Metaverse-Driven Mental Health Solutions: A Scoping Review of AI and Extended Reality Applications

Aliya Tabassum[1], Ibrahim Ghaznavi[2], Alaa Abd-Alrazaq[3], Junaid Qadir[1*]
[1*]Department of Computer Science and Engineering, College of Engineering, Qatar University, Doha, Qatar.
[2]Department of Computer Science, Information Technology University (ITU), Lahore, Pakistan.
[3]AI Center for Precision Health, Weill Cornell Medicine - Qatar, Doha, Qatar.
*Corresponding author(s). E-mail(s): jqadir@qu.edu.qa;

Abstract

Background: The global shortage of trained mental health professionals has created a significant barrier to accessible mental healthcare, with patient-to-clinician ratios falling short even in well-resourced countries like the United States. Artificial Intelligence (AI) and Extended Reality (XR), when integrated within immersive metaverse environments, offer innovative approaches to augment mental healthcare delivery and extend its reach beyond traditional clinical settings.

Objective: This scoping review aims to explore the state-of-the-art applications of AI and XR in metaverse frameworks for mental healthcare, assessing their capabilities and limitations while identifying key ethical challenges and societal risks. Specifically, it examines governance gaps related to data privacy, patient-clinician dynamics, algorithmic biases, digital inequality, and psychological dependency.

Methods: A systematic search was conducted across five electronic databases for peer-reviewed literature published from 2014 to October 2024. The search incorporated terms related to XR, mental healthcare, psychotherapy, and the metaverse. Studies were screened against predefined eligibility criteria by two independent reviewers, with relevant data extracted and synthesized through a narrative review approach.

Results: Of the 1,288 articles identified, 48 studies met the inclusion criteria. The reviewed literature highlighted diverse AI applications, including emotion recognition, therapy chatbots, and decision-support systems, alongside XR-enabled therapeutic interventions. Key ethical concerns included inadequate transparency in AI algorithms, data privacy vulnerabilities, and risks of AI-induced biases in therapeutic decisions. XR-driven interventions showed promise in enhancing therapy accessibility and engagement but raised concerns about psychological dependency and the exclusion of underprivileged populations due to digital inequality. While these technologies demonstrated potential efficacy in controlled settings, many studies relied on single-institution datasets and lacked longitudinal validation.

Conclusions:AI-XR technologies hold transformative potential for addressing global mental healthcare challenges, but their adoption must be guided by ethical considerations. Future research must prioritize the development of inclusive, transparent, and equity-driven frameworks for responsible integration. Multicentre collaborations, public datasets, and rigorous regulatory standards will be essential to ensure sustainable innovation that balances scalability with patient safety and societal well-being.

Keywords: Metaverse; Augmented Reality (AR); Mental Health; Extended Reality (XR); Virtual Reality (VR); Psychotherapy; XR Ethics.

Introduction

Overview of AI and XR in Mental Health

The accelerating convergence of Artificial Intelligence (AI) and Extended Reality (XR) technologies heralds a paradigm shift in mental health care, transforming it from static, clinician-driven interventions to dynamic, immersive therapeutic experiences. Extended Reality is an umbrella term that encompasses immersive technologies such as Augmented Reality (AR) and Mixed Reality (MR), which are transforming mental health care [1]. Augmented Reality overlays digital content, like images and sounds, onto the physical world through devices such as smartphones or AR glasses, enhancing the user's real-world experience [2]. Mixed Reality goes further by blending the physical and digital worlds, allowing real and virtual elements to interact in real-time, creating immersive environments that engage users more deeply [3]. Together, AR, MR, and XR technologies offer significant potential to advance mental health treatments by providing adaptable, personalized, and accessible interventions.

These technologies, as part of the broader XR spectrum, enable the creation of interactive, sensory rich therapeutic experiences where virtual environments can be crafted to support emotional and cognitive engagement. Virtual Reality (VR) is another key component of XR, but it differs from AR and MR in that it creates fully immersive, computer-generated environments that completely replace the physical world. Using VR headsets, users can enter a 3D virtual space to interact with digital objects and surroundings, typically through motion-tracking sensors and controllers [4] [5]. VR technologies enable patients to step into meticulously designed psychotherapeutic spaces, where teleportation, avatar embodiment, and sensory rich cues, including 360-degree visuals, haptic feedback, ambient soundscapes, and even olfactory stimuli, can foster profound emotional and cognitive engagement[1]. The emergence of the metaverse—a sprawling, interconnected digital ecosystem—further augments this potential, offering a scalable platform for delivering therapy to those isolated by geography, socioeconomic barriers, or stigma [6].

The promise of AI-XR in mental health lies not merely in its technical capabilities but in its ability to democratize access to care. By transcending the constraints of traditional in-person therapies, these technologies open pathways for addressing mental health disparities, particularly in underserved regions [7–9]. They provide

1 https://www.civicspace.tech/technologies/augmented-reality-virtual-reality/

patients with tailored interventions that adapt to individual needs, creating an unprecedented synergy between technological precision and therapeutic intent [9, 10–12]. Studies highlight the value of AI-XR technologies in enhancing mental health treatments by creating adaptive, immersive environments that foster emotional regulation and support cognitive therapies [2]. Frameworks like schema therapy have demonstrated how virtual settings promote deeper emotional and cognitive engagement, helping users confront and adapt to challenging scenarios [13]. These technologies have also been shown to hold promise in addressing psychiatric care through metaverse platforms, which facilitate personalized psychiatric tools and virtual medical training, albeit with challenges related to accessibility and socioeconomic disparities [2][38].

Emerging Opportunities and Challenges

Moreover, MR environments have been explored for their capacity to create adaptable therapeutic contexts that encourage social interactions and emotional regulation, potentially addressing gaps in conventional therapies [3]. In a research, photorealistic virtual avatars are employed by Ghaznavi et al. [14] to enhance the efficacy of Self-attachment psychotherapy to treat patients with symptoms of chronic anxiety and depression [15]. Another study by Wei et al. [16] investigates the effects of virtual avatar's facial expressions and animations on paranoid patients. An automated immersive virtual reality platform is also employed by Freeman et al. [5] to treat patients with symptoms of persecutory delusions. Further advancements in telerehabilitation using AI-XR technologies are also underway, with studies suggesting that they may soon rival or surpass traditional rehabilitation methods. However, the need for long-term validation remains pressing [17].

The integration of AI-XR technologies in the metaverse has also opened up new possibilities for treating mental health conditions, including autism and PTSD [18, 19]. These technologies offer promising new approaches but remain under exploration regarding their long-term effects, such as the risk of dependency, cognitive impacts, and sustainability within immersive therapeutic environments [20]. By incorporating multiple sensory stimuli—such as sight, touch, sound, and smell—into therapeutic VR and AR settings, these applications engage various brain regions, which in turn can improve emotional regulation, cognitive processing, and even neural reprogramming. Neuroplasticity is leveraged to help patients rewire their neural circuits, fostering healthier reactions to trauma and stress [65].

These technological advances not only enhance user experiences but also signal the advent of Web 3, where the human-machine interface is envisioned as a seamless integration of physical, digital, and spatial dimensions. The metaverse thus holds a dual significance for mental health. On the one hand, it presents unparalleled opportunities to expand therapeutic access and foster collaborative care across cultural and geographic boundaries. On the other hand, it poses risks, including potential harm to vulnerable populations and the erosion of critical human elements in mental health care [20]. Despite the growing body of research on AIXR's applications, there remains a conspicuous gap in actionable regulatory frameworks that address these challenges. Equally underexplored are the long-term societal impacts of integrating immersive technologies into mental health care, including their influence on patient-therapist dynamics and the broader ethics of care delivery. Addressing these gaps requires a multidisciplinary lens that unites technological innovation with philosophical rigor.

Research Questions

This paper presents a scoping review combined with a state-of-the-art analysis of AI-XR technologies in mental health. It examines their benefits, challenges—across ethical, legal, technological, and societal dimensions—and potential future directions. By identifying research gaps, it provides actionable recommendations for the ethical integration of AI-XR and metaverse technologies into psychotherapeutic practices, aiming to support researchers, policymakers, and practitioners in maximizing their potential while minimizing risks.

The review's primary objectives are threefold: first, to explore current applications of AI-XR in mental health and assess their benefits; second, to identify the key ethical, societal, and technological barriers to their widespread adoption; and third, to evaluate the regulatory frameworks required for their equitable integration, addressing concerns such as data privacy, algorithmic bias, and equitable access to these transformative technologies.

This scoping review addresses the following research questions:
RQ1 *What are the key applications and associated benefits of AI-XR technologies in mental health and psychotherapy?* This question seeks to explore how AI-XR (Artificial Intelligence and Extended Reality) technologies are being utilized across various therapeutic domains, including exposure therapy for phobias, cognitive behavioral therapy (CBT) for depression, anxiety, post-traumatic stress disorder (PTSD), and obsessive-compulsive disorder (OCD), as well as pain management, schizophrenia treatment, social skills training (e.g., autism and social anxiety), mindfulness techniques, trauma recovery, neurofeedback, and cognitive rehabilitation. It also examines how these applications leverage the unique capabilities of AI-XR to improve therapeutic outcomes and address diverse mental health challenges.

RQ2 *What are the ethical and societal challenges and barriers to adopting AI-XR technologies in mental health therapy?* These challenges include the digital divide, cybersickness, data privacy, and security concerns, informed consent issues, potential effects on the therapeutic relationship, and long-term psychological effects on mental health outcomes.

RQ3 *What regulatory frameworks are necessary to ensure ethical and equitable integration of AI-XR technologies in mental health treatment?* This includes addressing challenges such as equitable access, data privacy, and algorithmic bias, while safeguarding the humanistic elements essential to psychotherapy. Emphasis is

placed on integrating comprehensive governance models, clinical oversight, and ethical principles to maintain care integrity and uphold human dignity.
The structure of the paper is as follows: Section 2 details the scoping review methodology, including study selection criteria, search strategies, and data extraction and synthesis methods. Section 3 categorizes the findings based on the benefits, ethical dilemmas, and societal impacts of AI-XR technologies. In Section 4, the results are analyzed in relation to the research questions, emphasizing research gaps in areas such as long-term mental health outcomes, equity, and governance frameworks. Finally, limitations of the study are outlined in Section 5 and Section 6 concludes the article by summarizing key findings and suggesting future research directions for the ethical integration of AI-XR technologies in mental health.

Methods

To conduct this scoping review, we followed the guidelines outlined in the Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) [21]. The PRISMA-ScR checklist relevant to this review is provided in Multimedia Appendix 1 [21]. The subsequent subsections offer a detailed explanation of the methods employed in this review.

Search Strategy

The search was conducted across five academic databases to identify relevant literature on integrating AIXR and metaverse technologies in mental health treatments. These databases included PubMed, PsycINFO, IEEE Xplore, Scopus, and Google Scholar. The search process began on 16 September 2024, using a variety of keywords, and was iteratively refined based on the retrieval of studies. This search was concluded on 25 November 2024. Our search criteria comprised three main categories of terms — metaverse-related terms (VR, AR, XR, and immersive technologies), psychotherapy-related terms (anxiety, depression, stress, cognitive, Post-Traumatic Stress, schizophrenia, eating disorder, phobia, autism) and outcome-related terms (benefit, advantage, effective, impact, improve, efficacy, challenge, barrier, privacy, safety, regulate). The detailed search queries used for each database are provided in Multimedia Appendix 2.

Study Eligibility Criteria

This review focused on the studies that investigated the application of metaverse and XR technologies in mental health treatment, with particular emphasis on therapeutic applications such as psychotherapy, cognitive behavioral therapy, or treatments for psychological disorders. Articles using AI (including machine learning and deep learning) and XR technologies (virtual reality, augmented reality, mixed reality, or metaverse-based systems) in mental health treatment. However, we have excluded articles that solely focus on AI technologies, ensuring that our review includes studies that explore the intersection of AI with VR, XR, AR, or the metaverse. The inclusion and exclusion criteria are shown in Table 1.

In terms of predicted outcomes, our focus was centered on three key categories. Firstly, articles that discussed the benefits and applications of the metaverse, VR, XR, and AR in mental health treatments, highlighting their potential to enhance therapeutic outcomes. Secondly, the studies addressing the risks, social problems, and broader implications arising from the intersection of these technologies with mental health practices. Finally, the research that emphasized the importance of developing regulatory frameworks, strategies, and future guidelines to ensure these immersive technologies are utilized safely and effectively, minimizing negative implications while maximizing their benefits. However, studies focused solely on technical aspects of AI or XR without direct application to mental health treatment were excluded. Likewise, we eliminated the articles that addressed non-mental-health-related applications of AI-XR technologies, such as general healthcare, education, or gaming.

This study adopts a scoping review methodology to systematically explore the intersection of immersive technologies—such as the metaverse, virtual reality, extended reality, and augmented reality—with mental health applications. To ensure comprehensiveness, we included peer-reviewed articles, theses, dissertations, experimental studies, systematic reviews, and conference papers published in English, with constraints on publication year from 2014-2024. Non–peer-reviewed articles, preprints, reviews, opinion pieces, editorials, case studies, conference abstracts, and protocols were excluded. This approach allows for a structured synthesis of current evidence and insights, emphasizing themes and gaps critical to advancing this interdisciplinary field.

Study Process

The study selection process was carried out in three phases. First, duplicates were removed from the retrieved studies using a simple Excel duplicate removal tool. The titles and abstracts of the remaining articles were then screened for relevance. In the final phase, the full texts of the shortlisted studies were thoroughly evaluated. The selection process was conducted independently by two reviewers, disagreements resolved through discussions. Studies were excluded if they did not specifically focus on AI and XR in mental health or psychotherapy. Priority was given to studies that provided a clear discussion of benefits, outcomes, ethical considerations related to patient data, privacy, and therapeutic use of immersive technologies. Articles that did not adequately address benefits, outcomes, ethical or privacy concerns were excluded.

Data Extraction

The data extraction process involved two independent reviewers using a detailed form in Microsoft Excel to gather key information from the studies. Extracted data included study characteristics such as author, publication year, type, and methodology, alongside details of settings (e.g., virtual environments, clinics, or online platforms) and findings related to the applications and impacts of AI, XR, VR, AR, and Metaverse technologies in mental health. Additional data encompassed the technologies and tools used (e.g., AI-powered chatbots, virtual therapy platforms), AI methods like machine learning and NLP, and the studies' focus on psychotherapy and broader mental health treatments, including conditions such as anxiety, depression, and PTSD. Discrepancies between the reviewers were resolved through discussion.

Table 1: Inclusion and exclusion criteria for selecting eligible articles.

Inclusion Criteria:
Research Articles: Peer-reviewed articles only.
Publication Date: Published post-2014 for relevance.
Mental Health Context: Focus on mental health treatments and psychotherapy.
Ethical Issues: Includes ethical considerations of AI/XR in mental health.
Exclusion Criteria:
Non-Research Articles: Excludes non-peer-reviewed sources like editorials, abstracts, policy statements.
Irrelevant Focus: Not centered on mental health and the metaverse.

Data Synthesis

The data synthesis process utilized a narrative approach to analyze and present the findings from the included studies. Key aspects, such as study metadata (e.g., publication year and country), were outlined, and the distribution of studies across research questions was calculated and visually represented using figures and tables. The synthesis delved into the applications of AI and XR technologies in mental health, highlighting their roles in enhancing psychotherapy approaches, including AI-driven cognitive behavioral therapy and the use of AI-powered psychotherapy chatbots. Ethical and social implications, such as privacy concerns, bias, fairness, and cybersickness, were also explored. Furthermore, regulatory frameworks were reviewed, emphasizing transparency, data protection, and security in the use of AI-XR technologies for mental health. The synthesis provided insights into both the benefits and challenges of these technologies, incorporating data on clinical outcomes, patient engagement, therapist feedback, and evaluation metrics, while proposing strategies and guidelines for their effective and ethical integration into mental healthcare practices.

Results

Search Results

This section presents the findings from the review. Figure 1 presents the search process and results from the preselected databases, yielding a total of 1,288 records. Articles published before 2014 were excluded, resulting in 898 records. Of these, we refined articles for only English language which resulted in elimination of 135 articles. Then from these, 355 duplicates were identified and removed using Excel. Screening the titles and abstracts of the remaining 408 articles led to the exclusion of 295. Out of the 113 articles remaining, 9 full texts were unavailable. Following a detailed full-text review of the accessible articles, 41 were excluded for various reasons outlined in Figure 1. Additionally, 7 articles were identified through backward and forward referencing. In total, 48 articles were included in this review which are shown in Table 6.

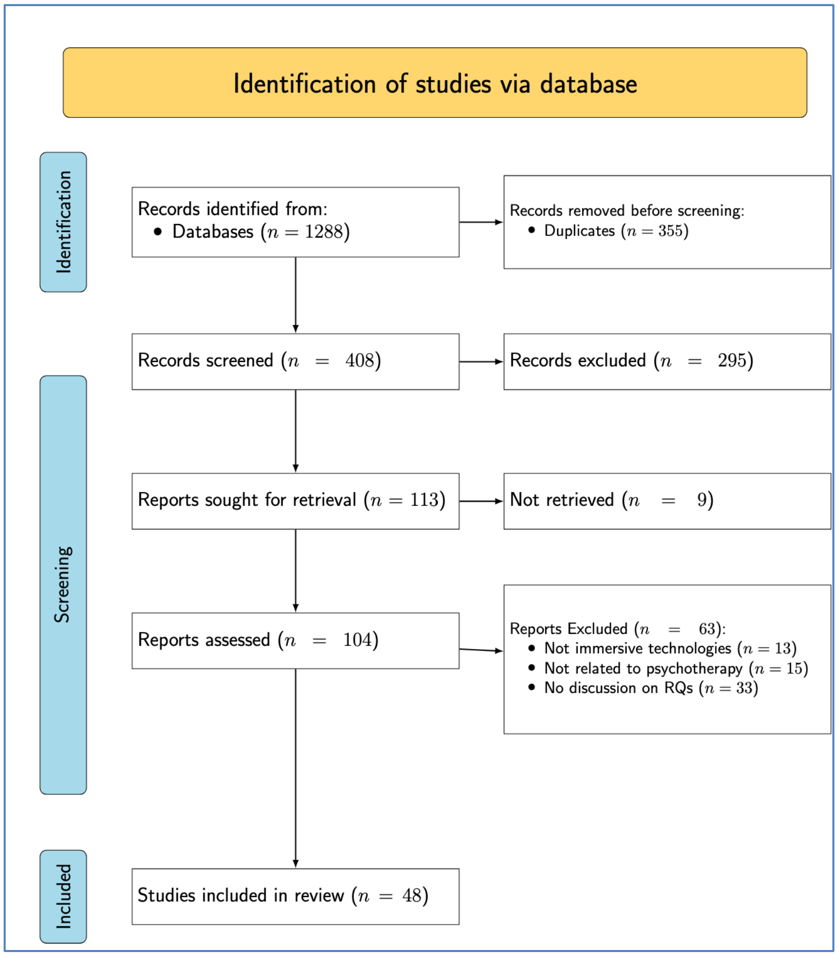

Fig. 1. PRISMA chart illustrating the article selection process.

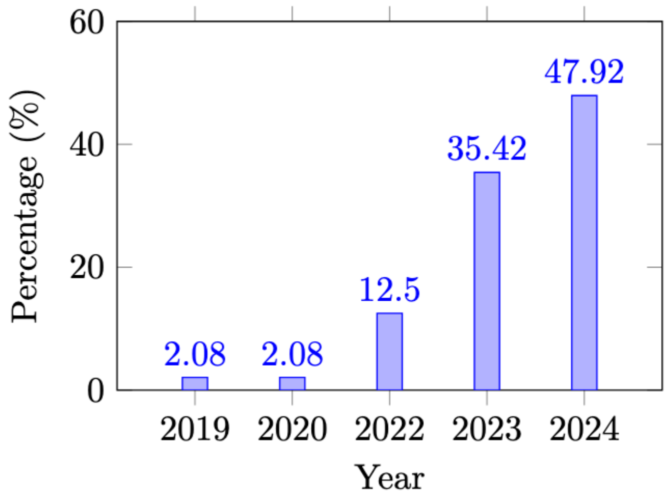

Fig. 2: Distribution of Studies by Year: A bar graph illustrating the percentage of studies conducted from 2014 to 2024, highlighting a significant increase in studies in 2024 compared to earlier years.

Characteristics of Included Studies

Table 6 provides an in-depth summary of the 48 articles reviewed in this study, focusing on the applications of AI and XR technologies in mental health treatment. The table categorizes the studies by type—empirical, conceptual, or systematic13 review—and highlights their main contributions and key findings. Each article's relevance to the research questions (RQs) is mapped in the final column, showing how the study contributes to understanding the benefits (RQ1), risks (RQ2), or regulatory frameworks and strategies (RQ3) associated with these technologies.

The studies included in this review were published between 2014 and 2024, with the largest number appearing in 2024 (n=23, 47.95%) and 2023 (n=16, 35.42%), as shown in Figure 2. Furthermore, the majority of the studies are journal articles (n=43, 89.58%), with conference papers (n=5, 10.42%) comprising a smaller proportion, underscoring a predominant focus on journal publications. Figure 3 illustrates the distribution of studies across the three primary research questions in the field of AI-XR technologies in mental health. RQ1 focuses on exploring the benefits of AI-XR technologies, particularly in mental health and psychotherapy. As shown, a substantial majority of the studies (35 in total) are dedicated to this area, highlighting a strong academic interest in how AI-XR can enhance therapeutic practices and improve mental health outcomes. This reflects a growing recognition of the potential positive impact of these technologies on treatment efficacy and patient engagement.

RQ2 delves into the ethical implications of using AI-XR technologies in mental health care. With 23 studies dedicated to this topic, it is clear that while ethical concerns such as data privacy, informed consent, and algorithmic bias are of significant importance, they are comparatively less explored than the benefits discussed in RQ1. The focus of these studies ranges from the challenges posed by AI decision-making processes to the broader implications for patient autonomy and confidentiality within virtual environments.

RQ3 examines regulatory frameworks surrounding the implementation of AI-XR technologies in mental health. This research area, with only 12 studies, is the least explored of the three. Despite its critical importance, it appears that there is a need for further exploration of how existing regulations can adapt to the rapid advancement of AI and XR technologies. This gap suggests an opportunity for more comprehensive discussions on policy-making, accountability, and standards in AI-XR mental health interventions.

Similarly, Figure 4 presents a percentage breakdown of the number of studies addressing each research question, offering a clearer comparison of the distribution of scholarly attention across the topics. As noted earlier, RQ1 (50%) dominates with the most studies, while RQ2 (33%)and RQ3(17%) receive comparatively less focus, particularly the latter. Finally, Table 2 complements these figures by detailing the benefits, challenges, and regulatory guidelines for implementing AI-XR technologies in mental health treatments. This table provides an organized overview of the key insights from the studies, summarizing the positive impacts, potential obstacles, and the current state of regulatory frameworks, contributing to a holistic understanding of the field's progress and the hurdles that still need to be overcome.

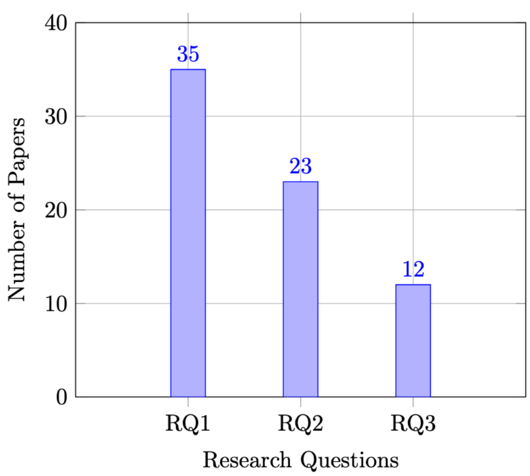

Fig. 3: Number of studies discussing each research question

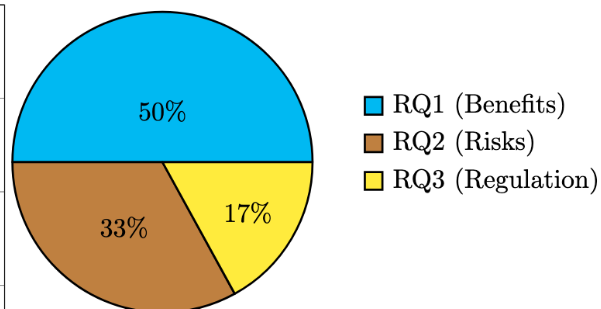

Fig. 4: Percentage distribution of studies across research questions: RQ1 focuses on the benefits of AI-XR technologies in mental health, RQ2 explores ethical implications, and RQ3 examines regulatory frameworks

RQ1: Benefits of AI-XR Applications in Mental Health Treatment

The integration of artificial intelligence and extended reality technologies has ushered in transformative advancements in mental health treatment [64] [24]. These technologies offer distinct advantages in treatment accessibility, personalization, and immersive experiences, addressing a wide range of mental health conditions. This section explores the applications of AI-XR in mental health therapy, focusing on the key benefits, including increased engagement, customization, and the ability to tackle specific psychological conditions effectively.

Table 2: Insights from scoping review: key benefits, challenges, and regulatory strategies for AI-XR applications in mental health.

Category	Details (with References)
Benefits	*Therapeutic Interventions:* • Exposure Therapy [22, 23] • Cognitive Restructuring [2, 24] • Dance Movement Therapy (DMT) [25, 26] • Psychotic Disorders Treatment [3, 24, 27, 87] • Cognitive Rehabilitation [3] • Anxiety and Depression Reduction [28] *Support and Recovery:* • Trauma Recovery [29] • Addiction Recovery [3, 28, 30] • Emotional Regulation [4, 31] • Shared Coping Strategies [27, 31] • Real-Time Feedback [1, 28, 32] *Enhanced Personalization:* • Adaptive Personalized Experiences [3, 13] • Personalized Therapy Models [11], [32] • Real-Time Progress Tracking [33] • AI-Powered Chatbots [34] • Embodied AI-Driven CBT [7][22] *Technological Advancements:* • Immersive Interactive Experiences [28] • Immersive Diagnostics [35] • Gamified Therapy [17] • Accessibility in Remote Regions [4, 28, 31, 36] • Scalable Interventions [28]
Challenges	*Privacy and Security:* • Data Privacy and Data Breaches [37–40] • Identity Verification, Trust and Confidentiality [37–40] • Informed Consent Challenges [41] *Ethical and Societal Concerns:* • Algorithmic Bias and Lack of Transparency [2] • Digital Divide [41–43] • Accessibility Issues and Data Inequality [41–43] • Inclusive Design Gaps [41, 42] *Psychological Risks:* • Emotional Overload and Over-reliance[13, 20, 41, 44] • Cognitive Strain and Emotional Dependency [13, 41, 44] • Maladaptive Escapism [45–48] • Identity Dissociation and Dual Representations [45–48] *Technological Limitations:* • Virtual Harassment and Offensive Behaviors [20, 44] • Trust and Regulatory Issues [44, 78] • Cybersickness and Screen mediated Connection Issues [22, 49]

	• Autonomy & Patient-Therapist Relationship Disruptions [34, 43, 50–52] • Cognitive and Epistemic Limitations of LLMs [2, 39, 53, 54, 90] • Intuition and Empathy Loss [2, 54, 78]
Regulatory guidelines and frameworks.	*Ethical Standards:* • Fairness, Transparency, and Accountability [55, 56] • Standards for Explainable AI (XAI) [56, 66] *Privacy and Consent:* • Data Privacy Regulations [57, 58] • Re-Consent Protocols in Long-Term AI-XR Treatments [54, 59] *Content and Accreditation:* • Age-Based Content Filters and Disclaimers [60] • Standardization of Clinician Accreditation [61] *Technological Strategies:* • Blockchain and Distributed Ledger Technology [56, 62] • Long-Term Impact Assessments [52, 66] *Global Policies:* • Digital Health Policies from the WHO Global Strategy on Digital Health (2020–2025) [63]

Therapeutic Applications of AI-XR in Mental Health

AI-XR is particularly effective in *exposure therapy*, enabling patients to confront and manage phobias (e.g., arachnophobia, social anxiety, PTSD) in controlled virtual environments [22, 23, 88]. These platforms facilitate *cognitive restructuring* and deliver personalized, engaging cognitive behavioral therapy experiences, showing promising outcomes in reducing symptoms of anxiety, depression, OCD, and trauma-related disorders [2, 12, 24]. AI-XR also aids in *chronic and acute pain management* by enhancing relaxation and providing psychological distraction. In addition, it supports the treatment of *psychotic disorders (e.g. schizophrenia)* through virtual environments and interactive avatars, offering safe spaces for individuals with autism and social anxiety to practice social skills [3, 24, 27, 57, 87]. AI-XR systems are further effective in fostering mindfulness, emotional regulation, and stress management, benefiting patients with anxiety, depression, and general stress [1]. VR-based interventions are particularly impactful in *trauma recovery*, providing controlled environments for processing traumatic memories and reducing PTSD symptoms [29]. Cognitive rehabilitation through AR is also revolutionizing treatments for ADHD, dementia, and brain injuries, while AI-XR tools support *addiction recovery* and improve mental well-being in elderly populations [3, 23, 28, 30, 57].

AI-XR tools promote *mindfulness and emotional regulation* through immersive, personalized environments [4, 12, 31, 35]. For instance, XR combined with Brain-Computer Interfaces (BCIs) enhances emotional self-regulation and relaxation, particularly for patients with depression, anxiety, and stress disorders [30] [65]. VR platforms like MUVR create shared spaces for practicing coping strategies, fostering emotional stability and therapeutic engagement [27, 31]. Embodied AI systems and immersive metaverse environments further enhance therapeutic outcomes by enabling adaptive, personalized experiences for complex conditions like personality disorders [3, 13]. AI-enhanced avatars and digital twins also improve treatment precision through real-time progress tracking and predictive analytics [33]. Structured VR platforms, such as TRIPP, utilize immersive visuals and soundscapes to reduce anxiety and depression, highlighting XR's ability to replicate therapy sessions with high levels of presence and customization [28].

Immersion, Personalization and Accessibility of Therapies

AI-powered XR environments enable personalized therapy models tailored to individual needs [11, 32]. For example, RelaXRoom integrates meditation, peer-to-peer, and group therapy with mood-based customization, while biosensing technologies (e.g., eye-tracking, electrodermal activity monitoring) provide real-time feedback to dynamically adjust therapy [1, 28, 32]. Innovations like Dance Movement Therapy (DMT) in XR settings have shown notable reductions in anxiety and depression symptoms, making mental health treatment more accessible, particularly in remote or underserved regions [25, 26]. VRET (Virtual Reality Exposure Therapy) has emerged as a powerful alternative to traditional in vivo exposure therapy, addressing logistical and accessibility challenges. Applications like 3MDR leverage multisensory inputs for trauma-focused therapies, demonstrating VR's efficacy in reducing psychological distress [4, 22, 28, 31, 36].

AI-powered psychotherapy chatbots (e.g., Woebot, Wysa, Joyable) leverage natural language processing (NLP) to deliver structured, evidence-based interventions like CBT, addressing barriers such as financial, geographical, and societal constraints [34]. These chatbots provide accessible, confidential, and timely support, with tools like Woebot aiding in managing *depression, anxiety, addiction, and loneliness* through pattern recognition in text and emojis [2]. Integrating chatbots within XR environments enhances their therapeutic potential by creating immersive, interactive experiences, positioning them as vital components in extending mental healthcare into the Metaverse[3] [58]. However, challenges remain, such as generating real-time, contextually appropriate responses in immersive environments that require multimodal understanding of voice, gestures, and avatar behaviors, which can impact user trust and therapeutic efficacy.

Table 3 summarizes how different AI-XR technologies are being applied in various therapeutic areas. Figure 5 presents a visual representation of the interaction between tools, applications, and users in the context of AI and immersive technologies for mental health interventions. On the left, essential tools such as AI-powered chatbots, virtual reality, and XR are depicted as the foundational technologies driving these advancements. On the right, specific therapeutic applications, including CBT and VRET, are highlighted as outcomes enabled by these tools. Positioned centrally on the left, the "Users" node represents the interaction of patients and clinicians with these technologies, connected by arrows to emphasize their active participation. We have listed the benefits that AI-XR applications bring

2 https://woebothealth.com/
3 https://www.civicspace.tech/technologies/augmented-reality-virtual-reality/

to mental health and psychotherapy treatments in Table 2. While AI-XR technologies hold immense promise for mental health care, several challenges must be addressed to fully realize their potential. Issues such as **cybersickness**, the need for **therapist training**, and the **ethical handling of sensitive biofeedback data** remain critical barriers. Future research should prioritize investigating **long-term therapeutic outcomes** and optimizing user experience to enhance the efficacy and accessibility of these interventions.

Table 3: RQ1: AI-XR Applications in Mental Health Treatments

AI-XR Technology	Application	Mental Health Domain	Reference
VR	Virtual Reality Exposure Therapy	Treatment for phobias, PTSD, social anxiety, and trauma recovery.	[22], [23], [29], [88].
VR	Multi-User Virtual Reality (Group Sessions)	Eating disorders, social anxiety, group therapy, and collaborative coping strategies.	[3], [3], [87], [91].
VR	Immersive Relaxation Environments (e.g., TRIPP)	Anxiety reduction, stress management, mindfulness, and relaxation.	[28], [58], [64]. [65].
VR	Structured Therapeutic Platforms (e.g., CBT Modules)	OCD, schizophrenia, addiction treatment, and targeted therapeutic goals.	[2], [7], [57].
Metaverse	Personalized Digital Avatars and Twins	Depression, social anxiety, and personalized therapy through anonymity and emotional engagement.	[2], [24], [64].
Metaverse	Immersive Environments	Integrative platforms for meditation, peerto-peer, and group therapy.	[13], [28].
Metaverse	Adaptive Schema Mode for Immersive Social Interaction	Personality disorder treatment through immersive experiences.	[13], [87].
AR	AR-based Tools for Neurofeedback and Cognitive Rehabilitation	ADHD, dementia, and brain injury recovery.	[3] [23].
AR	Interactive AR Environments	Social skills training for autism and social anxiety.	[27], [89].
AI-XR Chatbots	Chatbots (e.g., Woebot, Wysa, Joyable) in XR	Delivering CBT and emotional support in interactive XR settings.	[28], [71].
XR	Biosensing Technologies (Eye Tracking, PPG, EDA, Wearables)	Real-time stress management, trauma recovery, and neurofeedback therapy.	[1], [65].
XR	Dance Movement Therapy (XR-	Non-pharmacological interventions for anxiety and depression.	[25].

	DMT)		
XR	Multi-Modular Tools (e.g., 3MDR)	PTSD and trauma recovery leveraging multisensory inputs for memory processing.	[28].
MR	Gamified XR and AI-driven Therapy	Tele-rehabilitation and cognitive training.	[17] [23].
MR	DigitalTechnologies for Self-Regulation	Meditation, mindfulness, addiction recovery, and emotional regulation. Chronic pain and acute pain episodes	[30].
MR	Embodied AI-Driven CBT	Therapeutic interventions leveraging embodied AI.	[7] [9][22].
MR	Immersive Healthcare Diagnostics with AI and XR	Diagnostics and therapeutic advancements in immersive environments.	[35].

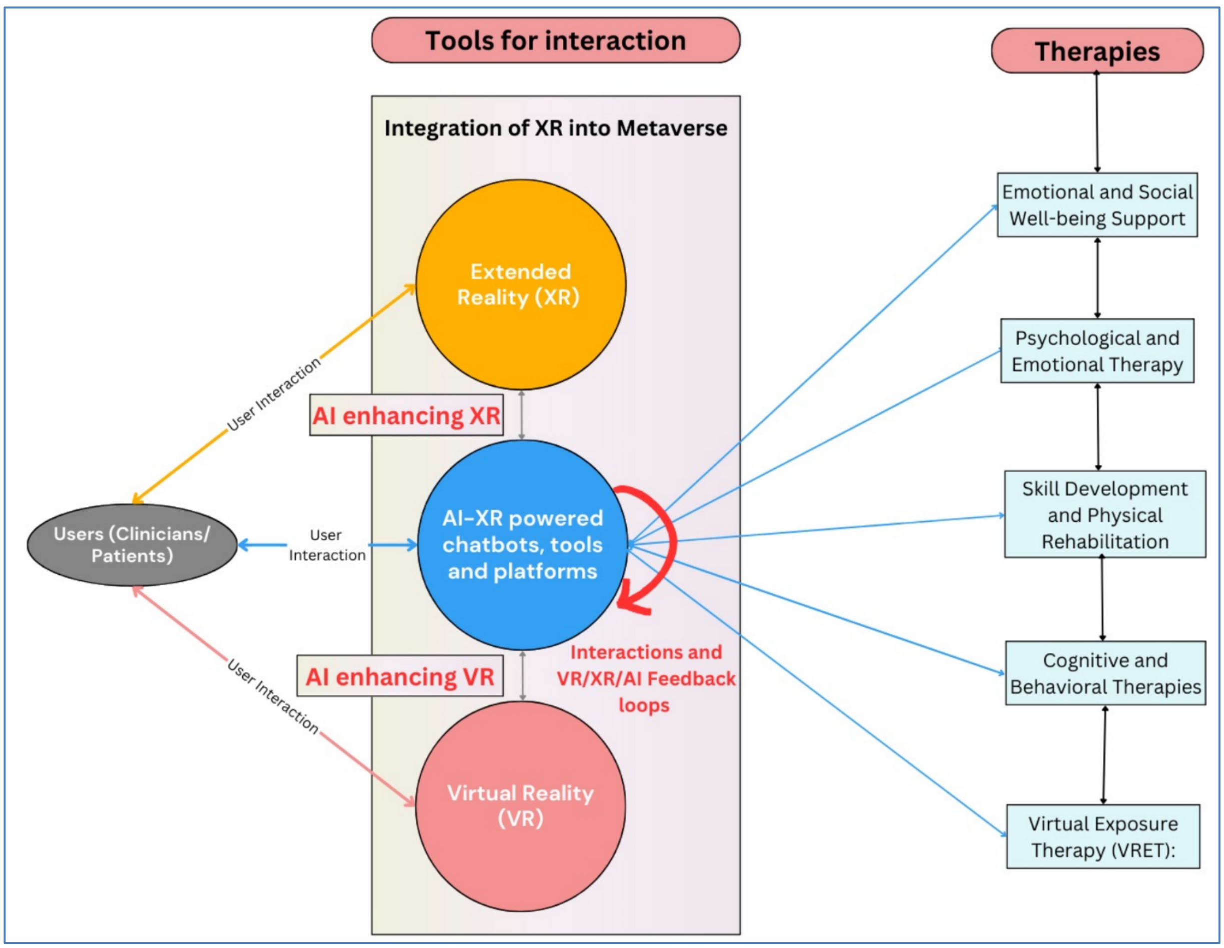

Fig. 5: Conceptual diagram showing interactions between users, AI and Immersive Technologies for Mental Health Therapies

In summary, AI-XR technologies represent a transformative approach to mental health therapy, with diverse applications including exposure therapy, mindfulness, and personalized interventions. However, addressing existing challenges and advancing research are essential to ensure their effective integration into clinical practice and to maximize their therapeutic impact.

RQ2: Challenges of adopting AI-XR in mental health

The integration of AI-XR technologies into psychotherapy and mental health care presents significant challenges that must be addressed to ensure their responsible and effective use. These challenges span ethical, psychological, relational, and cognitive domains, as outlined below.

Ethical Challenges in AI-XR Psychotherapy

AI-XR systems in psychotherapy raise critical ethical concerns, particularly regarding *data privacy* and *identity verification*. The collection of sensitive health data in immersive environments creates complex data flows that challenge regulatory oversight, increasing risks of data breaches and unauthorized access [37–39]. The use of avatars and AI-driven systems further complicates user identity verification, potentially enabling impersonation and undermining therapeutic integrity [20, 39]. Secure authentication protocols, such as multi-factor authentication and biometric verification, are essential to safeguard patient privacy and build trust in these platforms [40]. *Algorithmic bias* is another pressing ethical issue, as AI systems trained on non-representative datasets can produce biased treatment outcomes, disproportionately affecting marginalized groups and exacerbating health inequities [2]. Ensuring informed consent is also more complex in AI-XR therapies, as patients must understand both the technology and its potential biases. Clinicians must prioritize transparency during the consent process to enable informed decision-making [41]. The *digital divide* further complicates AI-XR adoption, as individuals in low-income or rural areas often lack access to necessary technologies [42]. Additionally, AI-XR systems frequently fail to address the diverse needs of populations with disabilities or elderly users, highlighting the need for inclusive design to prevent exacerbating existing inequalities [43]. However, many inclusive features remain underdeveloped, leaving certain populations underserved [41, 42].

Risk of Overload, Emotional Exposure, and Psychological Vulnerability

Overuse of AI-XR in mental health care can lead to emotional overload and cognitive strain, particularly for individuals with anxiety or phobias. Excessive reliance on virtual environments for emotional regulation may result in isolation and negative psychological effects [13], [41], [44]. Effective therapy protocols must include boundaries and recovery periods to prevent *overstimulation* and ensure patients integrate virtual experiences into real-world contexts. The potential for maladaptive escapism in AI-XR environments poses risks for users with mental health issues, as excessive personalization could reinforce unhealthy coping mechanisms. Additionally, avatars and digital personas may exacerbate body image issues or identity dissociation, leading to *psychological harm* [45, 46, 47]. Careful management of personalization is essential to mitigate these risks while maintaining therapeutic benefits [48]. Concerns around offensive behaviors and *virtual assaults* in digital ecosystems are heightened in the metaverse due to its realism and interactivity[4] [44]. Furthermore, *cybersickness*, caused by *sensory incongruences* in XR environments, can lead to *nausea, dizziness, and discomfort*, underscoring the need for advancements in XR design to minimize these effects [49] [22].

Issues in Preserving Patient-Therapist Relationships

While AI-driven therapies can improve accessibility, they risk undermining the human connection essential to effective psychotherapy. Automation may reduce the quality of care by replacing empathetic, nuanced human interactions with automated systems. Although AI tools like chatbots offer consistent support, they cannot replicate the emotional depth of human therapists [43, 50, 51]. Balancing technological advancements with the preservation of therapeutic relationships is crucial to ensure patient well-being [34, 52, 53]. Ultimately, while technology can enhance accessibility and offer valuable support, maintaining the genuine human connection at the heart of therapy remains challenging. As AI and automation become more integrated into mental health care, it is crucial to focus on preserving the emotional and relational aspects of psychotherapy to ensure patients' well-being is not compromised.

Cognitive and Epistemic Limitations of LLMs and AI-XR technologies

Large Language Models (LLMs) like GPT-based systems, despite their ability to simulate empathy, lack the cognitive and epistemic depth of human therapists. Their responses are based on statistical patterns rather than genuine understanding, which may mislead patients or provide inadequate emotional support [38, 39, 53]. Researchers recommend using LLMs as support tools rather than replacements for human therapists, emphasizing the need for regulatory frameworks to ensure their safe and effective use in mental health care [2, 54, 90].

---

4 https://www.technologyreview.com/2021/12/16/1042516/the-metaverse-has-a-groping-problem/

From the above exploration, we understand that the ethical and social implications of AI-XR technologies in mental health treatment are multifaceted, encompassing issues of *data privacy, algorithmic bias, digital inequality, informed consent, psychological impact, inclusivity, and regulatory oversight*, as illustrated in Table 4. Ethical governance frameworks prioritizing patient privacy, transparency, and inclusivity are necessary to guide the deployment of these technologies in a way that fosters trust and benefits all patients equitably [37]. We will explore these in the following subsections.

RQ3: Regulatory Frameworks for AI-XR in Mental Health Treatment

Ensuring AI-XR technologies uphold human dignity and clinical oversight is crucial in mental health care [10]. Regulatory frameworks must address key ethical concerns, including equitable access, data privacy, and algorithmic bias, while safeguarding care quality and maintaining professional accountability [67].

Integrating Fairness and Transparency

Developing trustworthy AI-XR systems requires addressing both technical and ethical complexities. Mitigating algorithmic bias necessitates diverse and inclusive data collection to prevent disproportionate outcomes [55] [82]. Regular audits can identify biases, recalibrate decision-making models, and enhance accountability. Explainable AI (XAI) principles are essential for clinician and patient trust, ensuring that AI-driven decisions remain interpretable and transparent [56][78]. Simplified consent mechanisms, incorporating visual aids and plain language, improve patient comprehension [22, 54, 59]. Continuous re-consent protocols uphold patient autonomy in long-term AI-XR treatments, addressing ethical concerns around ongoing informed consent. AI-driven exposure therapies offer personalized interventions that dynamically adjust to patient responses, optimizing therapeutic outcomes [22, 54]. Transparency and interpretability in AI models, particularly in conversational agents, are critical to fostering trust among patients and clinicians [54]. Intuitive interface designs and user guides help reduce technological barriers for non-expert users. Empathy remains a challenge in AI-XR psychotherapy, as sentiment analysis tools require continuous refinement to improve contextual accuracy and sensitivity [68] [78] [82]. Lexicon-based and rule-based systems must adapt to evolving linguistic patterns to enhance reliability [69]. Data quality and representation are fundamental, as pre-trained models must be rigorously adapted to mental health contexts to prevent biases and ethical violations [32, 70].

Table 4: RQ2: Ethical and Social Issues in AI-XR Technologies for Mental Health Applications

Mental Health Domain	Ethical and Social Issues	References
General Mental Health	Ethical concerns with user data, inclusivity in access, digital inequality	[30], [72]
Psychiatric Care	Algorithmic bias leading to inequitable care, challenges with informed consent in immersive	[2], [41]

	environments	
Therapy	Challenges in preserving patient-therapist relationships, lack of empathy in AI therapy, risks of screen-mediated therapy feeling impersonal	[50], [51]
LLMs in Therapy	Issues with bias, misinformation, over-reliance on AIgenerated insights, and inability to address nuanced emotional or psychological needs	[2], [34], [37], [90]
Immersive Experiences	Psychological impacts such as maladaptive escapism, dependency, virtual traumas, dual representations, cybersickness	[45, 46, 47, 48]
Digital Wellbeing	Privacy concerns, digital inequalities, risks of harmful content and over-reliance on virtual therapies	[65]
Virtual Medical Consultations	Data security concerns, challenges with identity verification in virtual environments	[66]
Public Health and Remote Therapy	Data privacy risks, potential digital addiction, over-reliance on AI-driven virtual therapy platforms	[23]
Surgical Precision and Remote Care	Risks of digital divides, data security, privacy concerns, inequality in accessing advanced technologies	[33]

Inclusive Design and Accessibility in Metaverse-Based Psychotherapy

Standardizing clinician accreditation for immersive therapies and enforcing stringent data security measures will foster public trust. Real-time auditing, dynamic informed-consent processes, and interdisciplinary collaborations are essential for creating robust governance models for AI-XR in mental health care [61]. Inclusive design is vital for enhancing accessibility in AI-XR mental health technologies. Customizable features accommodate individuals with disabilities and diverse cultural backgrounds, improving engagement and therapeutic outcomes [71]. Automated therapy delivery systems can lower costs and expand access, particularly in underserved regions. Adaptive AI personalizes interventions based on patient needs, improving treatment effectiveness. AI-powered Medical Technology (MeTAI) provides data-driven treatment recommendations for mental health conditions such as depression and PTSD [42, 72]. However, rigorous regulatory oversight is necessary to ensure transparency, mitigate bias, and protect patient confidentiality. The digital divide continues to hinder equitable AI-XR deployment, necessitating targeted interventions such as subsidized infrastructure and culturally adaptive content to ensure inclusivity [63].

Regulatory Oversight

Regulatory frameworks must extend beyond data privacy concerns to enforce ethical standards and clinical oversight [39]. Certification programs and ethical review boards can validate AI-XR platforms for mental health care, ensuring accountability for developers and equitable access to effective treatments [10] [60]. Embodied AI therapies, including avatars and virtual counselors, require continuous monitoring to prevent over-reliance and emotional dependency [45, 46, 73]. Long-term studies are necessary to assess their effects and ensure a balanced integration of virtual and real-world interactions [52]. Age-based content filters and disclaimers should be implemented to enhance safety and transparency in AI-driven psychotherapy tools [60]. Collaboration with safety experts is crucial for developing adaptive regulatory policies. Drawing from the WHO Global Strategy on Digital Health (2020–2025), regulatory frameworks should incorporate ethical, legal, and operational safeguards [63].

Data Protection and Security Concerns

AI-XR technologies process sensitive mental health data, raising significant concerns regarding confidentiality and security [57, 58]. Secure data management is critical, with blockchain and Distributed Ledger Technology (DLT) offering tamper-proof records to safeguard patient privacy [56]. These technologies enhance transparency and accountability while minimizing the risks of data breaches [62]. Table 5 outlines the key areas of mental health treatment that require regulatory oversight, highlighting the need for comprehensive strategies to protect patient welfare.

The legal and regulatory landscape for digital healthcare technologies remains fragmented, particularly those used in immersive psychotherapy (e.g., the Metaverse, AR, VR, and XR). As with telemedicine, these technologies face challenges related to cross-border governance, patient confidentiality, liability, and compliance with data protection laws. The European Union's frameworks, such as the GDPR and the Cross Border Healthcare Directive, aim to standardize digital health services but often fall short in addressing the unique nuances of mental health care [63]. This lack of harmonized policies creates uncertainty for practitioners and erodes patient trust, potentially hindering the adoption of these innovative approaches. National-level variations in regulation, funding, and legal interpretations further exacerbate these issues, leaving clinicians with insufficient guidance for safe and effective practice [74].

Training Mental Health Professionals

AI-XR platforms also play a transformative role in training psychiatrists, psychologists, and therapists. Immersive environments like PATIENT allow for interactive and practical learning experiences, enabling clinicians to practice diagnosing and treating patients in simulated scenarios [32]. These virtual tools can reduce training costs and increase accessibility, though equitable access must be ensured to prevent knowledge disparities [37].

The metaverse can simulate various psychological conditions, enabling clinicians to practice diagnosing and treating patients in a safe, controlled environment. Ford et al. emphasize the importance of equitable access to these virtual educational tools to prevent a knowledge gap between those with and without access to these advanced technologies [75]. By incorporating metaverse-based training, mental health professionals can gain hands-on experience treating patients through AI-XR systems, preparing them for real-world applications [89].

Lastly, virtual group therapy platforms allow patients to participate in therapeutic sessions with others from diverse geographical locations, fostering a sense of community [7]. However, these platforms also present challenges, such as maintaining patient confidentiality in a group setting and ensuring that the technology doesn't unintentionally exclude individuals due to socioeconomic factors [76]. Regulatory frameworks must address these concerns by setting clear guidelines for data protection, identity verification, and accessibility. Additionally, legal protections must ensure that patients maintain control over their data and participation in virtual group therapy sessions. Systemically addressing barriers through coordinated global efforts that align with ethical standards and human rights principles is needed. Strategies such as subsidizing infrastructure development, fostering digital literacy, and ensuring culturally relevant therapeutic content are crucial for creating a genuinely equitable digital mental health ecosystem. With comprehensive frameworks and intentional design, immersive technologies can transform mental health care, delivering innovative and compassionate solutions to diverse populations worldwide.

Table 5: RQ3: Integrated Strategies, Regulatory Frameworks, and Ethical Issues for AI-XR Technologies in Mental Health and Psychotherapy

Strategy/Issue	Implications and Regulatory Frameworks	Mitigation Strategy / References
Expand tele-rehabilitation through AI-XR	Long-term regulatory monitoring and efficacy assessments; Risk of data privacy breaches, misuse, or cyber-attacks.	Security measures, confidentiality protocols, and data anonymization [17] [52]
Incorporate blockchain for secure healthcare	Regulatory frameworks for data privacy and secure transactions; Address digital inequality and ensure equitable access.	Subsidies for disadvantaged groups, diverse AI training, and transparent regulations [8] [56]
Develop embodied AI as complementary therapy	Ethical frameworks for AI risk assessments; Bias in AI algorithms leading to discrimination.	Diverse, unbiased, and transparent training of AI models [7][89]
Address accessibility in metaverse psychiatric care	Regulations ensuring equitable access; Psychological impact such as anxiety and over-reliance.	Clear consent forms, patient education, and regular psychological assessments [2] [36]
Revolutionize medical education with metaverse tech	Inclusivity, data privacy, and security regulations; Inclusivity issues related to cultural sensitivity.	Developing culturally sensitive content, inclusive practices, and clear patient communication [10][42]
Enhance healthcare with MeTAI (Medical Technology and AI)	International regulations on AI and blockchain in healthcare; Regulatory gaps in unverified wellness content.	Accountability measures, oversight, and validated platforms for ethical compliance [72] [36]
Healthcare (Secure Data)	XR, Blockchain, DLT: Unregulated use, lack of policies.	Oversight and validated platforms for secure healthcare data [8] [56]
Leverage digital twins in surgical precision	Regulatory collaboration for safe and equitable use of digital twins; Risk of digital inequality.	Equitable access through subsidies and transparent data use [33]
Regulate XR-based gro	Legal protections for privacy,	Developing inclusive

therapy	control, and equitable access; Risk of bias and exclusion in virtual therapy.	practices and unbiased, transparent regulations [3]

Discussion

The findings from this review underscore the transformative potential of AI-XR technologies in mental health treatment while highlighting the ethical, social, and regulatory challenges that must be addressed to unlock their full benefits. Each research question explored in this study contributes to a holistic understanding of the opportunities and risks associated with AI-XR applications.

Issues in Expanding Therapeutic Potential through AI-XR Technologies

AI-XR technologies show promise in revolutionizing mental health treatment by offering immersive, personalized environments that enhance flexibility and adaptability, surpassing traditional methods. Studies indicate that XR tools like VR and AR, combined with large language models, create real-time, responsive therapeutic settings [9] [77]. However, there is a critical need for further empirical studies to assess their longterm impact. Kurata et al. [17] highlight the lack of long-term data on VR-based rehabilitation, hindering our understanding of its sustained effectiveness.

While preliminary research is promising, the scalability and reproducibility of AI-XR solutions remain uncertain. Moodley et al. [8] cite technical challenges, scalability issues, and regulatory hurdles that impede broader adoption. For example, despite offering significant advantages, VR interventions are still constrained by high hardware demands, financial limitations, and usability challenges [24]. Akhtar [11] emphasizes the scarcity of real-world data, particularly for long-term treatments, and ethical concerns surrounding the use of AI with vulnerable populations.

Moreover, many studies suffer from limitations such as small sample sizes, short intervention durations, and overlooked side effects like cybersickness. Mitsea et al. [30] recommend using standardized tools like the Simulator Sickness Questionnaire to assess these adverse effects. Existing literature also shows limited replication of findings, with inadequate exploration of emotional regulation and stress management [78] [79]. Studies often focus on executive functions and socialization, neglecting the potential risks of excessive exposure or emotionally taxing interactions like cyberbullying [48]. This highlights the need to safeguard patients and establish ethical virtual therapy practices.

An over-reliance on technology in mental health care presents additional challenges, as individual needs require more adaptive, nuanced treatments. Automated VR cognitive therapy (VRCT) models, such as those tested by Freeman et al. [71], offer consistent therapy delivery independent of therapist variability. The THRIVE trial and initiatives like gameChange[5] illustrate the scalability and real-world potential of automated VR interventions. However, despite their scalability, these models must

5 https://www.gamechangevr.com/lander

undergo rigorous validation across diverse populations and settings to ensure their effectiveness.

Challenges in Remote and Chatbot-Based Therapies

While AI-driven psychotherapy chatbots show promise, they currently face limitations in adapting to immersive and multimodal virtual environments. A significant challenge is the over-reliance on text-based communication, which restricts user engagement in rich, interactive settings. Existing systems exhibit low compliance rates due to their static nature, which does not align with users' expectations for dynamic communication [34]. Future chatbots must incorporate multimodal capabilities, such as voice interaction, motion recognition, and haptic feedback, to create more immersive therapeutic experiences. These enhancements would enable chatbots to interpret users' intentions more accurately and improve engagement, satisfaction, and compliance. Furthermore, extending conversational depth and providing highly personalized therapy plans can significantly enhance long-term effectiveness and treatment outcomes [32].

Another pressing issue is the lack of robust danger detection mechanisms. In virtual environments, where users may exhibit heightened emotional states due to anonymity or detachment, chatbots must be equipped to recognize signs of self-harm or suicidal ideation [68]. Emotion recognition algorithms, capable of analyzing text, voice, and other multimodal inputs, can detect risky behaviors and initiate timely interventions. Misunderstandings caused by ambiguous inputs or insufficient contextual awareness further compound the challenge. By integrating contextual understanding from multimodal data—such as body language or avatar expressions—chatbots can deliver more accurate and empathetic responses, improving therapeutic outcomes.

Remote and immersive technologies also offer transformative potential in extending healthcare accessibility, particularly in underserved or crisis-affected areas. Innovations such as remote surgeries, telemedicine, and robot-assisted therapies (RAT) have proven effective in addressing medical needs in rural regions with limited resources [43] [73]. Virtual psychotherapy platforms, enhanced by AI-XR technologies, provide scalable solutions for trauma and mental health treatment in conflict-affected regions such as Gaza, Lebanon, Yemen, Sudan, and Ukraine [80]. These solutions can be vital for humanitarian organizations like UNICEF, enabling them to deliver mental health support in remote, resource-constrained environments. However, for these technologies to reach their full potential, inclusivity and accessibility must be prioritized. Systemic barriers such as digital inequality, language limitations, and cultural sensitivity need to be addressed to ensure equitable access for all individuals, regardless of socioeconomic status or geography [81]. Additionally, these technologies should be positioned as complementary tools for healthcare professionals, rather than replacements, to maintain high-quality, human-centered care globally.

Addressing Ethical and Social Challenges

The integration of AI and XR technologies into mental health care presents significant ethical concerns, particularly around data privacy, autonomy, and long-term psychological effects [52]. Extended use of virtual environments, such as the metaverse, may impact self-perception, identity, and real-world relationships, especially among vulnerable populations, such as adolescents and individuals with anxiety or psychosis [29, 33]. These risks necessitate further research into the psychological impact of prolonged virtual immersion. Cultivating supportive virtual spaces that foster positive interactions, extending beyond the virtual realm, is essential to mitigate potential negative outcomes [48].

The integration of AI-driven avatars in therapy introduces concerns about patients becoming overly dependent on virtual therapists, which could diminish the critical role of human connection in the healing process [41]. To preserve the humanistic essence of psychotherapy, it is essential to establish clear boundaries that limit the scope of virtual interactions. For instance, virtual therapists could be positioned as supplemental tools to human therapists rather than replacements, ensuring that they assist with specific tasks such as skill-building or preliminary assessments while reinforcing the primacy of human-led sessions. These boundaries help minimize over-reliance on AI-driven avatars by maintaining a balanced therapeutic approach that prioritizes genuine human connection.

Social equity also requires addressing the digital divide to ensure marginalized communities can access AI-XR-driven mental health care. Bridging gaps related to internet access, cultural content relevance, and device design is crucial for equitable participation across diverse populations [2, 42]. Decentralized platforms in the metaverse could empower these communities, allowing for more culturally sensitive approaches to mental health care.

Regulatory frameworks for AI-XR technologies in mental health care are still in development and must evolve to address emerging risks. Establishing standards for data privacy, transparency, and fairness will be key to fostering trust in these systems [74, 82]. These regulations should also ensure that AI-XR technologies complement, rather than replace, the clinician-patient relationship [77]. Furthermore, establishing trusted digital identities and safeguards against misuse will help preserve patient safety in virtual environments.

A critical regulatory need is the establishment of standards for AI-XR content in virtual therapeutic settings. As AI-XR integration deepens, it is essential to enforce guidelines that prioritize evidence-based practices, real-time content monitoring, and culturally appropriate material [66, 83]. Additionally, fostering metaverse literacy through training initiatives can balance virtual and physical realities and maintain ethical use of these technologies [48].

Engaging patients in the co-creation of their virtual therapeutic environments—such as adjusting sensory elements—can significantly enhance their sense of control and self-efficacy. This active participation in the therapeutic process not only improves motivation but also fosters deeper engagement, potentially leading to better therapeutic outcomes. By allowing patients to tailor their experience to their individual needs, co-creation empowers them as active partners in their healing journey. As AI-XR technologies continue to evolve, regulatory frameworks must be adaptive. Traditional regulations are unlikely to remain effective as these technologies advance, highlighting the need for a dynamic, collaborative approach between technologists, healthcare providers, and policymakers [36]. Future regulatory models could incorporate real-time policy adjustment tools to address emerging AI-XR applications and ensure compliance with patient safety standards [83].

Human supervision remains essential in immersive environments to mitigate risks and uphold the principle of non-maleficence [79]. Developers must ensure that the benefits of using immersive technologies, such as VR and AI-driven chatbots, outweigh the potential risks, particularly in sensitive areas like depression intervention [84]. Ultimately, rigorous testing, continuous updates, and human oversight are essential to ensuring that metaverse-based psychotherapy systems prioritize safety, equity, and positive therapeutic outcomes [74]. Transparent communication about AI limitations and active mitigation of dependency risks will safeguard patient interests in virtual therapeutic contexts.

Limitations

This scoping review has some limitations. First, we did not include the Web of Science (WoS) database in our search strategy, which may have resulted in the omission of relevant studies indexed exclusively in this source. Although the WoS database was not included in our search strategy, the databases we selected—PubMed, PsycINFO, IEEE Xplore, Scopus, and Google Scholar—were specifically chosen to provide broad interdisciplinary coverage of health sciences, psychology, and technology literature. We note that many journals indexed in WoS are also covered by Scopus, PubMed, and PsycINFO, and the inclusion of Google Scholar provided additional breadth by capturing grey literature and non-indexed studies. Thus, while we believe our search strategy minimized the potential impact of excluding WoS by capturing the scope and diversity of relevant studies, we acknowledge that our methodology is limited and not exhaustive.

Second, our approach to duplicate screening and data analysis was conducted using simple tools, such as Microsoft Excel. Pragmatic choices were made, with the second reviewer sampling the data and the bulk of data extraction performed by the first author. This may have limited the efficiency and comprehensiveness of the screening process, and more sophisticated systematic review tools could have enhanced the accuracy and rigor of our methodology.

Additionally, while our search strategy included studies addressing mental health in general and psychotherapy in particular, it did not specifically target studies focusing on various types of mental health disorders or their corresponding treatments. As a result, this review may not fully capture the breadth of AI and XR applications tailored to specific mental health conditions, such as depression, anxiety, bipolar disorder, or PTSD. Future reviews should aim to adopt a more granular focus on disorder-specific interventions to provide a more detailed understanding of the scope and impact of these technologies in mental healthcare.

Conclusions

This review is the first to comprehensively evaluate AI and XR applications in metaverse-driven mental health care, focusing on therapeutic potential, ethical concerns, and societal impacts. Guided by PRISMA guidelines, we examined 48 peer-reviewed articles published post-2014, selecting studies that addressed clinical applications, ethical issues, and social implications of AI-XR in mental health, while excluding non-research articles and irrelevant or inaccessible sources. Organized around three research questions—*Therapeutic Potential* (RQ1), *Ethical Challenges* (RQ2), and *Societal Implications* (RQ3)—this synthesis offers a nuanced understanding of AI-XR technologies' transformative potential in mental health treatment, alongside the ethical and social frameworks necessary for responsible integration. For RQ1, we found that AI-XR applications like virtual reality therapy and AI-driven assessment tools show promise in enhancing accessibility and patient engagement, particularly in resource-limited settings, though further large-scale, longitudinal studies are needed to validate their effectiveness across diverse conditions and populations [85]. In addressing RQ2, the review revealed substantial ethical concerns, especially regarding privacy, data security, and fairness, with sensitive biosignal data raising issues of data ownership and misuse, and algorithmic bias posing risks of inequitable treatment for marginalized groups, highlighting the need for governance frameworks that prioritize transparency, consent, and privacy protections specific to immersive mental health solutions [86]. For RQ3, societal implications emerged, including potential shifts in patient-therapist dynamics, risks of digital exclusion, and psychological dependency on virtual spaces, underscoring the necessity of inclusive regulatory policies to mitigate long-term social and relational impacts in AI-XR therapy. Through this paper, we provide a foundational overview of the capabilities, challenges, and future directions for AI-XR technologies in mental health care, offering critical insights to guide researchers, practitioners, and policymakers in advancing the responsible and equitable integration of these innovations.

Acknowledgements

The authors would like to acknowledge support from Qatar University High Impact Internal Grant QUHICENG23/24- 127.

Appendix

The **PRISMA-ScR Checklist** can be found in Multimedia Appendix 1. The **search strategy** and **search queries** for all databases are detailed in Multimedia Appendix 2, along with a complete overview of the article screening process across all stages. The **data extraction form**, which includes extracted characteristics, methodology, technology, and other attributes, is available in Multimedia Appendix 3.

Table 6: Summary of Studies in our Review: Focusing on AI-XR Technologies and Their Applications in Mental Health Treatment.

No.	Author, Year & Reference	Key Findings	Relevance
1	Akhtar (2024) [11]	AI-XR for personalized mental health. Shortened therapy cycles with VR	RQ1: Tailored treatment
2	Al Dweik et el. (2024) [37]	Technological challenges and need of human oversight in digital mental health interventions, platforms, and modalities.	RQ1, RQ3: Improved clinical outcomes and need for regulatory frameworks.
3	Ali et al. (2024) [56]	Metaverse system for doctor-patient interactions and blockchain for secure data	RQ1: Secure healthcare in metaverse

4	Arnfred et al. (2022) [22]	CBT via group therapy vs VR as potential alternative to in vivo CBT.	RQ1: Iterative design of VR scenarios with feedback.
5	Benrimoh et al. (2022) [23]	Impact of metaverse on mental health and teletherapy risks	RQ2: VR risks in mental health
6	Bhugaonkar et al. (2022) [33]	VR dependency and privacy risks in therapy	RQ2: Long-term mental health outcomes.
7	Blackmore et al. (2024) [1]	Stress and cognitive load due to use of biosignals in healthcare and education.	RQ1: Tracks patient states
8	Cerasa et al. (2024) [87]	Metaverse for psychiatry and disorder-specific treatments	RQ1: Disorder treatment in metaverse.
9	Chengoden et al. (2023) [35]	Metaverse adoption in healthcare with future solutions	RQ2, RQ3: Accessibility barriers
10	Cho et al. (2023) [58]	Higher satisfaction metaverse vs in-person counseling	RQ1, RQ3: Therapeutic relationship benefits
11	Coghlan et al. (2023) [68]	Potentials and pitfalls of chatbots in healthcare: issues and recommendations.	RQ1, RQ2: Benefits and Ethical concerns of chatbots.
12	Costanzo (2024) [44]	Conscious use of therapy, psychology and technology to avoid the risks of addiction or emotional desensitization.	RQ3: Guidelines to implement AR VR in therapy.
13	Distor et al. (2023) [20]	Balances benefits and risks in immersive environments in therapy	RQ1, RQ3: Risks and benefits in therapy
14	Usmani et al. (2022) [45]	Future visions and applications of the Metaverse in health delivery.	RQ1, RQ2: Potential of metaverse for healthcare and limitations.
15	Fajnerova et al. (2024) [24]	Application of multiuser virtual environments (MUVEs) and VR in mental health vs in-person interventions.	RQ1, RQ2: Challenges and Benefits associated with VR, including technical demands, usability barriers, and cybersickness.
16	Ferrario et al. (2024) [39]	Ethical, technical, clinical challenges and Recommendations for responsible design and deployment	RQ2: Risks of integrating LLMs.
17	Fiske et al. (2024) [7]	Ethical insights on AI in mental health and decision-making	RQ2, RQ3: Ethical challenges and guidelines.
18	Ford et al. (2023) [2]	Metaverse applications, accessibility and ethical risks in psychiatric care.	RQ1, RQ2: Therapeutic benefits and challenges
19	Freeman et al. (2019) [71]	Automated VR cognitive therapy using an avatar coach for psychological therapy.	RQ1: Patients find VR environments less intimidating.
20	Guest et al. (2023) [64]	Identified substantial digital solutions for diagnostics, monitoring, and treatment.	RQ1, RQ2: Benefits and limitations.
21	Jiang et al. (2024) [32]	Automated therapist quality assessment and client adherence prediction based on adherence to CBT protocols.	RQ1: Role of AI in CBT.
22	Kruse et al. (2023) [66]	Design for virtual psychotherapy and avatar realism	RQ1, RQ2, RQ3: Design and therapy benefits, challenges and guidelines.
23	Kurata et al. (2023) [17]	AI-XR in tele-rehabilitation outperforms traditional methods	RQ1: Enhances therapeutic outcomes
24	Meinlschmidt et al. (2023) [48]	Framework identifying four mental health conditions related to the metaverse.	RQ1 and RQ2: Benefits and risks of immersive treatments.
25	Mitsea et al. (2023) [30]	Effectiveness of XR digital technologies in Healthcare: self-regulation, mental well-being, and productivity.	RQ1: AI-XR technologies role in enhancing mental well-being.
26	Moldoveanu et al. (2023) [88]	PhoVR Phobia therapy through Virtual Reality for treating acrophobia, claustrophobia, and public speaking anxiety.	RQ1: VR therapy for treating acrophobia, claustrophobia, and public speaking anxiety.
27	Moodley et al. (2023) [8]	XR and blockchain for secure health data and privacy in remote counseling	RQ1, RQ2: Secures patient data and risks.

28	Nadarasa et al. (2024) [10]	Framework for metaverse in healthcare to preserves human dignity	RQ2: Patient-centered ethics.
29	Namkoong et al. (2024) [89]	Nonverbal immediacy in AI counseling to enhance rapport in virtual settings	RQ3: User experience in AI counseling.
30	Navas-Medrano et al. (2024) [3]	XR-based therapeutic environments to enhances social interaction	RQ1: Social benefits in therapy.
31	Obremski et al. (2024) [54]	Autonomous virtual reality exposure therapy to improve the psychotherapeutic supply	RQ1, RQ3: Benefits and ethical principles measurable for developers.
32	Olukayode et al.(2024) [36]	Transformative advancements of mental health, AI therapy and the need of robust regulations.	RQ2, RQ3: Risks and needed regulatory frameworks.
33	Patel et al. (2024) [41]	Ethics, privacy and consent in VR for mental health.	RQ2, RQ3: Ethical considerations and guidelines.
34	Pavlopoulos et al. (2024) [90]	Potential of LLMs in managing anxiety and depression, but ethical issues remain.	RQ1, RQ2: Effectiveness, accessibility, and risks of AI tools in mental health care
35	Radanliev et al. (2024) [25]	XR framework for Dance Movement Therapy (DMT) diversifies mental health therapies	RQ1: XR in alternative therapies
36	Radanliev et al. (2024) [26]	Potential of integrating DMT within XR environments.	RQ1: Convergence of DMT, XR, and AI in mental health treatment.
37	Riches et al. (2024) [65]	VR for stress reduction and wellness intervention access	RQ1: Stress and wellness benefits
38	Saudagar et al. (2024) [42]	AR/VR privacy and access issues in medical education	RQ2: Access and equity issues.
39	Schoenberg (2023) [57]	Digital disparities and schizophrenia issues in mental health	RQ2: Challenges
40	Sestino et al. (2023) [78]	Avatar use for empathy and trust in healthcare.	RQ1, RQ2: Benefits and Challenges
41	Shao et al. (2023) [29]	Risks of unregulated wellness content and privacy.	RQ2: Challenges
42	Soares et al. (2020) [91]	Potential psychotherapy using AI and emerging technologies.	RQ1: Benefits and trends in psychotherapy.
43	Taccgin (2023) [28]	A compact virtual therapy environment: IVR prototype called RelaXRoom.	RQ1: Application of VR therapy.
44	Wang et al. (2022) [72]	MeTAI revolutionizes healthcare infrastructure	RQ1: Technological advancement
45	Xu and Zhuang (2022) [34]	Guided and standardized chatbots must be used in psychotherapy.	RQ1, RQ2, RQ3: Benefits, risks and guidelines.
46	Yin et al. (2022) [13]	Framework for metaverse in schema therapy to support behavioral change	RQ1: Immersive therapy benefits
47	Zafar (2023) [43]	Metaverse's potential to revolutionize the healthcare.	RQ1: Use cases of metaverse in healthcare and therapy.
48	Zhang et al. (2024) [52]	The long-term efficacy of AI-based therapies is questionable.	RQ1, RQ2: Potential advantages and challenges.

References

1. Blackmore KL, Smith SP, Bailey JD, Krynski B. Integrating biofeedback and Artificial Intelligence into extended reality training scenarios: a systematic literature review. Simulation & Gaming; 2024(55):445–478.

2. Ford T, Buchanan DM, Azeez A, Benrimoh DA, Kaloiani I, Bandeira ID, Hunegnaw S, Lan L, Gholmieh M, Buch V, et al. Taking modern psychiatry into the metaverse: integrating augmented, virtual, and mixed reality

technologies into psychiatric care. Frontiers in Digital Health; 2023(5):1146806.

3. Navas-Medrano S, Soler-Dominguez JL, Pons P. Mixed reality for a collective and adaptive mental health metaverse. Frontiers in Psychiatry; 2024(14):1272783.

4. Hur JW, Shin H, Jung D, Lee HJ, Lee S, Kim GJ, Cho CY, Choi S, Lee SM, Cho CH. Virtual reality–based psychotherapy in social anxiety disorder: fMRI study using a self-referential task. JMIR Mental Health; 2021(8):25731.

5. Freeman D, Lister R, Waite F, Galal U, Yu LM, Lambe S, Beckley A, Bold E, Jenner L, Diamond R, et al. Automated virtual reality cognitive therapy versus virtual reality mental relaxation therapy for the treatment of persistent persecutory delusions in patients with psychosis (thrive): a parallel-group, single-blind, randomised controlled trial in England with mediation analyses. The Lancet Psychiatry; 2023(10):836–847.

6. Kislev E. Relationships 5.0: how AI, VR, and robots will reshape our emotional lives. Oxford University Press, Oxford, UK; 2022. ISBN:0195176332.

7. Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied Artificial Intelligence in psychiatry, psychology, and psychotherapy. Journal of Medical Internet Research; 2019(21):13216. Available at: https://doi.org/10.2196/13216

8. Moodley J, Meiring GA, Mtetwa N, Motuba O, Mphephu O, Maluleke M, Balmahoon R. Beyond reality: an application of extended reality and blockchain in the metaverse. IEEE International Conference on Omni-layer Intelligent Systems (COINS); 2023:1–4.

9. Chen E, Prakash S, Janapa Reddi V, Kim D, Rajpurkar P. A framework for integrating Artificial Intelligence for clinical care with continuous therapeutic monitoring. Nature Biomedical Engineering; 2023:1–10.

10. Nadarasa A. Social prescribing in the metaverse: a new frontier for primary care practice. Global Health Journal; 2024(8):32–35.

11. Akhtar Z. Biomedical engineering (BME) and medical health science: an investigative perspective. Quantum Journal of Medical and Health Sciences; 2024(3):1–24.

12. Bansal G, Rajgopal K, Chamola V, Xiong Z, Niyato D. Healthcare in metaverse: a survey on current metaverse applications in healthcare. IEEE Access; 2022(10):119914–119946.

13. Yin B, Wang YX, Fei CY, Jiang K. Metaverse as a possible tool for reshaping schema modes in treating personality disorders. Frontiers in Psychology; 2022(13):1010971.

14. Ghaznavi I, Gillies D, Nicholls D, Edalat A. Photorealistic avatars to enhance the efficacy of self-attachment psychotherapy. IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR); 2020:60–67.

15. Edalat A. Self-attachment: a holistic approach to computational psychiatry. Computational Neurology and Psychiatry; 2017:273–314.

16. Wei S, Freeman D, Harris V, Rovira A. A randomised controlled test in virtual reality of the effects on paranoid thoughts of virtual humans' facial animation and expression. Scientific Reports; 2024(14):1–10. Available at: https://doi.org/10.1038/s41598-024-67534-4

17. Kurata T, Ogata K, Kanazawa S, Imamura Y, Sato A, Ogiso S, Kobayashi Y, Ichikari R, Nakae S, Tada M, et al. Project overview on multimodal XR-AI platform for tele-rehab and the reciprocal care coupling with health guidance. Authorea Preprints; 2023.

18. Rizzo A, Cukor J, Gerardi M, Alley S, Reist C, Roy M, Rothbaum BO, Difede J. Virtual reality exposure for PTSD due to military combat and terrorist attacks. Journal of Contemporary Psychotherapy; 2015(45):255–264.

19. Rizzo AS, Hartholt A, Mozgai S. Settling the score: virtual reality as a tool to enhance trauma-focused therapy for PTSD. Handbook of Media Psychology: the Science and the Practice; 2024:187–213. Springer, New York, NY.

20. Distor C, Dhaou SB, Nielsen M. Metaverse vs. metacurse: the role of governments and public sector use cases. CEUR Workshop Proceedings; 2023.

21. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Moher D, Peters MD, Horsley T, Weeks L, et al. PRISMA extension for scoping reviews (PRISMA-SCR): checklist and explanation. Annals of Internal Medicine; 2018(169):467–473.

22. Arnfred B, Bang P, Hjorthøj C, Christensen CW, Moeller KS, Hvenegaard M, Agerskov L, Gausboel UK, Soe D, Wiborg P, et al. Group cognitive behavioural therapy with virtual reality exposure versus group cognitive behavioural therapy with in vivo exposure for social anxiety disorder and agoraphobia: a protocol for a randomised clinical trial. BMJ Open 12(2):051147 (2022)

23. Benrimoh D, Chheda FD, Margolese HC. The best predictor of the future—the metaverse, mental health, and lessons learned from current technologies. JMIR Mental Health 9(10):40410 (2022)

24. Fajnerova I, Hejtmanek L, Sedlak M, Jablonska M, Francova A, Stopkova P. The journey from nonimmersive to immersive multiuser applications in mental health care: Systematic review. Journal of Medical Internet Research 26:60441 (2024)

25. Radanliev P. Dance as a mental health therapy in the metaverse: exploring the therapeutic potential of dance movement therapy as a non-pharmacological treatment in the metaverse. Frontiers in Computer Science 6:1334027 (2024)

26. Radanliev P. Dance sport movement therapy in the metaverse: A new frontier for alternative mental health therapies. Digital Health 10:20552076241258273 (2024)

27. Matsangidou M, Otkhmezuri B, Ang CS, Avraamides M, Riva G, Gaggioli A, Iosif D, Karekla M. "Now I can see me" designing a multi-user virtual reality remote psychotherapy for body weight and shape concerns. Human–Computer Interaction 37(4):314–340 (2022)

28. Tacgın Z. Designing and developing a compacted immersive virtual therapy environment: Relaxroom. Acıbadem Universitesi Sağlık Bilimleri Dergisi 14(3):462–469 (2023)

29. Shao L, Tang W, Zhang Z, Chen X. Medical metaverse: Technologies, applications, challenges and future. Journal of Mechanics in Medicine and Biology 23(02):2350028 (2023)

30. Mitsea E, Drigas A, Skianis C. Digitally assisted mindfulness in training self-regulation skills for sustainable mental health: A systematic review. Behavioral Sciences 13(12):1008 (2023)

31. Pisalski D, Hierhager M, Stein C, Zaudig M, Bichlmeier C. Influencing the affective state and attention restoration in VR-supported psychotherapy. In: 2020 IEEE Conference on Games (CoG), pp. 646–649 (2020). IEEE

32. Jiang M, Zhao Q, Li J, Wang F, He T, Cheng X, Yang BX, Ho GW, Fu G. A generic review of integrating artificial intelligence in cognitive behavioral therapy. *arXiv preprint* arXiv:2407.19422 (2024).

33. Bhugaonkar K, Bhugaonkar R, Masne N. The trend of metaverse and augmented & virtual reality extending to the healthcare system. *Cureus* 14(9) (2022).

34. Xu B, Zhuang Z. Survey on psychotherapy chatbots. *Concurrency and Computation: Practice and Experience* 34(7):6170 (2022).

35. Chengoden R, Victor N, Huynh-The T, Yenduri G, Jhaveri RH, Alazab M, Bhattacharya S, Hegde P, Maddikunta PKR, Gadekallu TR. Metaverse for healthcare: A survey on potential applications, challenges, and future directions. *IEEE Access* 11:12765–12795 (2023).

36. Olukayode SM, Arogundade B, Adebanjo AO, Kalu OC. Digital mental health interventions and AI in therapy.

37. Al Dweik R, Ajaj R, Kotb R, Halabi DE, Sadier NS, Sarsour H, Elhadi YAM. Opportunities and challenges in leveraging digital technology for mental health system strengthening: A systematic review to inform interventions in the United Arab Emirates. *BMC Public Health* 24(1):2592 (2024).

38. Tortora L. Beyond discrimination: Generative AI applications and ethical challenges in forensic psychiatry. *Frontiers in Psychiatry* 15:1346059 (2024).

39. Ferrario A, Sedlakova J, Trachsel M. The role of humanization and robustness of large language models in conversational artificial intelligence for individuals with depression: A critical analysis. *JMIR Mental Health* 11:56569 (2024).

40. ENGAGE. Spatial computing artificial intelligence. Available at: https://engagevr.io/

41. Patel NJ. Exploring the implications of the metaverse: Opportunities and challenges for dance movement therapy. *Body, Movement and Dance in Psychotherapy*, 1–12 (2024).

42. Saudagar AKJ, Kumar A, Khan MB. Mediverse beyond boundaries: A comprehensive analysis of AR and VR integration in medical education for diverse abilities. *Journal of Disability Research* 3(1):20230066 (2024).

43. Zafar Y. An exploration of metaverse applications in the health sector and their limitations. In: *2023 IEEE Engineering Informatics*, pp. 1–10 (2023). IEEE.

44. Costanzo T. New psycho-technological horizons: The integration of virtual reality and augmented reality in the treatment of psychopathologies and phobias. *Journal of Current Trends in Computer Science Research* 3(5):1–7 (2024).

45. Usmani SS, Sharath M, Mehendale M. Future of mental health in the metaverse. *General Psychiatry* 35(4) (2022).

46. Dellazizzo L, Sert O, Potvin S, Breton R, Pelletier J-F, Renaud P, Dumais A. Avatar therapy for persistent auditory verbal hallucinations: A case report of

a peer research assistant on his path toward recovery. *Psychosis* 10(3):213–219 (2018).

47. Aardema F, O'Connor K, Côté S, Taillon A. Virtual reality induces dissociation and lowers sense of presence in objective reality. *Cyberpsychology, Behavior, and Social Networking* 13(4):429–435 (2010).

48. Meinlschmidt G, Herta S, Germann S, Chee Pui Khei C, Klöss S, Borrmann M. Mental health and the metaverse: Ample opportunities or alarming threats for mental health in immersive worlds? In: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–5 (2023).

49. Weech S, Kenny S, Barnett-Cowan M. Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology* 10:158 (2019).

50. Upton SR, Renshaw TL, Morice A. Sacred vs. secular mindfulness meditation: The influence of presentation priming on therapeutic effectiveness. *Mental Health, Religion & Culture* 25(7):682–695 (2022).

51. Trub L. To the screen, and beyond. *European Journal of Psychotherapy & Counselling* 26(1-2):180–195 (2024).

52. Zhang Z, Wang J. Can AI replace psychotherapists? Exploring the future of mental health care. *Frontiers in Psychiatry* 15:1444382 (2024).

53. Felin T, Holweg M. Theory is all you need: AI, human cognition, and causal reasoning. *Preprint*, submitted April 4 (2024).

54. Obremski D, Wienrich C. Autonomous VR exposure therapy for anxiety disorders using conversational AI—ethical and technical implications and preliminary results. *Journal of Computer and Communications* 12(4):111–138 (2024).

55. Mensah GB. Artificial intelligence and ethics: A comprehensive review of bias mitigation, transparency, and accountability in AI systems. *Preprint*, November 10 (2023).

56. Ali S, Abdullah, Armand TPT, Athar A, Hussain A, Ali M, Yaseen M, Joo M-I, Kim HC. Metaverse in healthcare integrated with explainable AI and blockchain: Enabling immersiveness, ensuring trust, and providing patient data security. *Sensors* 23(2):565 (2023).

57. Schoenberg PL. Welcoming the "metaverse" in integrative and complementary medicine: Introductory overview. *OBM Integrative and Complementary Medicine* 8(4):1–11 (2023).

58. Cho S, Kang J, Baek WH, Jeong YB, Lee S, Lee SM. Comparing counseling outcomes for college students: Metaverse and in-person approaches. *Psychotherapy Research*, 1–14 (2023).

59. Thomason J. Metahealth—How will the metaverse change health care? *Journal of Metaverse* 1(1):13–16 (2021).

60. Khawaja Z, Bélisle-Pipon J-C. Your robot therapist is not your therapist: Understanding the role of AI-powered mental health chatbots. *Frontiers in Digital Health* 5:1278186 (2023).

61. Tabassum A, Elmahjub E, Qadir J. Ethics and regulation in metaverse: Risks, current approaches, and way forward. *Understanding the Metaverse: Applications, Challenges, and the Future*, 265–290 (2024).

62. Gorbunova M, Masek P, Komarov M, Ometov A. Distributed ledger technology: State-of-the-art and current challenges. *Computer Science and Information Systems* 19(1):65–85 (2022).

63. Marinelli S, Basile G, Zaami S. Telemedicine, telepsychiatry and COVID-19 pandemic: Future prospects for global health. In: *Healthcare*, vol. 10, p. 2085 (2022). MDPI.

64. Guest PC, Vasilevska V, Al-Hamadi A, Eder J, Falkai P, Steiner J. Digital technology and mental health during the COVID-19 pandemic: A narrative review with a focus on depression, anxiety, stress, and trauma. *Frontiers in Psychiatry* 14:1227426 (2023).

65. Riches S, Williams G, Saidel S. The metaverse as a therapeutic space: Opportunities for public health benefits and clinician-industry partnerships. *International Journal of Human–Computer Interaction*, 1–4 (2024).

66. Ullah H, Manickam S, Obaidat M, Laghari SUA, Uddin M. Exploring the potential of metaverse technology in healthcare: Applications, challenges, and future directions. *IEEE Access* 11:69686–69707 (2023). Available at: https://doi.org/10.1109/ACCESS.2023.3286696.

67. Ishtaiwi A, Hadi W, Al-Qerem A. Realizing the potential of metaverse to revolutionize the future of healthcare. In: *2024 2nd International Conference on Cyber Resilience (ICCR)*, pp. 01–05 (2024). IEEE.

68. Coghlan S, Leins K, Sheldrick S, Cheong M, Gooding P, D'Alfonso S. To chat or bot to chat: Ethical issues with using chatbots in mental health. *Digital Health* 9:20552076231183542 (2023).

69. Cunningham PB, Gilmore J, Naar S, Preston SD, Eubanks CF, Hubig NC, McClendon J, Ghosh S, Ryan-Pettes S. Opening the black box of family-based

treatments: An artificial intelligence framework to examine therapeutic alliance and therapist empathy. *Clinical Child and Family Psychology Review* 26(4):975–993 (2023).

70. Blease C, Torous J. ChatGPT and mental healthcare: Balancing benefits with risks of harms. *BMJ Mental Health* 26(1) (2023).

71. Freeman D, Lister R, Waite F, Yu L-M, Slater M, Dunn G, Clark D. Automated psychological therapy using virtual reality (VR) for patients with persecutory delusions: Study protocol for a single-blind parallel-group randomized controlled trial (THRIVE). *Trials* 20:1–8 (2019).

72. Wang G, Badal A, Jia X, Maltz JS, Mueller K, Myers KJ, Niu C, Vannier M, Yan P, Yu Z, et al. Development of metaverse for intelligent healthcare. *Nature Machine Intelligence* 4(11):922–929 (2022).

73. Holohan M, Fiske A. "Like I'm talking to a real person": Exploring the meaning of transference for the use and design of AI-based applications in psychotherapy. *Frontiers in Psychology* 12:720476 (2021).

74. Galatzer-Levy IR, Aranovich GJ, Insel TR. Can mental health care become more human by becoming more digital? *Dædalus* 152(4):228–244 (2023).

75. Barak A, Grohol JM. Current and future trends in internet-supported mental health interventions. *Journal of Technology in Human Services* 29(3):155–196 (2011).

76. Kneeland ET, Hilton BT, Fitzgerald HE, Castro-Ramirez F, Tester RD, Demers C, McHugh RK. Providing cognitive behavioral group therapy via videoconferencing: Lessons learned from a rapid scale-up of telehealth services. *Practice Innovations* 6(4):221 (2021).

77. Thieme A, Belgrave D, Doherty G. Machine learning in mental health: A systematic review of the HCI literature to support the development of effective and implementable ML systems. ACM Transactions on Computer-Human Interaction (TOCHI); 2020;27(5):1–53.

78. Sestino A, D'Angelo A. My doctor is an avatar! The effect of anthropomorphism and emotional receptivity on individuals' intention to use digital-based healthcare services. Technological Forecasting and Social Change; 2023;191:122505.

79. Eryılmaz A, Başal A. Rational AIs with emotional deficits: ChatGPT vs. counselors in providing emotional reflections. Current Psychology; 2024:1–16.

80. Diab M, Veronese G, Jamei YA, Kagee A. The interplay of paradigms: Decolonizing a psychology curriculum in the context of the siege of Gaza. Nordic Psychology; 2020;72(3):183–198.

81. Amro N. Post-traumatic stress disorder among nursing students at Palestine Polytechnic University during the Gaza war and the attack on the healthcare system. Middle East Current Psychiatry; 2024;31(1):68.

82. Qayyum A, Butt MA, Ali H, Usman M, Halabi O, Al-Fuqaha A, Abbasi QH, Imran MA, Qadir J. Secure and trustworthy artificial intelligence-extended reality (AI-XR) for metaverses. ACM Computing Surveys; 2024;56(7):1–38.

83. Tabassum A, Elmahjub E, Qadir J. Pathway to prosocial AI-XR metaverses: A synergy of technical and regulatory approaches. In: 2023 International Symposium on Networks, Computers, and Communications (ISNCC); 2023. p. 1–8. IEEE.

84. Denecke K, Gabarron E. The ethical aspects of integrating sentiment and emotion analysis in chatbots for depression intervention. Frontiers in Psychiatry; 2024;15:1462083.

85. Ebrahimzadeh F, Safa R. Unlocking the potential of the metaverse for innovative and immersive digital care. arXiv preprint arXiv:2406.07114; 2024.

86. Magee P, Ienca M, Farahany N. Beyond neural data: Cognitive biometrics and mental privacy. Neuron; 2024;112(18):3017–3028.

87. Cerasa A, Gaggioli A, Pioggia G, Riva G. Metaverse in mental health: The beginning of a long history. Current Psychiatry Reports; 2024:1–10.

88. Moldoveanu A, Mitrut O, Jinga N, Petrescu C, Moldoveanu F, Asavei V, Anghel AM, Petrescu L. Immersive phobia therapy through adaptive virtual reality and biofeedback. Applied Sciences; 2023;13(18):10365.

89. Namkoong M, Park Y, Lee S, Park G. Effect of AI counsellor's nonverbal immediacy behaviour on user satisfaction: Serial mediation of positive expectancy violation and rapport. Behaviour & Information Technology; 2024:1–17.

90. Pavlopoulos A, Rachiotis T, Maglogiannis I. An overview of tools and technologies for anxiety and depression management using AI. Applied Sciences; 2024;14(19):9068.

91. Soares EE, Thrall JN, Stephens TN, Rodriguez Biglieri R, Consoli AJ, Bunge EL. Publication trends in psychotherapy: Bibliometric analysis of the past 5 decades. American Journal of Psychotherapy; 2020;73(3):85–94.