



Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), septiembre-octubre 2025,
Volumen 9, Número 5.

https://doi.org/10.37811/cl_rcm.v9i5

ANÁLISIS LINGÜÍSTICO ASISTIDO POR INTELIGENCIA ARTIFICIAL: ESTUDIO PILOTO DE CARACTERIZACIÓN LINGÜÍSTICA Y PSICOPATOLÓGICA EN PACIENTES PSIQUIÁTRICOS EN SEGUIMIENTO CLÍNICO

LINGUISTIC ANALYSIS ASSISTED BY ARTIFICIAL
INTELLIGENCE: PILOT STUDY ON LINGUISTIC AND
PSYCHOPATHOLOGICAL CHARACTERIZATION IN
PSYCHIATRIC FOLLOW-UP OUTPATIENTS

Carlos Armando Herrera-Huerta

Asociación Iberoamericana de Neurociencias y Psiquiatría AILANCYP

Angélica Gándara-López

Asociación Iberoamericana de Neurociencias y Psiquiatría AILANCYP

Pedro Moreno-Gea

Editor de psiquiatria.com

DOI: https://doi.org/10.37811/cl_rcm.v9i5.19844

Análisis Lingüístico Asistido por Inteligencia Artificial: Estudio Piloto de Caracterización Lingüística y Psicopatológica en Pacientes Psiquiátricos en Seguimiento Clínico

Carlos Armando Herrera-Huerta¹drherrerasaludmental@gmail.com<https://orcid.org/0000-0001-9677-2344>Asociación Iberoamericana de Neurociencias y
Psiquiatría AILANCYP
México**Angélica Gándara-López**contacto@psicangelicagandara.com<https://orcid.org/0009-0000-2868-0050>Investigador Independiente
Asociación Iberoamericana de Neurociencias y
Psiquiatría AILANCYP
México**Pedro Moreno-Gea**pmoreno@psiquiatria.comEditor de psiquiatria.com
España

RESUMEN

Se presenta un estudio piloto orientado a evaluar la utilidad y precisión diagnóstica de la inteligencia artificial (IA) en psiquiatría mediante el análisis lingüístico asistido por la grabadora de aplicaciones inteligentes AIREC. Se incluyeron 10 pacientes en seguimiento clínico en Guaymas, Sonora, México, con diagnósticos categoriales DSM-5-TR. Los registros de audio fueron transcritos por el dispositivo y procesados con algoritmos de IA (ChatGPT-5, Google Speech-to-Text, DeepSearch). La concordancia diagnóstica con el clínico humano fue del 78%. Las discrepancias se concentraron en la interpretación de prosodia emocional, neologismos y fragmentación narrativa. Este trabajo surge ante la creciente necesidad de herramientas tecnológicas que optimicen los procesos de evaluación psiquiátrica en contextos donde existe escasez de personal especializado. El estudio describe cómo la integración de sistemas de IA permite identificar patrones idiolécticos y lingüísticos asociados a distintos cuadros psicopatológicos, aportando información complementaria al juicio clínico tradicional. Los resultados obtenidos resaltan la importancia de adaptar los algoritmos a factores culturales y contextuales específicos, así como de perfeccionar la detección de matices emocionales en el lenguaje. Los hallazgos sugieren que la IA puede constituir un recurso de apoyo diagnóstico complementario, con potencial de mejorar la precisión y eficiencia en la atención psiquiátrica, especialmente en regiones con recursos limitados y alta demanda de servicios de salud mental.

Palabras clave: inteligencia artificial, psiquiatría, análisis lingüístico, diagnóstico clínico

¹ Autor principal

Correspondencia: drherrerasaludmental@gmail.com

Linguistic Analysis Assisted by Artificial Intelligence: Pilot Study on Linguistic and Psychopathological Characterization in Psychiatric Follow-Up Outpatients

ABSTRACT

This pilot study aims to evaluate the diagnostic accuracy of artificial intelligence (AI) in psychiatry through linguistic analysis using the AIREC smart recording device. 10 patients under clinical follow-up in Guaymas, Sonora, México were included, all with DSM-5-TR categorical diagnoses. Audio recordings were transcribed and processed using AI algorithms (ChatGPT-5, Google Speech-to-Text, DeepSearch). Diagnostic concordance with human clinicians was 78%. Discrepancies were mainly observed in interpreting emotional prosody, neologisms and narrative fragmentation. The present study addresses the growing need for technological tools to enhance psychiatric assessment in settings with limited specialized personnel. Findings show that integrating AI systems can help identify idiolectal and linguistic patterns linked to various psychopathological conditions, providing complementary insights to traditional clinical judgement. Results highlight the importance of adapting algorithms to cultural and contextual factors and improving the detection of emotional nuances in speech. AI shows potential as a complementary diagnostic support tool to improve accuracy and ²efficiency in mental health care, particularly in resource-limited regions with high demand for psychiatric services.

Keywords: artificial intelligence, psychiatry, linguistic analysis, clinical diagnosis

*Artículo recibido 02 setiembre 2025
Aceptado para publicación: 29 setiembre 2025*



INTRODUCCIÓN

La inteligencia artificial (IA), es reconocida por la Organización Mundial de la Salud (OMS, 2021) como una herramienta con amplio potencial para mejorar la salud, pero que demanda de un enfoque ético y principios normativos que garanticen la seguridad y equidad del paciente, así como una responsabilidad intachable en términos de diseño y uso. Esta tecnología aún es un terreno de relativa incertidumbre, ya que, gracias a los persistentes aportes de la ciencia, día con día nos encontramos con novedosos datos y hallazgos deslumbrantes, lo cual nos pone en evidencia el gran potencial de este Modelo de Grande Lenguaje (LLM).

Esta potencial, creciente y aún joven tecnología, surgió como una idea de ciencia ficción, según la revisión histórica realizada por Sarangi y Sharma (2018), en la cual se reflejaba el interés de generar una “entidad” capaz de desempeñar tareas humanas, como en la obra de teatro de 1920 “Los Robots Universales de Rossum” o R.U.R. por sus siglas en inglés, del escritor checo Karel Čapek en donde se acuña el concepto en idioma checo “robota”, que significa “servidumbre”. Luego, en 1950 pasó de ser solo una concepción arquetípica, a posicionarse como un modelo teórico en el artículo “Computing Machinery and Intelligence” escrito por Alan Turing, el cual abre con el cuestionamiento “Can machines think?” (Turing, 1950). En este manifiesto hacia la IA, realiza una narrativa sobre su experimento llamado “The imitation game” (en español, el juego de la imitación), el cual consistió en que un individuo, con función de interrogador (A), interactuara de manera indirecta (sin poder observar) con dos respondientes, otro individuo (B), y una computadora (C); se buscó que el individuo A, a partir de la intercomunicación, pudiera distinguir cuál de los interrogados era la máquina y el ser humano; el resultado fue sorprendente: el interrogador no pudo discernir entre ambos (B, C). A raíz de esto, emergió un interés filosófico, y sobre todo ontológico, y se le comenzó a darle importancia a este nuevo fenómeno, dándole apertura a nuevos cuestionamientos, y sobre todo, en el campo de la ciencia, a investigaciones que hoy en día, han marcado el curso de la humanidad.

Desde el siglo XX y hasta el día de hoy, esta innovación que comenzó como una extravagante representación, ha ido evolucionando de manera amplificada, veloz y eficaz, de tal manera que su estudio y aplicación ha ido ganando popularidad en diversas áreas, sobre todo en el campo de la salud, comenzando a perfilarse como una potencial herramienta médica durante las décadas de los 50’s y 60’s,



aunque los hallazgos o “avances” eran meramente conceptuales, pero sentaron las bases para la invención de INTERNIST-1, en 1971, un sistema desarrollado por la Universidad de Pittsburgh, capaz de analizar signos y síntomas, generar presunciones diagnósticas y así guiar la toma de decisiones clínicas (Pople et al., 1982). Después, la Universidad de Stanford, en 1976, desarrolló MYCIN, sistema experto en diagnosticar infecciones bacterianas graves, demostrando la viabilidad de la IA como herramienta diagnóstica (Shortliffe et. Al., 1976).

Ésta tecnología, ha estado presente en el área de la psicología y psiquiatría desde los años 60’s, comenzando con los sistemas basados en reglas enfocados al procesamiento de información, toma de decisiones y resolución de problemas simples, como ELIZA (Weizenbaum, 1966), el primer dispositivo basado en la interacción humano-máquina, encargado de simular un estilo terapéutico conversacional “Rogeriano” (haciendo alusión al psicólogo estadounidense Carl Rogers). Después tenemos a Diagno I (Spitzer et al., 1970) y a PARRY (Colby, 1972), el primero como un sistema diagnóstico estructurado y el segundo como simulador de un paciente con diagnóstico de esquizofrenia paranoide, incorporando variables emocionales y cognitivas básicas.

En las décadas de los 80’s y 90’s, se desarrolla la “Computerized Diagnostic Interview Schedule (C-DIS)”, una entrevista clínica estructurada, basada en los criterios del DSM-III, para fines diagnósticos; aunado a eso, surgen los primeros intentos de utilizar la IA para interpretación de resultados de escalas como el MMPI y el BDI. Hasta este punto, los dispositivos de Inteligencia Artificial mostraban la eficiencia esperada en función de su diseño, pero solo se caracterizaban por tener respuestas predefinidas ante reglas fijas o específicas, dejando de lado el aprendizaje autónomo, presentando limitaciones en temas de adaptación y procesamiento de información nueva, siendo una limitación determinante para el área del diagnóstico en salud mental, ya que la complejidad de los casos clínicos no se basa en el diagnóstico, si no en los factores variables y volátiles del comportamiento humano (Kaul, 2020).

Llega el siglo XXI junto con el desarrollo de dos áreas sustanciales dentro del campo de la IA: el “Machine Learning (ML)” y el “Deep Learning “DL” . El ML es una subdivisión de la IA enfocada únicamente en el desarrollo de algoritmos que permiten a los dispositivos aprender patrones a partir de datos, sin la necesidad de una programación previa y su funcionamiento es limitado, ya que depende en



gran medida de expertos. El DL, es una subdivisión del ML, que utiliza redes neuronales profundas (Deep Neural Networks o DNNs, ver figura 1) para aprender representaciones más complejas de datos, y a diferencia de las ML, esta sub-área presenta una capacidad de aprendizaje automático. (Goodfellow, Bengio y Courville, 2016).

Los datos o información de entrada recibidos por inputs (que cumplen una función dendrítica, como receptores y conectores), son procesados por medio de capas ocultas (procesamiento jerárquico y distribución de la información), donde cada conexión tiene un peso o score (peso sináptico) que se ajusta durante el entrenamiento (plasticidad neuronal); en cada una de las capas o layers, los valores pasan por una función de activación (integración/impulso neuronal no lineal), permitiendo a la red, aprender patrones complejos. (Goodfellow et al., 2016; LeCun et al., 2015). El Deep Learning, aprende de forma autónoma, similar a cómo el cerebro procesa la información, generando una salida final sin instrucciones explícitas, teniendo así, un aprendizaje eficiente por medio de algoritmos como el Spiking Neural Networks (SNNs) inspirados en los pulsos eléctricos y la normalización dendrítica (estabilización de la actividad cerebral) antes de ajustar conexiones, promoviendo un aprendizaje rápido y robusto. (Oxford University, 2024) (Bird et al., 2021; Chavlis & Poirazi, 2025).

En conclusión, las ML aprenden con ejemplos, instrucciones y reglas creadas por un individuo, y las DL funcionan como un cerebro que aprende directamente de la experiencia y hace conclusiones y descubrimientos por sí mismo, sin una instrucción específica. Por consecuencia, nace la IA predictiva y medicina de precisión, que cuentan con sistemas de IA con la capacidad de aprender y procesar grandes volúmenes de datos, permitiendo la creación de modelos predictivos-diagnósticos basados en el análisis de historiales clínicos, patrones conductuales, marcadores lingüísticos asociados a diagnósticos psiquiátricos, etc., teniendo el desarrollo de chatbots terapéuticos como “Wysa” (2016), “Woebot” (2017) y “Tess” (2018) (Fulmer, et.al., 2018) o algoritmos capaces de predecir recaídas, monitorear la adherencia al tratamiento e incluso analizar patrones discursivos de pacientes para detectar indicios tempranos de depresión, psicosis o demencia (Holmes et al., 2021). Todo esto, ha contribuido a las bases de la IA como herramienta capaz de detectar patrones ocultos y correlaciones complejas, como lo demostró la compañía Microsoft (Time, 2024), a partir del desarrollo de modelos



de IA que han demostrado superar la precisión diagnóstica de médicos en tareas específicas, incluyendo la interpretación de información relacionada con salud mental dentro de evaluaciones clínicas.

Cuando hablamos de diagnóstico en psicología o psiquiatría, también hablamos de una problemática coexistente: los sesgos diagnósticos. Se ha documentado que 7 de cada 10 fallos clínicos derivan de procesos cognitivos automáticos, como el sesgo de disponibilidad o el anclaje terapéutico (Croskerry et al., 2013; O'Sullivan & Schofield, 2018; Saposnik et al., 2016), sumándole a estos factores la fatiga clínica que afecta de manera directa, la capacidad de discernimiento en contextos de alta demanda asistencial (Doherty & Balzer, 2020), generando diagnósticos erróneos o inespecíficos, lo cual impacta de manera directa la atención y el tratamiento adecuado y digno, y refuerza la necesidad de modelos asistenciales precisos y eficaces (Newson et al., 2021).

En este contexto, el lenguaje ha tomado un lugar determinante como biomarcador del estado mental (y desde hace décadas); por ello, el análisis de componentes lingüísticos como lo son la prosodia, los neologismos, las disgregaciones narrativas o la coherencia ideo-afectiva han dado pie a establecer correlaciones entre patrones lingüísticos particulares o específicos (dialécticos) entre patrones psicopatológicos.

Según Meta-Análisis recientes, se han evidenciado patrones de deficiencias en la percepción de la prosodia emocional en individuos con esquizofrenia (Lin et al., 2018), así como anormalidades lingüísticas persistentes en personas con el mismo diagnóstico (Ehlen et al., 2023; Parola et al., 2022). Asimismo, estudios con muestras de voz, han revelado la capacidad de la IA para discriminar síntomas de depresión con los de deterioro cognitivo, con una precisión aproximada de 90% (Mirheidari et al., 2019; Pan et al., 2023). Y cuando hablamos del uso de agentes conversacionales o “chat bots”, se manifiesta un efecto positivo relacionado con el bienestar percibido y la detección temprana de ideación suicida, sin dejar de lado la obligatoriedad ética y legal de la supervisión clínica (Li et al., 2023; Wang et al., 2024). No obstante, la evidencia nos afirma la inexistencia de patrones lingüísticos universales relacionados con la psicopatología, eso nos lleva a querer considerar de manera muy minuciosa las particularidades dialécticas y regionales para evitar falsos positivos o conclusiones diagnósticas sesgadas (Parola et al., 2023; Zhu & Jurgens, 2021). Por lo tanto, resulta de extensa relevancia cuando hablamos de regiones con diversidad lingüística, en donde las expresiones locales o modismos pueden

confundirse con alteraciones psicopatológicas (Granados, 2017; Instituto Nacional de los Pueblos Indígenas, 2020), lo cual pone en duda la incorporación de la IA como herramienta clínica analítica del lenguaje, ya que puede objetivas y sistematizar estos rasgos.

El potencial de la IA en el área de la salud mental, específicamente en psiquiatría de precisión es amplio: desde la predicción de respuesta a tratamientos (Chekroud et al., 2021; Curtiss & DiPetro, 2025) hasta la identificación de patrones de riesgo, tomando en cuenta los datos clínicos, contextuales y lingüísticos.

JUSTIFICACIÓN

Las limitaciones diagnósticas en psiquiatría no son solo un problema metodológico, sino también social. En regiones como Guaymas, Sonora, el déficit de personal médico y de especialistas en salud mental es una barrera crítica para la atención. Con una población superior a los 150 mil habitantes y con un crecimiento sostenido en la última década, la demanda supera la capacidad instalada, generando sobrecarga de trabajo, fatiga clínica y mayor probabilidad de sesgos en la práctica diaria (Secretaría de Economía, 2020).

A esto se suma la diversidad lingüística y cultural de la región, donde coexisten el español, variantes locales del noroeste y lenguas indígenas como el yaqui (Instituto Nacional de los Pueblos Indígenas, 2020; Buitimea Valenzuela et al., 2016). Expresiones coloquiales o idiomáticas, documentadas en estudios dialectológicos, pueden ser interpretadas como neologismos o alteraciones del pensamiento si no se considera el contexto sociolingüístico (Granados, 2017; Lope Blanch et al., 1990).

Frente a estas problemáticas, el análisis de componentes lingüísticos asistido por inteligencia artificial representa una alternativa prometedora. La IA ofrece ventajas como rapidez en el procesamiento, estandarización de criterios y reducción del impacto de la fatiga o los sesgos individuales. No obstante, para que esta herramienta sea realmente útil en contextos como el de Guaymas, debe adaptarse a la realidad lingüística y cultural de la región, incorporando corpus locales y marcos éticos que garanticen la protección de datos sensibles. Organismos internacionales como la Organización Mundial de la Salud (2024) y la Unión Europea (2024) subrayan la importancia de establecer regulaciones claras, priorizando la seguridad, la transparencia y la no discriminación en el uso de estas tecnologías en salud. El presente estudio piloto se plantea como un esfuerzo por explorar la viabilidad del análisis lingüístico con apoyo de IA en un contexto de alta relevancia social y clínica.



Al evaluar la concordancia diagnóstica entre el juicio humano y el análisis automatizado se busca cimentar un precedente en la integración de nuevas tecnologías en la práctica psiquiátrica mexicana, aportando evidencia sobre su utilidad, limitaciones y pertinencia cultural.[1]

Objetivo General del estudio

La necesidad de medidas objetivas en el ámbito psiquiátrico ha estimulado el interés por indicadores alternativos que complementen la evaluación clínica tradicional.

El análisis lingüístico del habla –es decir, el estudio minucioso de los patrones lingüísticos únicos de cada paciente– surge como una vía prometedora para vincular lo subjetivo con lo objetivo en la caracterización de los trastornos mentales.

Evidencia reciente sugiere que la aplicación de inteligencia artificial (IA) al lenguaje puede aportar información valiosa para afinar diagnósticos y pronósticos: una revisión sistemática identificó cuatro grandes ámbitos de aplicación del análisis del habla asistido por IA en psiquiatría (clasificación diagnóstica, evaluación de la severidad, detección temprana de la aparición de la enfermedad y estimación de pronóstico o respuesta terapéutica).

Por ejemplo, en el caso de la depresión mayor se ha logrado distinguir a los pacientes por su forma de hablar con más de un 80% de precisión, dado que tienden a un discurso más lento, monótono y de contenido negativo, asimismo, ciertos marcadores semánticos del discurso (como la coherencia y complejidad narrativa) han permitido predecir el riesgo de psicosis a dos años vista con una exactitud cercana al 100% en poblaciones vulnerables. Frente a este panorama, el objetivo general de nuestro estudio piloto es caracterizar de manera sistemática los patrones lingüísticos de pacientes psiquiátricos en seguimiento clínico, empleando herramientas de aprendizaje automático para correlacionar dichas características del habla con sus perfiles psicopatológicos. En última instancia, se busca explorar el potencial de esta aproximación para complementar el proceso diagnóstico tradicional y enriquecer el seguimiento clínico con indicadores lingüísticos objetivos, contribuyendo así a una psiquiatría más precisa y personalizada

MATERIALES Y MÉTODOS

Ubicación geográfica del estudio: Zona costera del sur del estado de Sonora, al Noroeste de México.

Participantes

Población: muestra no probabilística de tipo intencional, conformada por 10 pacientes. Como criterios de selección, se incluyeron adultos dentro de un rango de 18-64 años de edad con un diagnóstico categorial conforme al DSM-5-TR confirmado mediante entrevista clínica semi-estructurada y observación clínica., ≥ 4 sesiones, estabilidad clínica y consentimiento informado y se excluyeron los individuos que presentaron alteraciones orgánicas del habla, discapacidad intelectual moderada/grave y descompensación aguda.

Observadores: dos profesionales de la salud mental con formación en psicología y psiquiatría. Ambos con especialización en evaluación y diagnóstico psicopatológico empleando los criterios diagnósticos propuestos por el DSM-5-TR.

Diseño y tipo de estudio: estudio de carácter mixto; a partir de la clasificación de los diferentes tipos de estudio sugerida por Méndez et. Al. (2004), se optó por el protocolo 1, de carácter observacional, retrospectivo/transversal, siendo útil para establecer las bases para estudios posteriores (de réplica/seguimiento)[2] en un nuevo conjunto de sujetos.

Criterios de selección

Inclusión: mediante entrevista clínica semiestructurada; Entre 4 y 6 sesiones previas, con un promedio de tratamiento de 7 a 12 meses; Estabilidad clínica y tiempo mínimo de grabación por sesión ≥ 35 minutos; Consentimiento informado.

Exclusión: Alteraciones orgánicas del habla; Discapacidad intelectual moderada o grave; Descompensación aguda de la patología de base.

Instrumentos

Formato de historia clínica estándar: conformado por ficha de identificación, antecedentes personales patológicos y no patológicos, antecedentes heredofamiliares, salud mental previa, estado clínico actual, clinimetría o psicometría empleada, integración diagnóstica con base en el sistema categorial del DSM-5 TR, plan terapéutico farmacológico y no farmacológico.

Grabadora digital inteligente: compatible con smartphones. Cuenta con 500 horas de grabación continua e integra un micrófono de alta precisión que registra el audio WAV, lo que garantiza la conservación y protección de frecuencias, timbre, prosodia y microvibraciones de la voz, resultando significativo para el estudio de componentes del lenguaje como la prosodia emocional y el dialecto; asimismo, cuenta con un procesamiento automatizado del habla, a través de algoritmos de Inteligencia Artificial basados en: Google Speech-to-text, ChatGPT-4.0, DeepSearch y Claude. Éstas características en conjunto, dan pie a una grabación inmediata (característica determinante para contextos clínicos), transcripción y traducción multilingüe instantánea (favoreciendo la interacción en contextos multiculturales), permiten sintetizar aspectos relevantes y esquemas gráficos y también la modificación y almacenamiento de la información. La exactitud y sensibilidad en cuanto a la transcripción automática es de un 35%, por lo tanto, es indispensable en análisis clínico- comparativo de los textos generados.

Procedimiento

Se procedió a una selección no probabilística de expedientes con suficiente densidad narrativa; las sesiones se registraron con AIREC, se transcribieron de manera automática bajo resguardo de anonimato y se segmentaron en unidades discursivas. El corpus resultante se analizó con apoyo de ChatGPT 5 mediante un prompt estandarizado para la codificación de rasgos idiolécticos —léxico, prosodia, patrones de interrupción, neologismos y coherencia ideoafectiva—, siguiendo un esquema mixto que combinó revisión clínica y procesamiento de lenguaje natural para identificar regularidades sintáctico-semánticas, evaluar la organización narrativa y explorar correspondencias con el diagnóstico nosológico a lo largo del seguimiento. En dicho prompt se instruyó: (i) preprocesar la transcripción anotando errores de ASR, regionalismos y segmentos inaudibles, con segmentación por turnos y normalización mínima; (ii) codificar las dimensiones idiolécticas indicadas y clasificar los neologismos como neologismo probable / regionalismo probable / error de ASR; (iii) sintetizar una matriz de severidad (0–3) por dimensión; (iv) realizar correlación clínica e integración diagnóstica categorial auxiliar (no concluyente), mapeando rasgos → DSM-5-TR con nivel de apoyo, argumentos lingüísticos y contraevidencias; (v) aplicar control cultural y educativo previo al etiquetado patológico; y (vi) entregar salidas estructuradas (resumen idioléctico, tabla de severidad, ejemplos con marca temporal, mapa de riesgos y recomendaciones para verificación humana y para la próxima sesión).



Figura 1. Esquema operativo de análisis Linguístico-Psicopatológico

1) Historia clínica: ficha/antecedentes/estado mental/clinimetría/plan/alarmas]
2) Entrevista semiestructurada: Guía + sondeo/fenomenología/formulación cultural (CFI)]
3) Integración diagnóstica categorial (DSM-5-TR): hipótesis/diferenciales/severidad–comorbilidad/decisiones iniciales]
4) Grabación (durante la entrevista): AIREC (WAV)/prosodia–microseñales/consentimiento–resguardo]
5) Análisis de transcripción: ASR→ limpieza/segmentación/normalización/codificación idioléctica (léxico–prosodia– interrupciones–neologismos–coherencia ideoafectiva–disgregación) /salidas (ficha idioléctica–mapa de riesgos–correlación diagnóstica)]

Consideraciones Éticas

La grabadora de aplicaciones inteligentes AIREC no cumple con las normativas de protección de datos clínicos como HIPPA (Ley de Portabilidad y Responsabilidad del Seguro Médico), que es una ley federal de EE. UU. aprobada en 1996 que establece normas nacionales para proteger la privacidad y seguridad de la información médica personal, tampoco cuenta con la GDPR (Reglamento General de Protección de Datos), una legislación europea que establece normas estrictas sobre cómo las empresas y organizaciones deben recopilar, procesar y proteger los datos personales de los ciudadanos de la Unión Europea, Aunque el dispositivo no cuenta con las certificaciones antes mencionadas, el uso que se le dio fue bajo un marco clínico-exploratorio y con fines de investigación, excluyendo el uso de la herramienta en contextos clínicos regulados. Asimismo, el acceso a los datos solo está disponible bajo autorización, y se encuentra almacenado en dispositivos cifrados.

Según el Manual de usuario de la grabadora de aplicaciones inteligentes AIREC, se especifica que el dispositivo cumple con el apartado 15 de las normas de la FCC (Comisión Federal de Comunicaciones) de EE.UU.; el objetivo de dicha norma es específicamente regular y controlar la interferencia electromagnética (EMI) de los dispositivos electrónicos, tanto intencionales (teléfonos, WIFI o Bluetooth) como no intencionales (computadoras o microondas), para garantizar un funcionamiento seguro y confiable.

El consentimiento informado se presenta con un formato exhaustivo, completo y detallado, estableciendo y describiendo todos los aspectos relacionados con el procedimiento, confidencialidad, seguridad, el uso de la IA, el manejo de datos, y la opción de eliminación de datos sensibles o de relevancia personal si el paciente lo desea.

Normativas vigentes en México

Los protocolos metodológicos de recolección y manipulación de datos, así como con lo que respecta con la protección de los derechos de los participantes, se realizaron conforme a la Ley General de Salud, el Reglamento de la Ley General de Salud en Materia de Investigación para la Salud, la NOM-012-SSA-2012 y la Ley Federal de Protección de Datos Personales en Posesión de los Particulares.

Elementos Éticos

Se trabajó partiendo de las propuestas de Emmanuel EJ, et. Al. (2000), en las cuales se describen siete puntos a considerar para valorar si una investigación clínica cuenta con bases y prácticas éticas. Las propuestas son las siguientes: 1) Valor social o científico, 2) Validez científica, 3) Selección equitativa del sujeto, 4) Proporción favorable de riesgo-beneficio 5) Evaluación independiente 6) Consentimiento información 7) Respeto a los sujetos inscritos. También, el desempeño de la investigación partió de los principios éticos fundamentales, como la beneficencia, la justicia y el respeto a las personas.

RESULTADOS

Se analizaron los registros de 10 pacientes psiquiátricos en seguimiento clínico, con un rango de edad entre 25 y 55 años. El número de consultas documentadas osciló entre 4 y 13 por paciente, con una media de 7.7 consultas.

El análisis idioléctico mediante inteligencia artificial (IA) mostró una concordancia diagnóstica del 78% con el juicio clínico humano, mientras que la confiabilidad interobservador entre dos clínicos alcanzó el 82%.

Las discrepancias más notorias se observaron en la interpretación de:

Prosodia emocional, donde la IA identificó rasgos de impulsividad o afectividad negativa que los clínicos no consideraron patológicos.

Neologismos y disgregaciones, interpretados por la IA como psicopatológicos, mientras que los clínicos los atribuyeron en algunos casos a regionalismos o expresiones culturales locales.



Fragmentación narrativa, categorizada por la IA como signo de disociación cognitiva, pero interpretada por los clínicos como un efecto del bajo nivel educativo en ciertos pacientes.

En términos diagnósticos, los pacientes dentro del espectro de la psicosis y trastorno afectivo bipolar mostraron mayor frecuencia de neologismos, disgregaciones y alteraciones de la coherencia ideoafectiva. En contraste, los pacientes con trastornos depresivos y ansiosos presentaron un patrón lingüístico más estable, con predominio de alteraciones en prosodia y afectividad.



Tabla 1. Características clínicas y sociodemográficas de los pacientes (N=10)

Paciente	Edad	Sexo / Estado civil	Escolaridad	Diagnósticos principales	Tratamiento actual	N° consultas
1	54	Femenino / casada	Secundaria	Tr. de síntomas somáticos, TDM en remisión, duelo	Escitalopram 20 mg, Clonazepam 0.5 mg	9
2	27	Femenino / soltera	Lic. Ingeniería Industrial	Tr. depresivo recurrente en remisión, TDAH adulto, Tr. personalidad dependiente	Desvenlafaxina 100 mg, Metilfenidato 54 mg, Alprazolam 0.5 mg	11
3	32	Femenino / casada	Preparatoria	Esquizoafectivo bipolar en remisión	Litio 1200 mg, Risperidona 1 mg, Trazodona PRN	13
4	54	Femenino / soltera	Maestría	TAG, duelo complicado, rasgos afectividad negativa	Paroxetina 20 mg, Clonazepam 2 mg	7
5	25	Femenino / soltera	Lic. Contaduría Pública	TAG, Tr. depresivo recurrente en remisión	Escitalopram 10 mg, Clonazepam gotas	5
6	50	Femenino / casada	Preparatoria	Tr. por atracón, TDM leve, ansiedad no especificada	Mysimba, Clonazepam 2 mg (retiro gradual)	5
7	49	Femenino / soltera	Lic. Administración	TAG + ansiedad por enfermedad, estado postmenopáusico	Paroxetina 20 mg, Sulbutiamina 200 mg	4
8	55	Femenino / casada	Preparatoria	TAG, dependencia benzodiacepinas, TDM en remisión	Aripiprazol, Fluvoxamina, Alprazolam	6
9	32	Femenino / casada	Lic. Turismo	TAG, episodio depresivo leve	Escitalopram 15 mg, Bupropión 150 mg, Alprazolam 0.25 mg	7
10	44	Femenino / casada	Preparatoria	T. personalidad límite, TDM leve, dependencia BZD	Lamotrigina, Alprazolam, Quetiapina	10

Tabla 2. Codificación lingüística de rasgos idiolécticos (N=10)

Paciente	Rasgos idiolécticos predominantes	Manifestaciones clínicas	Correlato DSM-5-TR
1	Léxico afectivo, prosodia elevada, disgregaciones, hipercorrecciones	Regulación emocional parcial, duelo en proceso, ansiedad por desempeño	F43.21, F41.1, F34.1
2	Ambivalencia afectiva, silencios, somatización verbal	Evitación de conflicto familiar, duelo prolongado, estrés postraumático	F34.1, F43.22, F43.10
3	Repeticiones, prosodia irregular, disonancia narrativa, neologismos	Negación de síntomas, adherencia irregular, trauma complejo	F31.73, F43.10, F31.71
4	Neologismos (“cronosofá”), pausas largas, disociación, hipercorrecciones	Duelo complicado, ansiedad, fenómenos disociativos	F43.23, F41.1, F48.1
5	Repeticiones ansiosas, neologismos, hipercorrecciones, disociación narrativa	Ansiedad generalizada, somatización, distimia/adaptación	F41.1, F34.1, F45.41
6	Muletillas, ambivalencia afectiva, neologismos, pausas largas	Evasión emocional, ansiedad social, atracones, impulsividad	F60.7, F34.1, F44
7	Prosodia reflexiva, diminutivos, disociación, impulsividad alimentaria	Ansiedad generalizada, fobia específica, fatiga, depresión residual	F41.1, F40.29, F33.0
8	Neologismos (“bajones”), prosodia irregular, hipercorrecciones	Minimización de recaídas, ansiedad social, episodios depresivos	F33.x, F41.1, F40.10
9	Repeticiones, ambivalencia, hipercorrecciones	Estancamiento vital, anhedonia, ansiedad, problemas conyugales	F34.1, F32.9, F41.0
10	Léxico idiosincrático, ambivalencia, autolesiones, hipercorrecciones	Catastrofización, disonancia afectiva, impulsividad autodestructiva	F40.2, F60.3, F34.1



DISCUSIÓN

Nuestros hallazgos muestran que un esquema operativo de bajo costo (captura con AIREC, transcripción automática y análisis con modelos de IA generalistas) puede sintetizar narrativas clínicas de ~40 minutos y alcanzar una concordancia diagnóstica del 78% respecto al juicio clínico humano, con una confiabilidad interobservador del 82% como punto de referencia interno. Para un estudio piloto, estas cifras son operativamente alentadoras: si se considera que la fiabilidad test–retest de varios diagnósticos categoriales en los field trials del DSM-5 fue moderada, la proximidad de la IA a la fiabilidad humana sugiere viabilidad como herramienta auxiliar en el proceso diagnóstico (Regier et al., 2013). En otras palabras, cuando el entrevistador humano opera como “patrón oro contextual”, la IA comienza a perfilarse como un segundo lector estandarizado que tamiza lo idiosincrásico y aporta consistencia a la valoración clínica.

Este rendimiento resulta conceptualmente coherente con la literatura reciente que posiciona el lenguaje como psicobiomarcador: la prosodia, la coherencia narrativa y los desajustes semánticos han mostrado capacidad discriminativa y/o pronóstica en depresión, psicosis y demencia (Low, Bentley & Ghosh, 2020; Parola et al., 2022; Ehlen et al., 2023; Mirheidari et al., 2019; Pan et al., 2023). En nuestra muestra, la mayor densidad de neologismos, disgregaciones y alteraciones ideo-afectivas en cuadros del espectro esquizofrénico y bipolar, frente a un perfil prosódico-afectivo en trastornos depresivos y ansiosos, replica el gradiente descrito en esas revisiones. En términos clínicos, esta convergencia respalda la centralidad lingüística como puente entre la fenomenología (lo que el paciente dice y cómo lo dice) y la nosología (cómo lo clasificamos), y legitima el uso del análisis idioléctico computacional como “biomarcador blando” para afinar hipótesis diagnósticas y monitorear trayectorias.

Las ventajas observadas —rapidez, estandarización y reducción de los sesgos por fatiga o por estilo del entrevistador— son igualmente consistentes con marcos más amplios sobre IA clínica (Dziadkowiec et al., 2020; Nagendran et al., 2020; Liu et al., 2019). La estandarización no sustituye el juicio clínico, pero homogeneiza el umbral de alerta para ciertos rasgos (p. ej., fragmentación narrativa), y permite comparabilidad longitudinal dentro del mismo caso.



Con todo, las limitaciones son relevantes. Primero, la exactitud de transcripción ($\approx 35\%$) compromete la captura de micro-marcadores (interjecciones, disfluencias finas, solapamientos), con impacto probable en la detección de neologismos y en métricas de coherencia local. Segundo, la sensibilidad cultural de los modelos resultó insuficiente: regionalismos del noroeste y rasgos dialectales del español de Sonora —incluida la variación sociolingüística documentada para /s/ y léxico local— pueden simular psicopatología si no se contextualizan (Granados, 2017; Núñez-Méndez, 2022; Lope Blanch et al., 1990). Este punto justifica entrenar con corpus locales y acoplar la evaluación idioléctica a marcos de formulación cultural (CFI del DSM-5), cuya utilidad ha sido recientemente sintetizada (Kayrouz et al., 2024). Tercero, el riesgo de sesgos propios de los LLM debe monitorizarse con marcos de evaluación específicos para soporte clínico (Templin, Boehme & Rieckmann, 2025) y con lineamientos éticos y de gobernanza para LLM en salud (World Health Organization, 2024) y el AI Act europeo (European Parliament & Council, 2024).

En cuanto a perspectivas, el siguiente ciclo metodológico debiera: (a) ampliar la muestra y migrar de un diseño retrospectivo/transversal a cohortes prospectivas con puntos de corte predefinidos; (b) optimizar la ASR (modelos acústicos para acentos regionales y domain adaptation clínica) con verificación humana selectiva; (c) entrenar y validar modelos con corpus locales anonimizados, incorporando metadatos culturales y educativos; (d) integrar el CFI de manera sistemática como “capa cultural” del pipeline; (e) explorar métricas transdiagnósticas alineadas con RDoC asistidas por LLM (McCoy & Perlis, 2024), y (f) evaluar impacto clínico real (tiempo de consulta, tasa de re-entrevista, cambios en precisión diagnóstica y decisiones terapéuticas). Bajo estas condiciones, la IA idioléctica puede convertirse en un co-evaluador semántico-prosódico que robustece la clínica, especialmente en contextos con déficit de personal y alta diversidad lingüística.

CONCLUSIONES

Un análisis idioléctico asistido por IA, aun con transcripción subóptima ($\sim 35\%$), alcanzó 78% de concordancia con el diagnóstico humano y se aproximó a la fiabilidad interobservador (82%), lo que avala su uso como herramienta diagnóstica auxiliar en entornos de recursos limitados.



Los patrones lingüísticos observados por diagnóstico (neologismos y disgregación en espectros psicóticos/afectivos; predominio prosódico-afectivo en depresión/ansiedad) convergen con la evidencia internacional y sostienen el lenguaje como psicobiomarcador útil para refinar hipótesis clínicas y monitorear procesos.

La estandarización y rapidez del esquema operativo ofrecen ganancias técnicas (tamizaje semántico-prosódico, comparabilidad longitudinal), sin reemplazar el juicio clínico ni la formulación cultural.

La sensibilidad a regionalismos y la baja exactitud de ASR son cuellos de botella abordables mediante corpus locales, adaptación acústica a acentos y verificación humana focalizada.

La integración ética y regulatoria —alineada con guías OMS para LMM en salud y marcos normativos como el AI Act— es condición necesaria para su despliegue clínico responsable.

Futuras fases deben ampliar muestra, prospectivar el diseño, incorporar CFI de rutina y medir impacto clínico (precisión diagnóstica, decisiones terapéuticas, eficiencia), consolidando a la IA idioléctica como co-evaluador que complementa la entrevista psiquiátrica y fortalece la precisión diagnóstica.

Agradecimientos

A los pacientes y colaboradores que contribuyeron al estudio.

Declaración de conflictos de interés

Ninguno declarado.

Fuentes de financiamiento

Sin financiamiento externo.

Declaración de contribuciones

Carlos Armando Herrera-Huerta: concepción y diseño del estudio, análisis de resultados, revisión crítica.

Angélica Gándara-López: revisión de material y recolección de datos, transcripción, análisis lingüístico, redacción del manuscrito.

Pedro Moreno-Gea: supervisión académica, revisión crítica, aportes teóricos en IA y psiquiatría digital.



REFERENCIAS BIBLIOGRAFICAS

- Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S. W., Karthikesalingam, A., King, D., Ashrafian, H., & Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging: A systematic review and meta-analysis. *NPJ Digital Medicine*, 4(1), 65. <https://doi.org/10.1038/s41746-021-00438-z>
- Bird, A. D., Jedlicka, P., & Cuntz, H. (2021). Dendritic normalisation improves learning in sparsely connected artificial neural networks. *PLoS computational biology*, 17(8), e1009202. <https://doi.org/10.1371/journal.pcbi.1009202>
- Buitimea Valenzuela, C., Estrada Fernández, Z., Grageda Bustamante, A., & Silva Encinas, M. C. (2016). Diccionario yaqui de bolsillo: Jiak noki español, español Jiak noki. Universidad de Sonora. https://letrasylinguistica.unison.mx/wp-content/uploads/2018/09/2016Diccionario_Yaqui_Espanol.pdf
- Cámara de Diputados del H. Congreso de la Unión. (2025). Ley Federal de Protección de Datos Personales en Posesión de los Particulares (Nueva Ley publicada en el Diario Oficial de la Federación el 20 de marzo de 2025). <https://www.diputados.gob.mx/LeyesBiblio/pdf/LFPDPPP.pdf>
- Carnegie Mellon University Libraries. (2020, 6 de julio). The imitation game: Rare Alan Turing article at CMU Libraries. <https://www.library.cmu.edu/about/news/2020-07/imitation-game-rare-alan-turing-article-cmu-libraries>
- Chavlis, S., & Poirazi, P. (2025). Dendrites endow artificial neural networks with accurate, robust and parameter-efficient learning. *Nature Communications*, 16(1), 1193. <https://doi.org/10.1038/s41467-025-56297-9>
- Chekroud, A. M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., Cohen, Z., Belgrave, D., DeRubeis, R., Iniesta, R., Dwyer, D., & Choi, K. (2021). The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry*, 20(2), 154–170. <https://doi.org/10.1002/wps.20882>
- Colby, K. M. (1972). PARRY: A computer simulation of paranoid behavior. Stanford University.



- Comisión Nacional de Bioética; Consejo Nacional de Ciencia y Tecnología (CONACYT). (2015). Ética de la investigación: integridad científica [Libro]. https://www.conbioetica-mexico.salud.gob.mx/descargas/pdf/Libro_Etica_de_la_Investigacion_gratuito.pdf
- Congreso de los Estados Unidos Mexicanos. (1994). Ley General de Salud (última reforma 07-06-2024). <https://www.diputados.gob.mx/LeyesBiblio/pdf/LGS.pdf>
- Congreso de los Estados Unidos Mexicanos. (2014). Reglamento de la Ley General de Salud en materia de investigación para la salud (Reg_LGS_MIS). https://www.diputados.gob.mx/LeyesBiblio/regley/Reg_LGS_MIS.pdf
- Crespo Cobo, Y., Esteban, F. J., Iglesias Parro, S., Paredes Olay, C., & Molina López, R. (2025). La escritura como biomarcador del estado de salud mental. *Seminario Médico*, 65(1), 11–22.
- Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: Origins of bias and theory of debiasing. *BMJ Quality & Safety*, 22(Suppl 2), ii58–ii64. <https://doi.org/10.1136/bmjqs-2012-001712>
- Curtiss, J., & DiPietro, B. C. (2025). Machine learning in the prediction of treatment response for emotional disorders: A systematic review and meta-analysis. *Clinical Psychology Review*, 120, 102593. <https://doi.org/10.1016/j.cpr.2025.102593>
- Dell’Armo, K., & Tassé, M. J. (2023). Diagnostic overshadowing of psychological disorders in people with intellectual disability: A systematic review. *American Journal on Intellectual and Developmental Disabilities*. Advance online publication. https://www.aaid.org/docs/default-source/prepressarticles/diagnostic-overshadowing-of-psychological-disorders-in-people-with-intellectual-disability-systematic-review.pdf?sfvrsn=1fd20221_0
- Dikaios, K., Rempel, S., Dumpala, S. H., Oore, S., Kiefe, M., & Uher, R. (2023). Applications of speech analysis in psychiatry. *Harvard Review of Psychiatry*, 31(1), 1–13. <https://doi.org/10.1097/HRP.0000000000000356>
- Doherty, T. S., & Balzer, W. K. (2020). Cognitive and affective processes in clinical decision-making: Implications for ethical practice. *AMA Journal of Ethics*, 22(9), E742–E748. <https://doi.org/10.1001/amajethics.2020.742>



- Dziadkowiec, O., Callahan, K. E., O'Connor, S., & Williams, J. S. (2020). Artificial intelligence in health care: Benefits and challenges of implementation. *HCA Healthcare Journal of Medicine*, 1(0), 1–9. <https://doi.org/10.36518/2689-0216.1010>
- Ehlen, F., Montag, C., Leopold, K., & Heinz, A. (2023). Linguistic findings in persons with schizophrenia: A review of the current literature. *Frontiers in Psychology*, 14, 1287706. <https://doi.org/10.3389/fpsyg.2023.1287706>
- Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA*. 2000 May 24-31;283(20):2701-11. doi: 10.1001/jama.283.20.2701. PMID: 10819955.
- European Parliament, & Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending certain Union legislative acts (Artificial Intelligence Act). *Official Journal of the European Union*, L, 12.7.2024, 1–144. <http://data.europa.eu/eli/reg/2024/1689/oj>
- Fernandino, L., Binder, J. R., Desai, R. H., Humphries, C. J., Gross, W. L., Conant, L. L., & Seidenberg, M. S. (2022). Decoding the information structure underlying the neural representational code of lexical semantics in heteromodal cortex. *Proceedings of the National Academy of Sciences*, 119(52), e2207398120. <https://doi.org/10.1073/pnas.2207398120>
- Franco Trujillo, E. D. (2020). Reflexiones en torno a la variación léxica en el ámbito hispánico y la traducción. *Estudios de Lingüística Aplicada*, 38(71), 145–176. <https://doi.org/10.22201/enallt.01852647p.2020.71.905>
- Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized controlled trial. *JMIR Mental Health*, 5(4), e64. <https://doi.org/10.2196/mental.9782>
- Fusar-Poli, P., Davies, C., Karpenko, O., Christodoulou, N., Ramalho, R., Nandha, S., Damiani, S., Provenzani, U., Esposito, C. M., Mensi, M. M., Borgatti, R., Stefana, A., McGuire, P., & Minichino, A. (2019). Preventive treatments for psychosis: Umbrella review. *Frontiers in Psychiatry*, 10, 764. <https://doi.org/10.3389/fpsyg.2019.00764>



- Future Care Capital. (2023). Analysis of speech patterns can diagnose mental illness. <https://futurecarecapital.org.uk/latest/analysis-of-speech-patterns-can-diagnose-mental-illness/>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. <https://www.deeplearningbook.org>
- Granados, A. (2017). “¡Oye, tú, bato!, ¿Dónde están los plebes y tu morra?” Aproximación lexicológica, mediante la dialectología perceptual, al habla de Bahía de Kino y Guaymas, Sonora. *Káñina, Revista de Artes y Letras, Universidad de Costa Rica*, 41(1), 59–68. <https://revistas.ucr.ac.cr/index.php/kanina/article/view/29379>
- Greden, J. F., Parikh, S. V., Rothschild, A. J., et al. (2019). Effect of pharmacogenomic testing for drug-gene interactions on medication selection and remission of symptoms in major depressive disorder. *JAMA Psychiatry*, 76(1), 19–28. <https://jamanetwork.com/journals/jama/fullarticle/2794053>
- Heim, E., Karatzias, T., & Maercker, A. (2022). Cultural concepts of distress and complex PTSD: Future directions. *Clinical Psychology Review*, 93, 102143. <https://doi.org/10.1016/j.cpr.2022.102143>
- Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7), 1527–1554. <https://doi.org/10.1162/neco.2006.18.7.1527>
- Holmes, E. A., Ghafur, S., & Shafran, R. (2021). Artificial intelligence in mental health: The state of the art and future prospects. *Nature Medicine*, 27(5), 766–781. <https://doi.org/10.1038/s41591-021-01362-5>
- IBM. (s. f.). Large language models. IBM. Recuperado el 24 de septiembre de 2025, de <https://www.ibm.com/think/topics/large-language-models>
- Instituto Mexicano del Seguro Social. (2012). Norma Oficial Mexicana NOM-012-SSA3-2012, para la vigilancia epidemiológica, prevención, y control de las infecciones nosocomiales (versión publicada). https://www.imss.gob.mx/sites/all/statics/profesionalesSalud/investigacionSalud/normativaNac/4_NOM-012-SSA-3-2012.pdf



- Instituto Nacional de los Pueblos Indígenas. (2020). Lengua yaqui [Atlas de los Pueblos Indígenas de México]. Gobierno de México. <https://atlas.inpi.gob.mx/yaquis-lengua/>
- Jiang, J., Liu, R., & Demner-Fushman, D. (2023). Large language models for the mental health community: Framework for translating code to care. *NPJ Digital Medicine*, 6, 150. <https://doi.org/10.1038/s41746-023-00849-2>
- Johnson, B. T., Low, T. L., & MacDonald, G. (2019). Advances in meta-analysis for health sciences: Research synthesis methods and applications. *Research Synthesis Methods*, 10(2), 126–145. <https://doi.org/10.1002/jrsm.1337>
- Kaul, V. (2020). History of artificial intelligence in medicine. *Gastrointestinal Endoscopy*, 91(3), 523–532. <https://doi.org/10.1016/j.gie.2019.10.028>
- Kaul, V. (2020). History of artificial intelligence in medicine. *Gastrointestinal Endoscopy*, 91(3), 523–532. <https://doi.org/10.1016/j.gie.2019.10.028>
- Kayrouz, R., Karin, E., Staples, L., et al. (2024). A quantitative systematic review to evaluate the favorability of the DSM-5 CFI. *Journal of Cross-Cultural Psychology*, 55(8), 835–864. <https://doi.org/10.1177/00220221241269994>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105. <https://doi.org/10.1145/3065386>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lefkowitz, M. (2019, 25 de septiembre). Professor's perceptron paved the way for AI – 60 years too soon. *Cornell Chronicle*. <https://news.cornell.edu/stories/2019/09/professors-perceptron-paved-way-ai-60-years-too-soon>
- Li, H., Wong, C. K. H., & Cheung, C. L. (2023). Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *NPJ Digital Medicine*, 6, 180. <https://doi.org/10.1038/s41746-023-00849-2>



- Li, R., Wang, X., Lawler, K., Garg, S., Bai, Q., & Alty, J. (2022). Applications of AI to aid early detection of dementia. *Journal of Biomedical Informatics*, 127, 104030. <https://arxiv.org/abs/2104.14073>
- Lin, Y., Ding, H., & Zhang, Y. (2018). Emotional prosody processing in schizophrenia: A meta-analysis. *Journal of Clinical Medicine*, 7(10), 363. <https://doi.org/10.3390/jcm7100363>
- Liu, X., Faes, L., Kale, A. U., et al. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *The Lancet Digital Health*, 1(6), e271–e297. [https://doi.org/10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2)
- Lope Blanch, J. M., Alcalá Alba, A., Ávila, R., et al. (1990). *Atlas lingüístico de México. Tomo I: Fonética (Vol. 1)*. El Colegio de México.
- Low, D. M., Bentley, K. H., & Ghosh, S. S. (2020). Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investigative Otolaryngology*, 5(1), 96–116. <https://doi.org/10.1002/lio2.354>
- Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, 91, 5–27. <https://doi.org/10.1016/j.jml.2016.03.009>
- Manuales+. (s. f.). Manual de usuario de la grabadora de aplicaciones inteligentes AIREC [Manual]. <https://manuals.plus/es/airec/smart-app-recorder-manual>
- Mbuagbaw, L., Lawson, D. O., Puljak, L., Allison, D. B., & Thabane, L. (2020). A tutorial on methodological studies: The what, when, how and why. *BMC Medical Research Methodology*, 20, 226. <https://doi.org/10.1186/s12874-020-01140-4>
- McCoy, T. H., & Perlis, R. H. (2024). Characterizing RDoC symptoms among psychiatric inpatients using LLMs. *Journal of Mood & Anxiety Disorders*, 8, 100079. <https://pubmed.ncbi.nlm.nih.gov/40655912/>
- Méndez Ramírez, I., Namihira Guerrero, D., Moreno Altamirano, L., & Sosa de Martínez, C. (1990). *El protocolo de investigación: Lineamientos para su elaboración y análisis (2.ª ed., reimpresión 1996)*. México: Trillas.



- Mirheidari, B., Blackburn, D., Walker, T., Reuber, M., & Christensen, H. (2019). Dementia detection using automatic analysis of conversations. *Computer Speech & Language*, 53, 65–79. https://eprints.whiterose.ac.uk/id/eprint/136339/8/CSL_2018_Revised.pdf
- Nagendran, M., Chen, Y., Lovejoy, C., Gordon, A. C., Komorowski, M., Harvey, H., Topol, E. J., ... & Maruthappu, M. (2020). Artificial intelligence versus clinicians: Systematic review of design, reporting standards, and claims of deep learning studies in medical imaging. *The BMJ*, 368, m689. <https://doi.org/10.1136/bmj.m689>
- Newson, J. J., Pastukh, V., & Thiagarajan, T. C. (2021). Poor separation of clinical symptom profiles by DSM-5 disorder criteria. *Frontiers in Psychiatry*, 12, 775762. <https://doi.org/10.3389/fpsyt.2021.775762>
- Noriega Jacob, M. A. (Comp.). (2016). Estudios del lenguaje en el noroeste de México. Centro Regional de Formación Docente e Investigación Educativa del Estado de Sonora. https://creson.sonora.edu.mx/images/INVESTIGACION/LIBROS/estudios_del_lenguaje_en_el_noroeste_de_mexico_a125387f5b_compressed.pdf
- Núñez-Méndez, E. (2022). Variation in Spanish /s/: Overview and new perspectives. *Languages*, 7(2), 77. <https://doi.org/10.3390/languages7020077>
- O’Sullivan, E. D., & Schofield, S. J. (2018). Cognitive bias in clinical medicine. *Journal of the Royal College of Physicians of Edinburgh*, 48(3), 225–232. <https://doi.org/10.4997/JRCPE.2018.306>
- Organisation for Economic Co-operation and Development. (2019). OECD principles on artificial intelligence. OECD. <https://oecd.ai/en/ai-principles>
- Organización Mundial de la Salud. (2021). La OMS publica el primer informe mundial sobre inteligencia artificial (IA) aplicada a la salud y seis principios rectores relativos a su concepción y utilización. <https://www.who.int/es/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use>
- Oxford University. (2024, January 3). The way the brain learns is different from AI systems: prospective configuration mechanism. <https://www.ox.ac.uk/news/2024-01-03-study-shows-way-brain-learns-different-way-artificial-intelligence-systems-learn>



- Paass, G. (2021, 21 de abril). Deep learning: How do deep neural networks work? Lamarr Institute.
<https://lamarr-institute.org/blog/deep-neural-networks/>
- Pan, Y., Zhang, T., Wu, L., et al. (2023). Diagnostic accuracy of deep learning using speech samples in depression: A systematic review and meta-analysis. *Psychological Medicine*, 53(10), 4562–4576. <https://doi.org/10.1017/S0033291723001854>
- Parola, A., et al. (2023). Voice patterns as markers of schizophrenia: A systematic review. *Schizophrenia Bulletin*, 49(Suppl.), s. p.
- Parola, A., Simonsen, A., Bliksted, V., & Fusaroli, R. (2022). Language abnormalities in schizophrenia: Binding core symptoms through contemporary empirical evidence. *Schizophrenia Bulletin*, 48(1), 180–192. <https://doi.org/10.1093/schbul/sbac152>
- Pearson, A. L., Tribby, C., Brown, C. D., Yang, J.-A., Pfeiffer, K., & Jankowska, M. M. (2024). Systematic review of best practices for GPS data usage, processing, and linkage in health, exposure science and environmental context research. *BMJ Open*, 14, e077036. <https://doi.org/10.1136/bmjopen-2023-077036>
- Pople, H. E., Jr., Miller, R. A., & Myers, J. D. (1982). Internist-1: An experimental computer-based diagnostic consultant for general internal medicine. *New England Journal of Medicine*, 307(8), 468–476. <https://doi.org/10.1056/NEJM198208193070804>
- Quaak, M., van de Mortel, L., Thomas, R. M., & van Wingen, G. (2021). Deep learning applications for the classification of psychiatric disorders using neuroimaging data: Systematic review and meta-analysis. *NeuroImage: Clinical*, 30, 102584. <https://doi.org/10.1016/j.nicl.2021.102584>
- Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 field trials in the United States and Canada, Part II: Test-retest reliability of selected categorical diagnoses. *American Journal of Psychiatry*, 170(1), 59–70. <https://doi.org/10.1176/appi.ajp.2012.12070999>
- Reid, M. C., Carpenter, J. S., & Naylor, M. D. (2019). Diagnostic error in mental health: A review. *General Hospital Psychiatry*, 60, 12–19. <http://qualitysafety.bmj.com/>
- Robins, L. N., Helzer, J. E., Croughan, J. L., & Ratcliff, K. S. (1981). National Institute of Mental Health diagnostic interview schedule: Its history, characteristics, and validity. *Archives of*



- General Psychiatry, 38(4), 381–389. <https://doi.org/10.1001/archpsyc.1981.01780290015001>
- Rony, M. K. K., Das, D. C., Khatun, T., et al. (2025). Artificial intelligence in psychiatry: A systematic review and meta-analysis of diagnostic and therapeutic efficacy. *Digital Health*, 11, 1–19. <https://doi.org/10.1177/20552076251330528>
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- Ryan, K., Yang, H., Kim, B., & Kim, J. P. (2025). Assessing the impact of AI on physician decision-making for mental health treatment in primary care. *npj Mental Health Research*, 4, 16. <https://doi.org/10.1038/s44184-025-00016-1>
- Salinas, M. P., Sepúlveda, J., Hidalgo, L., Peirano, D., Morel, M., Uribe, P., Rotemberg, V., Briones, J., Mery, D., & Navarrete-Dechent, C. (2024). A systematic review and meta-analysis of artificial intelligence versus clinicians for skin cancer diagnosis. *NPJ Digital Medicine*, 7(1), 125. <https://doi.org/10.1038/s41746-024-01103-x>
- Saposnik, G., Redelmeier, D. A., Ruff, C. C., & Tobler, P. N. (2016). Cognitive biases associated with medical decisions: A systematic review. *BMC Medical Informatics and Decision Making*, 16, 138. <https://doi.org/10.1186/s12911-016-0377-1>
- Sarangi, S., & Sharma, P. (2025). *Inteligencia artificial: evolución, ética y políticas públicas* (S. M. Olivares Bari, Trad.). Trillas. (Obra original publicada en 2018). ISBN 978-607-17-4862-1
- Secretaría de Economía. (2020). *Guaymas: Perfil sociodemográfico y económico*. DataMéxico. <https://www.economia.gob.mx/datamexico/es/profile/geo/guaymas>
- Shen, H., Zhang, L., Xu, C., Zhu, J., Chen, M., & Fang, Y. (2018). Analysis of misdiagnosis of bipolar disorder in an outpatient setting. *Shanghai Archives of Psychiatry*, 30(2), 93–101. <https://doi.org/10.11919/j.issn.1002-0829.217080>
- Shortliffe, E. H., Buchanan, B. G., & Cohen, S. N. (1976). MYCIN: An early expert system for bacterial infection diagnosis and antibiotic recommendations [Tesis doctoral, Universidad de Stanford]. Stanford University.
- Spitzer, R. L., & colaboradores. (1970). *Psychiatric Status Schedule*. New York State Psychiatric Institute.



- Templin, T., Boehme, P., & Rieckmann, N. (2025). Evaluation framework for bias in large language models for clinical decision support. *NPJ Digital Medicine*, 8, 123. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12234702/>
- Time. (2024, agosto 7). Microsoft's AI is now better than doctors at diagnosing patients. *Time Magazine*. <https://time.com/7299314/microsoft-ai-better-than-doctors-diagnosis/>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Wang, S. (2025, agosto 13). Capabilities of GPT 5 on multimodal medical reasoning (arXiv:2508.08224 v2) [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2508.08224>
- Wang, Y., Chen, H., Patel, R., et al. (2024). Evaluating BERT-based and large language models for suicide detection, prevention, and risk assessment: A systematic review. *Frontiers in Psychiatry*, 15, 138902. <https://doi.org/10.3389/fpsyt.2024.138902>
- 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. <https://doi.org/10.1145/365153.365168>
- World Health Organization. (2024). Ethics and governance of artificial intelligence for health: Guidance on large multimodal models. Geneva: WHO. <https://www.who.int/publications/i/item/9789240084759>
- Wu, Z., Wang, J., Zhang, C., Peng, D., Mellor, D., Luo, Y., & Fang, Y. (2024). Clinical distinctions in symptomatology and psychiatric comorbidities between misdiagnosed bipolar I and bipolar II disorder versus major depressive disorder. *BMC Psychiatry*, 24, 352. <https://doi.org/10.1186/s12888-024-05810-3>
- Zhang, Y., Ji, Y., & Papafragou, A. (2023). Representation of event boundedness in English and Mandarin speakers. *Journal of Memory and Language*, 132, 104373. <https://doi.org/10.1016/j.jml.2022.104373>
- Zhu, J., & Jurgens, D. (2021). Idiosyncratic but not arbitrary: Learning idiolects in online registers reveals distinctive yet consistent individual styles. In M.-F. Moens, X. Huang, L. Specia, & S. W.-t. Yih (Eds.), *Proceedings of the 2021 Conference on Empirical Methods in Natural*



Language Processing (pp. 279–297). Association for Computational Linguistics.

<https://doi.org/10.18653/v1/2021.emnlp-main.25>

