]"* (G3 Interview 2 07122023, Pos. 5).

Interestingly, their motivation to recommend the app to patients was higher at the beginning:

> *"[…] because we started the module and [app name] was 100 % free. So, therapists, they really loved that and recommended the app to our users. When [app name] started to have this premium model that there are some free features and other features that are paid – the recommendation from mental health professionals dropped"* (G3 Interview 2 07122023, Pos. 5).

Additionally, both respondents mentioned to incorporate approved therapeutic standards, like CBT into product design and to consider current research findings. However, both CEOs objected to transform their products into medical devices. Instead, CEO 2 proposed a middle course to steer between wellness apps and digital therapeutics:

> *"I see that there are three main categories of mobile health apps […] And there is the second layer that will be an app that's backed by science, which will help prevent or even*

> *improve mental health symptoms like depression and anxiety […]"* (G3 Interview 2 07122023, Pos. 59).

Likewise, CEO 1 highlighted the significance of voluntary certifications for promotion

> *"[…] one for Europe and one for the US. But we plan to to get more certifications. We are also working with one European project, funded by the European Union, and they are building […] an European health label for health apps. […] We are one of the only 24 apps in Europe that decided to pilot this program. […] And as soon as this certification label would become available, we will be the first ones"* (G3 Interview 1 05122023, Pos. 59, 61).

11. DISCUSSION

This thesis examined viewpoints and opinions of three stakeholder groups concerning the utilization of AI in mental health apps. Data indicate that chatbots represent a crucial AI technology in mental healthcare, having evolved into demanded tools with an increasing user base over the years. As compared to conventional healthcare, eHealth may engage a larger population: Presently, the most popular chatbots for mental health services are assisting over a million people each (see Table 1) – a figure unattainable by any human therapist (Kelders and Howard, 2018, p. 50).

Nonetheless, the public and non-users frequently express suspicion towards these applications, alleging that they are commercial and lack transparency: In a slim majority, 37.7 % of all coded segments in G_1, people generally disapproved of AI in mental healthcare apps, whereas 11.2 % of all coded segments G_1 dealt with distrust in AI or associated app business (cf. Figures 10 and 11). Non-users, in addition to holding idealized views of in-person therapy, are particularly inclined to compare AI solutions to human therapists. Critics also contend that AI solutions lack empathy and human emotions, and cannot replace human therapists. Only a few non-users on X and Reddit offered ideas to advance mental health chatbots, such as integrating AI solutions into dolls or cuddly animals to provide a more lifelike interaction.

Despite the majority of non-users to see them with a negative lens, users perceive chatbots as beneficial in scenarios when human therapists are inaccessible or prohibitively expensive, often viewing them as a complement to in-person treatment rather than a replacement. Users typically praised AI use in mental healthcare in 41.1 % of all coded segments in G_2, 35.8 % of coded segments concerned the positive effect of app usage, which users perceived (cf. Figures 12 and

13). Unlike critics in G_1, users of such applications (both in G_1 and G_2) typically claim the lack of a human touch as the principal motivation for their usage. They seek opportunities for disclosure without being concerned about judgment, regardless of time or place. A multitude of users perceive chatbots for mental health as an adjunct to in-person therapy. In addition, this reflects the intentions of the providers: Boucher (2021, p. 45) asserts that providers did not intend to create chatbots as replacements for human therapists, a position reinforced by remarks from the two CEOs in G_3 and several X-tweets in G_1 (see Figure 8, where *support* is the primary phrase). The restricted accessibility and significant expense of in-person therapy drive affected individuals to seek digital assistance. A certain level of suffering seems to make people less reluctant to engage with these commercial products, which can provide ambiguous information about their technology or therapeutic approach, as exemplified by the app VOS. Despite such concerns expressed by critics in G_1, app users showed a lower awareness of these problems.

Furthermore and in line with existing literature, users may form strong attachments to mental health chatbots, comparable to friendship or even pet ownership due to numerous affectionate descriptions. In 560 coded segments in G_2, users indicated a relationship, the majority of which was with the chatbot – suggesting that these ties are an important UX component of mental health chatbots and that these products have progressed beyond their infancy. The findings confirm that several factors are required for the establishment of such *parasocial interactions*: Users value the ability to select both the length and timing of discussions without pressure. Another requirement is that the implemented AI is both innovative and entertaining, in addition to being responsive, possessing both pragmatic and hedonic traits, as noted by Følstad and Brandtzæg (2020, p. 6).

However, in addition to strong ties with the chatbot, there is the phenomenon of attrition, which is mostly caused by pricing models that hide the most features behind paywalls. Attrition can also impact existing relationships with the chatbot: Previously enthusiastic users may interpret these departures as a rupture of friendship, as my findings show. The 425 coded segments in G_2 underline the relevance of price considerations for app users. Due to the unregulated and unsubsidized nature of the market for mental health chatbots, customers generally face costs for their usage. New clients are attracted to test various features due to the freemium model. Data indicates that in order to make their goods profitable, providers must carefully consider which and how many functions to conceal behind a paywall. With regard to Eysenbach's (2005, p. 5) *law of attrition*, users tend to explore all features immediately after downloading, at no expense. Many users brought up this issue and expressed disapproval of the current approach in a number of apps. This finding is significant and aligns with the research of Lorenzo-Luaces

et al. (2024, pp. 3–5), who indicates a pronounced price sensitivity regarding mental health apps among all ethnic groups in the United States. Alqahtani and Orji (2020, pp. 2058, 2060), on the other hand, regard it as a minor concern, as they discovered just a small number of users expressing complaints. The varying results in prior research could be attributed to the fact that providers have changed their price policies in recent years: According to CEO 2 from G_3, their product was free until they implemented a freemium model (G3 Interview 2 07122023, Pos. 5). In 2022, Ahmed et al. (2022, p. 8) reported that the majority of such apps was free, while *"[...] people did not take well to in-app purchases in some cases"*.

In case of critical reviews, users often contrasted the current app with previous versions. According to the pattern observed, many users responded unfavorably primarily to developments in Youper and Woebot in 2023 and 2024. The reasons behind these critical appraisals were a changed freemium model in the case of Youper and a new access policy in the case of Woebot, as well as competition from free platforms such as YouTube and, notably, ChatGPT (version: GPT-4). Other studies have already noticed usability factors, such as payments and updates; nevertheless, the comparison with other applications and technology was a new reason for negative reviews that I identified in this work. For this user group, generative AI has become the benchmark for mental health chatbots, whereas less complex AI-driven systems are increasingly losing in acceptance since they appear less intelligent. Previous research on mental health chatbots did not reveal this trend because it was mostly conducted before ChatGPT's introduction in 2023. Chaudhry and Debi (2024, p. 12) admit that LLMs provide a more tailored UX, but they make no mention of rising user expectations as a result of this technology. Despite scientific reservations about employing ChatGPT and generative AI for mental health, stakeholders from G_1, G_2, and G_3 praised the tool's versatility and ease of use as compared to mental health chatbots. One of the two providers in G_3, currently employing this technology, emphasized the enhanced conversational capabilities of generative AI. There is a strong possibility that more providers will incorporate this technology into their products, leading to more adaptive chatbots.

In addition to pricing strategies and technology choices, mental health app providers have to define their targeting approach: In general, providers have not distinguished between user groups so far (Alattas et al., 2021, pp. 805–806). Woebot is now the exception in this regard, as it revised its access policy in 2023 to only admit users who had previously established a connection with a human therapist. At present, it is unclear if other providers might follow Woebot's step, which also includes increased reliance on physicians as distribution partners and less possibilities for the AI component to learn due to a drop in new consumers.

The choice of which user groups to target is related to the style of product communication. Given the similarity of product characteristics and intensifying competition among providers, it is reasonable to assume that suppliers of mental health applications will increasingly emphasize their product's unique selling proposition in the future.

12. CONCLUSION

The findings here reinforce the idea that mobile applications, including chatbots, possess special qualities that differentiate them from traditional mental healthcare. My thesis argues that, unlike previous studies, the efficacy of mental health chatbots should be evaluated not solely on medical evidence but also on users' self-assessments. The majority of investigated users reported positive sentiments, even supportive relationships, after utilizing such products. This result is consistent with existing literature, which focuses on the benefits of mHealth for consumers: Mental health chatbots can provide a meditative and cathartic experience, as Hanley and Wyatt (2022, p. 83) state, that fits users' need to communicate feelings without feeling embarrassed or judged. Self-diagnoses cannot substitute for professional medical assessments; however, they illustrate that AI has the potential to lessen the demands on conventional mental healthcare. Researchers have also identified persistent challenges, such as the lack of transparency in these applications; yet, these concerns do not deter users from utilizing such products.

My results also provide practical implications. When it comes to mental health chatbots, the lack of a human touch is a major characteristic for both users and non-users, although their opinions differ. Non-users see the absence of real emotions as a drawback, whilst users like it for being non-judgmental. So far, relevant app developers have stressed their product's constant availability and the scientific evidence behind them, but they may modify their promotional strategies to reach skeptics as potential users. One approach is to highlight the advantages of an emotionless app: Unlike human therapists, it is less prone to errors in diagnosis, while it better adheres to therapy standards such as CBT and ACT (Minerva and Giubilini, 2023, pp. 813–814; Sharma et al., 2023, p. 6).

According to my findings, the majority of mental health chatbot users tends to be open-minded and technologically adept. Given the fierce rivalry among suppliers, this attitude opens up new opportunities for product development, such as personal sensing or user content analysis (D'Alfonso, 2020, pp. 112–114). However, developers may consider a modular structure for

their products, incorporating such features as separate components to pick from, in addition to the AI chatbot. From a marketing standpoint, this might create a more distinct product identity and increase customer retention. Consistent with prior studies, effective mental health chatbots ought to be developed as informal conversational partners rather than medical instruments, as the majority of users favor enjoyable applications that prioritize user-friendliness (Abd-Alrazaq et al., 2021, pp. 8–9; Haque and Rubya, 2023b, p. 9). Contrary to Følstad et al. (2019, p. 153), who view health chatbots as linear coaching instruments, this calls for the tools to have AI capabilities because rule-based mechanisms are unable to meet such requirements.

Payment-related concerns have become integral to the UX of mental health chatbots, and providers should reconsider the structure of their freemium models. Recent decisions caused significant criticism from users, particularly loyal ones, which may impact product reputation. Another justification for a highly modular product structure is that users usually expect a range of free features with notable benefits before deciding to pay for additional features.

My work clearly has some limitations. First, I have only looked into English-language reviews, primarily from US users. The results might not apply to other markets and countries. Spanish and Japanese speakers are the target group for significant chatbots for mental health, such as Yana and Awarefy, however I did not consider these languages. German-language reviews of European apps, such as VOS, are also accessible, but I equally left them out. Second, care should be taken because the G_3 sample size was small and the findings may not apply to all mental health app providers. Thus, the image of how developers view AI in mental health apps is still incomplete.

The worldwide character of these apps should be the main focus of future research as app providers look to scale their offerings: Users' willingness to pay, their readiness to disclose to these applications, their willingness to develop relationships with them, and their impressions of AI should all be compared across different cultures. Another important topic for future research is why numerous providers do not use AI in their mental health apps, particularly when compared to those that do. This may further enhance our comprehension of appropriate AI applications in mental healthcare.

13. DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the author used QuillBot in order to improve readability and language. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

14. REFERENCES

- Abd-Alrazaq, A.A., Alajlani, M., Ali, N., Denecke, K., Bewick, B.M., Househ, M., 2021. Perceptions and Opinions of Patients About Mental Health Chatbots: Scoping Review. J Med Internet Res 23, e17828. https://doi.org/10.2196/17828
- Aboujaoude, E., 2018. Telemental health: why the revolution has not arrived. World Psychiatry 17, 277–278. https://doi.org/10.1002/wps.20551
- Aboujaoude, E., Gega, L., 2020. From Digital Mental Health Interventions to Digital "Addiction": Where the Two Fields Converge. Front. Psychiatry 10, 1017. https://doi.org/10.3389/fpsyt.2019.01017
- Adamopoulou, E., Moussiades, L., 2020. Chatbots: History, technology, and applications. Machine Learning with Applications 2, 100006. https://doi.org/10.1016/j.mlwa.2020.100006
- Ahmad, R., Siemon, D., Gnewuch, U., Robra-Bissantz, S., 2022. Designing Personality-Adaptive Conversational Agents for Mental Health Care. Inf Syst Front 24, 923–943. https://doi.org/10.1007/s10796-022-10254-9
- Ahmed, A., Ali, N., Aziz, S., Abd-alrazaq, A.A., Hassan, A., Khalifa, M., Elhusein, B., Ahmed, M., Ahmed, M.A.S., Househ, M., 2021. A review of mobile chatbot apps for anxiety and depression and their self-care features. Computer Methods and Programs in Biomedicine Update 1, 100012. https://doi.org/10.1016/j.cmpbup.2021.100012
- Ahmed, A., Aziz, S., Khalifa, M., Shah, U., Hassan, A., Abd-Alrazaq, A., Househ, M., 2022. Thematic Analysis on User Reviews for Depression and Anxiety Chatbot Apps: Machine Learning Approach. JMIR Form Res 6, e27654. https://doi.org/10.2196/27654
- Ahmed, A., Hassan, A., Aziz, S., Abd-alrazaq, A.A., Ali, N., Alzubaidi, M., Al-Thani, D., Elhusein, B., Siddig, M.A., Ahmed, M., Househ, M., 2023. Chatbot features for anxiety and depression: A scoping review. Health Informatics J 29, 1–17. https://doi.org/10.1177/14604582221146719
- Alagarsamy, S., Mehrolia, S., 2023. Exploring chatbot trust: Antecedents and behavioural outcomes. Heliyon 9, e16074. https://doi.org/10.1016/j.heliyon.2023.e16074

- Alattas, A., Teepe, G., Leidenberger, K., Fleisch, E., Car, L., Salamanca-Sanabria, A., Kowatsch, T., 2021. To What Scale Are Conversational Agents Used by Top-funded Companies Offering Digital Mental Health Services for Depression?: Presented at the Workshop on Scaling-Up Healthcare with Conversational Agents, SCITEPRESS - Science and Technology Publications, Vienna, Austria, pp. 801–808. https://doi.org/10.5220/0010413308010808

- Ali, O., Abdelbaki, W., Shrestha, A., Elbasi, E., Alryalat, M.A.A., Dwivedi, Y.K., 2023. A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. Journal of Innovation & Knowledge 8, 100333. https://doi.org/10.1016/j.jik.2023.100333

- Allouch, M., Azaria, A., Azoulay, R., 2021. Conversational Agents: Goals, Technologies, Vision and Challenges. Sensors 21, 8448. https://doi.org/10.3390/s21248448

- Alqahtani, F., Orji, R., 2020. Insights from user reviews to improve mental health apps. Health Informatics J 26, 2042–2066. https://doi.org/10.1177/1460458219896492

- American Psychiatric Association, 2023. Digital Mental Health 101: What Clinicians Need to Know When Getting Started [WWW Document]. URL https://www.psychiatry.org/getattachment/b250c6ff-d1f5-4c4f-8ad1-f478fba5773d/Resource-Document-Digital-Mental-Health-101.pdf (accessed 11/19/23).

- Appfigures, 2024a. Number of mHealth apps available in the Apple App Store from 1st quarter 2015 to 2nd quarter 2024 [WWW Document]. Statista. URL https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide/ (accessed 11/06/24).

- Appfigures, 2024b. Number of mHealth apps available in the Google Play Store from 1st quarter 2015 to 2nd quarter 2024 [WWW Document]. Statista. URL https://www.statista.com/statistics/779919/health-apps-available-google-play-worldwide/ (accessed 11/06/24).

- Aristidou, A., Jena, R., Topol, E.J., 2022. Bridging the chasm between AI and clinical implementation. The Lancet 399, 620. https://doi.org/10.1016/S0140-6736(22)00235-5

- Aung, Y.Y.M., Wong, D.C.S., Ting, D.S.W., 2021. The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. British Medical Bulletin 139, 4–15. https://doi.org/10.1093/bmb/ldab016

- Balan, R., Dobrean, A., Poetar, C.R., 2024. Use of automated conversational agents in improving young population mental health: a scoping review. npj Digit. Med. 7, 75. https://doi.org/10.1038/s41746-024-01072-1

- Ball, M.J., Lillis, J., 2001. E-health: transforming the physician/patient relationship. International Journal of Medical Informatics 61, 1–10. https://doi.org/10.1016/S1386-5056(00)00130-1

- Belk, R., 2022. Chatbots, in: The Routledge Handbook of Digital Consumption. Routledge, London, pp. 124–135. https://doi.org/10.4324/9781003317524-12

- Bendig, E., Erb, B., Schulze-Thuesing, L., Baumeister, H., 2022. The Next Generation: Chatbots in Clinical Psychology and Psychotherapy to Foster Mental Health – A Scoping Review. Verhaltenstherapie 32, 64–76. https://doi.org/10.1159/000501812

- Bickmore, T.W., Mitchell, S.E., Jack, B.W., Paasche-Orlow, M.K., Pfeifer, L.M., O'Donnell, J., 2010. Response to a relational agent by hospital patients with depressive symptoms. Interacting with Computers 22, 289–298. https://doi.org/10.1016/j.intcom.2009.12.001

- Bilquise, G., Ibrahim, S., Shaalan, K., 2022. Emotionally Intelligent Chatbots: A Systematic Literature Review. Human Behavior and Emerging Technologies 2022, 1–23. https://doi.org/10.1155/2022/9601630

- Birnbaum, G.E., Mizrahi, M., Hoffman, G., Reis, H.T., Finkel, E.J., Sass, O., 2016. What robots can teach us about intimacy: The reassuring effects of robot responsiveness to human disclosure. Computers in Human Behavior 63, 416–423. https://doi.org/10.1016/j.chb.2016.05.064

- Boegershausen, J., Datta, H., Borah, A., Stephen, A.T., 2022. Fields of Gold: Scraping Web Data for Marketing Insights. Journal of Marketing 86, 1–20. https://doi.org/10.1177/00222429221100750

- Bond, R.R., Mulvenna, M.D., Potts, C., O'Neill, S., Ennis, E., Torous, J., 2023. Digital transformation of mental health services. npj Mental Health Res 2, 13. https://doi.org/10.1038/s44184-023-00033-y

- Boucher, E.M., Harake, N.R., Ward, H.E., Stoeckl, S.E., Vargas, J., Minkel, J., Parks, A.C., Zilca, R., 2021. Artificially intelligent chatbots in digital mental health interventions: a review. Expert Review of Medical Devices 18, 37–49. https://doi.org/10.1080/17434440.2021.2013200

- Bowman, R., Cooney, O., Newbold, J.W., Thieme, A., Clark, L., Doherty, G., Cowan, B., 2024. Exploring how politeness impacts the user experience of chatbots for mental health support. International Journal of Human-Computer Studies 184, 103181. https://doi.org/10.1016/j.ijhcs.2023.103181

- Brandtzæg, P.B., Følstad, A., 2018. Chatbots: Changing User Needs and Motivations. interactions 25, 38–43. https://doi.org/10.1145/3236669

- Brandtzæg, P.B., Følstad, A., 2017. Why People Use Chatbots, in: Kompatsiaris, I., Cave, J., Satsiou, A., Carle, G., Passani, A., Kontopoulos, E., Diplaris, S., McMillan, D. (Eds.), Internet Science. Springer International Publishing, Cham, pp. 377–392. https://doi.org/10.1007/978-3-319-70284-1_30

- Cafiero, G., Rosa, M.D., Schiavone, F., 2023. Improving Healthcare by Designing AI-Human-Centered Systems, in: La Torre, D., Appio, F.P., Masri, H., Lazzeri, F., Schiavone, F. (Eds.), Impact of Artificial Intelligence in Business and Society. Routledge, London, pp. 250–269. https://doi.org/10.4324/9781003304616-16

- Caliandro, A., 2016. Ethnography in Digital Spaces: Ethnography of Virtual Worlds, Netnography, & Digital Ethnography, in: Denny, R.M., Sunderland, P. (Eds.), Handbook of Anthropology in Business. Routledge, pp. 658–679. https://doi.org/10.4324/9781315427850

- Chang, A., 2023. The Role of Artificial Intelligence in Digital Health, in: Meyers, A. (Ed.), Digital Health Entrepreneurship, Health Informatics. Springer International Publishing, Cham, pp. 75–85. https://doi.org/10.1007/978-3-031-33902-8
- Chaudhry, B.M., Debi, H.R., 2024. User perceptions and experiences of an AI-driven conversational agent for mental health support. mHealth 10, 22–22. https://doi.org/10.21037/mhealth-23-55
- Chaves, A.P., Gerosa, M.A., 2021. How Should My Chatbot Interact? A Survey on Social Characteristics in Human–Chatbot Interaction Design. International Journal of Human–Computer Interaction 37, 729–758. https://doi.org/10.1080/10447318.2020.1841438
- Cheng, S., Chang, C., Chang, W., Wang, H., Liang, C., Kishimoto, T., Chang, J.P., Kuo, J.S., Su, K., 2023. The now and future of ChatGPT and GPT in psychiatry. Psychiatry Clin Neurosci 77, 592–596. https://doi.org/10.1111/pcn.13588
- Ciechanowski, L., Przegalinska, A., Magnuski, M., Gloor, P., 2019. In the shades of the uncanny valley: An experimental study of human–chatbot interaction. Future Generation Computer Systems 92, 539–548. https://doi.org/10.1016/j.future.2018.01.055
- Cirucci, A.M., Pruchniewska, U.M., 2022. UX research methods for media and communication studies: an introduction to contemporary qualitative methods. Routledge, Taylor & Franics Group, New York.
- Commodity HQ & MITRE, 2023. How comfortable are you with the use of artificial intelligence (AI) for each of the following [WWW Document]. Statista. URL https://www-statista-com.uaccess.univie.ac.at/statistics/1449222/comfort-ai-use-daily-tasks-united-states/ (accessed 06/04/24).
- D'Alfonso, S., 2020. AI in mental health. Current Opinion in Psychology 36, 112–117. https://doi.org/10.1016/j.copsyc.2020.04.005
- Davies, A., Mueller, J., 2020. Developing Medical Apps and mHealth Interventions: A Guide for Researchers, Physicians and Informaticians, Health Informatics. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-47499-7
- Denecke, K., May, R., 2022. Usability Assessment of Conversational Agents in Healthcare: A Literature Review, in: Séroussi, B., Weber, P., Dhombres, F., Grouin, C., Liebe, J.-D., Pelayo, S., Pinna, A., Rance, B., Sacchi, L., Ugon, A., Benis, A., Gallos, P. (Eds.), Studies in Health Technology and Informatics. IOS Press. https://doi.org/10.3233/SHTI220431
- Dergaa, I., Fekih-Romdhane, F., Hallit, S., Loch, A.A., Glenn, J.M., Fessi, M.S., Ben Aissa, M., Souissi, N., Guelmami, N., Swed, S., El Omri, A., Bragazzi, N.L., Ben Saad, H., 2024. ChatGPT is not ready yet for use in providing mental health assessment and interventions. Front. Psychiatry 14, 1277756. https://doi.org/10.3389/fpsyt.2023.1277756
- Dingler, T., Kwasnicka, D., Wei, J., Gong, E., Oldenburg, B., 2021. The Use and Promise of Conversational Agents in Digital Health. Yearb Med Inform 30, 191–199. https://doi.org/10.1055/s-0041-1726510

- Doyle, C., 2016. freemium, in: A Dictionary of Marketing. Oxford University Press.
- Duplaga, M., 2021a. The transition from telemedicine and e-health to Health 4.0, in: Ćwiklicki, M., Duplaga, M., Klich, J. (Eds.), The Digital Transformation of Healthcare. Routledge, London, pp. 17–30. https://doi.org/10.4324/9781003144403-2
- Duplaga, M., 2021b. Patient empowerment in Health 4.0, in: Ćwiklicki, M., Duplaga, M., Klich, J. (Eds.), The Digital Transformation of Healthcare. Routledge, London, pp. 63–75. https://doi.org/10.4324/9781003144403-5
- Eysenbach, G., 2005. The Law of Attrition. J Med Internet Res 7, e11. https://doi.org/10.2196/jmir.7.1.e11
- Falzon, D., Temesgen, Z., Story, A., 2023. Digital Technology for Global Health, in: Raviglione, M.C.B., Tediosi, F., Villa, S., Casamitjana, N., Plasència, A. (Eds.), Global Health Essentials, Sustainable Development Goals Series. Springer International Publishing, Cham, pp. 383–387. https://doi.org/10.1007/978-3-031-33851-9_58
- Fan, X., Chao, D., Zhang, Z., Wang, D., Li, X., Tian, F., 2021. Utilization of Self-Diagnosis Health Chatbots in Real-World Settings: Case Study. J Med Internet Res 23, e19928. https://doi.org/10.2196/19928
- Følstad, A., Brandtzæg, P.B., 2020. Users' experiences with chatbots: findings from a questionnaire study. Qual User Exp 5, 3. https://doi.org/10.1007/s41233-020-00033-2
- Følstad, A., Nordheim, C.B., Bjørkli, C.A., 2018. What Makes Users Trust a Chatbot for Customer Service? An Exploratory Interview Study, in: Bodrunova, S.S. (Ed.), Internet Science. Springer International Publishing, Cham, pp. 194–208. https://doi.org/10.1007/978-3-030-01437-7_16
- Følstad, A., Skjuve, M., Brandtzæg, P.B., 2019. Different Chatbots for Different Purposes: Towards a Typology of Chatbots to Understand Interaction Design, in: Bodrunova, S.S., Koltsova, O., Følstad, A., Halpin, H., Kolozaridi, P., Yuldashev, L., Smoliarova, A., Niedermayer, H. (Eds.), Internet Science, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 145–156. https://doi.org/10.1007/978-3-030-17705-8_13
- Gamble, A., 2020. Artificial intelligence and mobile apps for mental healthcare: a social informatics perspective. AJIM 72, 509–523. https://doi.org/10.1108/AJIM-11-2019-0316
- Gemert-Pijnen, L. van, Kip, H., Kelders, S.M., Sanderman, R., 2018. Introducing eHealth, in: Gemert-Pijnen, L. van, Kelders, S.M., Kip, H., Sanderman, R. (Eds.), eHealth Research, Theory and Development: A Multi-Disciplinary Approach. Routledge, London, New York, pp. 3–26.
- Google Play, 2024. Youper - CBT Therapy Chatbot - Apps on Google Play [WWW Document]. URL https://play.google.com/store/apps/details?id=br.com.youper&hl=en&gl=US (accessed 12/08/24).

- Graham, S., Depp, C., Lee, E.E., Nebeker, C., Tu, X., Kim, H.-C., Jeste, D.V., 2019. Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. Curr Psychiatry Rep 21, 116. https://doi.org/10.1007/s11920-019-1094-0

- Greene, J.A., 2022. The Doctor Who Wasn't There: Technology, History, and the Limits of Telehealth. The University of Chicago Press, Chicago, London.

- Greenhill, A.T., Edmunds, B.R., 2020. A primer of artificial intelligence in medicine. Techniques and Innovations in Gastrointestinal Endoscopy 22, 85–89. https://doi.org/10.1016/j.tgie.2019.150642

- Hanley, T., Wyatt, C., 2022. What Do People Look for from Digital Support?, in: Wilson, H. (Ed.), Digital Delivery of Mental Health Therapies: A Guide to the Benefits and Challenges, and Making It Work. Jessica Kingsley Publishers, London ; Philadelphia, pp. 76–88.

- Haque, M.D.R., Rubya, S., 2023a. An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews. JMIR Mhealth Uhealth 11, e44838. https://doi.org/10.2196/44838

- Haque, M.D.R., Rubya, S., 2023b. An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews. JMIR Mhealth Uhealth 11, e44838. https://doi.org/10.2196/44838

- Hassenzahl, M., 2003. The Thing and I: Understanding the Relationship Between User and Product, in: Blythe, M.A., Overbeeke, K., Monk, A.F., Wright, P.C. (Eds.), Funology, Human-Computer Interaction Series. Springer Netherlands, Dordrecht, pp. 31–42. https://doi.org/10.1007/1-4020-2967-5_4

- Holm, A.B., Günzel-Jensen, F., 2017. Succeeding with freemium: strategies for implementation. JBS 38, 16–24. https://doi.org/10.1108/JBS-09-2016-0096

- Horvath, A.O., Del Re, A.C., Flückiger, C., Symonds, D., 2011. Alliance in individual psychotherapy. Psychotherapy 48, 9–16. https://doi.org/10.1037/a0022186

- Ipsos, 2023. WORLD MENTAL HEALTH DAY 2023 [WWW Document]. URL https://www.ipsos.com/sites/default/files/ct/news/documents/2023-10/Ipsos%20World%20Mental%20Health%20Day%202023_Global%20Report.pdf (accessed 05/12/24).

- Iqbal, J., Cortés Jaimes, D.C., Makineni, P., Subramani, S., Hemaida, S., Thugu, T.R., Butt, A.N., Sikto, J.T., Kaur, P., Lak, M.A., Augustine, M., Shahzad, R., Arain, M., 2023. Reimagining Healthcare: Unleashing the Power of Artificial Intelligence in Medicine. Cureus. https://doi.org/10.7759/cureus.44658

- ISO, 2019. ISO 9241-210:2019(en), Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems [WWW Document]. URL https://www.iso.org/obp/ui/en/#iso:std:iso:9241:-210:ed-2:v1:en (accessed 08/11/24).

- Istepanian, R.S.H., 2022. Mobile Health (m-Health) in Retrospect: The Known Unknowns. IJERPH 19, 3747. https://doi.org/10.3390/ijerph19073747

- Istepanian, R.S.H., Al-Anzi, T., 2018. m-Health 2.0: New perspectives on mobile health, machine learning and big data analytics. Methods 151, 34–40. https://doi.org/10.1016/j.ymeth.2018.05.015
- Johnston Roberts, K.J., 1999. Patient empowerment in the United States: a critical commentary. Health Expectations 2, 82–92. https://doi.org/10.1046/j.1369-6513.1999.00048.x
- Jones, L., Vigo, D., 2023. Mental Health and Substance Abuse, in: Raviglione, M.C.B., Tediosi, F., Villa, S., Casamitjana, N., Plasència, A. (Eds.), Global Health Essentials, Sustainable Development Goals Series. Springer International Publishing, Cham, pp. 197–201. https://doi.org/10.1007/978-3-031-33851-9
- Kaul, V., Enslin, S., Gross, S.A., 2020. History of artificial intelligence in medicine. Gastrointestinal Endoscopy 92, 807–812. https://doi.org/10.1016/j.gie.2020.06.040
- Kelders, S.M., Howard, M., 2018. Opportunities of technology to promote health and well-being, in: Gemert-Pijnen, L. van, Kelders, S.M., Kip, H., Sanderman, R. (Eds.), eHealth Research, Theory and Development: A Multi-Disciplinary Approach. Routledge, London, New York, pp. 48–69.
- Kim, J., Kim, Y., Kim, B., Yun, S., Kim, M., Lee, J., 2018. Can a Machine Tend to Teenagers' Emotional Needs?: A Study with Conversational Agents, in: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. Presented at the CHI '18: CHI Conference on Human Factors in Computing Systems, ACM, Montreal QC Canada, pp. 1–6. https://doi.org/10.1145/3170427.3188548
- King, D.R., Nanda, G., Stoddard, J., Dempsey, A., Hergert, S., Shore, J.H., Torous, J., 2023. An Introduction to Generative Artificial Intelligence in Mental Health Care: Considerations and Guidance. Curr Psychiatry Rep 25, 839–846. https://doi.org/10.1007/s11920-023-01477-x
- Kocaballi, A.B., Berkovsky, S., Quiroz, J.C., Laranjo, L., Tong, H.L., Rezazadegan, D., Briatore, A., Coiera, E., 2019. The Personalization of Conversational Agents in Health Care: Systematic Review. J Med Internet Res 21, e15360. https://doi.org/10.2196/15360
- Kozinets, R.V., 2020. Netnography. The Essential Guide to Qualitative Social Media Research, 3e ed. SAGE, Los Angeles.
- Kreutzer, R.T., Sirrenberg, M., 2020. Understanding Artificial Intelligence: Fundamentals, Use Cases and Methods for a Corporate AI Journey, Management for Professionals. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-25271-7
- Kruzan, K.P., Fitzsimmons-Craft, E.E., Dobias, M., Schleider, J.L., Pratap, A., 2022. Developing, Deploying, and Evaluating Digital Mental Health Interventions in Spaces of Online Help- and Information-Seeking. Procedia Computer Science 206, 6–22. https://doi.org/10.1016/j.procs.2022.09.081
- Law, J.L., 2016. scalability [WWW Document]. Oxford Dictionary of Business and Management. URL https://www.oxfordreference.com/display/10.1093/acref/9780199684984.001.0001/acref-9780199684984-e-7585 (accessed 11/12/23).

- Lewis Kausel, C., Hackett, P.M.W., 2015. Netnography. Possibilities and Resourcefulness, in: Hackett, P. (Ed.), Qualitative Research Methods in Consumer Psychology. Ethnography and Culture. Psychology Press, New York, pp. 252–261. https://doi.org/10.4324/9781315776378

- Li, H., Zhang, R., Lee, Y.-C., Kraut, R.E., Mohr, D.C., 2023. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. npj Digit. Med. 6, 236. https://doi.org/10.1038/s41746-023-00979-5

- Limbic, 2024. Limbic Care [WWW Document]. URL https://www.limbic.ai/care (accessed 12/08/24).

- Limpanopparat, S., Gibson, E., Harris, D.A., 2024. User engagement, attitudes, and the effectiveness of chatbots as a mental health intervention: A systematic review. Computers in Human Behavior: Artificial Humans 2, 100081. https://doi.org/10.1016/j.chbah.2024.100081

- Linardon, J., Fuller-Tyszkiewicz, M., 2020. Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. Journal of Consulting and Clinical Psychology 88, 1–13. https://doi.org/10.1037/ccp0000459

- Liu, H., Peng, H., Song, X., Xu, C., Zhang, M., 2022. Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness. Internet Interventions 27, 100495. https://doi.org/10.1016/j.invent.2022.100495

- Lorenzo-Luaces, L., Wasil, A., Kacmarek, C.N., DeRubeis, R., 2024. Race and Socioeconomic Status as Predictors of Willingness to Use Digital Mental Health Interventions or One-On-One Psychotherapy: National Survey Study. JMIR Form Res 8, e49780. https://doi.org/10.2196/49780

- Lovejoy, C.A., 2019. Technology and mental health: The role of artificial intelligence. Eur. psychiatr. 55, 1–3. https://doi.org/10.1016/j.eurpsy.2018.08.004

- Lu, T., Liu, X., Sun, J., Bao, Y., Schuller, B.W., Han, Y., Lu, L., 2023. Bridging the gap between artificial intelligence and mental health. Science Bulletin 68, 1606–1610. https://doi.org/10.1016/j.scib.2023.07.015

- Lucas, G.M., Gratch, J., King, A., Morency, L.-P., 2014. It's only a computer: Virtual humans increase willingness to disclose. Computers in Human Behavior 37, 94–100. https://doi.org/10.1016/j.chb.2014.04.043

- Luger, E., Sellen, A., 2016. "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. Presented at the CHI'16: CHI Conference on Human Factors in Computing Systems, ACM, San Jose California USA, pp. 5286–5297. https://doi.org/10.1145/2858036.2858288

- Luxton, D.D., 2020. Improving Mental Health Outcomes With Artificial Intelligence, in: Reger, G.M. (Ed.), Technology and Mental Health: A Clinician's Guide to Improving Outcomes. Routledge, New York, pp. 248–270. https://doi.org/10.4324/9780429020537

- Luxton, D.D. (Ed.), 2016. Artificial Intelligence in Behavioral and Mental Health Care. Elsevier. https://doi.org/10.1016/C2013-0-12824-3
- Malik, T., Ambrose, A.J., Sinha, C., 2022. Evaluating User Feedback for an Artificial Intelligence–Enabled, Cognitive Behavioral Therapy–Based Mental Health App (Wysa): Qualitative Thematic Analysis. JMIR Hum Factors 9, e35668. https://doi.org/10.2196/35668
- Malvey, D., Slovensky, D.J., 2014. mHealth: Transforming Healthcare. Springer US, Boston, MA. https://doi.org/10.1007/978-1-4899-7457-0
- Meskó, B., Drobni, Z., Bényei, É., Gergely, B., Győrffy, Z., 2017. Digital health is a cultural transformation of traditional healthcare. mHealth 3, 38–38. https://doi.org/10.21037/mhealth.2017.08.07
- Meyerowitz-Katz, G., Ravi, S., Arnolda, L., Feng, X., Maberly, G., Astell-Burt, T., 2020. Rates of Attrition and Dropout in App-Based Interventions for Chronic Disease: Systematic Review and Meta-Analysis. J Med Internet Res 22, e20283. https://doi.org/10.2196/20283
- Mindspa - #1 Self-Therapy App [WWW Document], 2024. . Mindspa. URL https://mindspa.me/en/ (accessed 09/21/24).
- Minerva, F., Giubilini, A., 2023. Is AI the Future of Mental Healthcare? Topoi (Dordr) 42, 809–817. https://doi.org/10.1007/s11245-023-09932-3
- Mitra, S., 2023. Advances in Mobile Health Technology: A Research Perspective, 1st ed, Chapman & Hall/CRC Studies in Informatics Series. CRC Press, Boca Raton.
- Morning Consult, 2023. Share of adults in the United States who are interested in artificial intelligence (AI) related products as of February 2023. [WWW Document]. Statista. URL https://www-statista-com.uaccess.univie.ac.at/statistics/1377783/us-adults-interest-in-artificial-intelligence-related-products/ (accessed 06/04/24).
- Nadarzynski, T., Miles, O., Cowie, A., Ridge, D., 2019. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. DIGITAL HEALTH 5, 205520761987180. https://doi.org/10.1177/2055207619871808
- Narasimhadevara, A., Rambhatla, R., 2020. Chapter 9: mHealth for Better Quality of Life, Healthier Lifestyles, and More Meaningful Lives, in: Murthy, P., Ansehl, A. (Eds.), Technology and Global Public Health. Springer International Publishing, Cham, pp. 143–159. https://doi.org/10.1007/978-3-030-46355-7_14
- Nass, C., Steuer, J., Tauber, E.R., 1994. Computers are Social Actors. Presented at the the SIGCHI conference, ACM Press, Boston, Massachusetts, United States, pp. 72–78. https://doi.org/10.1145/191666.191703
- Netflix Help Center, 2024. Plans and Pricing [WWW Document]. URL https://help.netflix.com/en/node/24926 (accessed 08/16/24).
- Newman, C.F., 2017. Cognitive Behavioral Therapy. The SAGE Encyclopedia of Abnormal and Clinical Psychology.

- Nwosu, A., Boardman, S., Husain, M.M., Doraiswamy, P.M., 2022. Digital therapeutics for mental health: Is attrition the Achilles heel? Front. Psychiatry 13, 900615. https://doi.org/10.3389/fpsyt.2022.900615
- Oshida, Y., 2022. Artificial Intelligence for Medicine: People, Society, Pharmaceuticals, and Medical Materials. De Gruyter, Berlin.
- Outbreaks Near Me & SurveyMonkey, 2023. Comfort in an AI therapy session in the U.S. 2023 [WWW Document]. Statista. URL https://www-statista-com.uaccess.univie.ac.at/statistics/1417569/comfort-in-an-ai-therapy-session-in-the-us/ (accessed 06/04/24).
- Palmatier, R.W., Steinhoff, L., 2019. Relationship Marketing in the Digital Age. London : Routledge, Taylor and Francis Group, New York.
- Prakash, A.V., Das, S., 2020. Intelligent Conversational Agents in Mental Healthcare Services: A Thematic Analysis of User Perceptions. PAJAIS 12, 1–34. https://doi.org/doi: 10.17705/1pais.12201
- Punugoti, R., Duggar, R., Dhargalkar, R.R., Bhati, N., 2023. Intelligent Healthcare: Using NLP and ML to Power Chatbots for Improved Assistance. Presented at the 2023 International Conference on IoT, Communication and Automation Technology (ICICAT), IEEE, Gorakhpur, India, pp. 1–6. https://doi.org/10.1109/ICICAT57735.2023.10263708
- Raaijmakers, S., Cremers, A., Krahmer, E., Westera, M., 2023. Editorial: Conversational AI. Front. Artif. Intell. 6, 1203910. https://doi.org/10.3389/frai.2023.1203910
- Rapp, A., Boldi, A., Curti, L., Perrucci, A., Simeoni, R., 2023. Collaborating with a Text-Based Chatbot: An Exploration of Real-World Collaboration Strategies Enacted during Human-Chatbot Interactions. Presented at the CHI '23: CHI Conference on Human Factors in Computing Systems, ACM, Hamburg Germany, pp. 1–17. https://doi.org/10.1145/3544548.3580995
- Ratheesh, A., Alvarez-Jimenez, M., 2022. The future of digital mental health in the post-pandemic world: Evidence-based, blended, responsive and implementable. Aust N Z J Psychiatry 56, 107–109. https://doi.org/10.1177/00048674211070984
- Reddy, S., 2024. Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. Implementation Sci 19, 27. https://doi.org/10.1186/s13012-024-01357-9
- Renfrew, M.E., Morton, D.P., Morton, J.K., Hinze, J.S., Przybylko, G., Craig, B.A., 2020. The Influence of Three Modes of Human Support on Attrition and Adherence to a Web- and Mobile App–Based Mental Health Promotion Intervention in a Nonclinical Cohort: Randomized Comparative Study. J Med Internet Res 22, e19945. https://doi.org/10.2196/19945
- Safi, Z., Abd-Alrazaq, A., Khalifa, M., Househ, M., 2020. Technical Aspects of Developing Chatbots for Medical Applications: Scoping Review. J Med Internet Res 22, e19127. https://doi.org/10.2196/19127

- Sainz-de-Abajo, B., de la Torre-Díez, I., Góngora-Alonso, S., López-Coronado, M., 2021. Privacy issues in eHealth and mHealth apps, in: Tzanou, M. (Ed.), Health Data Privacy under the GDPR: Big Data Challenges and Regulatory Responses, Routledge Research in the Law of Emerging Technologies. Routledge, New York, NY ; Milton Park, Abingdon, Oxon, pp. 71–82.
- Saldaña, J., 2021. The Coding Manual for Qualitative Researchers, 4E. Fourth edition. ed. SAGE, London.
- Saxena, A., Saxena, M., Deo, A., Rodríguez Llerena, A., 2021. Transformation of mHealth in Society, in: Saxena, A., Brault, N., Rashid, S. (Eds.), Big Data and Artificial Intelligence for Healthcare Applications. CRC Press, Boca Raton, pp. 165–193.
- Schueller, S.M., 2018. Mental health and eHealth technology, in: Gemert-Pijnen, L. van, Kelders, S.M., Kip, H., Sanderman, R. (Eds.), eHealth Research, Theory and Development: A Multi-Disciplinary Approach. Routledge, London, New York, pp. 91–110.
- Schuetzler, R.M., Grimes, G.M., Scott Giboney, J., 2020. The impact of chatbot conversational skill on engagement and perceived humanness. Journal of Management Information Systems 37, 875–900. https://doi.org/10.1080/07421222.2020.1790204
- Sciglimpaglia, D., Raafat, F., 2022. Freemium marketing: use of demand-side research in market segmentation strategy. Journal of Strategic Marketing 30, 667–690. https://doi.org/10.1080/0965254X.2020.1824013
- Shan, Y., Ji, M., Xie, W., Qian, X., Li, R., Zhang, X., Hao, T., 2022. Language Use in Conversational Agent–Based Health Communication: Systematic Review. J Med Internet Res 24, e37403. https://doi.org/10.2196/37403
- Sharma, S., Rawal, R., Shah, D., 2023. Addressing the challenges of AI-based telemedicine: Best practices and lessons learned. J Edu Health Promot 12, 338. https://doi.org/10.4103/jehp.jehp_402_23
- Shatte, A.B.R., Hutchinson, D.M., Teague, S.J., 2019. Machine learning in mental health: a scoping review of methods and applications. Psychol. Med. 49, 1426–1448. https://doi.org/10.1017/S0033291719000151
- Simon, G.J., Aliferis, C. (Eds.), 2024. Artificial Intelligence and Machine Learning in Health Care and Medical Sciences: Best Practices and Pitfalls, Health Informatics. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-031-39355-6
- Smith, B.M., Levin, M.E., Twohig, M.P., 2017. Acceptance and Commitment Therapy. The SAGE Encyclopedia of Abnormal and Clinical Psychology.
- Spruijt-Metz, D., Nilsen, W., Pavel, M., 2015. mHealth for Behavior Change and Monitoring, in: Adibi, S. (Ed.), mHealth Multidisciplinary Verticals. CRC Press, Boca Raton, pp. 119–132.
- The Lancet Digital Health, 2021. Digital tools for mental health in a crisis. The Lancet Digital Health 3, e204. https://doi.org/10.1016/S2589-7500(21)00045-5

- Tong, F., Lederman, R., D'Alfonso, S., Berry, K., Bucci, S., 2023. Conceptualizing the digital therapeutic alliance in the context of fully automated mental health apps: A thematic analysis. Clin Psychology and Psychoth 30, 998–1012. https://doi.org/10.1002/cpp.2851

- Torous, J., Bucci, S., Bell, I.H., Kessing, L.V., Faurholt-Jepsen, M., Whelan, P., Carvalho, A.F., Keshavan, M., Linardon, J., Firth, J., 2021. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 20, 318–335. https://doi.org/10.1002/wps.20883

- Torous, J., Nicholas, J., Larsen, M.E., Firth, J., Christensen, H., 2018. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Mental Health 21, 116–119. https://doi.org/10.1136/eb-2018-102891

- Tremain, H., McEnery, C., Fletcher, K., Murray, G., 2020. The Therapeutic Alliance in Digital Mental Health Interventions for Serious Mental Illnesses: Narrative Review. JMIR Ment Health 7, e17204. https://doi.org/10.2196/17204

- USC Annenberg, 2023. Robert Kozinets [WWW Document]. URL http://annenberg.usc.edu/faculty/robert-kozinets (accessed 01/18/24).

- VERBI Software, 2024. MAXQDA 24.

- VOS, 2024. VOS.health - Mental Health App - App [WWW Document]. URL https://vos.health/en/application (accessed 12/08/24).

- Woebot Health, 2024a. FAQ [WWW Document]. Woebot Health. URL https://woebothealth.com/faq/ (accessed 09/15/24).

- Woebot Health, 2024b. App Support [WWW Document]. Woebot Health. URL https://woebothealth.com/app-support/ (accessed 09/15/24).

- World Health Organization, 2023a. WHO Calls For Safe and Ethical AI for Health [WWW Document]. URL https://www.who.int/news/item/16-05-2023-who-calls-for-safe-and-ethical-ai-for-health (accessed 10/22/23).

- World Health Organization, 2023b. WHO outlines considerations for regulation of artificial intelligence for health [WWW Document]. URL https://www.who.int/news/item/19-10-2023-who-outlines-considerations-for-regulation-of-artificial-intelligence-for-health (accessed 10/22/23).

- Xu, H., Roberts, K., 2022. Natural Language Processing, in: Cohen, T.A., Patel, V.L., Shortliffe, E.H. (Eds.), Intelligent Systems in Medicine and Health, Cognitive Informatics in Biomedicine and Healthcare. Springer International Publishing, Cham, pp. 213–234. https://doi.org/10.1007/978-3-031-09108-7_7

- Xue, J., Zhang, B., Zhao, Y., Zhang, Q., Zheng, C., Jiang, J., Li, H., Liu, N., Li, Z., Fu, W., Peng, Y., Logan, J., Zhang, J., Xiang, X., 2023. Evaluation of the Current State of Chatbots for Digital Health: Scoping Review. J Med Internet Res 25, e47217. https://doi.org/10.2196/47217

- Yokotani, K., Takagi, G., Wakashima, K., 2018. Advantages of virtual agents over clinical psychologists during comprehensive mental health interviews using a mixed methods design. Computers in Human Behavior 85, 135–145. https://doi.org/10.1016/j.chb.2018.03.045

- Zafar, F., Fakhare Alam, L., Vivas, R.R., Wang, J., Whei, S.J., Mehmood, S., Sadeghzadegan, A., Lakkimsetti, M., Nazir, Z., 2024. The Role of Artificial Intelligence in Identifying Depression and Anxiety: A Comprehensive Literature Review. Cureus. https://doi.org/10.7759/cureus.56472

- Zamora, J., 2017. I'm Sorry, Dave, I'm Afraid I Can't Do That: Chatbot Perception and Expectations, in: Proceedings of the 5th International Conference on Human Agent Interaction. Presented at the HAI '17: The Fifth International Conference on Human-Agent Interaction, ACM, Bielefeld Germany, pp. 253–260. https://doi.org/10.1145/3125739.3125766

- Zhong, W., Luo, J., Zhang, H., 2024. The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. Journal of Affective Disorders 356, 459–469. https://doi.org/10.1016/j.jad.2024.04.057