THE ETHICS OF APPLYING ARTIFICIAL INTELLIGENCE AND CHATBOTS IN MENTAL HEALTH: A PSYCHOANALYTIC PERSPECTIVE

ÉTICA DA APLICAÇÃO DE INTELIGÊNCIAS ARTIFICIAIS E CHATBOTS NA SAÚDE MENTAL: UMA PERSPECTIVA PSICANALÍTICA

Paulo Victor dos Reis Silveira[1]

João Luiz Leitão Paravidini[2]

Abstract: This paper discusses the ethical implications of the use of Artificial Intelligences (AIs) or therapist chatbots in mental health through the method of narrative review of literature. It defines what these devices are, their benefits and limitations for clinical practice. Then, it examines the ethical problems that the result from these technologies from perspective of psychoanalysis. It highlights the risks of complex emotional harm, lack of validity and reliability of information, and absence of accountability of the agents involved. Finally, it presents some recommendations for the development of more ethical AIs and the questions that arise in this sense.

Keywords: Artificial Intelligence; Mental Health; Psychoanalysis; Ethics.

Resumo: Este artigo discute as implicações éticas do uso de Inteligências Artificiais (IAs) ou chatbots terapeutas na saúde mental através do método da revisão narrativa da literatura. Define-se o que são esses dispositivos, quais são seus benefícios e seus limites para a prática clínica. Em seguida, examinam-se quais problemas éticos que resultam dessas tecnologias pela perspectiva da psicanálise. Destacam-se os riscos de danos afetivos complexos, de falta de validade e confiabilidade das informações e de ausência de responsabilização dos agentes envolvidos. Por fim, apresentam-se algumas recomendações para o desenvolvimento de IAs mais éticas e as questões que se colocam nesse sentido

Palavras-chave: Inteligência Artificial; Saúde Mental; Psicanálise; Ética.

1 Introduction

The emergence of Artificial Intelligences (AIs) such as ChatGPT has enabled an unprecedented expansion of technologies based on human interaction with machines. As these technologies gain influence in various areas of knowledge, concerns also increase, both technical, which seek to explain the functioning and behavior of AIs, and ethical. In 2023, a letter was sent worldwide[3] requesting the suspension of training of AIs more advanced than GPT-4, which is version 4 of a machine learning model that employs supervised and unsupervised learning methods to understand and produce a language that mimics human (Radford *et al*. 2018). It includes the signatures of researchers from the

---

[1] Bachelor in Psychology at Centro Universitário do Triângulo (UNITRI). Master's Degree student at the Federal University of Uberlândia (UFU), Uberlândia-MG, Brazil. Email: pvreis.silveira@gmail.com
[2] PhD in Health Sciences (Mental Health) at the State University of Campinas (Unicamp). Professor at the Institute of Psychology - Federal University of Uberlândia (UFU), Uberlândia-MG, Brazil. Email: jlparavidini@gmail.com ; paravidini@ufu.br

largest technological laboratories in the world, demonstrating the level of alert with such advances.

The progress of AIs has already reached the field of mental health and shows significant advances. Programs using AIs in conjunction with psychiatrists can diagnose depression and schizophrenia more accurately than psychiatrists alone. In addition, these AIs can also predict the consequences of treatments for these diseases with greater accuracy. Furthermore, AIs are used to assist in work with elderly people with dementia, children on the autism spectrum, and to help patients confront their auditory hallucinations (Grodniewicz; Hohol, 2023).

It is important and necessary to reflect on the next steps of the advancement of technologies with AI in relation to psychotherapy. They question whether AIs are already or will be able to perform psychotherapy and if this would represent a break with traditional concepts on the subject. Currently, the closest to this model are the relationships between mental health specialist chatbots and human users (Grodniewicz; Hohol, 2023).

This work aims to present how chatbots and AIs are present in the field of mental health and what possible ethical implications related to the presence of these devices. To do so, we raise the following questions: how does the ethics of psychoanalysis articulate with the ethics of the development and use of AIs? How does Freud's idea that the analyst should not satisfy the analysand's demand apply to configurations and articulations at the social level between humans and AIs? How does the development of AIs increasingly capable of simulating consciousness affect the narcissistic structure of the subject?

The scientific and social importance of the discussion the article proposes to make is about "the ethical implications of the use of Artificial Intelligences (AIs) or therapeutic chatbots in mental health", taking into perspective what psychoanalysis warns us about the ethics of desire. Thus, it is a narrative review of the literature, which aims to describe and discuss the development or the "state of the art" of this theme, from a theoretical or contextual point of view, producing a comprehensive assessment of the literature (Rother, 2007).

1.1 Therapeutic Chatbots

Artificial Intelligence (AI) is an innovative way to expand therapeutic interactions without human mediation. Relational therapeutic agents, based on some principles of

psychology, such as in the therapist-patient relationship, can establish empathy, trust, and simulate a therapeutic alliance with users. Recent evidence (Lucas *et al*. 2014) shows that people tend to disclose more sensitive information during initial clinical interviews conducted by computers than when conducted by humans. The test performed with 97 participants compares user responses in interaction with humans and in interaction with virtual agents in which users reported being more capable of revealing information about what they were feeling and about sadness with virtual agents.

A strong therapeutic alliance can be formed even without face-to-face contact, including non-human interactions with applications or with a conversation agent (chatbots). Although most mental health applications are abandoned a few days after installation, meta-analytical evidence suggests chatbots can increase engagement and pleasure in digital mental health care (Prochaska *et al*. 2023).

For the most part, chatbots serve as an entry point for conventional therapy, indicating that users seek a human therapist (Coghlan *et al*. 2023). However, others also help users regarding mental health, passing exercises based on cognitive-behavioral therapy and emulating emotions and psychological connections. These interactions occur through direct conversation chat and are attractive because they are easily accessible, such as downloading on a smartphone, and because they are fully available to users.

New forms of digital therapies, such as therapeutic chatbots, show promising results in the treatment of various types of mental health problems. A study by Suharwardy *et al*. (2023) demonstrated the Woebot chatbot was effective in reducing symptoms of depression, even without human supervision. In another study, addicts who interacted with therapeutic agents showed a reduction in the use of substances such as alcohol and drugs. In addition, therapeutic chatbots are more effective than digital psychoeducation.

Woebot is a therapeutic chatbot which simulates various elements of a real therapy session. It uses rapport techniques, such as asking open questions and demonstrating interest, to establish a connection with the user. It also confirms if it understood correctly what the user said, normalizes the user's experience, and demonstrates empathetic language. In addition, Woebot suggests a type of therapeutic alliance and works psychoeducation with the user. Finally, Woebot applies key concepts of cognitive therapy, inviting the user to reflect and reframe certain patterns understood during the conversation (Grodniewiecz; Hohol, 2023).

Despite the numerous benefits of artificial intelligence (AI) technology in mental health care, there are also some limitations that need to be addressed. One of the main limitations is the lack of data on certain mental disorders. AI tools depend on large amounts of data to provide accurate and reliable diagnoses and the lack of data can limit the effectiveness of these tools. In addition, the lack of transparency and accountability about the use of AI in mental health care is also a concern. Patients need to be aware of how their data is being used and able to control its use (Rana; Singh, 2023).

Another limitation of AI-based mental health care is the potential for algorithmic bias. AI tools are as effective as the data they are trained on. Therefore, if the data is biased, the results will also be biased, leading to incorrect diagnoses and treatment recommendations. Ethical considerations around the use of AI in mental health care need to be addressed. This includes issues such as privacy, informed consent, and the potential of AI to replace human interaction (Rana; Singh, 2023).

1.2 Ethics and Chatbots

According to Coghlan *et al*. (2023), for a chatbot to be functional and accepted by the user, it needs to be safe, meet demands effectively, and have a form of feedback with a team of involved developers. Among the applications which failed in these fundamental requirements of users are those which, in a harmful and sometimes unethical way, became threats by being offensive, causing emotional damage, confusing user questions, and not fulfilling what they promised to deliver in advantages.

Other risks can be more serious and should even be discussed with greater rigor. In Belgium, a man committed suicide after six weeks of intense conversation with the ELIZA chatbot (Walker, 2023). The program simulated an interaction a person would have with a therapist from a Carl Rogers' person-centered approach, being able to deceive users into thinking they were really talking to a human therapist. Here the first ethical question is raised, as simulating something is a form of deception and the more deceived the user is, the more effective the program is (Grodniewicz; Hohol, 2023).

Tekin (2023) raises three important points of ethical concern in the use of chatbots in mental health. The first is what the author calls "the bot is not a therapist". For him, it is incorrect to call what chatbots do "therapy", or to call chatbots "therapists". Chatbots cannot provide the same level of care and attention as a human therapist and can

undermine the importance of building trust between patients and health professionals who could misinterpret the role of the psychologist in general.

The second point raised by Tekin (2023) is the "presumption of traceability", which assumes the point of view that users will be honest and truthful with their chatbot. However, people who are knowledgeable about private data and how this data can "leak", hesitate to reveal sensitive information to chatbots and may not feel comfortable. As a result, the chatbot would create inaccurate diagnoses or ineffective treatment recommendations. There is also a concern with data privacy, as chatbots collect sensitive personal information from users, and there is a risk this information will be misused or shared without the user's consent, as seen in companies like Facebook with the manipulation or sales of information that has become a widely publicized public concern.

Still regarding this second point, the research of Lucas *et al*. (2014) and Tekin (2023) present distinct but complementary points of view. Lucas *et al*. (2023) suggest users tend to reveal more sensitive information to virtual agents than to humans. On the other hand, Tekin (2023) argues the presumption that users will be honest with chatbots is an ethical problem, especially if users are aware of the privacy risks associated with disclosing sensitive information. Tekin's (2023) ethical survey is due to the growing wave of data leaks, which was not observed in the interaction of the initial interviews of the experiment conducted by Lucas *et al*. (2014) and may point to an expansion in the reliability of the interaction if there is clarity, transparency, and security about how the data of the interaction can be used or as data security is improving.

The third point of concern presented by Tekin (2023) is with the "evidence gap", which refers to the lack of empirical evidence supporting the effectiveness of chatbots in diagnosing and treating mental health conditions. Although some studies have shown promising results, there is still a need for more rigorous research to determine whether chatbots can provide effective and safe diagnoses and treatments for mental health conditions. Without this evidence, it is difficult to justify the use of chatbots in mental health care and there is a risk that patients may be harmed by relying on unproven or ineffective treatments.

According to Sedlakova and Trachsel (2022), there is the problem of not having a consensus of ethical precautions nor guides and criteria in the development of chatbots. There is also no regulation or body to analyze whether such programs are being developed ethically. Furthermore, there is a concern, as a chatbot does not have the same level of

responsibility and duties as a human therapist since an AI is not a "moral subject" and there are no clear legal regulations regarding the bad outcomes of this relationship.

There is concern with harm prevention, a field of ethics in technology that deals with malfunction, operation in an unforeseen way, or theft of information by hackers or unauthorized monitoring (Fiske; Henningsen; Buyx, 2019). In addition, care must be taken with government agencies or health services replacing health services with cheaper tools such as therapeutic chatbots, resulting in less availability of the already existing resources to the detriment of AI technologies.

Fiske, Henningsen, and Buyx (2019) point out that all AI technology works with algorithms and databases and both are subject to ethical problems, such as human biases reproducing and reinforcing forms of existing social inequality, which can generate racist and sexist chatbots, and the lack of transparency required by market competitiveness makes it difficult to act punctually on the problem and identify possible problems, leaving all responsibility in the hands of the programmers themselves.

1.3 Ethics in Psychoanalysis

Freud's and Lacan's psychoanalytic theory can add some density to the discussion on ethics both in the production of AIs and in the development of more ethical Ais, as it thinks about the social bond and the social field from the relationship which applies, by extension, to the relationship between subject and machines. Furthermore, it is possible to question the norms imposed by culture, society, and the dominant discourse, giving space to think through an ethics of singularity.

Since its origin, psychoanalysis has addressed ethics as a way of reading sociocultural phenomena. According to Rosa, Carignato, and Berta (2006), this reading highlights the mode of social bond that constitutes culture, whether in the form of consumption, profit, or suffering. According to Palumbo (2016), psychoanalysis can contribute to a more complex social understanding, as the subject is not just a product of society, but also a desiring subject who can position himself critically in relation to social norms and values. Therefore, the ethics of psychoanalysis reintroduces the social bond, "inside out", as it invites the subject to assume his desire and reflect on his subjective participation in the social bond.

When addressing ethics in psychoanalysis, Freud (1919/1996, p.214) argues analytic therapy should not satisfy the patient's desires, but rather keep him in a state of

deprivation or abstinence, since frustration constituted a stimulus for the patient to change, and the satisfaction of needs could function as a substitute for symptoms and prevent their elaboration and transformation. Moreover, analytic therapy should not impose on the patient the ideals of the analyst, nor try to mold him in his image. This would be a form of violence and authoritarianism, which would prevent the patient from developing his own personality and autonomy. This model also serves for the social, because, still according to Freud (1921/1996, p.81), from the beginning individual psychology is at the same time a social psychology.

Freud (1913/1990a, p.164) states "the extraordinary diversity of the psychic constellations involved, the plasticity of all mental processes and the richness of the determining factors oppose any mechanization of the technique" renouncing regulations and rules in favor of the singular manifestations of each individual, being impossible any attempt to formulate a unique and true technique. The goal of psychoanalysis should be to assist the subject in reconciling with his desire and find a way to live with it satisfactorily.

Lacan (1959-60/1992, p. 314) resumes such Freudian perspective by stating that "the ethics of psychoanalysis is the ethics that suits our science, insofar as it is the science of the subject". For Lacan, the subject is the one who constitutes himself in the field of language, marked by lack and desire. The desire of the subject is not a simple appetite or a biological need, but an incessant search for a lost object that is never found. Desire is what moves the subject in his analysis, leading him to confront himself with the unconscious, with the Other, and with jouissance. Jouissance is a form of paradoxical and painful satisfaction, which surpasses the principle of pleasure and implies a transgression of the limits imposed by the symbolic law.

Jouissance is what resists analysis, as it is what escapes signification and interpretation. According to Lacan (1959-60/1992, p. 366), the ethics of psychoanalysis consists in "not giving up on his desire", that is, in not renouncing his desire in the name of an ideal or a moral norm. The ethics of psychoanalysis aims at a sublimation of jouissance in a creative and singular act, which affirms the difference of the subject. There is, therefore, an ethics of desire that differs from social ethics.

According to Possati (2023), ethics cannot be reduced to a code and, at the same time, as technology advances, it becomes more necessary to apply ethical analysis capabilities to AI systems. The discussion of ethics should start from the point of view of design and respond to three levels: ethics in design, ethics by design, and ethics for design.

Ethics by design is a way of integrating ethical principles and values in the conception, implementation, and evaluation of AI systems. The idea is that ethics should not be seen as an obstacle or a restriction, but as an essential and guiding element of the AI development process.

A research perspective inspired by sociology and psychoanalysis enhances the analysis of the classic problems of AI, as it considers the human and social aspects involved. Psychoanalysis allows us to raise a new point of view on the ethical issues raised by AI in a more relational way, considering the interests and perspectives not only of the agents but also of the users (Possati, 2023). From this perspective, it is possible to analyze the classic problems of AI not only from the viewpoint of a single system but as part of a community composed of humans and non-humans. Ethical problems are seen in their historical and social context. In addition, psychoanalysis helps us to accept and understand the limits of responsibility, recognizing that there are situations in which it is not possible to completely eliminate responsibilities and problems. If an AI system demonstrates discriminatory behavior towards women, this bias should be analyzed as a collective phenomenon, a "habitus" that the machine assimilated from a human-non-human context. To explain that bias, it would be necessary to reconstruct the unconscious communication of the habitus to the machine in a specific social context. And to correct that bias, it would be necessary to create new conditions of communication in this social field.

2 Ethical Recommendations in the Development of Chatbots

Coghlan *et al*. (2023) state that, for a chatbot to be functional and accepted by the user, it must follow the ethical principles of non-maleficence, beneficence, autonomy, justice, and explicability (transparency and accountability). To do so, they must assess the risks and benefits of using chatbots, seek and disclose evidence that supports their efficacy and safety, respect the privacy and consent of users, ensure the quality and accessibility of services, and monitor and evaluate the results of chatbots. They should also consider the possibility of revealing crimes to authorities if they detect emergencies or imminent risk to users or third parties. In addition, they should seek greater participation and consultation from users in the creation and regulation of chatbots, respecting diversity, particularities, and the experiences of those affected.

Fiske, Henningsen, and Buyx (2019) recommend the creation of clear guidelines on the regulation, supervision, and consent of embedded AI, as well as training health professionals and consulting users about the use of AI. The use of AI in mental health should be an addition and not a replacement for the resources present in the instances of mental health. It is still necessary to research the direct and indirect impacts of embedded AI on the therapeutic relationship, on other human relationships, and on self-awareness, agency, and identity of individuals.

Tawfeeq, Awqati, and Jasim (2023) affirm that, to protect the user, it is necessary to regulate the use of AI systems like ChatGPT, establishing rules to prioritize the well-being of the user, protecting them from possible damages caused by the emotional connection or dependence on the system. They should also ensure transparency and accountability of developers in relation to the design and training of AI systems, holding them accountable for any harmful or inappropriate content produced by the chatbot. Another fundamental point is the increase in education and awareness among developers, users, and the general public about the ethical implications of chatbots in general. Also important is the need to establish adequate governance and regulation for AI systems, including regulatory supervision and adherence to ethical standards. Finally, it is important to integrate relevant ethical issues with chatbots with existing ethical structures and guidelines for research and use of AI.

But is establishing norms to be followed enough to maintain an ethical practice? A controlled study on the individual decisions of software engineers reveals the guidelines have almost no impact on the behavior of technology professionals (Hagendorff, 2020). Participants were subjected to eleven ethical decision scenarios related to software, covering topics such as responsibility to report problems and errors, user data collection, intellectual property, code quality, honesty with the client, and time and personnel management. The study involved 63 students and 105 professional software developers. Researchers evaluated whether the code of ethics of computing professionals had any impact on ethical decisions in six of these scenarios and concluded that there was no statistically significant difference in responses between individuals who saw or did not see the code of ethics, both for students and professionals.

Ethics faces limitations both at the individual level and in society as a whole (Hagendorff, 2020). In the competitive market, many companies seek to profit from AI in various applications, following an economic logic not conditioned by an ethics based on values or principles. In turn, engineers and developers do not receive training or

support to deal with ethical issues. In the business context, speed is essential and ignoring ethics is the easiest way. Thus, the practice of development, implementation, and use of AI applications has little to do with the values and principles proposed by ethics. A lot of money is invested in the development and commercial use of machine learning-based systems, while ethical considerations are used mainly for public relations purposes.

For effective AI ethics, it is necessary to go beyond verification guidelines and adopt a situation-sensitive approach, which values personal virtues and dispositions, expansions of knowledge, responsible autonomy, and freedom of action of moral actors, not seeking to generalize cases under individual principles, but adapting to specific situations and technical configurations. This approach requires a balance between the focus on the technological details of AI methods and technologies and machine learning, to bring ethics closer to technical discourses, and the focus on social and personal aspects, to emancipate moral actors from problematic routines and promote individual responsibility. AI ethics deals less with AI itself than with forms of deviation or distancing of programmers from what is produced (Hagendorff, 2020).

At this point, in addition to understanding and analyzing the ethical guidelines that govern the development of AIs, it is important to analyze which aspects impact culture in general, highlight these aspects with greater depth and show the way these bonds operate in the social subject. This bond is determinant to understand the human and non-human relationship. Thinking about ethics only from the point of view of technology implies thinking about a neutral program that does not imply in the relationship and is not affected by the subject, which is not possible to think from the perspective of psychoanalysis.

Thinking about the ethical education of software engineers is thinking about education from the point of view of the ethics of desire. For Costardi and Endo (2013), education is committed to a civilizing project and makes a moral demand on the subject, who must insert himself in the scope of collective ideals without giving up his singularity. This insertion implies a cost to the subject's desire, which must renounce part of his satisfaction. Derived from psychoanalysis, the ethics of desire does not oppose this process, but problematizes it, as far as it recognizes the resistances of singularities to the homogenizing proposal of an education. These resistances are expressions of the subject of the unconscious, which escapes the didactic techniques and specialized knowledge directed to education, whether it is carried out within the school or in the broader sense of social experience. The ethics of psychoanalysis affects education, insofar as it has to

do with the ability to deal with fundamental evil by other means than goodness, to face a gap in knowledge intended to join means and ends, to put into play a new and particular knowledge that, sometimes, needs to be invented by the subjects who are implicated in the educational endeavor.

From Lacan's (1959/60) ethical perspective, the subject's desire is the desire of the Other and, upon entering the relationship with the law, tradition, and morality, he must renounce part of his enjoyment, that is, his instinctual satisfaction, to identify with collective ideals. This renunciation is symbolic castration, being the condition for access to the field of desire. It is important to recognize there is a remainder of enjoyment that escapes symbolic castration and manifests itself as an excess, a lack, or a transgression. This remainder or lack is what Lacan calls object a, which is the object cause of desire, but which can never be fully achieved or satisfied. The object a is the point of articulation between the symbolic and the real, between the law and enjoyment. The ethics of psychoanalysis consists in not giving up on one's desire, that is, in not renouncing the object a in the name of adaptation to social norms or an illusion of completeness. The ethics of psychoanalysis implies assuming the division of the subject, the lack in the Other, and the impossibility of absolute knowledge about desire.

Considering this Lacanian perspective of desire, the production of more ethical AIs should highlight the incongruities of the pretensions of universality, rationality, and transparency that are often attributed to the ethical production of such technologies. AIs are not just logical and computational systems but can also be understood as symbolic and affective systems, which relate to human subjects and their desires. Chatbots cannot ignore the dimensions of the unconscious, enjoyment, and fantasy which are present in human interactions. It is necessary to think about how these technologies will be able to deal with the complexity and ambiguity of ethical situations, without being reduced to algorithms or predefined protocols, and how or if they will be responsible for their actions and their effects, without exempting themselves from the consequences of their decisions.

Possati (2023) proposes that the development of AI should be based on the responsibility of agents (users, developers, programmers, designers, owners, and software) and patients (who use AI and interact with it). For him, there can be no real responsibility without the relationship between the agent and the patient, as the agent cannot act ethically without the patient's demand, regardless of all possible moral principles or virtues. Addressing the ethical dimension requires considering the individuals who conceive, build, and apply artificial intelligence in their community.

Ethics, however, does not align with the creation of globally applicable regulations nor with purely subjective obligations. Instead, in the context of interactions between humans and AI, it implies understanding AI systems as subjects and objects of ethical evaluation in the midst of a complex network of relations between individuals and non-humans.

Oliveira and Corrêa (2023) affirm that the technological virtuality affects and modifies the formation of social bonds. The new reality discards disturbing and unwanted aspects and the virtual reminder of a bodyless relationship, present in the relationship by an AI, highlights and points at the same time the possible and impossible bonds of this human and non-human relationship, or as proposed earlier, between agents and patients.

Thinking about the problems generated by the expansion of AI technologies which encompass the field of mental health from the psychoanalytic bias allows inflections about the nature of the relationship and how discourses can articulate with the uses of these technologies. Psychoanalysis contributes to the discussions of the natures of relationships and, therefore, is fundamental in the discussion of this work.

3 Conclusion

In this paper, we sought to analyze the ethical implications of the use of chatbots and artificial intelligences (AIs) in the field of mental health, from the perspective of psychoanalysis. We consider psychoanalysis can offer a relevant contribution to understand the relations between humans and AIs, as well as the subjective and moral impacts originated from this relationship.

We revisit Freud's idea that the analyst should not satisfy the analysand's will and expand it to the configurations and articulations at the social level between humans and AIs. We argue this logic is present in the human and non-human relationship, as the AI can both meet the user's demands, offering solutions to their psychic contents, anxieties, and pains, but can also frustrate or challenge the user, leading him to question or modify his patterns of thought and behavior. We exemplify this logic with concrete cases of chatbots that use psychotherapeutic techniques, such as Woebot and ELIZA, and discuss how they can affect not only the subject's relationship with the AI, but also with the subject's other social bonds. We question what the damage of access to a tool that is in total availability to the user meeting all their demands are. Would the AI be able to embody the demand to the Other, thus satisfying to sustain obedience to social ideals?

One of the emerging questions is how the responsibility of the creator (developer, owner, designer, and software) of the possible failures and damages caused by their creation (chatbots and relational AIs) is theorized. The discussions about the responsibility of parents for the subjective manifestations of their children are present in psychoanalytic discussions and serve as a basis to see how this problem is more serious than it is presented by the researchers cited in this work. What would be under responsibility, what kind of organization is necessary to maintain a body responsible for establishing such criteria and based on which point of view are some of the possible contradictions evidenced during the writing of this work.

The development of artificial intelligences increasingly capable of simulating a consciousness and thus being an artifact of desire seems to cross the subject in a way that throws him into an instance of narcissistic potency, considering the capacity to expand the limits of the subject himself in his relationship with reality, as well as to shake the narcissistic structure causing fear and discomfort, as if in some way the advancement of these technologies were increasingly evidencing human incapacities. At this point, the construction based on the ethics of AIs must be guided by articulation with psychoanalytic ethics, to allow new ways of thinking about such artifacts to exist, enabling an advance in the development of technologies articulated in the relations between subject and machines.

It is not possible to think of an ethics not involving the subject in the very ethical relationship in which he articulates himself and it is not possible to think of a total artificial neutrality when faced with human programmers and a relationship with a non-neutral subject who is the human. There is no exemption in a relationship, whether it is artificial or not.

References

COGHLAN, S.; LEINS, K.; SHELDERICK, S.; CHEONG, M.; GOODING, P.; D'AlFONSO, S. To chat or bot to chat: Ethical issues with using chatbots in mental health. **Digital Health**, Londres, v. 9, n.1 p. 1-10, jun. 2023.Disponível em: https://doi.org/10.1177/20552076231183542. Acesso em: 22 set. 2023.

CORDEIRO, A. M.; OLIVEIRA, G. M. de; RENTERÍA, J. M.; GUIMARÃES, C. A. Revisão sistemática: uma revisão narrativa. **Revista do Colégio Brasileiro de Cirurgiões**, Rio de Janeiro, v. 34, n. 6, p. 428–431, nov. 2007. Disponível em: https://doi.org/10.1590/S0100-69912007000600012. Acesso em: 29 set. 2023.

COSTARDI, G. G.; ENDO, P. C. Ética da psicanálise, educação e civilização. **Estilos clin**., São Paulo, v. 18, n. 2, p. 327-341, ago. 2013.   Disponível em

http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1415-71282013000200008&lng=pt&nrm=iso. Acesso em: 29 set. 2023.

ELIAS, C. S. R.; SILVA, L. A.; MARTINS, M. T.; RAMOS, N. A.; SOUZA, M. G.; HIPÓLITO, R. L. Quando chega o fim?: uma revisão narrativa sobre terminalidade do período escolar para alunos deficientes mentais. **Rev. Eletrônica Saúde Mental e drogas**, Ribeirão Preto, v. 8, n. 1, p. 48-53, abr. 2012. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1806-69762012000100008&lng=pt&nrm=iso. Acesso em: 30 jan. 2024

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**, Toronto v. 21, n. 5, p. e13216, maio. 2019. Disponível em: https://doi.org/10.2196/13216. Acesso em: 22 set. 2023.

FREUD, S. Sobre o início do tratamento: novas recomendações sobre a técnica da Psicanálise. In: FREUD, S. O caso Schreber, artigos sobre técnicas e outros trabalhos. **Edição standard brasileira das obras psicológicas completas de Sigmund Freud Vol. 12**. Tradução de J. Salomão. Rio de Janeiro: Imago, 1990a. p. 74-89. (Trabalho original publicado em 1913).

FREUD, S. Linhas de progresso na terapia psicanalítica. In: FREUD, S. História de uma neurose infantil e outros trabalhos. **Edição Standard Brasileira das Obras Completas de Sigmund Freud Vol. 17**. Tradução de J. Salomão. Rio de Janeiro: Imago, 1996. p. 98-104. (Trabalho original publicado em 1919).

FREUD, S. Psicologia de grupo e análise do ego. In: FREUD, S. Além do princípio do prazer. **Edição Standard Brasileira das Obras Completas de Sigmund Freud Vol. 18**. Tradução de J. Salomão. Rio de Janeiro: Imago, 1996. p. 44-90. (Trabalho original publicado em 1921).

GRODNIEWICZ, J. P.; HOHOL, M. Waiting for a digital therapist: three challenges on the path to psychotherapy delivered by artificial intelligence. **Frontiers in Psychiatry**, Laussane, v. 14, n.1 p. 01, maio. 2023. Disponível em: https://doi.org/10.3389/fpsyt.2023.1190084. Acesso em: 22 set. 2023.

HAGENDORFF, T. The ethics of AI ethics: an evaluation of guidelines. **Minds & Machines**, Cham v. 30, n.1 p. 99-120, fev. 2020. Disponível em: https://doi.org/10.1007/s11023-020-09517-8. Acesso em: 29 set. 2023.

LACAN, J. **Seminário 7 – A Ética da psicanálise**. Tradução de A. V. Milani. Rio de Janeiro: Jorge Zahar Editor, 1997. (Trabalho original publicado em 1959-60).

LUCAS, G. M.; GRATCH, J.; KING, A.; MORENCY, L-P. It's only a computer: virtual humans increase willingness to disclose. **Computers in Human Behavior**, Amsterdam, v. 37, n.1, p. 94-100, ago. 2014. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S0747563214002647?via%3Dihub. Acesso em: 22 set. 2023.

OLIVEIRA, G. D. F.; CORREA, H. C. S. Entre encontros faltosos e excessivos: laços amorosos e uso de tecnologias para pensar o sujeito. **Tempo psicanalítico**, Rio de Janeiro, v. 55, n. 1, p. 32-56, jun. 2023. Disponível em http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0101-48382023000100002&lng=pt&nrm=iso. Acesso em: 27 set. 2023.

POSSATI, L. M. **Unconscious Networks**: Philosophy, Psychoanalysis, and Artificial Intelligence. New York, NY: Routledge, 2023. Disponível em: https://doi.org/10.4324/9781003345572. Acesso em: 23 set. 2023.

PROCHASKA, J. J.; VOGEL, E. A.; CHIENG, A.; BAIOCCHI, M.; PAJARITO, S.; PIRNER, M. DARCY, A.; ROBINSON, A. A relational agent for treating substance use in adults: Protocol for a randomized controlled trial with a psychoeducational comparator. **Contemporary Clinical Trials**, Amsterdam. v. 127, n1. P. 01, abr. 2023. Disponível em: https://doi.org/10.1016/j.cct.2023.107125. Acesso em: 22 set. 2023.

RADFORD, A.; NARASIMHAN, K.; SALIMANS, T.; SUTSKEVER, I. Improving language understanding by generative pre-training. **Preprint**. 1-12. 2018. Disponível em: https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf. Acesso em: 25 jan. 2024.

RANA, U.; SINGH, R. The Role of Artificial Intelligence in Mental Health Care. **SocArXiv Papers**, College Park. p.1. set. 2023 Disponível em: https://doi.org/10.31235/osf.io/r4umy . Acesso em: 22 set. 2023.

ROSA, M. D.; CARIGNATO, T. T.; BERTA, S. L. Ética e política: a psicanálise diante da realidade, dos ideais e das violências contemporâneos. **Ágora**: Estudos em Teoria Psicanalítica, Vitória, v. 9, n. 1, p. 35-48, jan. 2006. Disponível em: https://doi.org/10.1590/S1516-14982006000100003. Acesso em: 22 set. 2023.

ROTHER, E. T. Revisão sistemática X revisão narrativa. **Acta Paulista de Enfermagem**, São Paulo. v. 20, n. 2, p. v–vi, abr. 2007. Disponível em: https://www.scielo.br/j/ape/a/z7zZ4Z4GwYV6FR7S9FHTByr/#. Acesso em: 09 fev. 2024.

SEDLAKOVA, J.; TRACHSEL, M. Conversational Artificial Intelligence in Psychotherapy: A New Therapeutic Tool or Agent? **The American Journal of Bioethics**, Londres, v. 23, n .5, p .4 -13, maio. 2023.Disponível em: https://doi.org/10.1080/15265161.2022.2048739.

SUHARWARDY, S.; RAMACHANDRAN. M.; LEONARD S. A.; GUNASEELAN A.; LYELL D. J.; DARCY A.; ROBINSON A.; JUDY A. Feasibility and impact of a mental health chatbot on postpartum mental health: a randomized controlled trial. **AJOG Global Reports**, Amsterdam, v. 3, n. 3, p. 1, ago. 2023.Disponível em: https://doi.org/10.1016/j.xagr.2023.100165. Acesso em: 29 set. 2023.

TAWFEEQ, T. M.; AWQATI, A. J.; JASIM, Y. A. The Ethical Implications of ChatGPT AI Chatbot: A Review. **Journal of Modern Computing and Engineering Research**, Mosul, v.2023, n.1 p .49 -57, jul. 2023.Disponível em: https://jmcer.org/research/the-ethical-implications-of-chatgpt-ai-chatbot-a-review/. Acesso em: 25 set. 2023.

TEKIN, Ş. Ethical issues surrounding artificial intelligence technologies in mental health: psychotherapy chatbots In: ROBSON, G. J.; TSOU, J. (ed.). **Technology ethics**: A philosophical introduction and readings. Londres: Routledge, 2023. p. 152 -159. Disponível em: https://doi.org/10.4324/9781003189466.

WALKER, L. Belgian man dies by suicide following exchanges with chatbot. **The Brussels Times**, Bruxelas. 28 mar .2023. Disponível em: https://www.brusselstimes.com/430098/belgian-man-commits-suicide-following-exchanges-with-chatgpt. Acesso em: 22 set. 2023.