

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

Aliya Tabassum, Ibrahim Ghaznavi, Alaa Abd-Alrazaq, Junaid Qadir

Submitted to: Journal of Medical Internet Research
on: February 09, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	46
Figures	47
Figure 1.....	48
Figure 2.....	49
Multimedia Appendixes	50
Multimedia Appendix 1.....	51
Multimedia Appendix 2.....	51
Multimedia Appendix 3.....	51
TOC/Feature image for homepages	52
TOC/Feature image for homepage 0.....	53

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

Aliya Tabassum¹; Ibrahim Ghaznavi²; Alaa Abd-Alrazaq³; Junaid Qadir¹

¹ Qatar University Doha QA

² Information Technology University Lahore PK

³ Weill Cornell Medical College in Qatar Doha QA

Corresponding Author:

Junaid Qadir

Qatar University
Al Tarfa,
Doha
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.

(JMIR Preprints 09/02/2025:72400)

DOI: <https://doi.org/10.2196/preprints.72400>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <https://www.jmir.org/>

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <https://www.jmir.org/>

Original Manuscript

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

Aliya Tabassum¹, Ibrahim Ghaznavi², Alaa Abd-Alrazaq³, Junaid Qadir^{1*}

¹*Department of Computer Science and Engineering, College of Engineering, Qatar University, Doha, Qatar.

²Department of Computer Science, Information Technology University (ITU), Lahore, Pakistan.

³AI Center for Precision Health, Weill Cornell Medicine - Qatar, Doha, Qatar.

*Corresponding author(s). E-mail(s): jqadir@qu.edu.qa;

Abstract

Background: Mental health systems globally face unprecedented strain due to rising psychological distress, limited access to care, and an insufficient number of trained professionals. Even in high-income countries, the ratio of patients to providers remains inadequate to address demand. Emerging technologies such as Artificial Intelligence (AI) and Extended Reality (XR) are being explored to improve access, engagement, and scalability of mental health interventions. When integrated into immersive metaverse environments, these technologies offer the potential to deliver personalized and emotionally responsive mental healthcare.

Objective: This scoping review explores the state-of-the-art applications of AI and XR technologies in metaverse frameworks for mental health. It identifies technological capabilities, therapeutic benefits, and ethical limitations, focusing on governance gaps related to data privacy, patient-clinician dynamics, algorithmic bias, digital inequality, and psychological dependency.

Methods: A systematic search was conducted in five electronic databases—PubMed, Scopus, IEEE Xplore, PsycINFO, and Google Scholar—for peer-reviewed literature published between January 2014 and October 2024. Search terms included combinations of “AI,” “XR,” “VR,” “mental health,” “psychotherapy,” and “metaverse.” Studies were eligible if they (1) involved mental health interventions, (2) utilized AI and/or XR within immersive or metaverse-like environments, and (3) were empirical, peer-reviewed articles in English. Editorials, conference summaries, and articles lacking clinical or technical depth were excluded. Two reviewers independently screened titles, abstracts, and full texts using predefined inclusion and exclusion criteria, with Cohen’s kappa values of 0.85 and 0.80 indicating strong inter-rater agreement. Risk of bias was not assessed due to the scoping nature of the review. Data synthesis followed a narrative approach.

Results: Of 1,288 articles identified, 48 studies met the inclusion criteria. Included studies varied in design and scope, with most conducted in high-income countries. AI applications included emotion detection, conversational agents, and clinical decision-support systems. XR interventions ranged from VR-based cognitive behavioral therapy and exposure therapy to avatar-guided mindfulness. Several studies reported improvements in patient engagement, symptom reduction, and treatment adherence. However, many studies were limited by small sample sizes, single-institution settings, and lack of longitudinal validation. Ethical risks identified included opaque algorithmic processes, risks of psychological overdependence, weak data governance, and the exclusion of digitally marginalized populations.

Conclusions: AI and XR technologies integrated within metaverse settings represent promising tools for enhancing mental health care delivery through personalization, scalability, and immersive engagement. However, the current evidence base is limited by methodological inconsistencies and a lack of long-term validation. Future research should employ disorder-specific frameworks, adopt standardized efficacy measures, and ensure inclusive, ethical, and transparent development practices. Strong interdisciplinary governance models are essential to support the responsible and equitable integration of AI-XR technologies into mental healthcare.

Limitations: The narrative synthesis limits generalizability, and the absence of a risk of bias

assessment hinders critical appraisal.

Keywords: Augmented Reality (AR); Virtual Reality (VR); Psychotherapy; XR Ethics, digital therapeutics, cognitive behavior therapy, immersive technology, telehealth, mental wellness, virtual interventions.

Introduction

Overview of AI and XR in Mental Health

The convergence of Artificial Intelligence (AI) and Extended Reality (XR) technologies is revolutionizing mental health care, shifting it from static, clinician-driven interventions to dynamic, immersive therapeutic experiences. AI-XR technologies hold transformative potential in addressing the global mental health crisis, particularly in underserved and remote populations. AI-powered chatbots, digital triage systems, and virtual agents offer scalable, 24/7 support without the need for in-person consultations [4, 22, 27]. Simultaneously, XR platforms—particularly in the context of the metaverse—create digital therapeutic environments that enable real-time remote interactions between therapists and patients. These networked ecosystems bypass geographic and socioeconomic barriers, enabling broader access to care and peer support [6]. Notably, metaverse-based clinics and VR teletherapy rooms are now being explored as accessible alternatives to traditional clinical settings [18, 19].

XR serves as a gateway to the metaverse, encompassing immersive technologies such as Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). Each of these modalities offers distinct capabilities. AR overlays digital content, such as images and sounds, onto the physical world through devices like smartphones or AR glasses, enhancing real-world experiences [2]. Similarly, MR blends the physical and digital environments, allowing real and virtual elements to interact in real-time, fostering deep engagement [3]. However, VR differs from AR and MR in that it fully immerses users in a computer-generated environment, completely replacing the physical world.

Among XR technologies, *VR has emerged as the most widely used for digital psychotherapeutic purposes*, facilitating treatments such as exposure therapy, mindfulness training, and cognitive-behavioural interventions. Using VR headsets equipped with motion-tracking sensors and controllers, users can immerse themselves in richly designed three-dimensional therapeutic environments where they interact with digital objects and navigate scenarios through embodied experiences. These virtual spaces are meticulously crafted to deliver multi-sensory interventions—combining 360-degree visuals, ambient soundscapes, haptic feedback, and even olfactory cues—enabling avatar embodiment, teleportation, and emotional immersion that fosters deep cognitive and affective engagement [4,5].

AI plays a crucial complementary role in this ecosystem, powering tools such as *therapy chatbots, emotion recognition algorithms, and decision-support systems* [4, 22, 27, 28, 29]. The most notable advancements have been in *LLM-powered conversational agents*, which provide scalable, interactive, and accessible mental health support. Recent breakthroughs, including *GPT-based models, Llama, and others*, have significantly enhanced AI-driven mental health interventions, making them more responsive and contextually aware [92] [93]. By integrating *AI-driven applications within XR environments*, mental health interventions can become more personalized, adaptive, and accessible than ever before.

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 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]. In these environments, patients interact not only with AI agents and therapists, but also with avatars of peers, family, or caregivers. This social dimension adds a unique layer to therapy, enabling collaborative care and community engagement across cultural and geographic divides [6, 20]. 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

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 AI-XR'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.

Clarification of Key Terminology

To ensure consistency throughout this review, we define key terms as follows:

- **AI-driven:** Refers to applications leveraging artificial intelligence techniques—including machine learning, natural language processing, and computer vision—to enhance mental health interventions. These include AI-powered diagnostic tools, chatbots, and personalized treatment recommendations [7].
- **Virtual Reality:** A fully immersive digital environment that replaces the user's real-world surroundings. VR is typically experienced through head-mounted displays (HMDs) and is used in mental health for exposure therapy, relaxation training, and simulations of social scenarios [83].
- **Augmented Reality:** An interactive experience where digital elements are overlaid onto the real-world environment, usually via mobile devices or AR glasses [44]. AR in mental health can support therapeutic activities by enhancing real-world objects with visual or auditory cues.
- **Mixed Reality:** A hybrid form that merges the physical and digital worlds, allowing real-time interaction between real and virtual elements. MR is increasingly used in therapy to enable more naturalistic interactions in digitally enhanced environments [2,3].
- **XR-enabled:** Encompasses applications utilizing *extended reality* technologies, which include *virtual reality*, *augmented reality*, and *mixed reality* [3]. XR-enabled interventions offer immersive therapeutic experiences, exposure therapy, and skill-building exercises in interactive environments.
- **Metaverse-based:** Describes mental health interventions conducted within *immersive, networked virtual spaces* that enable real-time interactions [58]. These platforms integrate *AI and XR* to facilitate remote therapy, peer support, and social engagement, creating digital therapeutic ecosystems.
- **AI-driven Extended Reality (AI-XR):** A term used to describe the *integration of AI and XR technologies* for mental health applications. AI-XR solutions combine AI-driven personalization, real-time data analysis, and adaptive interventions with immersive XR environments to create intelligent, interactive, and responsive therapeutic experiences [82, 83].

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.

Hypothesis: AI-XR technologies have the potential to *enhance therapeutic outcomes, increase accessibility, and transform mental health treatment* through immersive, AI-driven interventions. However, their adoption poses *significant ethical, societal, and regulatory challenges* that must be addressed to ensure their responsible and equitable integration into psychotherapy.

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 selection of research questions was guided by *a systematic review of existing literature* and an analysis of *unresolved challenges* in AI-XR applications for mental health. While prior studies have explored *specific use cases* (e.g., VR-based exposure therapy or AI-driven diagnostics), a *comprehensive synthesis of AI-XR's broader therapeutic potential, ethical challenges, and governance needs remains lacking*.

- **RQ1 (Applications & Benefits)** was chosen to consolidate fragmented evidence on AI-XR's role across multiple therapeutic domains, as existing studies often focus on isolated interventions without considering their cross-cutting impact on different disorders.
- **RQ2 (Ethical & Societal Challenges)** addresses critical concerns such as privacy risks, data biases, and patient-therapist dynamics, which are frequently mentioned in studies but lack a systematic assessment of their implications for mental health treatment.
- **RQ3 (Regulatory Frameworks)** was formulated in response to gaps in governance strategies, as current regulatory discussions on AI ethics often overlook the immersive and deeply personal nature of mental health interventions in XR environments.

By answering these questions, the objective of this study is to *bridge theoretical insights with practical considerations*, offering actionable recommendations for researchers, clinicians, and policymakers.

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 analysed in relation to the research questions, emphasizing research gaps in areas such as long-term mental health outcomes, equity, and governance frameworks. The limitations of the study are outlined in Section 5 and Section 6 highlights several critical areas require further investigation. Finally, Section 7 concludes the article by summarizing key findings and call for action to key stakeholders 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 AI-XR 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 during the second and third phases were resolved through discussion. Cohen's kappa was calculated to assess inter-rater agreement, yielding a score of 0.85 for the "title and abstract" screening phase and 0.80 for the full-text reading phase. 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.

<i>Inclusion Criteria:</i> Research Articles: Peer-reviewed empirical and review articles only. Publication Date: Published post-2014 to ensure relevance to modern AI and XR advancements. Mental Health Context: Articles addressing AI or XR interventions for mental health treatment, psychotherapy, or digital therapeutics. Ethical Issues: Studies discussing ethical and regulatory implications of AI/XR in mental healthcare.
<i>Exclusion Criteria:</i> Non-Research Articles: Editorials, commentaries, policy papers, abstracts, or conference summaries without empirical findings. Irrelevant Focus: Studies primarily focused on AI or XR without a clear application to mental health or the metaverse. Low-Quality Sources: Preprints, non-peer-reviewed literature, or articles with insufficient methodological transparency.

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 behavioural 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 7.

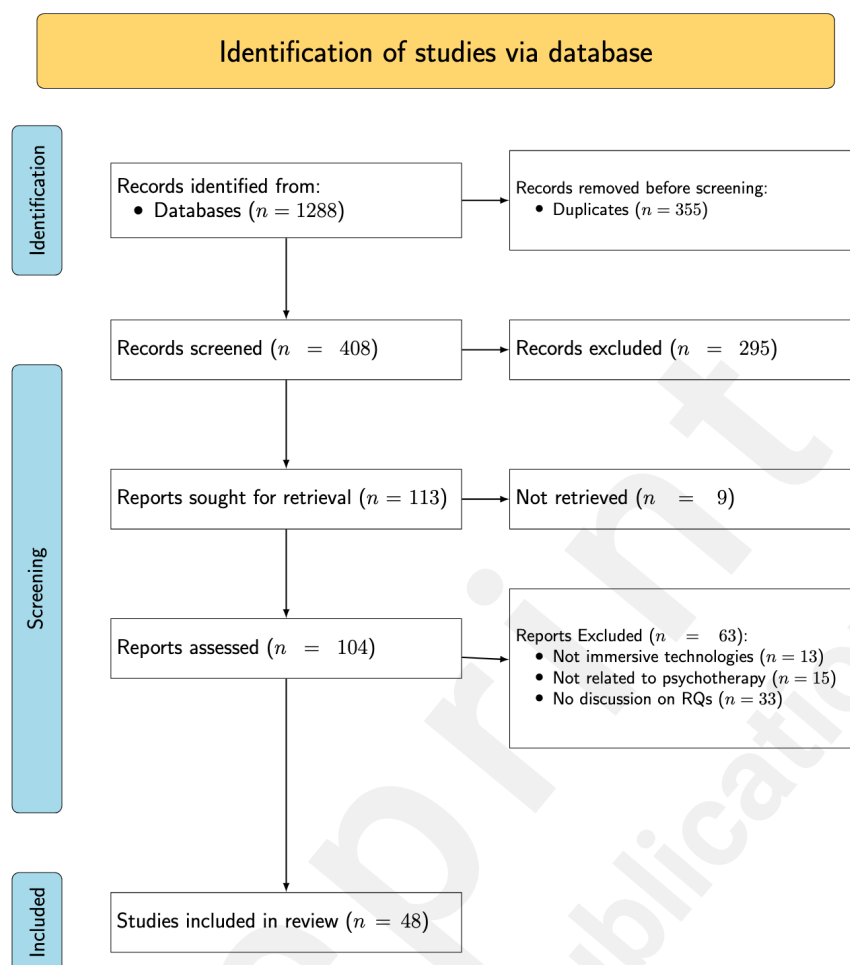


Fig. 1. PRISMA chart illustrating the article selection process.

Characteristics of Included Studies

Table 8 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 systematic 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.

Table 2 provides a comprehensive overview of the distribution of articles across various dimensions, focusing on the year of publication, country of publication, and the research questions addressed in the field of AI-XR applications in mental health. though the search was used from 2014-2024. In terms of publication year, 2024 dominates with 47.9% of references (23 publications), signaling a recent surge in AI-XR research in mental health. 2023 follows with 33.3%, indicating continued growth. Publications from 2022, 2021, 2020, and 2019 contribute smaller shares (2.1% each), reflecting the field's recent emergence.

In terms of country of publication, the United States leads with 27.1% (13 publications), followed by Switzerland at 22.9% (11 publications). The United Kingdom contributes 20.8% (10 references), underscoring global interest. Other countries such as Canada, China, and Italy account for smaller

yet notable shares, while a collective 12.5% comes from Hungary, India, Indonesia, Singapore, South Korea, Turkey, and Germany, highlighting the international scope.

Regarding research questions, RQ1 (benefits of AI-XR in mental health) dominates with 50%, focusing on the positive impacts of these technologies. RQ2 (challenges and limitations) follows with 32.86%, indicating significant attention to obstacles. RQ3 (ethical considerations and future directions) represents 17.14%, showing a lesser focus but still notable interest in the broader implications of AI-XR technologies.

Furthermore, the majority of the studies are journal articles (n=46, 95.8%), with conference papers (n=2, 4.2%) comprising a smaller proportion, underscoring a predominant focus on journal publications.

Finally, Table 3 details 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.

Table 2: Characteristics of the included studies (N=48).

Features	Studies	References
Year of publication, n(%)		
2024	23 (47.9%)	[1, 3, 7, 10, 11, 24, 25, 26, 28, 32, 36, 37, 39, 41, 42, 43, 44, 52, 54, 65, 87, 89, 90]
2023	16 (33.3%)	[2, 8, 20, 28, 29, 30, 35, 48, 56, 57, 58, 64, 66, 68, 78, 88]
2022	6 (12.5%)	[13, 22, 23, 33, 34, 45, 72]
2021	1 (2.1%)	[17]
2020	1 (2.1%)	[91]
2019	1 (2.1%)	[71]
Country of publication, n(%)		
Switzerland	13 (27.1%)	[2, 3, 13, 25, 26, 30, 34, 52, 56, 64, 66, 88, 90]
United States (US)	11 (22.9%)	[1, 8, 32, 33, 34, 35, 43, 57, 68, 87, 91]
United Kingdom (UK)	10 (20.8%)	[17, 22, 34, 41, 45, 58, 65, 71, 72, 89]
Canada	4 (8.3%)	[7, 23, 24, 39]
China	2 (4.2%)	[10, 54]
Italy	2 (4.2%)	44, 78
Others (6 countries)	6 (12.5%)	Hungary: [20] India: [36] Indonesia: [11] Singapore: [29] South Korea: [42] Turkey: [28] Germany: [30]
Publication type, n (%)		

Conference paper	2 (4.2%)	[43, 66]
Journal article	46 (95.8%)	[1, 2, 3, 7, 8, 10, 11, 13, 17, 20, 22, 23, 24, 25, 26, 28, 29, 30, 32, 33, 34, 35, 36, 37, 39, 41, 42, 44, 45, 48, 52, 54, 56, 57, 58, 64, 65, 68, 71, 72, 78, 87, 88, 89, 90, 91]
Research question distribution, n (%)		
RQ1	35 (50%)	[11, 37, 56, 22, 23, 1, 87, 35, 58, 68, 44, 20, 45, 24, 39, 7, 8, 32, 95, 30, 88, 8, 89, 54, 36, 25, 26, 65, 42, 57, 28, 72, 13, 52]
RQ2	23 (32.86%)	[23, 33, 2, 56, 20, 45, 44, 57, 66, 48, 72, 90, 32, 39, 68, 30, 89, 41, 91, 8, 36, 95, 72]
RQ3	12 (17.14%)	[37, 56, 7, 2, 71, 42, 90, 25, 65, 57, 32, 41]

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 3: 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, 95]
	Support and Recovery: • Trauma Recovery [29, 95] • 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	<i>Privacy and Security:</i> • Data Privacy and Data Breaches [37–40] • Identity Verification, Trust and Confidentiality [37–40] • Informed Consent Challenges [41]
	<i>Ethical and Societal Concerns:</i> • Algorithmic Bias and Lack of Transparency [2] • Digital Divide [41–43] • Accessibility Issues and Data Inequality [41–43] • Inclusive Design Gaps [41, 42]
	<i>Psychological Risks:</i> • 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]
	<i>Technological Limitations:</i> • 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, 95]
Regulatory guidelines and frameworks.	<i>Ethical Standards:</i> • Fairness, Transparency, and Accountability [55, 56] • Standards for Explainable AI (XAI) [56, 66]
	<i>Privacy and Consent:</i> • Data Privacy Regulations [57, 58] • Re-Consent Protocols in Long-Term AI-XR Treatments [54, 59]
	<i>Content and Accreditation:</i> • Age-Based Content Filters and Disclaimers [60] • Standardization of Clinician Accreditation [61]
	<i>Technological Strategies:</i> • Blockchain and Distributed Ledger Technology [56, 62] • Long-Term Impact Assessments [52, 66]
	<i>Global Policies:</i> • Digital Health Policies from the WHO Global Strategy on Digital Health (2020–2025) [63]

Therapeutic Applications, Interventions and Outcomes of AI-XR in Mental Health

AI-driven extended reality has proven particularly effective in exposure therapy, allowing

patients to confront and manage phobias such as arachnophobia, social anxiety, and PTSD within controlled virtual environments [23, 88]. Studies comparing AI-driven XR therapy (*CBT-in virtuo*) with traditional *CBT-in vivo* for individuals with social anxiety disorder and agoraphobia have demonstrated its effectiveness. Anxiety symptoms were evaluated using the Liebowitz Social Anxiety Scale (LSAS) and the Mobility Inventory for Agoraphobia (MIA), with scores standardized through the Percentage of Maximum Possible (POMP) method for direct comparison. *Post-treatment findings indicate that patients receiving AI-enhanced XR therapy exhibited lower anxiety symptoms than those undergoing conventional therapy*, underscoring the potential of AI-integrated XR interventions in mental health treatment [22]. Similarly, in neurosurgical planning, AR navigation techniques have been applied to improve surgical precision, particularly in spine fixation procedures. AR-assisted surgical navigation (ARSN) has led to higher accuracy in screw placement and a reduced incidence of cortical breaches, highlighting its potential to enhance surgical outcomes [11].

These platforms also facilitate *cognitive restructuring* and deliver personalized, engaging cognitive behavioral therapy experiences, showing promising outcomes in reducing symptoms of 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]. 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 systems are further effective in fostering mindfulness, emotional regulation, and stress management, benefiting patients with anxiety, depression, and general stress through immersive, personalized environments [1, 4, 12, 31, 35]. For instance, A study compared metaverse-based and in-person counseling in 60 South Korean university students, with 32 choosing metaverse and 28 in-person. Psychological symptoms were measured pre-, post-, and follow-up, while therapeutic relationships and satisfaction were assessed mid- and post-session. Both modalities showed similar symptom reduction, with metaverse counselling improving depression, anxiety, and social support deficits. Findings support its feasibility for tele-mental health, warranting further research on its broader applications [58]. 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-based interventions are particularly impactful in *trauma recovery*, providing controlled environments for processing traumatic memories and reducing PTSD symptoms [29].

VR platforms like Multi-user Virtual Reality (MUVR) or Multi-user Virtual Environment (MUVE) create shared spaces for practicing coping strategies, fostering emotional stability and therapeutic engagement [27, 31]. Studies on digital multiuser interventions in mental health care report positive therapeutic effectiveness, particularly in reducing symptoms of depression and anxiety. Primary outcome measures include symptom reduction, engagement, usability, and acceptability. Patient feedback highlights that digital platforms help overcome social barriers and foster a sense of community. While some studies show partial symptom reduction, randomized controlled trials consistently demonstrate that these interventions effectively alleviate symptoms, making them a promising alternative to traditional therapy [24]. 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]. These findings are summarized in Table 4.

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. VRET provides a safe, customizable, and immersive environment, demonstrating comparable efficacy in reducing symptoms of phobias, PTSD, and OCD while offering added benefits such as real-time customization and increased patient comfort [32, 94]. Applications like 3MDR leverage multisensory inputs for trauma-focused therapies, demonstrating VR's efficacy in reducing psychological distress [4, 22, 28, 31, 36].

Likewise, a study examined the impact of AI agents' nonverbal immediacy behavior (NIB) on user satisfaction in counseling settings within a metaverse consultation environment. Using 762 valid responses from participants recruited via Amazon Mechanical Turk, the study found that AI agents displaying NIB significantly enhanced user satisfaction. This effect was mediated by positive expectancy violation and rapport, with serial mediation further strengthening the relationship [89].

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 emojis¹. 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 Metaverse² [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. A study with 689 participants examined the impact of human-like interaction in digital healthcare services, comparing medical chatbots (low interaction) to doctors' avatars in the Metaverse (high interaction). Findings revealed that higher human-like interaction increased the intention to use digital healthcare services through perceived anthropomorphism. Emotional receptivity moderated this effect, with a significant impact only among participants with high emotional receptivity. A moderated-mediation analysis confirmed a positive indirect effect of human-like interactions on intention to use [78].

1 <https://woebothealth.com/>

2 <https://www.civicspace.tech/technologies/augmented-reality-virtual-reality/>

Table 4: AI-XR Interventions and Reported Outcomes

AI-XR Intervention	Sample Size	Outcome Measures	Reported Outcomes
AI-XR therapy for Social Anxiety & Agoraphobia	302 participants	LSAS, MIA (converted to POMP)	Symptom reduction, adherence rates TBD [22].
Enhanced exposure therapy using VR (VRET)	Not specified	Symptom reduction in phobias, PTSD, OCD, etc.	Comparable to traditional ET but with added benefits (e.g., safety, real-time customization) [32].
AR for neurosurgical planning	Not specified	Accuracy of screw placement, cortical breach rate.	Improved precision in spine fixation procedures [11] [94].
Metaverse-based vs. in-person counselling for psychological symptoms	60 students (32 metaverse, 28 in-person)	Psychological symptoms (pre-, post-, follow-up), therapeutic relationships (mid-session), satisfaction (post-session)	Metaverse counseling improved depression, anxiety, and social support deficits [58].
Virtual worlds	Not specified	Symptom reduction (depression, anxiety), engagement, usability, long-term benefits.	Positive overall; symptom reduction in videoconferencing/chat rooms [24].
AI-based metaverse consultation environment.	762 participants	User counseling satisfaction, Nonverbal Immediacy Behavior (NIB).	Enhanced user satisfaction [89].
Medical chatbots (low human-like interaction) vs. Doctors' avatars in the Metaverse (high human-like interaction)	689 participants	Intention to use digital-based healthcare services	Higher human-like interaction (doctors' avatars) increases intention to use via perceived anthropomorphism [78].

Table 5 summarizes how different AI-XR technologies are being applied in various therapeutic areas. Figure 2 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 to mental health and psychotherapy treatments in Table 3.

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 5: RQ1: Overall 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, peer-to-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	Non-pharmacological interventions	[25].

	Therapy (XR-DMT)	for anxiety and depression.	
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	Digital Technologies 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].

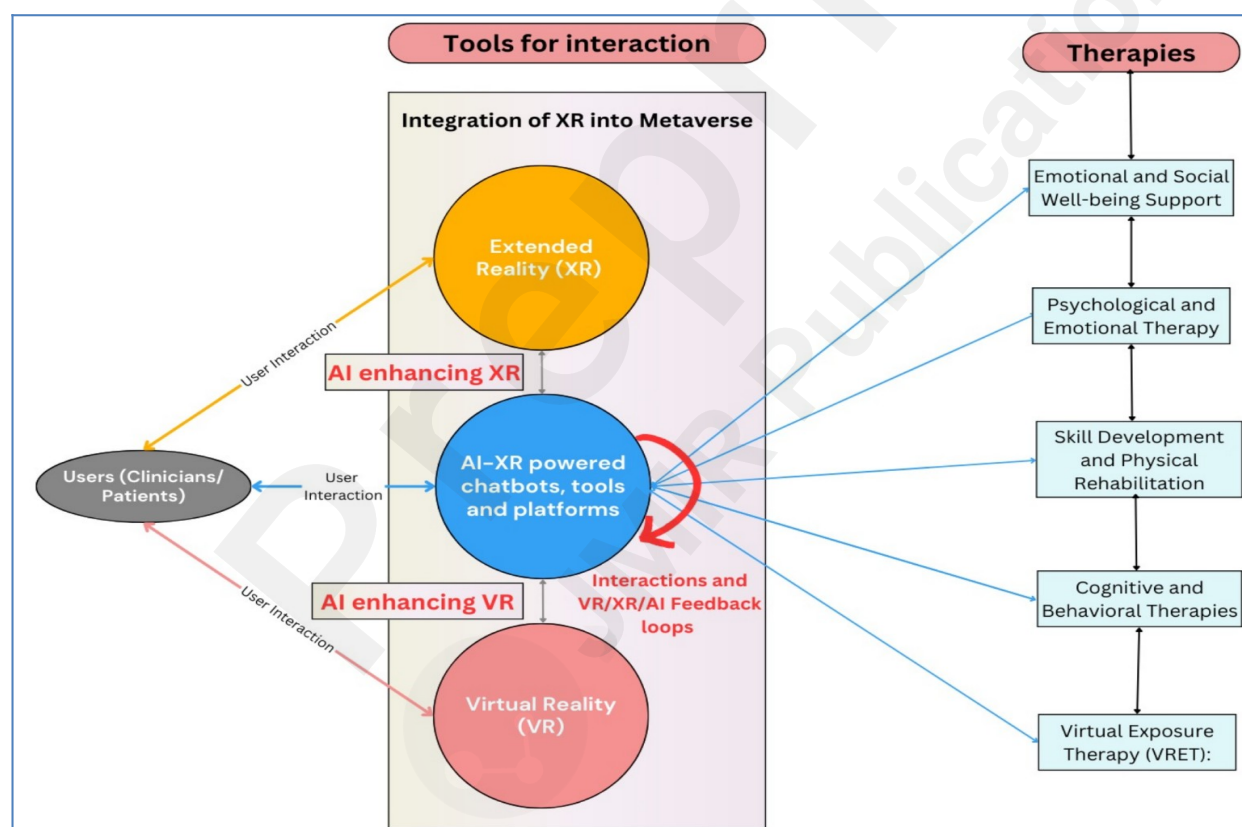


Fig. 2: Conceptual diagram showing interactions between users, AI and Immersive Technologies for Mental Health Therapies

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

One emerging concern in AI-XR mental health applications is *digital therapeutic dependency*, where patients become overly reliant on AI-driven interventions, potentially reducing their engagement in real-world therapeutic processes [65]. This risk is heightened when AI-XR tools offer high levels of personalization, leading to excessive reliance on virtual environments rather than real-life coping strategies. 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*³ [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

³ <https://www.technologyreview.com/2021/12/16/1042516/the-metaverse-has-a-groping-problem/>

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.

Clinician Perceptions and Resistance

Mental health professionals play a critical role in the adoption of AI-XR interventions, yet their acceptance is influenced by factors such as perceived efficacy, ethical concerns, and the potential disruption of traditional therapeutic models. Some clinicians express skepticism about AI-driven systems' ability to provide meaningful therapeutic interactions, fearing that automation may undermine the human elements essential to psychotherapy [55, 56]. Additionally, concerns about liability and professional accountability in AI-assisted care contribute to hesitancy among practitioners [57]. Addressing these concerns through clinician education, transparent AI decision-making processes, and ethical guidelines is essential for fostering trust in AI-XR applications.

Patient Trust and Adoption Barriers

Beyond clinician resistance, patient mistrust presents another significant challenge. Many individuals express scepticism toward AI-driven therapy due to concerns about data privacy, the impersonal nature of digital interventions, and the perceived lack of emotional depth in AI-generated responses [69]. For vulnerable populations, such as those with severe mental health conditions, the idea of engaging with AI-XR systems instead of a human therapist may raise concerns about misdiagnosis, lack of empathy, or inappropriate treatment recommendations [78, 79, 95]. Ensuring transparency in AI decision-making and offering hybrid care models—where AI-XR complements, rather than replaces, human therapists—can help mitigate these concerns [55].

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].

The Need for Cross-Disciplinary Collaboration

AI-XR adoption in mental health requires an interdisciplinary approach, where healthcare professionals, technologists, and policymakers work together to design, regulate, and implement effective solutions [30, 61]. Engineers must develop systems that are clinically validated and ethically sound, while policymakers must establish regulatory frameworks that ensure data security, algorithmic fairness, and equitable access. Ethical oversight, particularly in addressing algorithmic bias and ensuring inclusivity, is essential to prevent AI-XR technologies from exacerbating health disparities [83]. Collaborative efforts that integrate clinical expertise with technological advancements will be key to maximizing the benefits of AI-XR while addressing its limitations.

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 6. 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

As AI-XR technologies become integrated into mental health care, ensuring compliance with established regulatory frameworks is critical for safeguarding patient rights and maintaining ethical standards. Existing regulations such as the *General Data Protection Regulation (GDPR)* and the *Health Insurance Portability and Accountability Act (HIPAA)* provide guidelines for handling sensitive health data, enforcing privacy protections, and ensuring informed consent in digital healthcare settings. GDPR mandates explicit user consent, data minimization, and the right to erasure, while HIPAA focuses on safeguarding patient health information (PHI) from unauthorized access and misuse [62, 63]. However, the immersive nature of XR environments introduces new challenges, requiring updated policies that address real-time data collection, biometric tracking, and AI-driven decision-making. 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 6 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], [95]
LLMs in Therapy	Issues with bias, misinformation, over-reliance on AI-generated 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, Mental Health Ethics and AI-XR Governance

Beyond general data protection laws, mental health ethics guidelines, such as those from the *American Psychological Association (APA)* and the *World Health Organization (WHO)*, provide ethical principles for AI-assisted therapy. These frameworks emphasize patient autonomy, transparency in AI decision-making, and clinician oversight in AI-driven interventions [64]. Ensuring that AI-XR aligns with these principles is essential to maintaining therapeutic integrity and preventing unintended harm. 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 counsellors, require continuous monitoring to prevent over-reliance and emotional dependency [45, 46, 73]. Combining AI-driven interventions with human therapist supervision ensures balanced therapeutic engagement. Implementing session caps and scheduled breaks to encourage real-world application of therapeutic skills. Informing patients about AI's role as a supplement, not a replacement, for human therapy. Establishing guidelines to prevent excessive AI dependency, similar to existing mental health treatment protocols. 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 7 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 7: 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-enabled group therapy	Legal protections for privacy, control, and equitable access; Risk of bias and exclusion in virtual therapy.	Developing inclusive 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, despite promising early results, studies like those by Kurata et al. [17] and Akhtar [11] highlight a lack of long-term empirical data and real-world validation for XR-based interventions.

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 gameChange⁴ illustrate the scalability and real-world potential of automated VR interventions. However, despite their scalability, these models must 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]. To improve engagement and compliance, future chatbots must evolve beyond static, text-based models by integrating voice, gesture, and emotional recognition technologies. 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].

⁴ <https://www.gamechangevr.com/lander>

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].

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]. To maintain human-centered care, AI-XR technologies should be positioned as adjunct tools that support—rather than replace—clinicians, ensuring continuity of trust and emotional safety [77]. Furthermore, establishing trusted digital identities and safeguards against misuse will help preserve patient safety in virtual environments.

Clinical, Technological, and Policy recommendations

The findings from this review suggest several actionable insights for clinical practice, system design, and regulatory development.

Clinical Implications: AI-XR technologies hold transformative potential in clinical mental health care, particularly through telepsychiatry, virtual cognitive behavioral therapy, and immersive exposure interventions [63]. Their integration can help address professional shortages, reduce geographic barriers, and enhance continuity of care for patients in remote or underserved areas. However, clinical protocols must be adapted to the unique challenges of virtual care—such as managing cybersickness, assessing therapeutic alliance in immersive contexts, and determining the need for human intervention [24, 30].

Training healthcare professionals in the effective use of these technologies and establishing clear clinical guidelines will be essential. Human oversight remains critical in immersive environments to mitigate risks and uphold the principle of non-maleficence [79]. Developers must ensure that the benefits of immersive technologies, including VR and AI-driven chatbots, clearly outweigh potential risks—particularly in sensitive contexts like depression treatment [84]. Ultimately, rigorous testing, continuous updates, and human supervision are vital to ensuring that metaverse-based psychotherapy systems prioritize safety, equity, and therapeutic effectiveness [74].

Technological Considerations: From a system design perspective, AI-XR platforms must prioritize user-centered development—ensuring solutions are intuitive, accessible, and culturally sensitive. Addressing algorithmic bias is vital to avoid reinforcing existing health disparities. Developers should embed multimodal capabilities—such as voice, motion tracking, and emotional recognition—into AI chatbots and XR interfaces to enhance user engagement and therapeutic responsiveness.

Fostering metaverse literacy through targeted training can support ethical and informed use of these technologies. Patient involvement in co-creating their virtual therapeutic environments—by adjusting sensory elements, for example—can boost their sense of control, motivation, and self-efficacy. This active participation promotes deeper engagement and improved outcomes by empowering patients as partners in their healing journey.

Policy Recommendations: The ethical deployment of AI-XR in mental health care requires robust governance frameworks. Data protection policies must safeguard privacy and transparency, particularly for vulnerable populations. Regulatory alignment with frameworks such as the EU AI Act can help establish safety standards and accountability mechanisms [61]. It is also critical to define the legal boundaries of AI-augmented therapy—distinguishing clinical from non-clinical use—and to require real-time monitoring and auditability of AI decisions.

Investments in digital infrastructure, public education, and metaverse literacy initiatives can help bridge the digital divide and promote equitable access [83]. Transparent communication about AI limitations and active strategies to mitigate dependency risks will protect patient interests. Furthermore, standards for AI-XR content must ensure evidence-based practices, real-time moderation, and cultural appropriateness [66, 83].

As AI-XR technologies evolve, regulatory frameworks must remain adaptive. Traditional models may not suffice in addressing the rapid pace of innovation, necessitating a dynamic, collaborative approach among technologists, clinicians, and policymakers [36]. Future regulatory systems may

benefit from real-time policy adjustment tools to respond to emerging applications while safeguarding patient safety and care quality [83].

Limitations

This scoping review has several limitations. First, our search strategy did not include the *Web of Science (WoS)* database, which may have led to the omission of relevant studies indexed exclusively in this source. However, we selected *PubMed*, *PsycINFO*, *IEEE Xplore*, *Scopus*, and *Google Scholar* to ensure broad interdisciplinary coverage of health sciences, psychology, and technology literature. Many journals indexed in WoS are also covered by Scopus, PubMed, and PsycINFO, and Google Scholar provided additional breadth by capturing *grey literature* and non-indexed studies. Nonetheless, we acknowledge that Google Scholar's lack of advanced filtering and transparency in indexing may affect the reproducibility and methodological rigor of our search results. While we believe this approach minimized the impact of excluding WoS, we acknowledge that our methodology is not exhaustive.

Second, *the review's study selection and data analysis relied on manual processes*, with duplicate screening conducted using Microsoft Excel. Although a second reviewer sampled the data, most of the extraction was performed by the first author. The use of more *advanced systematic review tools* could have enhanced methodological rigor, improved efficiency, and reduced potential errors in data synthesis.

Third, our search strategy included studies addressing mental health broadly but did not systematically target disorder-specific interventions. As a result, this review may not fully capture the breadth of AI-XR applications tailored to specific conditions, such as *depression*, *anxiety*, *bipolar disorder*, or *PTSD*. Future research should adopt a *more granular approach* by assessing AI-XR interventions for distinct mental health disorders, allowing for a deeper understanding of their disorder-specific effectiveness.

Additionally, *the lack of standardized clinical efficacy metrics* in AI-XR research poses a challenge. The studies reviewed employed diverse assessment methods, making cross-study comparisons difficult. Establishing *uniform evaluation criteria* for AI-XR applications would improve the ability to assess their therapeutic impact reliably.

Furthermore, *longitudinal research on AI-XR interventions is scarce*. Most studies focus on short-term effects, leaving the long-term psychological, behavioral, and neurological implications of these technologies largely unexplored. Future research should prioritize *longitudinal studies* to assess sustained therapeutic benefits, potential risks such as *digital therapeutic dependency*, *cybersickness*, and *changes in patient outcomes over time*.

Finally, while this review highlights ethical concerns associated with AI-XR, empirical studies investigating *real-world ethical, societal, and clinical implications* remain limited. The impact of AI-driven interventions on *therapist-patient relationships*, *algorithmic biases*, and *unintended psychological consequences* warrants further investigation. Despite these limitations, this review provides a *comprehensive overview* of AI-XR applications in mental healthcare, identifying key research gaps and offering actionable recommendations for future studies.

Future Directions

To advance the field of AI-XR interventions in mental healthcare, future research should adopt more targeted and actionable approaches. While this review highlights the potential of AI-XR

applications, several critical areas require further investigation:

1. *Comparative Trials of AI-XR vs. Traditional Telehealth Approaches:*
Future studies should conduct randomized controlled trials (RCTs) comparing AI-driven XR interventions with conventional telehealth and in-person psychotherapy [21, 71]. These trials can assess the relative effectiveness, patient adherence, and clinical outcomes of AI-XR solutions in treating mental health conditions such as depression, anxiety, PTSD, and bipolar disorder.
2. *Disorder-Specific AI-XR Interventions:*
While AI-XR interventions have demonstrated promise across various mental health conditions, future research should *delve deeper into disorder-specific applications* to refine therapeutic approaches. Studies should investigate how AI-XR technologies can be tailored to distinct conditions such as *schizophrenia, OCD, borderline personality disorder (BPD), eating disorders, and substance use disorders*, which are currently underexplored in AI-XR literature. *Comparative studies* should assess the effectiveness of AI-XR interventions *within each disorder category*, ensuring that XR-based treatments align with disorder-specific cognitive, behavioral, and emotional characteristics.
3. *Human-AI Hybrid Therapy Models:*
Research should explore the development and evaluation of hybrid models that integrate AI-driven XR tools with human therapists. This includes investigating how AI can augment, rather than replace, human expertise in psychotherapy by providing real-time insights, personalized treatment recommendations, and automated therapeutic interventions [55]. Key areas of focus should include patient acceptance, therapist-AI collaboration, and ethical considerations.
4. *Scalability and Accessibility in Low-Resource Settings:*
The feasibility and effectiveness of AI-XR interventions in underserved and low-resource environments remain largely unexamined. Future studies should evaluate the adaptability of these technologies in settings with limited mental health infrastructure, internet connectivity, and digital literacy. Research should also examine cost-effectiveness, cultural considerations, and potential strategies to bridge disparities in access to AI-powered mental healthcare.
5. *Longitudinal Studies on Digital Therapeutic Dependency:*
While AI-XR technologies hold promise for enhancing mental health care, prolonged exposure may contribute to digital therapeutic dependency, cybersickness and maladaptive escapism where patients become overly reliant on virtual interventions [30, 45, 46]. Future research should investigate the psychological and behavioral implications of sustained AI-XR use, exploring strategies to mitigate dependency risks while maintaining therapeutic efficacy.
6. *Regulatory and Ethical Frameworks:*
The integration of AI and XR into mental healthcare necessitates robust regulatory guidelines to ensure patient safety, data privacy, and ethical AI use. Future studies should analyze how existing regulations (e.g., GDPR, HIPAA, EU Medical Device Regulation (EU-MDR 2017/745)) apply to AI-XR applications and propose frameworks to address emerging ethical concerns, such as algorithmic bias, informed consent in virtual environments, and clinician accountability [23, 62, 63].
7. *Personalization and Adaptive AI in Mental Healthcare:*
Future work should also explore how adaptive AI models can tailor XR interventions to individual patient needs, dynamically adjusting content based on real-time physiological and behavioural data [48]. This includes investigating machine-driven personalization

strategies that enhance user engagement, optimize treatment efficacy, and prevent adverse effects.

By addressing these key areas, future research can contribute to the responsible development and implementation of AI-XR interventions, ensuring their efficacy, accessibility, and ethical alignment in mental healthcare.

Conclusions

This review is the first to comprehensively evaluate *AI and XR applications in metaverse-driven mental healthcare*, focusing on their *therapeutic potential, ethical concerns, and societal impacts*. Following *PRISMA guidelines*, we analysed *48 peer-reviewed studies published post-2014*, selecting those that addressed clinical applications, ethical issues, and social implications of AI-XR, while excluding non-research articles and studies lacking relevance or accessibility.

Organized around three core research questions—*Therapeutic Potential (RQ1)*, *Ethical Challenges (RQ2)*, and *Societal Implications (RQ3)*—this synthesis highlights the *transformative potential* of AI-XR in mental health treatment and the *governance frameworks* needed for responsible deployment. For *RQ1*, our findings suggest that *AI-XR applications such as virtual reality therapy and AI-driven diagnostic tools enhance accessibility and patient engagement, particularly in resource-limited settings*. However, *large-scale, longitudinal studies are needed to validate their effectiveness across diverse mental health conditions and populations* [85].

Regarding *RQ2*, ethical concerns such as *privacy risks, data security, and algorithmic bias* emerged as key challenges. The *handling of sensitive biosignal data* raises questions of *data ownership and misuse*, while *bias in AI models* could result in *inequitable treatment for marginalized groups*. These findings underscore the need for *robust governance frameworks* that prioritize *transparency, informed consent, and privacy protections* tailored to immersive mental health solutions [86]. For *RQ3*, we identified significant *societal implications, including changes in patient-therapist dynamics, risks of digital exclusion, and psychological dependency* on virtual environments. These insights highlight the necessity of *inclusive regulatory policies to mitigate the long-term social and relational risks posed by AI-XR therapy*.

Call to Action: Bridging Innovation and Responsibility

To ensure *ethical, effective, and equitable adoption* of AI-XR in mental healthcare, targeted actions are needed from key stakeholders:

- *For Policymakers:* Establish *AI-XR-specific regulations*, aligning with existing frameworks (e.g., *GDPR, HIPAA, MDR*). Ensure *data security, bias mitigation, and patient safety* while supporting *standardized ethical guidelines for AI in mental health*.
- *For Clinicians and Mental Health Professionals:* Embrace *AI-XR tools as adjunctive rather than replacement therapies*. Push for *clinical validation, patient safety protocols, and AI literacy training* in medical and psychological practice.
- *For Developers and Researchers:* Design *transparent, bias-aware AI models* while prioritizing *privacy, security, and ethical AI governance*. Conduct *comparative trials* between AI-XR and traditional mental health interventions.

By collaborating across disciplines, we can harness AI-XR's potential while ensuring its safe, ethical, and inclusive integration into global mental healthcare systems.

Author Contributions

Aliya Tabassum: Conceptualization, Methodology, Literature Review, Data Curation, Formal Analysis, Resources, Validation, Writing – Original Draft.

Ibrahim Ghaznawi: Formal Analysis, Resources, Supervision, Validation, Writing – Review & Editing.

Alaa Abd Alrazaq: Formal Analysis, Resources, Supervision, Validation, Writing – Review & Editing.

Junaid Qadir: Conceptualization, Resources, Visualization, Supervision, Validation, Formal Analysis, Writing – Review & Editing.

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 8: 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 al. (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	Saredakis et al. (2023) [95]	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-enabled 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. Tacgin 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, Adebajo 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).

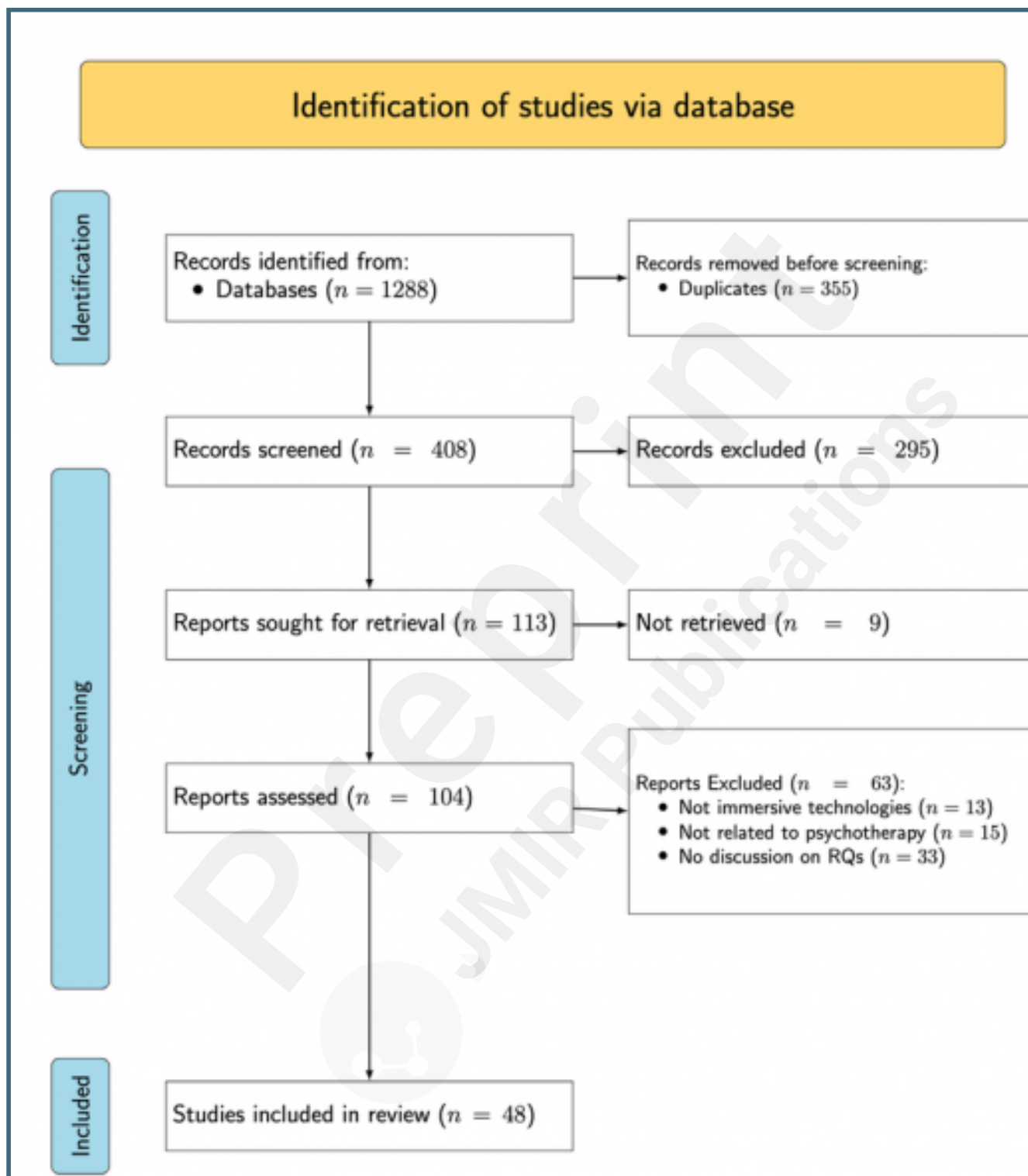
74. Galatzer-Levy IR, Aranovich GJ, Insel TR. Can mental health care become more human by becoming more digital? *Dædalus* 152(4):228–244 (2023).
75. Barak A, Grohol JM. Current and future trends in internet-supported mental health interventions. *Journal of Technology in Human Services* 29(3):155–196 (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. Eryilmaz 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.
92. Guo Z, Lai A, Thygesen JH, Farrington J, Keen T, Li K. Large language model for mental health: A systematic review. *arXiv preprint arXiv:2403.15401*. 2024 Feb 19.
93. Shah SM, Gillani SA, Baig MS, Saleem MA, Siddiqui MH. Advancing Depression Detection on Social Media Platforms Through Fine-Tuned Large Language Models. *arXiv preprint arXiv:2409.14794*. 2024 Sep 23.
94. Jameel L. *Virtual-reality assisted CBT for social difficulties: a feasibility study in early intervention for psychosis services* (Doctoral dissertation, King's College London).
95. Saredakis D, Keage HA, Corlis M, Loetscher T. Virtual reality intervention to improve apathy in residential aged care: protocol for a multisite non-randomised controlled trial. *BMJ open*. 2021 Feb 1;11(2):e046030.

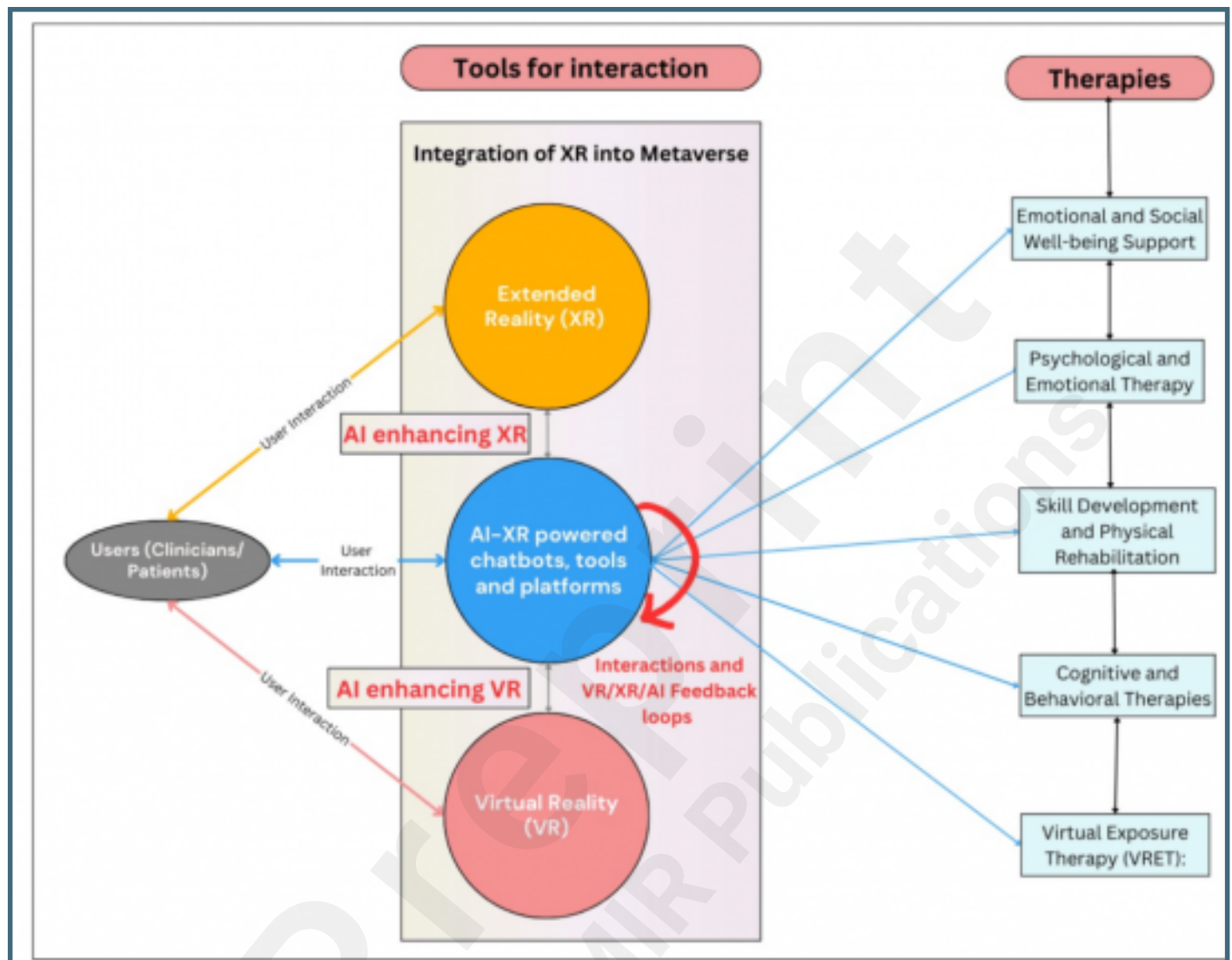
Supplementary Files

Figures

Prisma chart illustrating the process of article selection.



Conceptual diagram showing interactions between users, AI and immersive technologies for mental health therapies.



Multimedia Appendixes

Prisma-scr checklist.

URL: <http://asset.jmir.pub/assets/863ea9c337171cc4918b3d23c37c5bc5.docx>

Search queries and screening steps.

URL: <http://asset.jmir.pub/assets/d587955c37c9048a31dd77d1b2a6756a.docx>

Data extraction form with all details of data extracted from the included studies.

URL: <http://asset.jmir.pub/assets/70ceb1e059b75ca5fc2b30912c84d9db.docx>



TOC/Feature image for homepages

Diagram of interrelated ethical, social, and psychological challenges in AI-XR based psychotherapy treatments.

