



Universidad de Belgrano

***“Aplicación de la inteligencia artificial a la psicoterapia
cognitivo-conductual moderna”***

Alumno: Christian Uriel Wikinski

Matrícula: 21920

ID: 000-16-9376

Facultad: Humanidades

Carrera: Licenciatura en Psicología

Turno: Noche

Tutor: Mag. Juan Pablo Mora Penagos

Año académico: 2026

FIRMA DEL ALUMNO

Índice

Resumen.....	3
Introducción.....	4
Presentación del tema.....	4
Problema y pregunta de investigación.....	5
Objetivos generales y específicos.....	7
Objetivo general.....	7
Objetivos específicos.....	7
Alcances y límites.....	7
Antecedentes.....	8
Estado del arte.....	11
Marco teórico.....	13
Desarrollo metodológico.....	16
Procedimiento.....	16
Capítulo 1: La inteligencia artificial en la terapia cognitivo-conductual: herramientas digitales y aplicaciones clínicas.....	17
Capítulo 2: Desafíos éticos y prácticos de la inteligencia artificial en salud mental.....	17
Capítulo 1: La inteligencia artificial en la terapia cognitivo-conductual: herramientas digitales y aplicaciones clínicas.....	18
1.1 Chatbots terapéuticos basados en TCC: el caso de Woebot como herramienta de apoyo clínico.....	18
1.2 Chatbots terapéuticos orientados al bienestar emocional: el caso de Wysa y la intervención subclínica en TCC.....	19
1.3 Terapia de exposición con realidad virtual (VRET): aplicaciones de la realidad virtual en la terapia cognitivo-conductual.....	21
1.5. Intervenciones Adaptativas Justo a Tiempo (JITAI): un marco metodológico para la integración de la IA en la TCC.....	25
1.6. Inteligencia artificial en la gestión clínica: automatización administrativa y optimización del ejercicio profesional en psicología.....	27
Capítulo 2: Desafíos éticos y prácticos de la inteligencia artificial en salud mental.....	29
2.1 Privacidad, confidencialidad y protección de datos en el contexto terapéutico.....	29
2.2 Inteligencia artificial generativa, LLM y la problemática de la “caja negra”.....	29

2.3 Empatía, distancia psicológica y vínculo terapéutico mediado por IA.....	30
2.4 Autonomía, consentimiento informado y riesgos de manipulación.....	31
2.5 Sesgos algorítmicos y equidad en la toma de decisiones.....	31
2.7 Intervenciones JITAí y desafíos teóricos.....	32
Conclusión.....	34
Referencias bibliográficas.....	36

Resumen

La inteligencia artificial (IA) ha sido una de las tecnologías más disruptivas de los últimos tiempos, y su impacto se ha sentido en una amplia variedad de campos, desde la industria hasta la medicina. En este contexto, la psicología no ha sido ajena a esta revolución tecnológica, y en particular, la terapia cognitivo-conductual (TCC) ha comenzado a explorar las posibilidades que ofrece la IA para su complementación. Este trabajo se centra en la aplicación de la IA en el contexto de la TCC, examinando sus alcances, limitaciones y el impacto en la práctica terapéutica. El objetivo principal es considerar el aporte de la IA en distintos aspectos y etapas en la TCC y desarrollar las distintas herramientas que la IA puede ofrecer para trabajar sobre la autonomía del paciente en su proceso de TCC. Además, se destaca el potencial de la IA para complementar a los seres humanos en cuanto a tareas cognitivas desafiantes y se menciona que esto ha llegado a construir el desarrollo de algoritmos y sistemas más avanzados en el campo de la inteligencia artificial, sentando las bases para futuros avances en juegos y en una variedad de aplicaciones en la vida cotidiana. En general, la aplicación de la inteligencia artificial en la terapia cognitivo-conductual tiene como objetivo mejorar la eficacia del tratamiento y proporcionar nuevas herramientas para los profesionales de la salud mental, teniéndose en cuenta que la IA no puede reemplazar por completo la interacción humana en la TCC, y es fundamental encontrar el equilibrio adecuado entre la tecnología y la relación terapéutica, poniendo el foco en el uso correcto de los datos sensibles brindados por los pacientes.

Palabras clave: inteligencia artificial, terapia cognitivo-conductual, evaluación del progreso, autonomía del paciente, ética y privacidad.

Introducción

Presentación del tema

La inteligencia artificial (IA) se refiere a la simulación de procesos intelectuales humanos mediante la programación de sistemas informáticos para realizar tareas que normalmente requieren la participación de la mente humana (Simon & Newell, 1976). Estas tareas incluyen aprendizaje, resolución de problemas, toma de decisiones, reconocimiento de patrones, comprensión del lenguaje natural y adaptación a situaciones cambiantes. La IA se basa en algoritmos y modelos matemáticos que permiten a las máquinas procesar datos y aprender de ellos, lo que permite a las computadoras realizar tareas automáticamente y mejorar el rendimiento con el tiempo. Asimismo, desde un enfoque actual, la UNESCO provee una definición:

La Inteligencia Artificial son sistemas tecnológicos que demuestran un comportamiento inteligente analizando su entorno y actuando, con cierto grado de autonomía, para lograr objetivos específicos. (UNESCO, 2024).

La IA abarca una variedad de enfoques, como el aprendizaje automático, el procesamiento del lenguaje natural (PNL), la visión por computadora y la robótica, y se aplica en muchas industrias y campos que van desde la medicina, la producción hasta el servicio al cliente y la investigación científica (Russell & Norvig, 2021).

La psicoterapia, desde la perspectiva de la psicología cognitiva (TCC), se puede definir a través del enfoque de Beck (1979), uno de los fundadores de la terapia cognitiva. Para el autor, la psicoterapia es un proceso colaborativo entre terapeuta y cliente que tiene como objetivo identificar y cambiar patrones de pensamiento distorsionados y creencias irracionales que contribuyen a problemas emocionales y de conducta. La terapia cognitiva tiene como objetivo ayudar a las personas a reconocer y desafiar sus pensamientos negativos y automáticos, a través de un proceso de examen de evidencia positiva y estructurada para reemplazar patrones de pensamiento enredados y disfuncionales con un pensamiento más realista y adaptativo. Este enfoque tiene como objetivo cambiar las percepciones distorsionadas y promover el cambio cognitivo, que a su vez afecta a las emociones y comportamientos problemáticos. Básicamente, la psicoterapia, según Beck (1979), se basa en la reestructuración cognitiva para mejorar el bienestar mental y emocional de las personas.

Para Rizzo & Boucher (2023), la inteligencia artificial (IA) se ha convertido en un área de investigación y aplicación cada vez más importante en el campo de la psicología. Su importancia radica en su capacidad para analizar eficientemente grandes volúmenes de datos y extraer patrones sutiles, que pueden proporcionar nuevos enfoques y conocimientos para comprender los procesos cognitivos, emocionales y espirituales del comportamiento humano. Según los autores, la inteligencia artificial permite analizar grandes volúmenes de datos para identificar patrones emocionales y cognitivos, facilitando intervenciones adaptativas y el uso de agentes virtuales que reducen las barreras de acceso a la salud mental. Además, los chatbots y los agentes virtuales basados en IA pueden desempeñar un papel valioso a la hora de brindar apoyo emocional y cognitivo a las personas con dificultades emocionales, al tiempo que alivian las barreras percibidas para acceder a la atención terapéutica (Torous et al., 2021).

Problema y pregunta de investigación

Con la aparición de nuevas tecnologías y el rápido aumento de la población mundial que las utiliza, estamos viendo nuevas formas de operar y evolucionar en el entorno actual, eligiendo cada vez más herramientas que debido a su mayor velocidad, habilidad o conocimiento, se ven como complementarias en el sentido de que facilitan la vida de las personas. Como todo avance tecnológico de la historia, fue repentino e innovador respecto a los conocimientos previos, y el caso particular de la inteligencia artificial no es el primero, ni tampoco más distinto, sino aún más disruptivo que lo ya conocido.

Es en el contexto de la globalización y las nuevas tecnologías que comenzamos a ver su uso en muchos campos diferentes, ya sea para la recopilación de datos, el procesamiento rápido de información como los motores de búsqueda lineales o en la formación de imágenes virtuales derivadas de texto proporcionado por el usuario. El campo de la psicología no ha pasado desapercibido en este nuevo contexto, y es por ello que este trabajo reconoce la necesidad de introducir el uso de estas herramientas y complementar la utilización de éstas para la psicoterapia, en los casos que, más específicamente, pertenecen al marco teórico de la psicología cognitiva.

Teniendo en cuenta la problemática descrita surgen interrogantes acerca de: ¿cómo se involucra la integración de la inteligencia artificial en la psicoterapia cognitivo-conductual moderna? ¿Cómo podría la inteligencia artificial ayudar en el seguimiento y la evaluación del progreso de los pacientes a lo largo del tiempo en la terapia cognitiva? ¿De qué manera la inteligencia artificial podría influir en la autonomía del paciente en la toma de decisiones en su proceso de terapia cognitiva? ¿Cómo se le puede facilitar administrativamente el trabajo al

terapeuta? ¿Sabe el paciente cuándo está interactuando con un modelo generativo y cuándo con un humano, y qué uso se hace de los datos que el paciente brinda?

A pesar de los avances significativos en la integración de la inteligencia artificial (IA) en la psicoterapia cognitivo-conductual, persisten diversas cuestiones éticas que requieren un examen más profundo. En primer lugar, se necesita una mayor claridad en cuanto a la privacidad y la seguridad de los datos de los pacientes. La recopilación y el análisis de información sensible pueden plantear preocupaciones sobre la protección de la información confidencial, y es esencial establecer estándares rigurosos para garantizar la seguridad de estos datos.

Por último, la supervisión y la toma de decisiones éticas en el uso de la IA en psicoterapia son un área de investigación crítica. La falta de una regulación adecuada podría dar lugar a desafíos éticos relacionados con la responsabilidad, la toma de decisiones automatizada y la potencial falta de transparencia en los procesos de la IA. Los terapeutas y los desarrolladores de tecnología deben colaborar para garantizar que las decisiones terapéuticas sigan siendo éticas y humanas.

Relevancia de la temática

La integración de la inteligencia artificial en la terapia cognitivo-conductual supone un importante paso adelante para enriquecer y optimizar el proceso terapéutico. Al aprovechar la capacidad de la IA para analizar patrones físicos, cognitivos y emocionales complejos en grandes conjuntos de datos, se pueden determinar con mayor precisión y rapidez los pensamientos distorsionados y los patrones disfuncionales que contribuyen al trastorno afectivo. Esto ayuda a personalizar las intervenciones terapéuticas de forma más eficaz, adaptándolas a las necesidades individuales del paciente. Además, la IA permite un seguimiento continuo del progreso del tratamiento y proporciona apoyo en tiempo real, ampliando así las posibilidades de intervención y prevención. Es importante que el papel del psicólogo se actualice con nuevas herramientas que surgen en un mundo que sufre cambios cada vez más rápidos y complejos, en el que la rama de la psicología es un elemento fundamental relacionado con el desarrollo humano, ya sea desde aportes en el marco teórico o desde la inserción en el dispositivo terapéutico.

Asimismo, los aportes desplegados a lo largo del presente trabajo nos apoyarán y complementarán a actuales y futuros psicólogos en el uso de herramientas innovadoras para poder brindarle una mejor atención y utilizar efectivamente los recursos disponibles en el trabajo con el consultante, a partir de un enfoque teórico cognitivo conductual.

Objetivos generales y específicos

Objetivo general

Explorar las herramientas que provee la inteligencia artificial en la aplicación y seguimiento del progreso de los pacientes a lo largo de la terapia cognitiva.

Objetivos específicos

- Describir las aplicaciones de la inteligencia artificial en la terapia cognitivo-conductual.
- Describir los principales inconvenientes que representa el uso de la inteligencia artificial en la terapia cognitivo-conductual.
- Analizar los principales desafíos éticos y prácticos de la integración de la inteligencia artificial en la salud mental.

Alcances y límites

El siguiente trabajo de carácter exploratorio implica la investigación acerca de cómo se utiliza la inteligencia artificial en su aplicación en el contexto de terapia, en los que se tendrá en cuenta la utilización de chatbots terapéuticos, el monitoreo continuo y los análisis de datos.

Para el abordaje del plano psicológico, más específicamente el tipo de terapia donde se aplica, se tomarán autores con inclinación hacia lo cognitivo conductual y se dejarán de lado otros tipos de terapia, como lo pueden ser psicoanalíticas, humanista existencial o sistémica, por citar ejemplos.

El grueso de las investigaciones de la IA como herramienta de la terapia cognitivo conductual ocurren principalmente en Estados Unidos y a la vez está teniendo una creciente proyección en el resto de América, más específicamente América Latina y Argentina. En este aspecto se tendrá como alcance de la investigación geográfico a todo el continente americano.

En cuanto al tema de trabajo, éste estará limitado teniendo en cuenta que está en constante investigación por su carácter novedoso y de constante complejización en sus desarrollos, por ende se incluirán varias miradas que estarán centradas en su aplicación concreta en terapias cognitivo conductuales, ya que se reconoce que la IA es un campo en constante evolución y se debe reconocer que las tecnologías y enfoques pueden cambiar rápidamente.

Dicho esto, es posible que el cumplimiento de los objetivos planteados en esta

investigación queden limitados y están sujetos a confirmación y posterior desarrollo con el paso del tiempo.

Antecedentes

La terapia cognitivo-conductual (TCC) es un enfoque terapéutico que ha evolucionado a lo largo del tiempo. En la década de 1950, Albert Ellis desarrolla la Terapia Racional Emotiva Conductual (TREC), sentando las bases de la terapia cognitiva. Como se detalla en Ellis y Ellis (2019), los autores postulan que los pensamientos irracionales llevan a emociones negativas y, en última instancia, a conductas disfuncionales. Allí se describe como la TREC se enfoca en identificar y cambiar estos pensamientos irracionales. Es a partir de los años 60 que Beck (1979) introduce la Terapia Cognitiva, que se convierte en un enfoque importante para ese contexto. El autor sostiene que los pensamientos automáticos negativos y las creencias distorsionadas son fundamentales en la génesis de los trastornos emocionales. Desarrolla el Modelo Cognitivo, que se enfoca en la identificación y la reestructuración de estos patrones de pensamiento. Beck (1979), y a su vez otros terapeutas cognitivos, comienzan a desarrollar técnicas específicas, como la reestructuración cognitiva, la terapia de exposición, y la solución de problemas. Además, se empiezan a establecer protocolos de tratamiento estructurados para trastornos específicos, como la depresión y la ansiedad (Beck, 1979).

Desde 1980 la TCC se expande a un rango más amplio de trastornos psicológicos, y se enfatiza la importancia de la validación empírica de los enfoques terapéuticos (Beck, 1979). Se llevan a cabo numerosos estudios clínicos para evaluar la eficacia de la TCC donde la tecnología, como el uso de programas de computadora para la terapia cognitiva, se incorpora en el campo. Además, se producen cambios en la forma en que se practica la TCC, como la terapia en línea y la terapia de grupo. Será a partir del nuevo milenio que surge una nueva generación de terapias, como la Terapia de Aceptación y Compromiso y la Terapia de Mindfulness (Hayes et al., 2014). Aunque estos enfoques son diferentes de la TCC tradicional, comparten elementos de la terapia cognitiva y se utilizan en combinación con ella.

En cuanto a los primeros sistemas de IA, estos comenzaron a emerger en la década de 1960, con su continuo desarrollo hacia la década de 1990. Aunque no estaban específicamente diseñados para la psicoterapia cognitivo-conductual, sentaron las bases para futuros desarrollos:

En 1943, Warren McCulloch y Walter Pitts proponen el modelo de neuronas artificiales. Según su modelo, estas neuronas artificiales son unidades de procesamiento que pueden

recibir múltiples entradas binarias, realizar una suma ponderada de esas entradas y emitir una señal de salida si la suma excede un umbral (McCulloch & Pitts, 1943). Estas neuronas se utilizaron como un concepto fundamental en el desarrollo temprano de las redes neuronales artificiales y sentaron las bases para el procesamiento de información en sistemas computacionales inspirados en el cerebro. Asimismo, la "Prueba de Turing", llamada así por su creador Alan Turing en 1950, es un criterio para evaluar la inteligencia de una máquina. Consiste, según el autor, en una conversación entre un humano y una máquina, donde un juez humano debe determinar si las respuestas de la máquina son indistinguibles de las respuestas de un ser humano. Si la máquina puede engañar al juez se considera que ha superado la "prueba de Turing" y demuestra un nivel de inteligencia artificial (Turing, 1950).

Es de suma importancia rescatar y enunciar lo teorizado acerca de la teoría de la computación. Una máquina puede ser llamada inteligente si puede llevar a cabo cualquier tarea que un ser humano pueda realizar mediante el procesamiento de información (Turing, 1950). Esta teoría, según el autor, establece que la mente humana puede ser vista como una máquina de procesamiento de información, lo que en desarrollos posteriores llevó al desarrollo de la IA simbólica. La idea fundamental de su trabajo es que la inteligencia es el resultado de la manipulación de símbolos y reglas lógicas. Esta teoría ha influido en el enfoque clásico de IA, donde se utilizan reglas y representaciones simbólicas para resolver problemas.

Posteriormente, se crearon programas de "terapia por computadora" que brindaban intervenciones simples de apoyo emocional. La IA comenzó a utilizarse en la simulación de la conversación terapéutica. El programa ELIZA, desarrollado por Joseph Weizenbaum en la década de 1960, fue un precursor de los chatbots terapéuticos. Aunque no estaba basado en la psicoterapia cognitivo-conductual, sirvió como un hito en la interacción entre humanos y máquinas. Weizenbaum menciona:

ELIZA es un programa que simula la comunicación humana. Yo no afirmo que posea inteligencia en sentido ni en principio. Pero el hecho de que el simulacro sea lo suficientemente convincente para engañar a una persona lo suficientemente ingenua como para creerlo posee, en sí mismo, importancia. No porque no haya un principio que distinga entre lo que es y lo que no es inteligencia artificial, sino porque, hasta ahora, ningún principio de ese tipo ha sido descubierto. (Weizenbaum, 1966, p. 36)

"ELIZA", desarrollada en el MIT en los años sesenta, fue uno de los primeros programas informáticos diseñados para simular conversaciones terapéuticas. Aunque ELIZA, para Weizenbaum, no proporcionaba un tratamiento terapéutico genuino y no estaba basada en la terapia cognitivo-conductual, su diseño se centraba en mantener conversaciones con los usuarios y responder a sus entradas de manera reflexiva. El programa, que se utilizó

principalmente hasta comienzos del siglo XXI, utilizaba patrones simples de procesamiento del lenguaje natural para emular una especie de "terapeuta" que reflejaba las preguntas y declaraciones de los usuarios.

La contribución más significativa de ELIZA, según su autor, fue su demostración de que las personas a menudo estaban dispuestas a interactuar y compartir sus pensamientos y sentimientos con programas de computadora, incluso sabiendo que no estaban interactuando con un terapeuta humano real. Esto llevó a investigaciones posteriores sobre la "terapia por computadora" y, en última instancia, a la creación de chatbots terapéuticos más avanzados que utilizan técnicas más sofisticadas de procesamiento de lenguaje natural y aprendizaje automático para proporcionar apoyo emocional y psicológico a las personas. Asimismo, afirma en una entrevista que la inteligencia artificial no es un sustituto de la inteligencia humana, sino una herramienta que puede amplificar nuestras capacidades, ya que para el autor, la verdadera inteligencia reside en la sabiduría y la comprensión humanas (Weizenbaum, 1975).

Asimismo es de suma relevancia agregar la prueba "Deep Blue vs. Kasparov" en la que la computadora Deep Blue de IBM derrota a Garry Kasparov, en una partida histórica. En 1997, la supercomputadora Deep Blue desafió al campeón mundial de ajedrez, Garry Kasparov, en una serie de partidas de ajedrez. La máquina sorprendentemente derrotó a Kasparov en una de las partidas, marcando un hito en la historia de la inteligencia artificial al vencer a un campeón humano en un juego de estrategia altamente complejo. Como menciona Regan (2012), esta victoria provocó un aumento en el interés y la inversión en la inteligencia artificial y el aprendizaje automático, destacando el potencial de estas tecnologías para superar a los seres humanos en tareas cognitivas desafiantes. Además, comenta el autor, esto estimuló el desarrollo de algoritmos y sistemas más avanzados en el campo de la inteligencia artificial, sentando las bases para futuros avances en su aplicación en distintos ámbitos de la vida cotidiana.

Un autor relevante en este contexto particular es Herbert Simon, un destacado psicólogo clínico que ha contribuido significativamente al desarrollo y la promoción de tratamientos basados en la evidencia, incluida la terapia cognitivo-conductual. Fue un psicólogo cognitivo que contribuyó significativamente a la IA y la teoría de la toma de decisiones, y que junto con Allen Newell, desarrolló el programa "Logic Theorist", uno de los primeros programas de IA, y acuñó el término "inteligencia artificial" en 1956. En este sentido, afirmaba que "las máquinas se están volviendo más capaces y la IA se está volviendo más humana a medida que avanzamos" (Simon, 1969, p. 103).

Conjuntamente, Newell trabajó estrechamente con Herbert Simon y desarrolló el lenguaje de programación "IPL" (Information Processing Language) y el programa "General

Problem Solver" (GPS), que sentaron las bases para la resolución de problemas en la IA, afirmando que, en ese entonces, la inteligencia artificial se basaba en "el arte y la ciencia de formar máquinas capaces de realizar tareas que requieren inteligencia cuando son realizadas por personas" (Newell y Simon, 1976, p. 13).

A partir de la década del 2000, la terapia cognitivo-conductual basada en computadora comenzó a ganar popularidad inmensamente. El programa Beating the Blues (BtB) representó uno de los hitos más significativos en la transición hacia la psicoterapia digital durante la primera década del siglo XXI. Según Proudfoot et al. (2003), este sistema fue diseñado para ofrecer una intervención "basada en la evidencia, interactiva y multimedia que permitiera a los pacientes gestionar síntomas de ansiedad y depresión sin la presencia constante de un terapeuta" (Proudfoot et al., 2003, p. 384) . BtB se consolidó como un sistema de Terapia Cognitivo-Conductual Computarizada (TCCC) diseñado para el tratamiento de la depresión leve a moderada y la ansiedad basado en un protocolo estructurado de ocho sesiones interactivas. Los autores describen que el sistema emplea algoritmos de lógica de ramificación para personalizar la experiencia del usuario, integrando videos de casos clínicos, ejercicios de reestructuración cognitiva y tareas para el hogar. La relevancia histórica de BtB radica en su validación clínica ya que fue una de las primeras intervenciones tecnológicas en recibir el aval del National Institute for Health and Care Excellence (NICE). El sistema de salud británico validó oficialmente la tecnología al concluir que la herramienta es "clínicamente eficaz y debe estar disponible para los pacientes en la atención primaria para el manejo de la ansiedad y depresión" (NICE, 2006, p. 4).

La investigación y el trabajo de estos autores han sido fundamentales para comprender cómo la tecnología, incluida la inteligencia artificial, puede utilizarse de manera efectiva en la psicoterapia y cómo la terapia cognitivo-conductual puede ser adaptada y mejorada mediante enfoques digitales. Sus trabajos han contribuido a la evolución de la atención de la salud mental y han allanado el camino para la integración de la IA en la terapia cognitivo-conductual en las décadas siguientes.

Estado del arte

Actualmente, se están llevando a cabo investigaciones para evaluar la eficacia de la IA en la psicoterapia cognitivo-conductual. Esto implica la aplicación de algoritmos de aprendizaje automático para analizar el discurso y el comportamiento del paciente durante las sesiones terapéuticas, así como el desarrollo de aplicaciones y herramientas específicas para ayudar a los terapeutas en su trabajo. Zhou et al. (2022) exploran cómo la IA se ha utilizado en la psicoterapia, incluida la terapia cognitivo-conductual. Discuten aplicaciones específicas, como

chatbots terapéuticos y análisis de texto para evaluar el progreso del paciente, y es en este estudio, que los investigadores exploraron el estado actual de la aplicación de la inteligencia artificial en la psicoterapia y examinaron las técnicas y desarrollos futuros en este campo. A lo largo del estudio se centraron en cómo la inteligencia artificial se ha utilizado para mejorar la eficacia de la psicoterapia, con énfasis en la terapia cognitivo-conductual.

En el panorama actual de la salud mental, el estudio de Zhou et al. (2022) destaca cómo la inteligencia artificial ha dejado de ser una herramienta experimental para convertirse en un apoyo fundamental en la psicoterapia. El estudio concluye que la aplicación de estas tecnologías en la terapia cognitivo-conductual no solo facilita el acceso al tratamiento, sino que ofrece una visión más clara y continua del progreso del paciente, marcando el inicio de una era donde la tecnología y el acompañamiento humano trabajan de la mano para mejorar los resultados clínicos.

Otro ejemplo de investigación actual en este campo es el trabajo de investigadores como John Torous, quien es director del programa de psiquiatría digital en el Beth Israel Deaconess Medical Center y profesor en la Facultad de Medicina de Harvard. Torous (2021) ha estado investigando y desarrollando aplicaciones de inteligencia artificial para la evaluación y el tratamiento de trastornos mentales, abogando por la integración de la tecnología y la IA en la atención de la salud mental. En este sentido, el autor afirma que “la inteligencia artificial y la realidad virtual presentan nuevas oportunidades para la intervención remota” y que “se requieren nuevas técnicas en la ciencia de datos, como la inteligencia artificial y machine learning, para realmente combinar y utilizar grandes bases de datos para informar acerca de cuidados en salud mental” (Torous, 2021, p. 23). Sus investigaciones incluyen el uso de sensores y dispositivos móviles para recopilar datos sobre el comportamiento y el estado emocional de los pacientes, lo que permite, según el autor, un seguimiento continuo fuera de las sesiones terapéuticas tradicionales. Además, su trabajo ha explorado cómo la IA puede analizar estos datos para identificar patrones y proporcionar retroalimentación a los terapeutas.

Afirmaba Torous en una nota periodística::

La IA no impulsa el cuidado ni toma decisiones, pero puede ayudar a plantear nuevas sugerencias e ideas para la discusión. Combina lo que sabemos que funciona muy bien en el cuidado de la salud mental con la nueva tecnología. Y la Inteligencia Artificial ofrece a los pacientes lo mejor de ambos mundos. (citado en Beth Israel Deaconess Medical Center, 2023)

Otros autores relevantes en este contexto y que forman parte de los nuevos descubrimientos en el ambiente de la inteligencia artificial son Albert Rizzo y Giuseppe J. Kim,

dos psicólogos clínicos y profesores en la Universidad del Sur de California. Rizzo y Kim (2005) han estado involucrados en el desarrollo de aplicaciones de IA para la terapia de exposición virtual, una técnica efectiva en la psicoterapia cognitivo-conductual para tratar trastornos de ansiedad y trastorno de estrés postraumático. Según los autores, el trabajo desarrolla cómo la IA puede personalizar y mejorar la efectividad de las exposiciones virtuales, lo que puede ser especialmente valioso para veteranos militares y otras poblaciones con traumas. En este sentido, Rizzo afirma que “nada se lograría al tachar esta nueva tecnología de inútil si la entendemos como una ayuda más al servicio del terapeuta (...) la meta no es reemplazar a los humanos, sino darles herramientas” (en Bohannon, 2015 p. 250-251).

Estos autores y sus investigaciones representan ejemplos de cómo la inteligencia artificial se está utilizando en la psicoterapia cognitivo-conductual en la actualidad. Sus trabajos ayudan a avanzar en la comprensión y la aplicación de la IA en el campo de la salud mental, abriendo nuevas posibilidades para la mejora de la atención terapéutica y la personalización de los tratamientos.

Marco teórico

La inteligencia artificial (IA) es un campo multidisciplinario que abarca diversas ramas de la informática, la matemática, la psicología cognitiva y la filosofía. En este sentido, McCarthy et al. afirman que “en teoría, podría haber muchas cosas diferentes detrás de la inteligencia, y se necesita un gran número de enfoques diferentes para averiguar cómo podría funcionar (1955, p. 45).” La IA se centra en la creación de sistemas y programas informáticos capaces de realizar tareas que, si fueran realizadas por un ser humano, requerirían de inteligencia, implicando así, la capacidad de aprender, razonar, resolver problemas, comprender el lenguaje natural, percibir su entorno y adaptarse a nuevas situaciones (Korf, 2012). El autor también reafirma la postura en la cual la inteligencia artificial es el estudio de cómo hacer que las computadoras realicen cosas que, por el momento, los seres humanos hacen mejor (Knight y Rich, 1994).

Otro aporte importante es el del “aprendizaje automático”, que se basa en el concepto de que, según Andrew Ng “las máquinas pueden aprender de datos y experiencias para mejorar su rendimiento en tareas específicas” (2018, p. 24). El autor explica que los algoritmos de aprendizaje automático, como las redes neuronales artificiales y los algoritmos de aprendizaje profundo, son ejemplos de cómo las máquinas pueden aprender patrones a partir de datos y mejorar su desempeño con la práctica. Ng afirma:

El aprendizaje automático es una rama de la inteligencia artificial que se ha convertido en una herramienta poderosa para analizar datos y hacer predicciones en una variedad de campos, desde la medicina hasta la tecnología financiera y la ciencia de datos. (Ng, 2018, p. 74)

Asimismo, la Teoría de la toma de decisiones, otro aporte fundamental dentro del marco teórico, se enfoca en cómo las máquinas pueden tomar decisiones óptimas o razonables en situaciones complejas, según Simon (1955). Para el autor, esta teoría ha dado lugar al desarrollo de algoritmos de toma de decisiones, como los árboles de decisión y los algoritmos de planificación, que son fundamentales en la IA. Simon afirmaba que "los seres humanos no pueden considerar todas las alternativas posibles antes de tomar una decisión, su racionalidad está limitada por la información, el tiempo y la capacidad cognitiva" (1955, p. 94).

Retomando lo planteado previamente en el estudio de Zhou et al. (2022), los autores destacan la creciente popularidad de los chatbots terapéuticos, que son programas de inteligencia artificial diseñados para interactuar con los pacientes en un formato de conversación. Como los autores explican, estos chatbots pueden proporcionar intervenciones basadas en la terapia cognitivo-conductual, ofreciendo ejercicios y apoyo emocional. También se analizó cómo la inteligencia artificial se utiliza para el análisis de texto en el contexto de la psicoterapia. Esta técnica permite evaluar, según los autores, el progreso del paciente a lo largo del tiempo mediante el análisis de los cambios en el lenguaje y la expresión emocional, teniendo en cuenta que uno de los beneficios clave de la inteligencia artificial en la terapia cognitivo-conductual es su capacidad para personalizar las intervenciones. Asimismo, los programas de IA pueden adaptar las estrategias terapéuticas según las necesidades individuales de cada paciente. En este sentido, los autores afirmaban que "tecnologías usadas como machine learning fueron implementadas para resolver problemas de salud mental; los datos emergentes en este campo proveen mayor entendimiento sobre la aplicación de tecnologías de inteligencia virtual en tratamientos psicológicos" (Zhou et al., 2022, p. 2).

Según Ulric Neisser (1967), la ciencia cognitiva se establece como el estudio interdisciplinario de la mente y sus procesos, fundamentándose en la premisa de que la cognición es, esencialmente, computación sobre representaciones mentales. Para el autor, el término cognición se refiere a todos los procesos mediante los cuales la entrada sensorial es transformada, reducida, elaborada, almacenada, recuperada y utilizada. Este enfoque permite que la inteligencia artificial se integre de forma natural en la psicología, ya que ambas disciplinas comparten el objetivo de comprender cómo un sistema (biológico o artificial) procesa la información para resolver problemas y tomar decisiones.

De acuerdo con Ingram y Kendall (1986), esta concepción de la mente como un sistema de procesamiento de datos establece el puente hacia la psicología clínica, al proponer que el bienestar emocional va de la mano con la precisión y eficacia con la que el sujeto procesa la información. En este sentido, los autores describen que "la psicopatología cognitiva puede definirse como el estudio de las estructuras y procesos cognitivos que operan de manera desadaptativa [...] vinculando la arquitectura del procesamiento de información con el funcionamiento psicológico" (Ingram y Kendall, 1986, p. 4).

Por el lado de la psicología propiamente dicha, encontramos y nos basamos en la Psicoterapia Cognitivo-Conductual la cual, según Beck (1979), afirma ser un enfoque terapéutico que se basa en la premisa fundamental de que nuestros pensamientos, emociones y comportamientos están interconectados y desempeñan un papel crucial en nuestra experiencia psicológica. La terapia cognitiva se centra en la identificación y evaluación de los pensamientos y creencias que contribuyen a la angustia emocional y al comportamiento disfuncional, en el que el objetivo es reemplazar estas cogniciones negativas con pensamientos más racionales y adaptativos, lo que lleva a una reducción de la angustia y una mejora en la salud mental del paciente (Beck, 1979).

En este modelo, se presta una atención especial a la identificación y modificación de los pensamientos irracionales o distorsionados que pueden dar lugar a emociones negativas y comportamientos problemáticos, según afirma el autor. Beck (1979) propuso que las personas tienen sistemas de creencias cognitivas que influyen en la percepción de sí mismas, de los demás y del mundo, y que la psicopatología a menudo se origina en patrones de pensamiento disfuncionales.

La terapia cognitivo-conductual, por tanto para el autor, se centra en ayudar a los individuos a reconocer y reevaluar estos patrones de pensamiento negativo, reemplazándolos por pensamientos más realistas y saludables, lo que a su vez conduce a una mejora en el bienestar emocional y en la capacidad para afrontar los desafíos de la vida cotidiana. Además, para Beck (1979), este enfoque también implica la implementación de estrategias de cambio conductual para consolidar las modificaciones cognitivas, lo que contribuye a la reducción de síntomas y al logro de metas terapéuticas. La Psicoterapia Cognitivo-Conductual de Beck, entonces, se ha utilizado en el tratamiento de una amplia variedad de trastornos mentales, incluyendo la depresión, la ansiedad, los trastornos de la alimentación y otros trastornos relacionados con el estado de ánimo y la conducta.

Desarrollo metodológico

Procedimiento

Las principales ideas que se analizan en este trabajo son las siguientes: la aplicación de la inteligencia artificial en la psicoterapia cognitivo-conductual, los beneficios de la IA en la evaluación del estado del paciente, el incentivo en la autonomía del paciente para la toma de decisiones y la mejora de la eficacia del tratamiento. Asimismo, se analizan las implicaciones y limitaciones del uso de la inteligencia artificial en la terapia cognitivo-conductual y las perspectivas futuras para la investigación en este campo.

Esta investigación de revisión bibliográfica bajo carácter exploratorio se llevó a cabo a través de búsquedas bibliográficas utilizando la bases de datos de Google Académico, PubMed, y Scielo. Las palabras claves que se utilizaron para esta búsqueda fueron: inteligencia artificial, psicoterapia cognitivo-conductual, evaluación del progreso del paciente, autonomía del paciente, mejora de la eficacia del tratamiento, redes neuronales artificiales, procesamiento del lenguaje natural, sistemas expertos. La búsqueda referida se realizó en 2025 y hemos optado por seleccionar información de múltiples periodos para lograr una visión más amplia de los distintos tipos de aplicación que tuvo y tiene hoy en día la inteligencia artificial en la terapia cognitivo conductual.

Se trabajará principalmente retomando los conceptos de Aaron Beck en cuanto a lo teorizado sobre terapia cognitivo conductual, y se tomarán autores enfocados a la aplicación de inteligencia artificial sobre este tipo de terapia, como Weizenbaum, Zhao, Newell, Simon, Torous, Rizzo, y otros autores contemporáneos.

Índice comentado

Capítulo 1: La inteligencia artificial en la terapia cognitivo-conductual: herramientas digitales y aplicaciones clínicas

Este capítulo tiene como objetivo describir y analizar las principales aplicaciones de la inteligencia artificial en el marco de la terapia cognitivo-conductual (TCC). Se presenta un recorrido por las herramientas digitales más relevantes actualmente disponibles, incluyendo sistemas conversacionales, tecnologías inmersivas, modelos multimodales y plataformas de apoyo clínico y administrativo. Asimismo, se examina la evidencia empírica que respalda su utilización, así como su potencial para ampliar el acceso a la atención psicológica, optimizar los procesos terapéuticos y complementar el trabajo del profesional. El capítulo adopta una perspectiva descriptiva y analítica, orientada a comprender de qué manera estas tecnologías se integran a la práctica clínica contemporánea y cuáles son sus alcances dentro del modelo cognitivo-conductual.

Capítulo 2: Desafíos éticos y prácticos de la inteligencia artificial en salud mental

Este capítulo aborda de manera crítica los principales desafíos éticos, conceptuales y prácticos que emergen de la incorporación de la inteligencia artificial en el ámbito de la salud mental. Se analizan problemáticas vinculadas a la privacidad y confidencialidad de los datos, la transparencia y trazabilidad de los sistemas algorítmicos, la empatía y la calidad del vínculo terapéutico mediado por tecnología, la autonomía del paciente y los riesgos de manipulación, así como los sesgos algorítmicos y las limitaciones actuales de los sistemas automatizados. El capítulo pone énfasis en la necesidad de una supervisión humana adecuada y en el desarrollo de marcos éticos y teóricos que permitan una integración responsable de la IA en la práctica clínica, garantizando la protección de los derechos de los pacientes y la calidad del cuidado psicológico.

Capítulo 1: La inteligencia artificial en la terapia cognitivo-conductual: herramientas digitales y aplicaciones clínicas

1.1 Chatbots terapéuticos basados en TCC: el caso de Woebot como herramienta de apoyo clínico

Para comenzar con aquellas aplicaciones que aportan de gran manera, ya sea para el terapeuta y el paciente, está la reconocida Woebot, un chatbot de IA basado en terapia cognitivo-conductual (TCC), que ha irrumpido en el mundo de la salud mental ofreciendo una herramienta innovadora para apoyar tanto a terapeutas como a pacientes. Según Darcy et al. (2017), Woebot funciona como un complemento valioso de la terapia tradicional que permite automatizar tareas repetitivas, optimizando el tiempo del terapeuta para interacciones clínicas más complejas y mejorando la accesibilidad al tratamiento. Woebot puede automatizar tareas repetitivas como programar citas, enviar recordatorios y responder preguntas frecuentes, liberando tiempo valioso para que los terapeutas se enfoquen en interacciones más complejas con los pacientes.

Asimismo, puede monitorear el progreso de los pacientes entre sesiones, recopilando datos sobre sus estados de ánimo, síntomas y cumplimiento de tareas. Esta información brinda a los terapeutas una visión más completa del estado del paciente, permitiéndoles ajustar las intervenciones según sea necesario. Alison Darcy et al (2017), al respecto, comentan:

Los datos que hemos recopilado son un reflejo significativo, rico y ecológicamente válido de la experiencia vivida por los usuarios, es decir, un mapeo del mundo real de su experiencia. Nuestros científicos pueden identificar sistemáticamente factores, tanto cognitivos, emocionales, conductuales y sociales, que interfieren con su proceso terapéutico. (en Woebot Health, 2021)

Woebot puede ofrecer terapia a personas que de otra manera no podrían acceder a ella, como aquellas que viven en áreas remotas o que tienen dificultades para programar citas presenciales. Esto, para los autores, democratiza la atención de salud mental y la hace más inclusiva. El beneficio que otorga esta herramienta recae en su disponibilidad las veinticuatro horas del día, los siete días de la semana, lo que permite a los pacientes acceder a apoyo cuando más lo necesitan, incluso fuera del horario de oficina habitual. Woebot se puede usar desde cualquier lugar con una conexión a internet, lo que lo convierte en una opción conveniente para personas con agendas ocupadas o con dificultades para desplazarse (Woebot Health, 2023).

Una investigación realizada por Durden et al. (2023) se propone evaluar aplicaciones inteligentes que ofrecen apoyo en salud mental, mas específicamente Woebot, si es que están asociadas con cambios en el estrés, el agotamiento y la resiliencia durante ocho semanas. La exposición a Woebot, según la investigación, se asoció con reducciones significativas en el estrés percibido y el agotamiento, así como un aumento significativo en la resiliencia durante el período del estudio. Se observó en los resultados una mayor reducción en el estrés entre aquellos con síntomas de ánimo clínicamente elevados (es decir, puntuaciones ≥ 10 en el Cuestionario de Salud del Paciente-8 o en el Cuestionario de Ansiedad Generalizada de 7 ítems) al inicio en comparación con aquellos sin ellos.

Asimismo, Darcy comenta en una entrevista que necesitamos herramientas en la primera línea que liberen a los terapeutas humanos para los pacientes que realmente necesitan ver a una persona y suma que hay mucha evidencia que respalda esta práctica, que se llama atención por etapas (citado en Mullin, 2017). Esto está relacionado con el contexto actual del sistema de salud mental mundialmente, el cual Darcy reflexiona:

"Hay un problema de oferta en el sentido de que no hay suficientes clínicos. Al mismo tiempo, hay una creciente crisis de salud mental. Muchas personas están sufriendo, con recursos insuficientes para ayudarlas. Para abordar esa necesidad, sentí que era necesario contar con herramientas digitales más sólidas" (citado en Cassata, 2018).

1.2 Chatbots terapéuticos orientados al bienestar emocional: el caso de Wysa y la intervención subclínica en TCC

Otro chatbot que está en tendencia es Wysa. Fue lanzada en 2016 por su fundadora, Jo Aggarwal, quien, lidiando con su propia depresión, descubrió que era, según sus palabras, más fácil hablarle a un bot en esas situaciones. Según su creadora, Wysa es "un rastreador del estado de ánimo, un entrenador de mindfulness, un ayudante para situaciones de ansiedad y un colega para levantar el estado de ánimo, todo en uno" (Aggarwal et al., 2023). Según un estudio hecho por Oracle & Workplace Intelligence (2020), los datos han demostrado que el 82% de las personas prefieren revelar sus problemas de salud mental a un confidente no humano, es decir, un robot. Argumenta que aumentar la accesibilidad en un entorno privado y anónimo ayuda a superar las barreras creadas por el estigma en torno a la salud mental. Asimismo, el estudio reveló que el 29% de las personas encuestadas afirmó que prefiere la IA porque ofrece respuestas rápidas a preguntas sobre salud mental, algo que los sistemas tradicionales humanos no siempre pueden garantizar (Oracle & Workplace Intelligence, 2020).

Aggarwal et al (2023) a partir de esto, al apoyar el amplio espectro de atención en salud conductual, mencionan que Wysa evita la necesidad de intervención humana a nivel

subclínico, al tiempo que brinda apoyo a las poblaciones clínicas con mayores necesidades mediante un acceso ampliado a coaches y terapeutas humanos. En este sentido, el contexto que Aggarwal et al. (2023) reconocen a cerca de la salud mental implicaba “crear algo que abordara la salud mental, sabíamos que no hay suficientes terapeutas, por ejemplo, en la India hay alrededor de 5.000 terapeutas/psiquiatras en total para 1,3 mil millones de personas” (en Vempati, 2020).

Sobre los beneficios que brinda la aplicación, Wysa puede realizar un seguimiento del estado emocional de los pacientes a lo largo del tiempo, ya que los terapeutas pueden acceder a datos sobre cómo están sus pacientes entre sesiones, lo que les permite ajustar el enfoque terapéutico según sea necesario y ofrecer intervenciones más personalizadas (Aggarwal et al., 2023). Wysa puede enviar recordatorios y sugerencias para practicar técnicas de afrontamiento específicas, como la reestructuración cognitiva o la exposición gradual, ayudando así a los pacientes a consolidar y generalizar las habilidades aprendidas en terapia.

Cuando se le pregunta a Aggarwal et al. (2023) acerca de cuáles eran sus intenciones con respecto a su nuevo proyecto Wysa, reflexionan que el objetivo era “construir algo similar a Eliza, pero más sabio, algo que pudiera escuchar a una persona y guiarla de manera efectiva para reformular sus pensamientos negativos” (en Wysa, 2020). Asimismo, la creadora de Wysa argumenta que “descubrimos que las personas no querían que sus seres queridos o sus médicos supieran que podrían tener depresión o ansiedad, y no estaban listos para ver a un terapeuta” (en Kumar, 2021).

A través de este objetivo y teniendo en cuenta hacia quienes estaba dirigido la aplicación, Aggarwal tenía como misión trabajar sobre el contexto en el que ella veía inmersa la salud mental actual: ofrecer a cada individuo acceso equitativo al nivel más adecuado de atención en salud mental posible, generar resultados clínicos que reduzcan el gasto en atención en salud mental, y por último, reducir el absentismo y el presentismo que, en sus palabras sería “proporcionando a nuestros socios empleadores las herramientas y recursos necesarios para satisfacer las necesidades de salud mental en todos los niveles, a lo largo del espectro de atención” (en Wysa, 2022).

En cuanto a cómo es utilizada la aplicación por sus usuarios, el equipo de investigadores de Aggarwal llegó a la conclusión que:

En Wysa, reconocemos que los usuarios anhelan ser escuchados en lugar de que les hablen. Al proporcionar a los usuarios una herramienta de apoyo disponible las 24 horas del día que les permite sentirse escuchados y valorados, hemos observado un

crecimiento orgánico de 4 millones de usuarios, logrado sin ningún gasto en publicidad. (Wysa, 2022)

Es en este sentido que Aggarwal identificó una variable muy importante a la hora de la interacción entre un paciente y un terapeuta, y a la vez, entre un usuario y un chatbot terapéutico: la dependencia. La creadora de Wysa desarrolla:

Todavía es sorprendente pensar que muchas personas no están dispuestas a hablar con otro ser humano porque temen la dependencia, el juicio o la debilidad, pero están dispuestas a hablar con un robot, y si ese robot puede parecer compasivo, como está diseñado Wysa para hacerlo, eso puede marcar toda la diferencia. (citado en Forbes, 2021)

La misión social por parte de Aggarwal con su aplicación Wysa recae en la destigmización de la salud mental: desestigmatizar la búsqueda de ayuda emocional al ofrecer un medio accesible y discreto para recibir apoyo. Para la creadora, esto puede alentar a más personas a buscar ayuda cuando la necesiten, incluso si tienen barreras para acceder a la terapia tradicional. Aggarwal añade que los terapeutas “han comenzado a recomendar Wysa entre sesiones a sus pacientes, así también como las escuelas a sus estudiantes que pueden no querer acudir al terapeuta, pero al menos pueden chatear con el bot” (en Kumar, 2021).

1.3 Terapia de exposición con realidad virtual (VRET): aplicaciones de la realidad virtual en la terapia cognitivo-conductual

A la par del desarrollo de chatbots terapéuticos, la tecnología ha abierto paso a herramientas de realidad aumentada para la enseñanza. Un ejemplo clave es la realidad virtual: aunque existen distintos artefactos para acceder a ella, las gafas han liderado el sector como el dispositivo pionero y más avanzado hasta la fecha.

Distintas aplicaciones hacen uso de este dispositivo, en este sentido hablamos de VRET (VR exposure therapy) que por sus siglas hacen alusión a un tipo de terapia de exposición mediante la utilización de distintos tipos de realidad virtual. En este sentido, Boeldt et al. la describen como “la exposición individualizada, gradual, controlada e inmersiva que es fácil de implementar para los terapeutas y, a menudo, más aceptable para los pacientes que la exposición in vivo o imaginaria” (2019, p. 4). En este sentido, los autores argumentan que esto puede permitir a los usuarios practicar habilidades conductuales en un entorno seguro con el apoyo de un terapeuta. La VRET se ha utilizado para el tratamiento de una variedad de

trastornos, incluyendo la ansiedad social, el Trastorno de Estrés Postraumático (TEPT) y el trastorno de pánico.

Los distintos usos, la personalización, y la inmensa adaptabilidad que posee esta herramienta para la utilización en terapias de exposición hacen de ésta un gran sostén para las nuevas dinámicas terapéuticas (Boeldt et al., 2019). En este sentido, los autores se apoyan en que probablemente en un futuro, la realidad virtual podría ayudar a proporcionar capacitación clínica estandarizada en terapia de exposición, haciendo que la formación sea más fácil y accesible, así como también un entorno de entrenamiento en VR permitiría a los terapeutas proveer tratamientos basados en evidencia y contrarrestar los miedos y dificultades que les podría aparecer en una terapia. Asimismo, los autores argumentan que:

En la práctica, la realidad virtual (VR) es cada vez más accesible en términos de costos. El precio del software y hardware de VR sigue disminuyendo, al mismo tiempo que la cantidad y calidad del contenido de VR aumenta. La investigación sobre la Terapia de Exposición con Realidad Virtual ha proliferado, explorando su uso en una variedad de trastornos de salud mental, especialmente trastornos de ansiedad como fobias. (Boeldt et al, 2019, p. 1)

Su practicidad y su capacidad de control otorga herramientas que favorecen el trabajo del terapeuta. En este sentido, los autores desarrollan:

Con VRET, el terapeuta puede seleccionar el contenido de realidad virtual y personalizarlo para el paciente. Los terapeutas pueden guiar a los pacientes a través de la exposición en la consulta mientras supervisan y apoyan al paciente. Los pacientes se sienten comprometidos, y la experiencia se siente "real", pero es una forma segura para que un paciente practique antes de enfrentarse a un estímulo temido por su cuenta en un entorno del mundo real. (Boeldt et al, 2019, p. 10)

1.4. Inteligencia artificial multimodal aplicada a la evaluación y el seguimiento terapéutico en TCC

La Inteligencia Artificial Multimodal (IA-MM) representa una innovación relevante en el campo de la psicoterapia, al permitir una comprensión más profunda y holística del estado mental de los pacientes mediante la integración simultánea de diversas fuentes de datos provenientes de la interacción terapéutica (Kim et al., 2025). A diferencia de los sistemas conversacionales tradicionales basados únicamente en texto, la IA multimodal combina el análisis del lenguaje con señales no verbales y fisiológicas, lo que amplía las posibilidades de diagnóstico, monitoreo y personalización de las intervenciones terapéuticas.

En esta línea, Döllinger et al. (2021) sostienen que el reconocimiento de las emociones no verbales resulta esencial para una comunicación efectiva y constituye una competencia central en la práctica clínica, directamente relacionada con la capacidad del terapeuta para brindar una terapia eficaz.

Kim et al. (2025) desarrollaron la base de datos denominada Multi Modal–Cognitive Support Conversation (M2CoSC), que asocia cada diálogo generado por el modelo GPT-4 con una imagen que representa las expresiones faciales del cliente virtual. Según los autores:

Para reflejar mejor la psicoterapia real, en la que las expresiones faciales conducen a interpretar evidencias emocionales implícitas, proponemos un enfoque de razonamiento psicoterapéutico multinivel, que identifica e incorpora de manera explícita las evidencias sutiles. (Kim et al., 2025, p. 1).

Los resultados de la investigación realizada por Subin Kim, Hoonrae Kim, Heejin Do y Gary Geunbae Lee evidencian mejoras significativas en las capacidades terapéuticas de los Modelos de Lenguaje Visual (VLMs) al emplear la base de datos M2CoSC. Los autores concluyen que “la estrategia propuesta de razonamiento psicoterapéutico multinivel, que integra expresiones faciales, pensamientos y distorsiones cognitivas, demuestra un rendimiento superior al ofrecer a los clientes sugerencias racionales empáticas, lógicamente coherentes y específicas” (Kim et al., 2025, p. 9).

La personalización constituye un componente esencial de estas tecnologías. En este sentido, Alasmrai et al. destacan que “la personalización es crucial porque los sistemas de aprendizaje automático utilizan los datos del paciente para adaptar los programas de tratamiento a sus necesidades psicológicas y objetivos terapéuticos” (2025, p. 3). Además, enfatizan que los modelos de IA “aprenden constantemente de los pacientes adultos y mejoran sus tratamientos para optimizar los resultados” (Alasmrai et al, 2025, p. 3).

Para los autores, la Inteligencia Artificial Multimodal (IA-MM) amplía los alcances del análisis tradicional del lenguaje verbal, integrando no solo lo que el paciente expresa, sino también cómo lo expresa. Este enfoque posibilita una evaluación más completa de la congruencia emocional durante el proceso terapéutico. En este sentido, Chu et al. (2025) desarrollaron un conjunto de datos denominado MESC, enriquecido con anotaciones que incluyen emociones y estrategias, abarcando las modalidades de texto, audio y video. El MESC (Multimodal Emotional Support Conversation) es un conjunto de datos diseñado para entrenar sistemas de IA en el reconocimiento de emociones y la aplicación de estrategias de apoyo emocional mediante texto, audio y video. Dicho conjunto de datos no solo permite el reconocimiento emocional, sino también la provisión de apoyo emocional, aspectos clave para la integración de la IA conversacional en aplicaciones de salud mental (Chu et al., 2025).

Los autores también proponen el Marco General Secuencial Multimodal de Apoyo Emocional (SMES), que “aprovecha las fortalezas de los modelos fundamentales multimodales para extraer señales emocionales del video y el audio” (Chu et al., 2025, p. 2). Este marco utiliza un modelo de razonamiento basado en Modelos de Lenguaje Extensos (LLM), capaz de generar de manera secuencial los resultados de múltiples tareas, como el reconocimiento de emociones del usuario, la predicción de estrategias y emociones del sistema, y la generación de respuestas, mediante un proceso de entrenamiento multitarea de máxima verosimilitud, lo que optimiza de manera integral el diálogo terapéutico (Chu et al., 2025).

Asimismo, el SMES se complementa con sistemas de conversación emocional, compuestos por tareas como el Reconocimiento de Emociones en la Conversación (ERC), la conversación emocional y la conversación empática. Según Chu et al. (2025), el ERC tiene como objetivo identificar y rastrear automáticamente los estados emocionales de los hablantes durante los diálogos, a partir de señales multimodales, como las expresiones faciales, la entonación de la voz y los gestos. Estas señales permiten una comprensión más precisa del estado psicológico del usuario, incrementando la capacidad de la IA para simular un apoyo emocional análogo al humano. Los autores concluyen que “el SMES mejora significativamente las capacidades empáticas y estratégicas de la IA, estableciendo un nuevo referente para la inteligencia artificial conversacional en el apoyo a la salud mental” (Chu et al., 2025, p. 9).

La incorporación de datos provenientes de dispositivos portátiles, como relojes inteligentes y pulseras de actividad, permite un monitoreo continuo y objetivo del estado del paciente fuera de la sesión terapéutica. De acuerdo con Kargarandehkordi et al., “la integración de dispositivos mejora la Terapia Cognitivo-Conductual (TCC) al registrar las condiciones físicas en tiempo real” (2025, p. 1).

Los autores explican que el monitoreo de síntomas puede optimizarse mediante el registro constante de señales fisiológicas:

Los dispositivos portátiles suelen estar equipados con sensores de movimiento (acelerómetros y giroscopios), biosensores (frecuencia cardíaca, actividad electrodérmica), sensores ambientales (presión, temperatura) y tecnologías de comunicación como Wi-Fi o Bluetooth. (Kargarandehkordi et al., 2025, p. 1).

El procesamiento mediante IA de estos datos permite identificar patrones sutiles asociados a diferentes condiciones psicológicas, como episodios de ansiedad o recaídas en trastornos afectivos. Dicha información resulta útil en distintas etapas del proceso terapéutico: desde la detección temprana de factores de riesgo, pasando por la orientación y seguimiento del tratamiento, hasta el acompañamiento en la recuperación (Haberer et al., 2013).

Según Mohr et al. (2017), los algoritmos transforman los datos brutos (frecuencia cardíaca o movimiento) en marcadores conductuales que pueden reflejar estados de fatiga, estrés o placer. Estos marcadores se utilizan para inferir síntomas y predecir posibles estados clínicos, incrementando la precisión diagnóstica. Complementariamente, Goessl et al. (2017) sostienen que el *biofeedback*, al enseñar a los individuos a regular funciones fisiológicas involuntarias, puede fortalecer la autorregulación emocional y reducir efectos negativos.

Finalmente, Kargarandehkordi et al. (2025) destacan que la integración de IA con tecnologías de biosensado ha incrementado notablemente la aplicabilidad de los biosensores en salud mental. Los algoritmos de aprendizaje automático se utilizan actualmente para detectar el estrés a partir de la variabilidad de la frecuencia cardíaca, identificar la depresión mediante cambios en la actividad física e inferir estados emocionales a través de la actividad electrodérmica, consolidando así un nuevo paradigma de terapia digital basada en datos fisiológicos en tiempo real.

1.5. Intervenciones Adaptativas Justo a Tiempo (JITAI): un marco metodológico para la integración de la IA en la TCC

El avance en el desarrollo de la Inteligencia Artificial y su aplicación en la terapia ha encontrado un marco metodológico fuerte en las Intervenciones Adaptativas Justo a Tiempo (JITAI, por sus siglas en inglés) (Just-in-Time Adaptive Interventions). Este diseño de intervención representa la evolución de la asistencia terapéutica, enfocándose en la adaptación continua para proporcionar la cantidad y el tipo de apoyo adecuados en el momento preciso (Nahum-Shani et al., 2018).

El principal aspecto de las JITAI es la capacidad de amoldarse al estado contextual e interno del individuo, el cual puede cambiar de forma rápida, inesperada y dentro de su entorno natural, la disponibilidad de potentes tecnologías móviles y de detección es lo que permite que esta adaptación tenga lugar en tiempo real (Nahum-Shani et al., 2018). El éxito de una JITAI depende de un entendimiento claro de su teoría del cambio, la cual define cuándo es el momento oportuno para intervenir. Este enfoque se basa en la eliminación de la ineficiencia en el proceso terapéutico: el criterio no solo es entregar la ayuda correcta en el momento adecuado, sino también evitar el desperdicio o al no proporcionar ninguna intervención cuando el momento no es oportuno o al no ofrecer un tipo de apoyo correcto (Nahum-Shani et al., 2018). Para los autores, el desperdicio se concibe como cualquier acción que absorbe recursos (tiempo, esfuerzo) pero que no añade valor o que, incluso, podría interrumpir el proceso de cambio.

En el contexto de las intervenciones de comportamiento de salud, la necesidad de ayuda justo a tiempo se justifica en el hecho de que los estados de vulnerabilidad a eventos

adversos (como la recaída) o los estados de oportunidad para cambios positivos emergen de manera impredecible y fuera de los entornos clínicos tradicionales. En este contexto, la intervención tiene como objetivo brindar la ayuda necesaria en el momento oportuno, con el fin de romper la conexión entre los desencadenantes situacionales, el surgimiento del estado de vulnerabilidad y su progresión hacia una recaída sintomática completa (Nahum-Shani et al., 2018).

Dado que los estados emocionales y conductuales son altamente dinámicos y cambiantes, para los autores, resulta poco viable ofrecer apoyo oportuno a través de intervenciones presenciales tradicionales. En este contexto, la posibilidad de brindar asistencia justo a tiempo depende del uso de tecnologías móviles e inalámbricas, un campo denominado *mHealth* (salud móvil).

Los avances tecnológicos recientes han posibilitado la monitorización continua de los usuarios, permitiendo detectar de manera temprana la aparición de estados de vulnerabilidad o de oportunidad para la intervención. Para los autores, esta capacidad se sustenta en la integración de diferentes fuentes de información, que combinan datos objetivos y subjetivos (Nahum-Shani et al., 2018):

- Sensores de dispositivos portátiles (wearables): tales como monitores de actividad física y relojes inteligentes, que registran parámetros fisiológicos y de movimiento.
- Detección basada en el teléfono móvil: a través del uso de acelerómetros y sistemas de posicionamiento global (GPS), que permiten inferir patrones de comportamiento y ubicación.
- Huellas digitales: derivadas de la interacción del usuario con redes sociales, calendarios u otras plataformas digitales, que ofrecen indicios sobre su estado emocional o nivel de actividad.
- Autoinformes de bajo esfuerzo: mediante métodos como la Evaluación Ecológica Momentánea (EMA), que recopilan información autodeclarada en el contexto natural del individuo y minimizan la carga de respuesta.

Un ejemplo destacado de las Intervenciones Adaptativas Justo a Tiempo (JITAI) en el ámbito de la salud es ACHES (Addiction–Comprehensive Health Enhancement Support System), un sistema diseñado para acompañar la recuperación de personas con trastornos por consumo de alcohol.

En este sentido, los autores enuncian:

ACHESS combina la asistencia terapéutica digital con la detección contextual, y brinda acceso continuo (las 24 horas del día, los 7 días de la semana) a servicios de ayuda mediante un teléfono inteligente. Integra componentes de terapia cognitivo-conductual (TCC) computerizada, junto con recursos comunitarios y de acompañamiento social. (Nahum-Shani et al., 2018, p. 2)

Su inteligencia contextual utiliza el GPS para identificar la proximidad del usuario a ubicaciones de alto riesgo previamente definidas (por ejemplo, un bar al que iba con frecuencia). Cuando el sistema detecta que el individuo se aproxima a uno de estos lugares, envía una alerta personalizada que lo invita a reflexionar sobre su intención y le ofrece estrategias de afrontamiento o apoyo inmediato. De este modo, la intervención se produce en el momento preciso del riesgo potencial, lo que incrementa su eficacia y relevancia.

Este ejemplo ilustra cómo las tecnologías móviles permiten implementar intervenciones adaptativas con alta precisión temporal y contextual, lo que favorece una atención más dinámica y personalizada.

El marco teórico de las JITAI ofrece una base especialmente adecuada para la incorporación de la inteligencia artificial en la terapia cognitivo-conductual (Nahum-Shani et al., 2018). Los algoritmos de IA pueden procesar de manera continua los datos obtenidos de tecnologías mHealth, detectar patrones de vulnerabilidad emocional o cognitiva en tiempo real y desplegar microintervenciones personalizadas, como recordatorios de reestructuración cognitiva o ejercicios de mindfulness, en el momento en que resultan más necesarios.

De esta forma, las JITAI no constituyen únicamente un tipo de intervención, sino un paradigma de diseño que optimiza la eficacia, la personalización y la relevancia ecológica de la terapia digital. Representan un cambio de enfoque en la prestación del apoyo psicológico, orientado hacia la respuesta inmediata, contextualizada y ajustada a las necesidades del individuo en su vida cotidiana.

1.6. Inteligencia artificial en la gestión clínica: automatización administrativa y optimización del ejercicio profesional en psicología

En su rol administrativo, la inteligencia artificial asiste al profesional de la psicología mediante la automatización de tareas rutinarias, como la programación y gestión de citas, la organización de agendas y la administración de historias clínicas. Estas herramientas también pueden transcribir, resumir o registrar sesiones de manera automática, así como gestionar la facturación y documentación del consultorio. En este contexto, la IA cumple una función

análoga a la de un asistente técnico dentro del ámbito clínico, lo que permite al psicólogo disponer de más tiempo para el trabajo terapéutico y la atención directa de los pacientes.

Entre las soluciones disponibles en esta categoría se destacan Eholo, Mentalyc, Texta AI y Menta, que representan diferentes enfoques hacia la optimización de la gestión profesional y la integración de la tecnología en el ejercicio clínico.

Eholo es una plataforma desarrollada en España que se ha consolidado como una de las herramientas administrativas más utilizadas por psicólogos hispanohablantes, con más de 10.000 profesionales registrados, lo que refleja su alta aceptación entre quienes buscan optimizar la gestión de sus consultas.

Según sus fundadores, Sonia Torras y Eduard Martí (2025), el proyecto surgió al advertir que *“los psicólogos perdíamos demasiado tiempo con tareas administrativas, cuando lo que realmente queríamos era dedicar nuestro tiempo a atender mejor a nuestros pacientes”*. A partir de esta necesidad, Eholo integró tecnología de inteligencia artificial con el propósito de hacer la gestión diaria de los consultorios más fácil, eficiente y segura, centralizando todas las funciones en una única plataforma.

El software, ofrecido bajo el modelo Software as a Service (SaaS), incluye un conjunto amplio de funcionalidades orientadas a los profesionales de la salud mental. Además de las herramientas clásicas de agenda, facturación y administración de pacientes, Eholo permite gestionar historias clínicas, elaborar cuestionarios y contratos terapéuticos, y generar resúmenes automáticos de sesiones. También incorpora transcripción automática de audios y videollamadas, integrando IA de forma práctica en el flujo de trabajo clínico.

El crecimiento de Eholo ha sido notable. En pocos años, la startup pasó de no registrar ventas a atender más de 250 clínicas de psicología y alcanzar los 13.000 usuarios activos en su plataforma hacia junio de 2025, según lo informado por La Vanguardia (2025).

En el contexto latinoamericano, la plataforma argentina Menta no solo simplifica la gestión del consultorio, sino que además integra funciones basadas en IA que permiten conectar a pacientes con psicólogos y facilitar los procesos administrativos. Por ejemplo, en marzo de 2025 se reportó que Menta ya había registrado más de 1.000 usuarios consultantes, 755 profesionales y 545 asignaciones entre pacientes y psicólogos, datos que reflejan su rápida adopción en el mercado argentino (La Nación, 2025). La plataforma incorpora herramientas como agenda inteligente, historial clínico digital y sistema de pagos, junto con un asistente virtual potenciado por IA denominado “Menty”, que sirve como puente inicial al profesional humano (La Nación, 2025).

Capítulo 2: Desafíos éticos y prácticos de la inteligencia artificial en salud mental

2.1 Privacidad, confidencialidad y protección de datos en el contexto terapéutico

Ahora bien, no todo es perfecto en estos tiempos modernos. De acuerdo con un estudio realizado por De la Fuente et al. (2024), “una de las principales preocupaciones de los psicólogos al introducir la IA en la terapia es garantizar la privacidad de la información en el contexto terapéutico” (p. 14). Los autores sostienen que esta cuestión resulta especialmente relevante debido al carácter altamente sensible de la información recolectada durante el proceso terapéutico, ya que cualquier vulneración podría comprometer tanto el vínculo terapéutico como el cumplimiento del código de ética profesional (De la Fuente et al., 2024).

En este sentido, los autores destacan que herramientas como Woebot y Wysa “han implementado medidas claras de protección de datos, garantizando que la información personal de los usuarios sea tratada de acuerdo con las normativas de privacidad” (De la Fuente et al., 2024, p. 22). No obstante, en relación con otras herramientas orientadas a funciones administrativas, se han identificado brechas vinculadas a la falta de información detallada sobre la gestión de los datos. De la Fuente et al. (2024) señalan que, si bien plataformas como Eholo y Texta AI declaran cumplir con las normativas de privacidad en sus sitios web, no brindan información suficiente que permita demostrar dicho cumplimiento de manera clara, lo que puede generar dudas entre los profesionales interesados en su implementación.

2.2 Inteligencia artificial generativa, LLM y la problemática de la “caja negra”

En relación con los modelos de lenguaje de gran escala (LLM), Moore et al. (2025) advierten que una de las principales dificultades radica en la ausencia de un marco de evaluación interdisciplinario y técnicamente informado para las herramientas de salud mental basadas en este tipo de tecnologías. Esta carencia limita la posibilidad de evaluar adecuadamente su eficacia, seguridad y adecuación ética en contextos clínicos.

Desde la perspectiva de la integridad de la investigación, Eacersall et al. (2024) sostienen que el uso de herramientas de inteligencia artificial generativa (GenAI) plantea múltiples desafíos, particularmente en términos de precisión de los datos, seguridad, fiabilidad y autoría. Los autores argumentan que la propia naturaleza de estas tecnologías puede poner en riesgo la capacidad del investigador para cumplir con los estándares de privacidad exigidos para la aprobación ética o para la protección de información comercial confidencial.

Asimismo, Eacersall et al. (2024) destacan la tendencia de estas herramientas a producir contenido fácticamente incorrecto, fenómeno conocido como “alucinaciones”

(Mahyoob et al., 2023), lo que implica que los datos y textos generados pueden carecer de exactitud o de matices relevantes. A esto se suma la falta de transparencia sobre los procesos de entrenamiento de los modelos, señalada como una preocupación central (Dwivedi et al., 2023).

Estas limitaciones se vinculan con la denominada naturaleza de “caja negra” de la IA (Barr, 2023), la cual dificulta la replicabilidad de los hallazgos y la evaluación de la fiabilidad y actualidad de los datos utilizados para generar resultados, afectando la confianza en los sistemas basados en IA (Eacersall et al., 2024).

2.3 Empatía, distancia psicológica y vínculo terapéutico mediado por IA

Desde una perspectiva más psicológica, Salazar-Garcés y Velastegui-Hernandez (2024) plantean que el avance de la IA genera una intersección crítica con la psicología humana, abriendo nuevos horizontes pero también desafíos éticos y prácticos. Los autores introducen el concepto de “distancia psicológica”, entendida como la percepción de cercanía o lejanía que las personas experimentan frente a la tecnología, y cómo esta influye en la relación emocional y cognitiva con las máquinas.

Según estos autores, herramientas como asistentes virtuales y chatbots, que interactúan de manera aparentemente “humana”, pueden reducir dicha distancia al generar sensaciones de familiaridad y empatía. Cuando los usuarios perciben que la IA es justa, transparente y comprensible, aumenta la probabilidad de que confíen en sus decisiones. Sin embargo, si la IA es percibida como una “caja negra” incomprensible, la confianza puede verse erosionada, generando resistencia o rechazo (Salazar-Garcés & Velastegui-Hernandez, 2024).

Los autores también problematizan la capacidad real de la IA para brindar empatía genuina, planteando interrogantes clave sobre el rol de los “robots terapeutas” y su impacto en la relación terapéutica, la confidencialidad del paciente, la privacidad y la autonomía individual. En este contexto, advierten sobre el riesgo de dependencia excesiva de la tecnología y la posible relegación de la interacción humana genuina. En esta misma línea, los autores concluyen:

Es imperativo que, como sociedad, reflexionemos sobre el impacto psicológico de la IA. No solo en términos de sus aplicaciones directas en terapia y tratamiento, sino también en cómo afecta nuestra percepción de nosotros mismos, nuestra relación con la tecnología y nuestra interacción con el mundo que nos rodea. La IA no es simplemente una herramienta; tiene el poder de influir en nuestra psicología, nuestras emociones y nuestra identidad. (Salazar-Garcés & Velastegui-Hernandez, 2024, p. 32)

2.4 Autonomía, consentimiento informado y riesgos de manipulación

La posibilidad de que la IA influya o manipule las decisiones y emociones de las personas sin su conocimiento o consentimiento constituye otro eje central del debate ético. Salazar-Garcés y Velastegui-Hernandez (2024) señalan que, a pesar de su capacidad para simular interacciones humanas, la IA carece de empatía auténtica, lo que puede dar lugar a malentendidos, pérdida de confianza y una experiencia terapéutica empobrecida.

En esta misma línea, Hoose (2024) advierte que, aunque la IA puede asistir en la elaboración de diagnósticos y recomendaciones personalizadas, existe el riesgo de deshumanización de los usuarios y de una comprensión insuficiente de sus emociones y experiencias. Bernert et al. (2020) coinciden en que estas herramientas presentan limitaciones para demostrar emociones y brindar cuidado, cualidades fundamentales en el ejercicio terapéutico, especialmente en casos clínicos complejos que requieren interacción directa entre paciente y profesional.

Hoose (2024) subraya que los principales desafíos éticos asociados al uso de IA en salud mental incluyen el consentimiento informado, la privacidad de los datos y la posibilidad de influencias perjudiciales sobre las personas. En este sentido, el autor sostiene que estas problemáticas sólo pueden abordarse mediante el cumplimiento de estándares de explicabilidad, responsabilidad y conformidad con las normativas de protección de datos personales, alertando además sobre el potencial uso predatorio o coercitivo de las predicciones generadas por la IA. El autor concluye:

Las implicaciones éticas y sociales del uso de la IA en la atención de la salud mental no pueden ser ignoradas. Se introducen nuevos riesgos en la toma de decisiones cuando se aplica la IA, tales como el refuerzo de los prejuicios actuales de la sociedad o la incapacidad de reconocer la diversidad cultural de los pacientes. Adicionalmente, se ha observado que se requiere la protección de los datos personales, particularmente en lo que respecta al estado emocional y mental de los pacientes, debido a que la mayoría de las soluciones de IA dependen de la recopilación masiva de datos (*big data*) y del análisis posterior del conjunto de datos. (Hoose, 2024, p. 15)

2.5 Sesgos algorítmicos y equidad en la toma de decisiones

Otro aspecto crítico señalado refiere a la presencia de sesgos en los sistemas de IA. Shimada (2023) advierte que los instrumentos de diagnóstico y evaluación basados en IA pueden producir sesgos presentes en los datos de entrenamiento, lo que daría lugar a resultados inequitativos y a la perpetuación de desigualdades socioeconómicas preexistentes.

El autor enfatiza que, si los datos utilizados para entrenar los algoritmos están compuestos mayoritariamente por un solo grupo demográfico, pueden generarse recomendaciones terapéuticas sesgadas y resultados subóptimos para poblaciones diversas. Por ello, resulta fundamental garantizar conjuntos de datos inclusivos y representativos que reduzcan el riesgo de perpetuar desigualdades en salud (Shimada, 2023).

Asimismo, Shimada señala que el uso de grandes volúmenes de información sensible incrementa los riesgos vinculados a la privacidad y la seguridad de los datos, lo que exige la implementación de protocolos estrictos de protección y normativas claras que fomenten la confianza en estos sistemas.

2.6 Supervisión humana y limitaciones de los chatbots en salud mental

En relación con los chatbots de salud mental, Coghlan et al. (2023) sostienen que su funcionamiento continuo y seguro requiere supervisión humana constante, ya que estos sistemas pueden adquirir fallos a medida que aprenden y evolucionan. La falta de supervisión adecuada incrementa el potencial de daño para los usuarios. En este sentido, los autores argumentan que "existe el riesgo de que la recopilación masiva de datos (*big data*) por parte de estos sistemas sea utilizada para fines comerciales o de vigilancia, comprometiendo la confidencialidad" (Coghlan et al., 2023, p. 6).

Los autores advierten que la moderación de estos sistemas puede aumentar la carga laboral de los profesionales y reducir el tiempo destinado a la interacción presencial y al trabajo en equipo. Además, situaciones como cortes de energía o la ausencia de supervisores capacitados pueden comprometer gravemente la prestación del servicio.

En este sentido, Coghlan et al. (2023) concluyen que, pese a los avances tecnológicos, los chatbots actuales no logran recrear la complejidad de la alianza terapéutica humana. Aun cuando superan a sistemas pioneros como ELIZA en capacidad de aprendizaje, el apoyo que brindan suele ser genérico y comparable al de los libros de autoayuda, sin captar las sutilezas sociales, psicológicas y biológicas que influyen en la salud mental.

2.7 Intervenciones JITAI y desafíos teóricos

Por último, en relación con las intervenciones just-in-time (JITAI), Nahum-Shani et al. (2018) señalan que uno de los principales obstáculos para su desarrollo efectivo radica en la naturaleza estática de muchas teorías conductuales existentes y en la falta de especificidad temporal de aquellas que reconocen dinámicas más complejas.

Los autores argumentan que, si bien conceptos como el malestar emocional son dinámicos y requieren monitoreo continuo, las teorías actuales no especifican adecuadamente

cómo cambian estos mecanismos a lo largo del tiempo ni qué tipo de apoyo debería brindarse en cada momento. En consecuencia, la implementación eficaz de JITAI se beneficiaría de teorías dinámicas que orienten el tipo, el momento y la intensidad del apoyo necesario para mejorar la adherencia y la retención en las intervenciones. En este sentido, los autores concluyen:

A pesar del creciente uso y atractivo de las JITAI, existe una brecha importante entre las crecientes capacidades tecnológicas para ofrecer JITAI y la investigación sobre el desarrollo y la evaluación de estas intervenciones [...] A medida que entramos en una nueva era de capacidad tecnológica [...] Es fundamental que los investigadores desarrollen teorías de comportamiento de salud sofisticadas y matizadas capaces de guiar la construcción de tales intervenciones (Nahum-Shani et al., 2018, p. 446)

Conclusión

Para concluir con el recorrido de esta investigación, es fundamental sintetizar los hallazgos expuestos en los dos pilares que estructuran este trabajo. En el primer capítulo, se examinó el ecosistema de herramientas digitales basadas en el modelo cognitivo-conductual, destacando el rol de los chatbots terapéuticos como Woebot y Wysa, los cuales facilitan la accesibilidad y el monitoreo constante del paciente. Asimismo, se analizó el impacto de la realidad virtual (VRET) en los protocolos de exposición y la emergencia de la inteligencia artificial multimodal, capaz de integrar señales fisiológicas y no verbales para una evaluación clínica más profunda. Finalmente, se exploró el marco de las intervenciones adaptativas (JITAI), que permiten brindar apoyo terapéutico en el entorno natural del individuo y en el momento exacto de necesidad.

Por otro lado, el segundo capítulo se centró en la dimensión crítica de estas tecnologías, abordando los desafíos éticos y prácticos que su implementación conlleva. Se discutieron aspectos imperativos como la privacidad y protección de datos sensibles, así como las limitaciones técnicas de los modelos generativos, específicamente el fenómeno de las "alucinaciones" y la falta de transparencia en sus procesos. Finalmente, se reflexionó sobre la importancia de la supervisión humana para mitigar sesgos algorítmicos y garantizar que el vínculo terapéutico (punto importante de la TCC) no se vea deshumanizado, afirmando que la inteligencia artificial debe actuar como un complemento que potencie, y no que reemplace, la labor del profesional.

Es innegable que la incorporación de nuevas tecnologías basadas en el uso de inteligencia artificial produce y seguirá produciendo cambios rotundos en todas las áreas, en especial aquellas de gran utilidad como la salud mental. Desde el abordaje de los procesos psicológicos hasta el monitoreo constante, la IA se consolida como una herramienta más en el abanico del terapeuta de la nueva era. Como destaca una reciente publicación de Listín Diario (2026), los programas conversacionales han evolucionado de ser simples sistemas de respuestas rígidas a herramientas capaces de ofrecer interacciones personalizadas que recuerdan el contexto del usuario, sirviendo como una "primera línea" de defensa ante el estrés y la ansiedad. Esta posibilidad de la IA como refuerzo permite cubrir áreas donde el terapeuta no puede llegar, favoreciendo la autonomía del paciente y actuando con rapidez frente a circunstancias que, por tiempo y espacio, sin estas tecnologías serían mucho más difíciles de sobrellevar.

Asimismo, personalizar cada tratamiento es clave y en este sentido la inteligencia artificial no se queda atrás. Aprender y detectar los datos que presentan los pacientes permite

realizar una intervención que varía de persona a persona, identificando factores de riesgo y agregando valor en las intervenciones que ameriten del terapeuta. Además, la integración con wearables y sistemas de biofeedback ofrece información objetiva sobre indicadores fisiológicos, fortaleciendo la autorregulación y la adherencia a las estrategias de la TCC. No obstante, es fundamental reconocer que esta eficiencia tecnológica tiene límites claros. La nota del Listín Diario (2026) advierte que, a pesar de los avances, “un programa conversacional puede ayudar a alguien a superar una crisis de ansiedad o un momento de estrés, pero difícilmente podrá reemplazar la profundidad de una relación terapéutica humana y la capacidad de un psicólogo para leer entre líneas”.

Finalmente, no hay que pasar por alto las falencias actuales: la privacidad de datos es un tema serio al añadir esta tecnología al proceso clínico. Debe prevalecer una consideración ética que posicione siempre a la inteligencia artificial como un complemento que no sustituya el juicio clínico ni borre la relación empática terapeuta-paciente. Como concluye el análisis de la noticia referenciada, el éxito no radica en reemplazar al humano, sino en utilizar la tecnología para democratizar el acceso, entendiendo que la intuición y el vínculo emocional siguen siendo capacidades irremplazables. El uso responsable y la continua adecuación de estos avances deben realizarse siempre por el bien del paciente, y nunca a costa de su integridad o su privacidad.

Referencias bibliográficas

- Agencia EFE / Listín Diario. (2026, 25 de enero). *¿Podría la inteligencia artificial reemplazar a un sicólogo humano?*. Listín Diario. https://listindiario.com/la-vida/tecnologia/20260125/programa-conversacional-inteligencia-artificial-reemplazar-psicoterapeuta-humano_891211.html
- Alasmrai, M. A., Ismail, R. M., & Al-Abyadh, M. H. (2025). *Personalized cognitive behavioral therapy for adults using machine learning: A multi-factor, reinforcement-based approach*. *Fusion: Practice and Applications*, 20(2), 53–64.
- Asan, O., & Choudhury, A. (2021). *Artificial intelligence human factors in healthcare: A mapping review*. *JMIR Human Factors*, 8, e28236. <https://doi.org/10.2196/28236>
- Beck, A. T. (1979). *Cognitive therapy and the emotional disorders*. New American Library.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. Guilford Press.
- Bernert, R. A., Hilberg, A. M., Melia, R., Kim, J. P., Shah, N. H., & Abnoui, F. (2020). *Artificial intelligence and suicide prevention: A systematic review of machine learning investigations*. *International Journal of Environmental Research and Public Health*, 17, 5929. <https://doi.org/10.3390/ijerph17165929>
- Beth Israel Deaconess Medical Center. (2023, 27 de junio). *The future of AI in mental health care*. <https://www.bidmc.org/about-bidmc/news/2023/06/the-future-of-ai-in-mental-health-care>
- Boden, M. A. (2017). *Inteligencia artificial*. Turner.
- Boeldt, D. L., Casmar, P. V., & Kyanko, N. (2019). *Using Virtual Reality to Improve the Treatment of Anxiety Disorders: A Systematic Review*. *NPJ Digital Medicine*. <https://doi.org/10.1038/s41746-019-0077-7>
- Bohannon, J. (2015). *The synthetic therapist*. *Science*, 349(6245), 250–251. <https://doi.org/10.1126/science.349.6245.250>

- Canabal Berlanga, A., Keough Delgado, E. A., & Alonso González, M. (2024). *Inteligencia artificial en la salud mental: Oportunidades, dificultades y cuestionamientos*. INTELETICA, 1(2), 16–26.
- Cassata, C. (2018, 2 de julio). *Can a Chatbot Help Your Mental Health?*. Healthline. <https://www.healthline.com/health-news/can-a-chatbot-help-your-mental-health>
- Chu, Y., Liao, L., Zhou, Z., Ngo, C.-W., & Hong, R. (2025). *Towards multimodal emotional support conversation systems*. IEEE Transactions on Multimedia, 1–12. <https://doi.org/10.1109/TMM.2025.xxxxx>
- Coghlan, S., Leins, K., Sheldrick, S., Cheong, M., Gooding, P., & D'Alfonso, S. (2023). *To chat or bot to chat: Ethical issues with using chatbots in mental health*. Digital Health, 9(2). <https://doi.org/10.1177/20552076231183542>
- Darcy, A. L., 受, J., & Vierhile, M. (2017). *Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial*. JMIR Mental Health, 4(2), e19. <https://doi.org/10.2196/mental.7785>
- De la Fuente, D., & Armayones Ruiz, M. (2025). *AI in psychological practice: What tools are available and how can they help in clinical psychology?*. Papeles del Psicólogo / Psychologist Papers, 46(1), 18–24.
- Döllinger, L., Laukka, P., Högman, L. B., Bänziger, T., Makower, I., Fischer, H., & Hau, S. (2021). *Training emotion recognition accuracy: Results for multimodal expressions and facial micro-expressions*. Frontiers in Psychology, 12, 708867. <https://doi.org/10.3389/fpsyg.2021.708867>
- Durden, E., Pirner, M. C., Rapoport, S. J., Williams, A., Robinson, A., & Forman-Hoffman, V. L. (2023). *Changes in stress, burnout, and resilience associated with an 8-week intervention with relational agent "Woebot"*. Internet interventions, 33, 100637.
- Eacersall, D., Pretorius, L., Smirnov, I., Spray, E., Illingworth, S., Chugh, R., ... Howell, K. A. (2024). *Navigating ethical challenges in generative AI-enhanced research: The ETHICAL framework for responsible generative AI use*. arXiv. <https://arxiv.org/abs/2401.xxxxx>

- Ellis, A., & Joffe-Ellis, D. (2019). *Rational emotive behavior therapy*. American Psychological Association.
- Forbes. (2021). *How AI Is Helping To Bridge The Mental Health Gap*. Forbes Magazine. <https://www.forbes.com/sites/forbesbooksauthors/2021/05/18/how-ai-is-helping-to-bridge-the-mental-health-gap/>
- Goessl, V. C., Curtiss, J. E., & Hofmann, S. G. (2017). *The effect of heart rate variability biofeedback training on stress and anxiety: A meta-analysis*. Psychological Medicine, 47(15), 2578–2586. <https://doi.org/10.1017/S0033291717001009>
- Haberer, J. E., Trabin, T., & Klinkman, M. (2013). *Furthering the reliable and valid measurement of mental health screening, diagnoses, treatment and outcomes through health information technology*. General Hospital Psychiatry, 35(4), 349–353. <https://doi.org/10.1016/j.genhosppsych.2013.03.009>
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2014). *Acceptance and commitment therapy: The process and practice of mindful change*. Second Edition. Guilford Press.
- Hoose, S., & Králiková, K. (2024). *Artificial intelligence in mental health care: Management implications, ethical challenges, and policy considerations*. Administrative Sciences, 14(9), 227. <https://doi.org/10.3390/admsci14090227>
- Hutchison, A. N., & Gerstein, L. H. (2012). *What's in a face? Counseling trainees' ability to read emotions*. Training and Education in Professional Psychology, 6(2), 100–112. <https://doi.org/10.1037/a0027221>
- Ingram, R. E., & Kendall, P. C. (1986). *Information processing approaches to clinical psychology* Cognitive clinical psychology: Implications for theory, research, and practice. En R. E. Ingram (Ed.). Academic Press.
- Kargarandehkordi, A., Li, S., Lin, K., Phillips, K. T., Benzo, R. M., & Washington, P. (2025). *Fusing wearable biosensors with artificial intelligence for mental health monitoring: A systematic review*. Biosensors. 15(4), 202. <https://doi.org/10.3390/bios15040202>

- Kim, S., Kim, H., Do, H., & Lee, G. G. (2025). *Multimodal cognitive reframing therapy via multi-hop psychotherapeutic reasoning*. arXiv. <https://arxiv.org/abs/2502.xxxxx>
- Knight, K., & Rich, E. (1994). *Inteligencia artificial*. Second edition.. McGraw-Hill.
- Kumar, S. (2021, 15 de abril). *Wysa: The AI Chatbot Helping Millions Manage Their Mental Health*. The Politic. <https://thepolitic.org/an-interview-with-jo-aggarwal-co-inventor-of-wysa/>
- La Nación. (2025, 27 de marzo). *Los riesgos y los beneficios de aplicar IA al campo de la psicología*. La Nación. <https://www.lanacion.com.ar/economia/IA/los-riesgos-y-los-beneficios-de-aplicar-ia-al-campo-de-la-psicologia-nid25032025/>
- Larson, E. J. (2022). *El mito de la inteligencia artificial: Por qué las máquinas no pueden pensar como nosotros lo hacemos*. Shackleton Books.
- McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (1955). *A proposal for the Dartmouth summer research project on artificial intelligence*. AI Magazine, 27(4).
- McCulloch, W. S., & Pitts, W. (1943). *A logical calculus of ideas immanent in nervous activity*. Bulletin of Mathematical Biophysics, 5(4), 115–133.
- Mohr, D. C., Lyon, A. R., Lattie, E. G., Reddy, M., & Schueller, S. M. (2017). *Accelerating digital mental health research from early design and creation to successful implementation and sustainment*. Journal of Medical Internet Research, 19(5), e153. <https://doi.org/10.2196/jmir.7722>
- Mullin, E. (2017). *The chatbot will see you now*. MIT Technology Review. [O alternativamente si es de Wired]: Wired Magazine.
- Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). *Just-in-time adaptive interventions (JITAI) in mobile health*. Annals of Behavioral Medicine, 52(6), 446–462. <https://doi.org/10.1007/s12160-016-9830-8>
- National Institute for Health and Care Excellence. (2006). *Computerised cognitive behavioural therapy for depression and anxiety* (Technology appraisal guidance TA97). <https://www.nice.org.uk/guidance/ta97>

- Neisser, U. (1967). *Cognitive Psychology*. Appleton-Century-Crofts.
- Ng, A. (2018). *Machine learning yearning: Technical strategy for AI engineers*. DeepLearning.ai.
- Oracle, & Workplace Intelligence. (2020). *AI at Work: As mental health frontier moves to the workplace, employees look to AI for help*. Oracle. <https://www.oracle.com/news/announcement/ai-at-work-100720/>
- Otero Parga, M. (2023). *¿Puede la inteligencia artificial sustituir a la mente humana?* Anales de la Cátedra Francisco Suárez, 57, 39–61.
- Proudfoot, J., Swain, S., Widmer, S., Watkins, E., Goldberg, D., Marks, I., Mann, A., & Gray, J. A. (2003). *The development and beta-test of a computer-delivered cognitive-behavioural therapy programme for anxiety and depression: "Beating the Blues"*. Computers in Human Behavior, 19(3), 383–397. [https://doi.org/10.1016/S0747-5632\(02\)00072-6](https://doi.org/10.1016/S0747-5632(02)00072-6)
- Regan, K. W. (2012). *Deeper Blue: Human chess performance after Deep Blue*. Springer.
- Rizzo, A., & Boucher, P. (2023). *Virtual Reality and Artificial Intelligence in Clinical Psychology and Psychotherapy*. Springer Nature.
- Rizzo, A., & Kim, G. J. (2005). *A SWOT analysis of the field of virtual reality rehabilitation and therapy*. Presence: Teleoperators and Virtual Environments, 14(2), 119–146.
- Russell, S. J., y Norvig, P. (2021). *Inteligencia artificial: Un enfoque moderno*. Cuarta edición. Pearson Educación.
- Salazar Garcés, L. F., & Velastegui Hernández, D. C. (2024). *Inteligencia artificial y su impacto en la psicología humana: Mini revisión*. Mediciencias UTA, 8(1), 26–34.
- Shimada, K. (2023). *The role of artificial intelligence in mental health: A review*. Science Insights, 43(5), 1119–1127.

- Simon, H. A. (1955). *A behavioral model of rational choice*. Quarterly Journal of Economics, 69(1), 99–118.
- Simon, H. A. (1969). *The sciences of the artificial*. MIT Press.
- Simon, H. A., & Newell, A. (1976). *Computer science as empirical inquiry: Symbols and search*. Communications of the ACM, 19(3), 113–126.
- Torous, J. (2020). *¿Cómo mejorar nuestra salud mental con la IA?*. Metro World News.
- Torous, J., Bucci, S., Bell, I. H., Kessing, L. V., Faurholt-Jepsen, M., Aradhya, P., ... & Powell, A. C. (2021). *The growing field of digital psychiatry: Current evidence and future directions*. World Psychiatry, 20(3), 318-335.
- Turing, A. M. (1950). *Computing machinery and intelligence*. Mind, 59(236), 433–460.
- Vempati, R. (2020). *Mental health for all: The Wysa story and the scale of the challenge*. Wysa Health. [O si prefieres el artículo donde se recogió la entrevista]: BBC News. (2020). *The chatbot helping India's mental health*.
- Vives, J. (2023, 18 de julio). *La startup en salud mental Eholo acelera su crecimiento gracias a una ronda de financiación de 200.000 euros*. La Vanguardia.
- Weizenbaum, J. (1966). *ELIZA—A computer program for the study of natural language communication between man and machine*. Communications of the ACM, 9(1), 36–45.
- Woebot Health. (2021). *The Science behind Woebot: Clinical evidence and outcomes*. Woebot Health Media. <https://woebothealth.com/science/>
- Wysa. (2020). *Our Story: From ELIZA to AI-driven Emotional Intelligence*. Wysa Inc. <https://www.wysa.com/about-us>
- Wysa. (2022). *Wysa for Employers: Supporting the full spectrum of mental health*. Wysa Inc. <https://www.wysa.com/employers>
- Zhou, S., Zhao, J., & Zhang, L. (2022). *Application of artificial intelligence on psychological interventions and diagnosis: An overview*. Frontiers in Psychiatry, 13.