Ethical Dilemmas, Mental Health, Artificial Intelligence, and LLM-Based Chatbots

Johana Cabrera(✉), M. Soledad Loyola, Irene Magaña, and Rodrigo Rojas

University of Santiago de Chile, Santiago, Chile
{johana.cabrera,maria.loyola.f,irene.magana,
rodrigo.rojas.a}@usach.cl

Abstract. The present study analyzes the bioethical dilemmas related to the use of chatbots in the field of mental health. A rapid review of scientific literature and media news was conducted, followed by systematization and analysis of the collected information. A total of 24 moral dilemmas were identified, cutting across the four bioethical principles and responding to the context and populations that create, use, and regulate them. Dilemmas were classified according to specific populations and their functions in mental health. In conclusion, bioethical dilemmas in mental health can be categorized into four areas: quality of care, access and exclusion, responsibility and human supervision, and regulations and policies for LLM-based chatbot use. It is recommended that chatbots be developed specifically for mental health purposes, with tasks complementary to the therapeutic care provided by human professionals, and that their implementation be properly regulated and has a strong ethical framework in the field at a national and international level.

1 Introduction

Over time, many tasks once considered human have been gradually replaced by machines. This phenomenon has been widely observed since the Industrial Revolution [1]. Currently, we are said to be experiencing the Fourth Industrial Revolution, characterized by a series of technological, digital, physical, and even biological convergences, including the development of the internet, automation, robotics, blockchain, cloud computing, 3D printing, and artificial intelligence [2]. The technology that, in the last period, has caused great controversy is Artificial Intelligence (AI), which was defined by John McCarthy as "the combination of science and engineering to create intelligent devices for human welfare" [3].

The current debate is generated by certain types of AI chatbots based on large-scale language models (LLM) [4]. Among these, ChatGPT-3 and ChatGPT-4 [5] have had the greatest impact thus far because of their large capacity for both function and interaction with humans. Such is the evolution of this technology that Bill Gates has called "the most important technological breakthrough since the graphical user interface" [6].

It is important to consider that this scenario takes place in a delicate context for humanity; in recent years, the world has experienced a series of economic, health, and sociopolitical problems at the national and international levels, which will also have

I. Rojas et al. (Eds.): IWBBIO 2023, LNBI 13920, pp. 313–326, 2023.
https://doi.org/10.1007/978-3-031-34960-7_22

a reciprocal impact, affecting people's health and mental health as well as driving the progress of scientific-technological devices [2, 7, 8].

As artificial intelligence becomes more prominent in society, it is crucial to address ethical concerns and enable scientific studies on its development, regulation, and implementation [6, 9, 10].

A topic of relevance to the ethical field can be seen from the perspective of the "ethical dilemma" or "moral dilemma," which are situations in which decisions must be made that can have implications ranging from moral to paradoxical [11–13]. These situations can be challenging because they often involve competing values or interests [14].

These ethical problems and dilemmas are often addressed in the field of "AI ethics," which seeks to establish reflections and guidelines for the development and implementation of AI technology [15].

However, what specific considerations should these analyses include from the fields in which these technologies are applied?

In this sense, international organizations, and donors such as Bill Gates [6, 16] have announced plans for two areas of development and LLM chatbots: health and education in a global scenario. This implies an investment in digital infrastructure to implement both health and education of advanced human capital and improve the population's access to these types of tools, which would constitute part of the main actions to be developed in these areas.

One particularly sensitive field that requires special attention when addressing ethical concerns regarding AI applications is Mental Health. Mental health has been defined as "The ability of individuals to interact with each other and the environment, to promote subjective well-being, the development and optimal use of their psychological, cognitive, emotional, and relational potential, the achievement of their individual and collective goals, in accordance with justice and the common good [32]." In this field, professionals and researchers from various disciplines seek the well-being of people from different perspectives.

In this sense, AI can improve various counseling axes, professional functions, patient care, help with diagnostic accuracy, optimize treatment plans, create intervention design, monitor treatment, optimization, accompaniment, interaction, referral decisions, development of support material, and different potentials for research. For example, it has been shown that these tools achieve a 19% increase in conversational empathy when people use them to generate text [17–19].

In this regard, ethical dilemmas related to mental health and chatbots could include issues surrounding privacy, transparency, accountability, conflicts of interest, cultural differences and biases, involuntary interventions, and balancing patient autonomy with their safety, informed consent, and problems related to power dynamics and inequalities [20–25]. Others may be related to confidentiality, such as deciding whether to violate a patient's privacy to prevent harm to themselves or to others [26, 27].

Some experts argue that, in the meantime, these chatbots will continue to impact human professions, spreading the biases of their creators, advancing persuasive interaction in political and marketing fields, and directly impacting people's thoughts, emotions, and behaviors [8, 28].

Due to the above, immediate actions are required, such as ethical reflections and scientific development, to shed light on developments, uses, and regulations surrounding what is cataloged as "the new era of Artificial Intelligence" [6], which we know will continue to generate changes in our society.

Identifying and addressing ethical dilemmas are related to decision-making and, if done correctly, can contribute to strengthening moral acts and promoting the development of positive behaviors [29].

The objective of this article is to identify and reflect on the bioethical dilemmas (beneficence, non-maleficence, justice, and autonomy) involved in the use of chatbots based on large-scale language models (LLM) in professional practice related to mental health [30]. The results seek to provide relevant information and recommendations that serve as interdisciplinary inputs for the regulation, use, research, and training of these bots in mental health.

To achieve these objectives, it is necessary to conduct interdisciplinary research involving ethics experts, mental health professionals, AI developers, and policymakers. These collaborative efforts can help ensure that ethical concerns are adequately addressed, and effective guidelines for the responsible use of LLM chatbots in the field of mental health are developed.

2 Methodology

To achieve the main objective of this research, a rapid review [31] of scientific literature and news articles from the media was carried out. Subsequently, a systematization and analysis of the collected information were performed.

The type of evidence included in this document considers more than just scientific evidence, as science and technology do not advance at the same pace. Media outlets could offer more recent information and experience; therefore, a review of news that responds to different contexts with multidisciplinary and/or multisectoral approaches is also included.

Search and selection strategy - Scientific literature.

For the literature review, the academic search engine "Web of Science" was used to identify relevant abstracts. The keyword used to refine the search was: "chatbot mental health."

The abstracts were evaluated according to the following inclusion and exclusion criteria:

Inclusion criteria:

- Scientific articles investigating the relationship between chatbots and mental health.
- Articles published up to April 27, 2023.
- Articles focused on any type of chatbot.

Exclusion criteria:

- Documents published before 2020.

After conducting the search on the "Web of Science" engine, a careful review of the results was performed to review the abstracts and apply the inclusion and exclusion criteria. Finally, 33 abstracts were selected for a more detailed analysis.

Table 1 provides a complete list of the articles included in this document, along with relevant information related to each source.

Search and selection strategy for media news.

Using the same rapid review technique, a search was carried out in the media, through the "news segments" using the Microsoft Bing/Edge search engine, using the keyword "chatbots mental health."

This type of search engine was chosen because it incorporates ChatGPT into its new technology.

The inclusion and exclusion criteria were established to evaluate the retrieved articles as follows:

Inclusion criteria:

- News articles and blog entries published between November 2022 (public release milestone) and April 25, 2023
- Focus on any LLM chatbot.
- Inclusion of mental health in the text.
- News articles in English.

Exclusion criteria:

- News published before 2022.
- Websites not directly linked to a specific news entry.
- Documents without free access to the text.
- News that did not provide information about chatbots or mental health.
- Articles without information or mention of LLM-based chatbots.

After the initial search, 140 news articles were retrieved. Following the previous step, the titles and publication dates of each entry were reviewed. In cases where the title was unclear, a full news article was read to determine its relevance. Based on this evaluation, 13 articles were selected for inclusion.

Table 2 provides a complete list of the news articles included in this document, along with relevant information related to each source. These tables facilitate transparency and reproducibility in news search and selection processes.

Table 1. Listing of scientific abstracts included in the review.

Article Title	Source Title	DOI
Perceptions and Opinions of Patients About Mental Health Chatbots: Scoping Review	Journal Of Medical Internet Research	10.2196/17828
Designing Personality-Adaptive Conversational Agents for Mental Health Care	Information Systems Frontiers	10.1007/s10796-022-10254-9
Development of a Positive Body Image Chatbot (KIT) With Young People and Parents/Carers: Qualitative Focus Group Study	Journal Of Medical Internet Research	10.2196/27807
Development process of artificial intelligence based chatbot to support and promote mental wellbeing in sparsely populated areas of five European countries	European Psychiatry	10.1192/j.eurpsy.2022.446
Artificially intelligent chatbots in digital mental health interventions: a review	Expert Review Of Medical Devices	10.1080/17434440.2021.2013200
Identifying Potential Gamification Elements for A New Chatbot for Families with Neurodevelopmental Disorders: User-Centered Design Approach	Jmir Human Factors	10.2196/31991
The Challenges in Designing a Prevention Chatbot for Eating Disorders: Observational Study	Jmir Formative Research	10.2196/28003
A Chatbot for Perinatal Women's and Partners' Obstetric and Mental Health Care: Development and Usability Evaluation Study	Jmir Medical Informatics	10.2196/18607
Implementation of Cognitive Behavioral Therapy in e-Mental Health Apps: Literature Review	Journal Of Medical Internet Research	10.2196/27791
A Mental Health Chatbot for Regulating Emotions (SERMO)-Concept and Usability Test	Ieee Transactions On Emerging Topics In Computing	10.1109/TETC.2020.2974478
Development of a chatbot for depression: adolescent perceptions and recommendations	Child And Adolescent Mental Health	10.1111/camh.12627
AI-based chatbot micro-intervention for parents: Meaningful engagement, learning, and efficacy	Frontiers In Psychiatry	10.3389/fpsyt.2023.1080770
Engagement and Effectiveness of a Healthy-Coping Intervention via Chatbot for University Students During the COVID-19 Pandemic: Mixed Methods Proof-of-Concept Study	Jmir Mhealth And Uhealth	10.2196/27965
Co-developing a Mental Health and Wellbeing Chatbot with and for Young People	Frontiers In Psychiatry	10.3389/fpsyt.2020.606041
Chatbot-Based Assessment of Employees' Mental Health: Design Process and Pilot Implementation	Jmir Formative Research	10.2196/21678
Chatbots to Support Young Adults' Mental Health: An Exploratory Study of Acceptability	Acm Transactions On Interactive Intelligent Systems	10.1145/3485874
A Test Platform for Managing School Stress Using a Virtual Reality Group Chatbot Counseling System	Applied Sciences-Basel	10.3390/app11199071
A Chatbot to Support Young People During the COVID-19 Pandemic in New Zealand: Evaluation of the Real-World Rollout of an Open Trial	Journal Of Medical Internet Research	10.2196/38743
Participatory Development and Pilot Testing of an Adolescent Health Promotion Chatbot	Frontiers In Public Health	10.3389/fpubh.2021.724779
Emotional Reactions and Likelihood of Response to Questions Designed for a Mental Health Chatbot Among Adolescents: Experimental Study	Jmir Human Factors	10.2196/24343
Improving body image at scale among Brazilian adolescents: study protocol for the co-creation and randomized trial evaluation of a chatbot intervention	Bmc Public Health	10.1186/s12889-021-12129-1
Combined Use of Virtual Reality and a Chatbot Reduces Emotional Stress More Than Using Them Separately	Journal Of Universal Computer Science	10.3897/jucs.77237
Assisting Personalized Healthcare of Elderly People: Developing a Rule-Based Virtual Caregiver System Using Mobile Chatbot	Sensors	10.3390/s22103829
Developing, Implementing, and Evaluating an Artificial Intelligence-Guided-Mental Health Resource Navigation Chatbot for Health Care Workers and Their Families During and Following	Jmir Research Protocols	10.2196/33717

Table 2. Listing of the news included in the present review.

Title	Media Communication	Web
ChatGPT Therapy: Why People Ask AI For Mental Health Advice	Inquirer	https://technology.inquirer.net/123143/chatgpt-therapy-why-people-ask-ai-for-mental-health-advice
ChatGPT and health care: Could the AI chatbot change the patient experience?	Fox News	https://www.foxnews.com/health/chatgpt-health-care-could-ai-chatbot-change-patient-experience
Artificial empathy: the dark side of AI chatbot therapy	Cyber News	Artificial empathy: the dark side of AI chatbot therapy | Cybernews
AI bots have been acing medical school exams, but should they become your doctor?	ZDNET	AI bots have been acing medical school exams, but should they become your doctor? | ZDNET
ChatGPT: From Healthcare to Banking, These Sectors Can Benefit Most From AI Chatbots	News 18	https://www.msn.com/en-in/money/news/chatgpt-from-healthcare-to-banking-these-sectors-can-benefit-most-from-ai-chatbots/ar-AA19T4S0
Could Bing Chat replace my Doctor?	Windows Central	https://www.msn.com/en-us/health/other/could-bing-chat-replace-my-doctor/ar-AA1a9WR6
3 ways AI will stand out in healthcare delivery in 2023	Health Care Asia	https://www.msn.com/en-xl/news/other/3-ways-ai-will-stand-out-in-healthcare-delivery-in-2023/ar-AA19NSJX
Conversing with AI: Revolutionizing Mental Health Support through Chatbots	Linkedin	https://www.linkedin.com/pulse/conversing-ai-revolutionizing-mental-health-support-through-pandey/
World Health Day 2023: How AI and ChatGPT Are Revolutionizing Telemedicine and Remote Patient Care	Tecnopedia	World Health Day 2023: How AI and ChatGPT Are Revolutionizing Telemedicine and Remote Patient Care - Techopedia
AI Therapy Is Here. But the Oversight Isn't. (msn.com)	MSN News	https://www.msn.com/en-us/health/other/ai-therapy-is-here-but-the-oversight-isn-t/ar-AA19AJTa
Exploring the Limits of Artificial Intelligence in Mental Health: ChatGPT as a Therapist for a Day (thequint.com)	The Quint	Exploring the Limits of Artificial Intelligence in Mental Health: ChatGPT as a Therapist for a Day (thequint.com)
Conversing with AI: Revolutionizing Mental Health Support through Chatbots	The Strait times	https://www.straitstimes.com/world/united-states/people-are-using-ai-for-therapy-even-though-chatgpt-wasn-t-built-for-it

3 Data Extraction and Analysis Procedure

After selecting the articles and news based on the inclusion and exclusion criteria, manual coding was performed based on the researchers' inference criteria to identify the bioethical dilemmas present in each of the abstracts and selected media texts.

Following the rapid review and using the GPT-4 chatbot, studied in this work, the following prompt was requested from the tool: "Sort, classify, analyze, and explain ethical dilemmas from the bioethical principle of (selection of the ethical principle) based on the following information (list of codes created either from abstracts or news texts)."

This prompt was used for each of the bioethical principles mentioned previously.

4 Results

After using the prompt with Chatgpt 4, the following results were generated for each of the bioethical principles:

- Identification and classification of bioethical dilemmas
- Ethical analysis

To present the key findings, Table 3 shows the results of the sorting and identification of bioethical dilemmas in the review of scientific abstracts, whereas Table 4 provides an overview of the sorting and identification of bioethical dilemmas present in media news, both categorized according to the four bioethical principles.

Table 3. Identification of ethical dilemmas from the primary four bioethical principles present in scientific literature.

Autonomy	Beneficence	No Maleficence	Justice
Limited access to mental health care: Insufficient access to mental health services can raise an ethical dilemma in relation to the principle of autonomy, as patients may not have access to necessary medical care to make informed decisions about their treatment and care. Evaluation of conversational agents: The lack of standard measures for evaluating conversational agents can raise an ethical dilemma in relation to the principle of autonomy, as patients may not have access to accurate and reliable information about the benefits and risks of these agents. Vulnerable populations: The pediatric population and those with schizophrenic or bipolar disorders can raise ethical dilemmas in relation to the principle of autonomy, as these patients may have difficulty making informed decisions about their treatment and care due to their age or medical condition. Informed decision-making: The ethical dilemma of the users' ability to make informed decisions about their mental health when interacting with LLM-based chatbots, and whether these programs are designed to provide accurate and complete information. Equity in access: The ethical dilemma of equity in access to chatbots for mental health care, as not all users have access to the necessary technology to use them. Privacy and confidentiality: The ethical dilemma of privacy and confidentiality of user information when interacting with mental health chatbots, as these	The lack of access to mental health services and online medical care can be an ethical dilemma, as individuals are being deprived of receiving the treatment, they need to improve their well-being. The use of conversational agents and chatbots to address the shortage of healthcare providers can raise ethical questions about the quality and effectiveness of treatment, as it is unknown whether these agents can provide the same level of care as a healthcare professional. The inability of conversational agents and chatbots to capture dynamic human behavior can raise ethical questions about the privacy and confidentiality of users, as it is unknown whether the information shared with these agents will be used appropriately and securely protected. The need to design adaptive conversational agents and chatbots to personality can raise ethical questions about manipulation and influence on users' decision-making, as it is unknown whether these agents can influence users' decisions in an unethical way. The ethical dilemma of using mental health chatbots as a quick and easy solution instead of addressing the underlying causes of mental health problems, which may perpetuate the lack of access to adequate and personalized mental health care. The need to ensure the privacy and confidentiality of users when interacting with mental health chatbots, which can be a technical and legal challenge.	The ethical dilemma associated with non-maleficence in the use of chatbots in mental health care can cause harm to patients if not properly addressed. The ethical dilemma of using experimental therapies in terminally ill patients, as there may be unknown and potentially harmful risks to the patient. The ethical dilemma of using assisted reproductive techniques in couples with fertility problems, as there may be risks to the health of the mother and fetus, as well as ethical issues related to embryo selection and genetic manipulation. The ethical dilemma of using gene therapies to treat genetic diseases, as there may be unknown and potentially harmful risks to the patient, as well as ethical issues related to genetic manipulation and the selection of physical and mental characteristics. Confidentiality: The use of a chatbot to monitor a person's health can raise concerns about the confidentiality of the information collected, as it may be shared with third parties without the person's consent. Bias: The chatbot may be programmed with unconscious biases that can affect the accuracy of the results and the healthcare provided to the person. Stigma: The use of a chatbot to monitor mental health can raise concerns about the stigma associated with mental health and the perception that the person needs constant monitoring. Integrity: It is important to ensure that the chatbot is used ethically and not used for malicious purposes, such as discrimination or manipulation of the collected information. The need to ensure that chatbots are programmed with accurate and up-to-date information about mental health, and that they do not perpetuate stereotypes or	The need to ensure equity in access to mental health chatbots for all patients. Possible discrimination and bias in the programming of mental health chatbots, which can affect the quality of care and treatment for different groups of people. The need to ensure that mental health chatbots are culturally appropriate and sensitive to the needs of different groups of people. Lack of access to mental health services for certain groups of people due to a lack of access to the technology needed to use mental health chatbots. Possible lack of privacy and security of user data when interacting with mental health chatbots, which can compromise confidentiality and trust in the healthcare system. Lack of regulation and supervision of mental health chatbots, which can lead to the proliferation of ineffective or even harmful chatbots for users' mental health. Lack of equitable access to the technology needed to use mental health chatbots and health games in healthcare. Possible discrimination in the selection of patients who can use mental health chatbots to receive care. Lack of human contact and empathy in mental health care through chatbots, which can negatively affect the quality of care and the therapeutic relationship. Possible exploitation of participants using mental health chatbots for research and chatbot development without their adequate informed consent. Possible lack of transparency

Table 4. Identification of ethical dilemmas from the primary four bioethical principles present in communication news media.

Autonomy	Beneficence	No Maleficence	Justice
Autonomy in choice of help: Concerns about the quality of information provided by chatbots compared to human experts and whether people have enough information to make autonomous decisions about their mental health care.	Accessibility and speed of access: Concerns about the potential for chatbots to improve accessibility and provide more immediate access to mental health care for those in need.	Quality and accuracy of information: Risks of inadequate, outdated, or biased information provided by chatbots.	Ethical dilemma of quality and accuracy in chatbots: Risk of low-quality, outdated, or biased information affecting users.
Autonomy in emotional interaction: Concerns about social isolation and technological dependence that may result from sharing emotions with chatbots instead of human professionals.	Impact on the doctor-patient relationship: Concerns about the quality and strength of the therapeutic relationship between the patient and the chatbot, and whether this may negatively affect the effectiveness of mental health care.	Human vs. machine interaction: Questions about the ability of chatbots to empathize and genuinely understand complex emotional situations.	Ethical dilemma of human vs. machine: Concerns about the empathy and understanding of chatbots in complex situations.
Autonomy in seeking information: Issues around the quality and accuracy of information provided by chatbots and people's ability to make informed autonomous decisions.	Bias and discrimination: Risks of bias and discrimination in the programming and use of chatbots in mental health care, which may affect the quality of care and treatment of different groups of people.	Cost and accessibility: Possible dehumanization and excessive dependence on technology instead of seeking human professional help.	Ethical dilemma of cost and accessibility: Dehumanization of care and excessive dependence on technology in mental health.
Autonomy in interaction with chatbots and replacement of human professionals: Concerns about whether chatbots can provide the same empathy, understanding, and personalized attention as human professionals, and how this affects the valuation of mental health care.	Capacity for personalized care: Limitations on the ability of chatbots to provide personalized care and tailor treatment to the individual needs of patients.	Substitution of human professionals: Impact on the quality of care and therapeutic relationship between patients and professionals.	Ethical dilemma of substitution of professionals: Impact on the quality of care and therapeutic relationship.
Autonomy and privacy of information: Concerns about the privacy and security of data shared with chatbots, and whether people are making informed decisions about privacy and use of their data.	Technological limitations: Limitations on the ability of chatbots to handle crisis or emergency situations, which may jeopardize patient safety and well-being.	Privacy and data security: Concerns about protection and proper handling of sensitive information and the risk of misuse.	Ethical dilemma of privacy and data security: Adequate protection of sensitive information and risk of misuse.
Autonomy and responsibility in the use of chatbots: Responsibility of developers, healthcare providers, and users in case of errors or harm caused by information provided by chatbots, and whether people are adequately evaluating the quality and reliability of the information provided.	Supervision and monitoring: Need for adequate supervision and monitoring to ensure that patients are receiving appropriate care and that any ethical or safety issues arising from the use of chatbots in mental health care are addressed.	Responsibility and supervision: Challenges in establishing responsibilities and ensuring that chatbots follow ethical guidelines and provide high-quality support.	Ethical dilemma of responsibility and supervision: Establishing responsibilities and ensuring quality of information and support.
Autonomy and equitable access to mental health care: Inequalities in access to chatbots, especially for marginalized or disadvantaged populations, and whether the use of chatbots promotes autonomy and equitable access to care.	Accessibility: Limitations on the accessibility of mental health chatbots for people with visual or hearing impairments, which may negatively impact their ability to receive quality care.	Digital divide: Risk of exacerbating inequalities and excluding populations without access to digital resources.	Ethical dilemma of digital divide: Exclusion of marginalized populations without access to digital technologies.
Autonomy and reliability of chatbots: Issues around the quality and reliability of information provided by	Inequities in access: Inequalities in access to mental health chatbots due to factors such as lack of access to necessary technology or digital literacy.	Excessive dependence on technology: Negative impact on long-term mental health and emotional well-being.	Ethical dilemma of technological dependence: Negative impact on mental health and long-term emotional well-being.
	Ethics and responsibility of developers: Ethical and legal responsibility of mental	Research and validation: Need for rigorous research and validation of the efficacy and safety of chatbots in mental health.	Ethical dilemma of research and validation: Need for rigorous studies to evaluate the effectiveness and safety of chatbots.
		Regulation and policies: Importance of policies and regulations to ensure quality, safety, and user data protection.	Ethical dilemma of regulation and policies: Ensuring quality, safety, and data protection, preventing inadequate liability and protection.
		Balance between innovation and caution: Adapting the promotion of innovation and caution to address ethical concerns and protect user	Ethical dilemma of balancing innovation and caution: Balance between fostering innovation and protecting the well-being of users.
			Ethical dilemma of education and public awareness:

5 Ethical Analysis

With the results obtained, we proceeded to unify the content and conduct an ethical analysis of the review results at both the scientific and news levels. This process was carried out by ChatGPT 4 and the research group.

In this sense, a total of 24 moral dilemmas were identified, most of which are transversal to the four bioethical principles. Similarly, these bioethical dilemmas respond to the context and populations that create, use, and regulate them. To make sense of the collected information, bioethical dilemmas have been classified according to specific populations and their functions in mental health. These are as follows:

- Mental health technology developers
- Mental health beneficiaries of mental health services
- Mental health professionals
- Mental health researchers and developers
- Mental health regulators

Table 5 presents the synthesis of these results and the main bioethical dilemmas classified according to the user population.

Table 5. Main bioethical dilemmas classified according to the user population.

Mental health technology developers	
Chatbots as conversational agents and the quality of care:	
1	Uncertainty about whether conversational agents can provide the same level of attention as healthcare professionals.
2	Designing adaptive and personalized chatbots raises ethical dilemmas about manipulation and influences user decision-making.
3	Possibility of unconscious biases in chatbot programming, affecting the accuracy of results and the medical care provided.
4	The risk that chatbots are not equipped to handle crises or emergency situations endangers the safety and well-being of the user.
5	Need to ensure that chatbots are culturally appropriate and sensitive to the needs of different groups of people.
6	How can the use of chatbots complement and improve mental health care provided by human professionals, rather than replacing them?
7	Privacy and confidentiality of user information when interacting with mental health chatbots.
Mental health professionals	
Human responsibility and supervision:	
8	Proper accountability and oversight in the use of chatbots in mental health care. Challenges in establishing accountability and ensuring chatbots follow ethical guidelines and offer high-quality support.
9	Proper accountability and oversight in the use of chatbots in mental health care. Challenges in establishing accountability and ensuring chatbots follow ethical guidelines and offer high-quality support.
Mental health beneficiaries of mental health services	
Access and exclusión:	
10	Not all users may have access to the necessary technology to use Chat LLM.
11	Gaps in vulnerable populations such as the pediatric population, those with schizophrenic or bipolar disorders, or with intellectual or physical disabilities.
12	Limited access to medical care and accurate, reliable information on the benefits and risks of chatbots.
13	Users' ability to make informed decisions about their mental health when interacting with LLM chatbots.
Interacción emocional y dependencia tecnológica	
15	Limitations of chatbots in addressing complex mental health issues that require empathetic understanding.
16	Negative impact on long-term mental health and emotional well-being due to excessive dependence on technology.
Mental health researchers and developers:	
17	Need for solid and ethical research to determine safety, effectiveness, and potential risks of chatbots.
18	Ensuring that chatbots are programmed with accurate and up-to-date mental health information and do not perpetuate stereotypes or biases.
19	Risks of inadequate, outdated, or biased information provided by chatbots due to their design.
20	Possible exploitation of participants using mental health chatbots for research and chatbot development without their proper informed consent.
Mental health regulators	
21	Importance of policies and regulations to ensure quality, safety, and data protection for users.
22	Balancing innovation and caution: Adjusting the promotion of innovation and caution to address ethical concerns and protect user well-being.
23	How the use of chatbots can complement and improve mental health care provided by human professionals, rather than replacing them.
24	Education and public awareness: The need to inform about the benefits and risks of chatbots in mental health, and understand their role as a complementary resource.

6 Conclusions

Considering the results and analysis carried out, it can be interpreted that there are various bioethical dilemmas present in the relationship between the use of chatbots in mental health, which are mostly transversal and changing, depending on the creating, benefiting, and regulating population, as well as the context in which they are developed.

In conclusion, there are 4 major areas in which bioethical dilemmas are categorized, requiring further reflection and analysis for the development of the creation and use of LLM chatbots in the field of mental health:

7 LLM Chatbots, Quality of Care, Responsible Research, and Development in Mental Health

To improve the quality of care, research, and technological development in mental health provided by chatbots, it is important to consider the different actors involved in this process, such as creators, technologists, professionals, users, researchers, and decision-makers in mental health.

To reduce risks and increase benefits, it is necessary to incorporate clarity in both the scope and risks of the technology itself, as well as the inclusion of various theories and practical knowledge of mental health with which chatbots are trained.

Therefore, it is recommended to create chatbots specifically designed to address mental health issues, with purposes and tasks that do not replace therapeutic care provided by human professionals. For example, technology could be used to provide companionship and emotional support, but therapeutic care should be centered or guided by human professionals, who in turn must have clear protocols to proceed in cases of emergencies detected or created by chatbots in mental health settings. This is because LLM chatbots can complement and improve the mental health care provided by human professionals rather than replace them. Therefore, it is essential to establish appropriate regulatory and ethical frameworks that allow various users to have autonomy, justice, and well-being.

8 Access, Exclusion, and User Dependence on Chatbots

In this regard, it is important to expand access to vulnerable populations and reduce elitism in mental healthcare, taking into account gender, social class, age, and other biases. Public and private institutions should focus their efforts with special emphasis on identifying and supporting marginalized and/or displaced populations to facilitate and educate the population about the use of chatbots in mental health topics.

It is also important to consider the interaction and dependence that these types of tools can induce in users, either in the public seeking care or professionals using chatbots as auxiliary tools. It is necessary to anticipate problems related to the preference that some people may have for chatbots, which can be a risk factor for social isolation, contributing to greater mental health problems and associated social issues.

9 Responsibility and Human Supervision of Chatbots

As it is already known, chatbots are imperfect tools, which can be misused for manipulation or influence based on misinformation. In this sense, the research group proposed to enhance critical thinking and verify the results provided by AI to ensure effectiveness and safety in the care provided to the public.

Likewise, it is important to consider the need for continuous supervision of AI in mental healthcare. Thus, it is necessary to question the future role of professionals and researchers in the field of mental health and LLM chatbots. Some of these roles could incorporate "the supervision of AI, both from the technological and mental health fields," who should be highly trained to detect and address problems in the services provided by these chatbots.

There are still questions about the responsibility that each actor must have in this process to ensure that AI is used effectively and ethically in mental healthcare.

10 Regulation and Chatbot Usage Policies

To achieve greater benefits in mental health, it is essential to establish clear regulations at the institutional and governmental levels in national and international settings.

In this sense, it is important to define who will be responsible for regulating and solving problems that may arise in the mental health care provided by chatbots. Furthermore, it is necessary to consider the liability for damage caused by chatbots and the biopolitical control that can be exerted through them.

Public and private institutions responsible for mental health care and those areas of political decision-making should promote the idea of human supervisors, who must ultimately be responsible for ensuring the effectiveness and safety of the care provided.

In conclusion, it is necessary to address the issue of mental health care through chatbots from a broad and conscious perspective, considering the different actors involved, the risks and benefits associated with their use, and the importance of establishing appropriate regulatory and ethical frameworks. All this, with the aim of improving the quality of mental health care and ensuring the widest and most equitable possible access, provides safe, effective, and inclusive care for all.

11 Limitations of the Present Review

Several limitations were identified, which could have affected the results and conclusions. The limitations of this study are as follows.

Dependence on prompts: LLM-based chatbots operate using prompts provided by users, meaning that the quality and relevance of the generated responses are directly related to the accuracy and clarity of the input prompts. This dependence can generate inaccurate or incomplete responses if prompts are inadequate.

User influence: Because LLM-based chatbots learn from interactions with their users, it is likely that the information provided will be influenced by the opinions, knowledge, and biases of these users.

Technical issues: On some occasions, the chatbot did not function properly, forcing researchers to use other versions of the LLM chatbots. These technical issues may have affected the efficiency and effectiveness of the research process and the quality of the data collected.

Peer control: Following the rapid review technique, it is reported that the inferences were not passed through controls by other peers, which could have influenced the results of the present investigation.

References

1. Sutz, J.: Engenharia e preocupação social: rumo a novas práticas, vol. 14 (2019)
2. CEPAL: Repercusiones en América Latina y el Caribe de la guerra en Ucrania: ¿cómo enfrentar esta nueva crisis? (2022)
3. Rupali, M., Amit, P.: A review paper on general concepts of 'artificial intelligence and machine learning.' Int. Adv. Res. J. Sci. Eng. Technol. **4**(4), 79–82 (2017). https://doi.org/10.17148/IARJSET/NCIARCSE.2017.22
4. Eloundou, T., Manning, S., Mishkin, P., Rock, D.: GPTs are GPTs: an early look at the labor market impact potential of large language models. arXiv, Mar. 21, 2023. Accessed: Mar. 25, 2023. http://arxiv.org/abs/2303.10130
5. Open AI: GPT-4 (2023). https://openai.com/research/gpt-4
6. Gates, B.: A new era, the age of AI has begun. Gates Notes (2023). https://www.gatesnotes.com/The-Age-of-AI-Has-Begun
7. Sadasivan, V.S., Kumar, A., Balasubramanian, S., Wang, W., Feizi, S.: Can AI- generated text be reliably detected? arXiv Mar. 17, 2023. Accessed: Mar. 25, 2023. http://arxiv.org/abs/2303.11156
8. Burtell, M., Woodside, T.: Artificial influence: an analysis of AI-driven persuasion. arXiv, Mar. 15, 2023. Accessed: Mar. 25, 2023. http://arxiv.org/abs/2303.08721
9. Sandu, I., Wiersma, M., Manichand, D.: Time to audit your AI algorithms. Maandblad voor Accountancy en Bedrijfseconomie **96**(7/8), 253–265 (2022). https://doi.org/10.5117/mab.96.90108
10. Pereira, G.V., et al.: South American expert roundtable: increasing adaptive governance capacity for coping with unintended side effects of digital transformation. Sustainability **12**(2), 718 (2020). https://doi.org/10.3390/su12020718
11. Schofield, G., Dittborn, M., Selman, L.E., Huxtable, R.: Defining ethical challenge(s) in healthcare research: a rapid review. BMC Med. Ethics **22**(1), 135 (2021). https://doi.org/10.1186/s12910-021-00700-9
12. Gotowiec, S., Cantor-Graae, E.: The burden of choice: a qualitative study of healthcare professionals' reactions to ethical challenges in humanitarian crises. J. Int. Humanitarian Action **2**(1), 2 (2017). https://doi.org/10.1186/s41018-017-0019-y
13. Molyneux, S., et al.: Model for developing context-sensitive responses to vulnerability in research: managing ethical dilemmas faced by frontline research staff in Kenya. BMJ Glob. Health **6**(7), e004937 (2021). https://doi.org/10.1136/bmjgh-2021-004937
14. Van Bavel, J.J., Packer, D.J., Haas, I.J., Cunningham, W.A.: The importance of moral construal: moral versus non-moral construal elicits faster, more extreme, universal evaluations of the same actions. PLoS ONE **7**(11), e48693 (2012). https://doi.org/10.1371/journal.pone.0048693
15. Hagendorff, T.: The ethics of AI ethics: an evaluation of guidelines. Mind. Mach. **30**(1), 99–120 (2020). https://doi.org/10.1007/s11023-020-09517-8
16. Gates Foundation Announces $1.27B in Health and Development Commitments to Advance Progress Toward the Global Goals. Bill & Melinda Gates Foundation. https://www.gatesfoundation.org/ideas/media-center/press-releases/2022/09/gates-foundation-unga-global-fund-replenishment-commitment (accessed May 01, 2023)
17. Fiske, A., Henningsen, P., Buyx, A.: Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. J. Med. Internet Res. **21**(5), e13216 (2019). https://doi.org/10.2196/13216
18. Philip, A., Samuel, B., Bhatia, S., Khalifa, S., El-Seedi, H.: Artificial intelligence and precision medicine: a new frontier for the treatment of brain tumors. Life **13**(1), 24 (2022). https://doi.org/10.3390/life13010024

19. Goisauf, M., Cano Abadía, M.: Ethics of AI in radiology: a review of ethical and societal implications. Frontiers in Big Data **5** (2022). Accessed: May 01, 2023. https://www.frontiersin.org/articles/https://doi.org/10.3389/fdata.2022.850383
20. Renier, L., Mast, M., Dael, N., Kleinlogel, E.: Nonverbal social sensing: what social sensing can and cannot do for the study of nonverbal behavior from video. Front. Psychol. **12**, 2874 (2021). https://doi.org/10.3389/fpsyg.2021.606548
21. Sollini, M., Bartoli, F., Marciano, A., Zanca, R., Slart, R.H.J.A., Erba, P.A.: Artificial intelligence and hybrid imaging: the best match for personalized medicine in oncology. Eur. J. Hybrid Imaging **4**(1), 1–22 (2020). https://doi.org/10.1186/s41824-020-00094-8
22. Robillard, J.M., et al.: Scientific and ethical features of English-language online tests for Alzheimer's disease. Alzheimers Dement (Amst) **1**(3), 281–288 (2015). https://doi.org/10.1016/j.dadm.2015.03.004
23. Mallakin, M., Dery, C., Vaillancourt, S., Gupta, S., Sellen, K.: Web-based co-design in health care: considerations for renewed participation. Interact. J. Med. Res. **12**(1), e36765 (2023). https://doi.org/10.2196/36765
24. Reamer, F.G.: The evolution of social work ethics: bearing witness. Adv. Soc. Work **15**(1), 163–181 (2013). https://doi.org/10.18060/14637
25. Kacetl, J., Maresova, P.: Legislative and ethical aspects of introducing new technologies in medical care for senior citizens in developed countries. CIA **11**, 977–984 (2016). https://doi.org/10.2147/CIA.S104433
26. Pourvakhshoori, N., Norouzi, K., Ahmadi, F., Hosseini, M., Khankeh, H.: Nurse in limbo: a qualitative study of nursing in disasters in Iranian context. PLoS ONE **12**(7), e0181314 (2017). https://doi.org/10.1371/journal.pone.0181314
27. Khatiban, M., Falahan, S.N., Soltanian, A.R.: Professional moral courage and moral reasoning among nurses in clinical environments: a multivariate model. JMEHM **14** (2022). https://doi.org/10.18502/jmehm.v14i20.8180
28. Park, S.: Heterogeneity of AI-Induced Societal Harms and the Failure of Omnibus AI Laws. arXiv, Mar. 15
29. Taufiq, A., Saripah, I., Herdi, H.: The role of education and supervision toward the candidates of group counselor competencies. Presented at the 3rd Asian Education Symposium (AES 2018), pp. 118–122. Atlantis Press (2019). https://doi.org/10.2991/aes-18.2019.28
30. Córdoba, A., Mejía, L.F., Mannis, M.J., Navas, A., Madrigal-Bustamante, J.A., Graue-Hernandez, E.O.: Current global bioethical dilemmas in corneal transplantation. Cornea **39**(4), 529–533 (2020). https://doi.org/10.1097/ICO.0000000000002246
31. Goris, G., Adolf, S.J.: Utilidad y tipos de revisión de literature. Ene **9**(2) (2015). https://doi.org/10.4321/S1988-348X2015000200002
32. Ministry of Health of Chile, Mental Health National Plan 2027 – 2025. Gobierno de Chile (2017). https://www.minsal.cl/wp-content/uploads/2017/12/PDF-PLAN-NACIONAL-SALUD-MENTAL-2017-A-2025.-7-dic-2017.pdf