



Marco EEPA para la evaluación ética de la inteligencia artificial en el bienestar universitario

The EEPA framework for ethical evaluation of AI in university well-being

Emgelberth Vargas, PhD

<https://orcid.org/0000-0003-3098-0719>

Instituto Pedagógico de Caracas (IPC) / Universidad Pedagógica Experimental Libertador (UPEL)
Escuela de Psicología, Carver University

Alberto Luis Ramírez

<https://orcid.org/0009-0002-7678-6639>

Escuela de Ingeniería de Sistemas, Carver University

Fecha de recepción: 16 de abril de 2026

Fecha de aceptación: 14 de mayo de 2026

Fecha de publicación: 1 de julio de 2026

Favor de citar este artículo de la siguiente forma:

Vargas, E., y Ramírez, A. L. (2026). Marco EEPA para la evaluación ética de la inteligencia artificial en el bienestar universitario.

AULA Revista de Humanidades y Ciencias Sociales, 72 (2), <https://doi.org/10.33413/aulahcs.2026.72i2.484>

RESUMEN

La incorporación de sistemas de inteligencia artificial (IA) en el acompañamiento emocional y psicológico de estudiantes universitarios, ha desplazado lo que parecía ser un problema técnico hacia una cuestión ética de primer orden. Este artículo examina, desde una perspectiva de revisión crítica y análisis documental, la relación entre la mediación algorítmica del bienestar psicológico y la experiencia subjetiva de adultos jóvenes en educación superior (denominados en este trabajo estudiantes universitarios o, en sentido amplio, adolescentes tardíos conforme a la definición de la OMS, que extiende este período hasta los 24 años), a partir de literatura especializada publicada entre 2019 y 2025. El análisis identifica tres tensiones que estructuran el debate contemporáneo, aunque no con el mismo peso. La más explorada, y paradójicamente la peor planteada, concierne a la eficacia parcial de los sistemas de IA frente a sus limitaciones empáticas constitutivas. La segunda, menos visible pero igualmente urgente, es el riesgo de que la narrativa democratizadora del acceso encubra procesos de mercantilización del malestar estudiantil. La tercera, y a juicio de los autores la más subestimada en la literatura reciente, es la posibilidad de que el uso intensivo de estas herramientas erosione, más que fortalezca, la autonomía psicológica del sujeto en formación. Los resultados indican que la evidencia disponible muestra efectos positivos acotados a síntomas leves en el corto plazo, mientras que los riesgos de dependencia algorítmica, mercantilización de datos emocionales y desresponsabilización institucional permanecen sistemáticamente subestimados. Frente a ese estado del conocimiento, el artículo propone el marco EEPA, Eficacia, Equidad, Privacidad y Autonomía, como instrumento analítico para la evaluación crítica de cualquier sistema de IA orientado al bienestar estudiantil. Se concluye que la legitimidad ética de estas herramientas depende de condiciones institucionales, pedagógicas y regulatorias específicas, cuya construcción es una responsabilidad impostergable de la educación superior.

Palabras clave: *inteligencia artificial en educación, bienestar psicológico estudiantil, ética de la IA, salud mental digital, gobernanza universitaria, dependencia algorítmica*

ABSTRACT

The incorporation of artificial intelligence (AI) systems into the emotional and psychological support of university students has turned a technical problem into a first-order ethical concern. Drawing on a critical review and documentary analysis of specialized literature published between 2019 and 2025, this article examines the relationship between algorithmic mediation of psychological well-being and the subjective experience of adolescents in higher education.

The study adopts a documentary analytical review design (Fink, 2019; Grant & Booth, 2009), drawing on 37 primary sources retrieved from Scopus, PsycINFO, ERIC, and Web of Science, supplemented by grey literature searches in Google Scholar. The review period spans specialized literature published between 2019 and 2025.

The analysis identifies three tensions that structure the contemporary debate, though not with equal weight. The most discussed, and paradoxically the most poorly framed, concerns the partial efficacy of AI systems against their constitutive empathic limitations. The second, less visible but equally urgent, is the risk that democratizing access narratives may conceal processes of commodification of student distress. The third, and in the authors' view the most underestimated in the recent literature, is the possibility that intensive use of these tools erodes, rather than strengthens, the psychological autonomy of the student as a subject in formation. The findings indicate that available evidence shows positive effects limited to mild symptoms over the short term, while risks of algorithmic dependency, commodification of emotional data, and institutional disengagement remain systematically underestimated. In response, the article proposes the EEPA framework, Efficacy, Equity, Privacy, and Autonomy, as an analytical instrument for the critical evaluation of AI systems oriented toward student well-being. It concludes that the ethical legitimacy of these tools depends on specific institutional, pedagogical, and regulatory conditions, the construction of which is an urgent responsibility of higher education.

Keywords: *artificial intelligence in education, student psychological well-being, AI ethics, digital mental health, university governance, algorithmic dependency*

Introducción

Hay algo revelador en el modo en que la inteligencia artificial llegó al campo de la salud mental universitaria: no como resultado de una decisión deliberada de política educativa, sino como consecuencia acumulada de respuestas pragmáticas a una crisis que las instituciones no supieron anticipar. La demanda de atención psicológica en las universidades latinoamericanas y globales supera con creces la oferta disponible. Los servicios de orientación resultan insuficientes, los tiempos de espera se extienden semanas o meses, y los estudiantes que experimentan malestar emocional encuentran, con frecuencia creciente, que las plataformas digitales responden más rápido que los profesionales institucionales. En ese vacío no diseñado, no planificado, la IA ha llegado como sustituto de facto.

Este fenómeno merece más atención de la que habitualmente se le dispensa. La literatura especializada comienza a revelar tensiones que no se resuelven ni con optimismo tecno-

lógico ni con desconfianza reflexiva. Algunos estudios documentan efectos positivos de ciertos sistemas de IA en la reducción de síntomas ansiosos y en el acceso a recursos de salud mental (Inkster et al., 2018; Woebot Health, 2021). Otros señalan, con igual fundamento empírico, los riesgos de desplazar la relación terapéutica humana, de mercantilizar el malestar psicológico y de exacerbar vulnerabilidades en poblaciones que ya eran frágiles antes de abrir la aplicación (Stade et al., 2024; Bickmore et al., 2018). La pregunta relevante, entonces, no es si la IA puede hacer algo útil en este dominio. Es en qué condiciones ese algo útil no viene acompañado de consecuencias que todavía no estamos midiendo.

El presente artículo se propone examinar esa pregunta con el rigor que exige. La interrogante central que lo orienta es: ¿en qué condiciones puede la integración de sistemas de inteligencia artificial en el apoyo al bienestar psicológico universitario resultar ética-

mente legítima y pedagógicamente valiosa, y cuándo reproduce o amplifica los riesgos que pretende mitigar? De esa pregunta central se desprenden tres más acotadas: ¿qué dimensiones del bienestar psicológico admiten mediación algorítmica y cuáles requieren presencia humana insustituible?; ¿qué marcos éticos y regulatorios existen, o deberían existir, para gobernar el uso de IA en salud mental estudiantil?, y ¿cómo afecta esa integración al desarrollo de la autonomía psicológica del estudiante como sujeto en formación?

Para abordar estas cuestiones, el artículo persigue tres objetivos específicos: analizar críticamente la evidencia disponible sobre la eficacia y los riesgos de los sistemas de IA como mediadores del bienestar psicológico en adolescentes universitarios; examinar las tensiones éticas estructurales que emergen de esa incorporación, con énfasis en las dimensiones de equidad, privacidad y autonomía; y proponer un marco analítico que oriente tanto la investigación futura como las políticas universitarias en este campo.

El supuesto teórico que articula el análisis postula que la inteligencia artificial no es, en sí misma, ni solución ni problema para el bienestar estudiantil: su valor o su riesgo depende de las condiciones institucionales, éticas y pedagógicas en que se inserta, y de la posición que ocupa el estudiante en relación con la herramienta. Este supuesto conecta, por un lado, con la tradición de la psicología crítica y la pedagogía de la autonomía (Freire, 2005; Hooks, 1994; Cushman, 2011) y, por otro, con los debates más recientes sobre gobernanza algorítmica en educación (Prinsloo & Slade, 2017; Selwyn, 2019). La convergencia entre estas dos tradiciones resulta, como se verá, considerablemente más productiva de lo que sugiere su distancia disciplinar aparente.

Marco Teórico

Bienestar psicológico en la educación superior: por qué la complejidad importa

Conviene comenzar por una advertencia conceptual que con frecuencia se omite en los estudios sobre tecnología y salud mental. El bienestar psicológico en el contexto universitario no es un estado discreto que se tiene o no se tiene. Es una configuración dinámica de experiencias subjetivas atravesadas por condiciones estructurales que ninguna aplicación de móvil puede modificar por sí sola. El modelo multidimensional de Ryff (1989), aunque ya clásico, y precisamente por eso, sigue siendo el más adecuado para dar cuenta de esta complejidad: sus seis dimensiones, autoaceptación, propósito vital, dominio ambiental, crecimiento personal, relaciones positivas y autonomía, configuran una visión del florecimiento humano que trasciende con mucho la ausencia de síntomas o la mera regulación del estado de ánimo.

Esta perspectiva, conocida como bienestar eudaimónico, tiene una implicación metodológica que raramente se hace explícita: gran parte de los instrumentos de evaluación utilizados en los estudios sobre IA y salud mental universitaria capturan, en el mejor de los casos, uno o dos de esos componentes. Cuando Fitzpatrick et al. (2017) reportan reducciones en las puntuaciones de ansiedad autodeclarada tras el uso de Woebot, lo que miden es un efecto sobre el componente afectivo negativo del bienestar hedónico. No el sentido de propósito. No la calidad de los vínculos. No la capacidad del estudiante de gobernarse a sí mismo. Esta distinción no es trivial ni meramente académica: define exactamente qué estamos esperando de estos sistemas y qué no podemos pedirles.

A esa complejidad conceptual se añaden las determinaciones estructurales que caracterizan el bienestar en contextos latinoamericanos, y que la mayoría de las plataformas de IA, diseñadas en el norte global, no consideran. Como señalan Matos et al. (2022), para muchos estudiantes universitarios la primera generación en sus familias que accede a educación superior, la experiencia universitaria es simultáneamente una oportunidad de movilidad social y una

fuentes de disonancia identitaria que ninguna encuesta de bienestar recoge con fidelidad. La precariedad económica, el conflicto entre exigencias académicas y responsabilidades laborales, la escasa integración cultural en instituciones que siguen reproduciendo patrones pedagógicos coloniales en formas sutiles o evidentes; todo eso conforma el sustento en el que se instala el malestar. Ignorarlo al diseñar intervenciones tecnológicas no es un descuido metodológico menor; es un error epistemológico. La literatura crítica ha denominado a este error psicologismo individualista: la tendencia a tratar síntomas de malestar social como disfunciones individuales susceptibles de regulación técnica (Fernández-Berrocal et al., 2021).

Los datos epidemiológicos agravan la urgencia del problema, aunque no simplifican su naturaleza. La Organización Mundial de la Salud (OMS, 2022) estima que cerca del 20% de los estudiantes universitarios experimenta algún tipo de trastorno mental durante su trayectoria académica. Estudios específicos en población latinoamericana reportan índices de ansiedad académica que oscilan entre el 25% y el 45%, según el tipo de institución y el nivel de estudio (Barraza-Macías, 2022; Bermúdez-Díaz & Villanueva-Hernández, 2021). Estas cifras justifican plenamente la búsqueda de soluciones, justificar la búsqueda, sin embargo, no equivale a validar cualquier respuesta.

IA y mediación del cuidado emocional: una historia más larga de lo que parece

La fascinación, y la inquietud, que generan los chatbots terapéuticos actuales tienen, en realidad, una genealogía bastante más larga de lo que los materiales de marketing de las plataformas suelen reconocer. El programa ELIZA, desarrollado por Joseph Weizenbaum en el MIT a mediados de la década de 1960, ya ponía de manifiesto un fenómeno que su propio creador denunció con alarma visible: la tendencia de los usuarios a atribuir subjetividad, empatía e intención genuina a un sistema que, en rigor, se limitaba a refor-

mular las palabras del interlocutor mediante patrones de sustitución. Weizenbaum quedó perturbado, la palabra no es excesiva, por la facilidad con que sus colegas y su propia secretaria desarrollaban apego emocional hacia ELIZA, y escribió al respecto con una honestidad que todavía incomoda (Weizenbaum, 1966, como se citó en Turkle, 2011). Ese malestar del inventor ante su invención sigue siendo, casi seis décadas después, una advertencia epistemológica de primer orden.

La paradoja no ha desaparecido con el avance tecnológico; se ha profundizado. Los sistemas de lenguaje natural de última generación producen respuestas contextualmente apropiadas, emocionalmente calibradas y lingüísticamente fluidas que hacen perceptualmente imperceptible la frontera entre simulación y cuidado auténtico para muchos usuarios, particularmente para los más jóvenes (Luger & Sellen, 2016). Turkle (2015) ha argumentado con rigor que la comodidad que ofrecen estos sistemas puede, paradójicamente, empobrecer la tolerancia a la ambigüedad y a la imperfección que el contacto humano inevitablemente implica. Si el interlocutor siempre está disponible, nunca juzga y siempre responde de manera apropiada, el umbral para la frustración relacional humana puede bajar de forma sostenida. Esto no implica que los sistemas de IA sean dañinos por definición. Implica que el tipo de vínculo que propician merece examinarse con mucho más cuidado del que recibe habitualmente.

Desde la psicología clínica, la alianza terapéutica ha sido identificada como uno de los predictores más robustos de los resultados en psicoterapia, independientemente de la orientación teórica del terapeuta (Wampold, 2015; Norcross & Wampold, 2019). Ese constructo presupone algo que ningún sistema algorítmico puede proporcionar en sentido estricto: un ser humano que es genuinamente afectado por el sufrimiento del otro, que tiene historia propia, corporalidad y la capacidad de ser sorprendido. Ahora bien, esta afirmación no cierra el debate empírico. Algunos estudios controlados han documentado que usuarios con síntomas leves de ansiedad o depresión

se benefician, al menos a corto plazo, de intervenciones basadas en IA que integran principios de terapia cognitivo-conductual (Fitzpatrick et al., 2017; Ly et al., 2017). Lo que señalan Stade et al. (2024) en una revisión sistemática reciente es que esa evidencia es suficiente para justificar la investigación continuada, pero insuficiente para legitimar la implementación clínica amplia sin supervisión profesional. La distinción entre ambas posiciones no es semántica: tiene consecuencias prácticas muy concretas para cómo las instituciones deberían regular estos sistemas.

Ética de la IA en educación: un campo que se construye mientras se usa

Uno de los rasgos más llamativos del campo de la ética de la IA en educación es que se ha desarrollado, en buena medida, a la sombra de implementaciones que ya estaban ocurriendo. Las universidades no esperaron a que existiese un marco normativo consolidado antes de introducir herramientas de análisis de aprendizaje, sistemas de tutoría inteligente o plataformas de apoyo emocional. Esa inversión del orden habitual, primero implementar, luego reflexionar, ha dejado una deuda ética que Holmes et al. (2022), Williamson y Eynon (2020) y Selwyn (2019) han contribuido a visibilizar con rigor creciente.

Los principios clásicos de la bioética de Beauchamp y Childress (2019), beneficencia, no maleficencia, justicia y autonomía, ofrecen un punto de partida útil, aunque su transposición al contexto educativo requiere adaptaciones que no siempre se hacen explícitas. En el aula universitaria, la asimetría de poder entre institución y estudiante, la vulnerabilidad específica de la población adolescente y la naturaleza formativa de la experiencia universitaria configuran condiciones que la bioética clínica no estaba pensada para contemplar. Cuando una universidad implementa un chatbot de salud mental, no solo ofrece un servicio. Comunica también una visión sobre qué tipo de cuidado merece el estudiante, quién es responsable de proveerlo y qué datos pueden recogerse en el proceso. Esa dimensión simbólica e institucional del

gesto tecnológico raramente aparece en las evaluaciones de eficacia.

El principio de autonomía merece atención particular porque, en el contexto de la IA, adquiere una dimensión que la bioética clásica no anticipaba: la amenaza de lo que podría llamarse heteronomía algorítmica. Si un estudiante aprende a gestionar su angustia a través de una interfaz, delegando progresivamente en el sistema los recursos de regulación que debería ir construyendo como propios, podría estar experimentando lo que Carr (2010) denominó efecto de desintermediación cognitiva: la tecnología no solo realiza una función, sino que atrofia la capacidad humana que esa función demandaba. Llevado al dominio psicológico, ese efecto configura lo que este artículo propone denominar dependencia algorítmica afectiva: un patrón de regulación emocional que requiere la mediación permanente de sistemas externos para funcionar. Es una hipótesis que necesita investigación longitudinal rigurosa y que la literatura actual, todavía demasiado orientada al corto plazo, no está en condiciones de confirmar ni de descartar.

Cabe señalar, además, que la brecha regulatoria entre el norte y el sur global es en este campo especialmente marcada. Mientras algunas universidades anglosajonas comienzan a desarrollar marcos de inteligencia artificial responsable para el cuidado estudiantil (Stanford HAI, 2023; Oxford Internet Institute, 2023), la mayoría de las instituciones latinoamericanas carece de políticas específicas que regulen el uso de IA en salud mental. Esa ausencia no es neutral: es el espacio en que se consolidan prácticas que serán muy difíciles de revertir una vez normalizadas.

Adolescencia tardía en la universidad: plasticidad, vulnerabilidad y mediación tecnológica

La transición a la universidad no es solo un cambio institucional. Para la mayoría de los jóvenes, es uno de los períodos de mayor plasticidad y mayor vulnerabilidad psicológica simultáneas. Arnett (2000) acuñó el término adultez emergente para describir preci-

samente ese intervalo entre la adolescencia y la adultez plena, caracterizado por la exploración de identidades, la inestabilidad afectiva y la búsqueda de un sentido de dirección vital que todavía no se ha consolidado. Es en ese estado de apertura radical donde muchos estudiantes universitarios comienzan a usar plataformas de IA para gestionar el malestar. La pregunta que esto plantea, y que la investigación no ha resuelto, es si esa gestión algorítmica en un momento de alta plasticidad configura hábitos de regulación que perduran.

La evidencia sobre el impacto de la mediación tecnológica en el bienestar adolescente, conviene reconocerlo, considerablemente más incierta de lo que el debate público suele sugerir. Twenge (2017) y Choukas-Bradley et al. (2022) han documentado asociaciones entre el uso intensivo de plataformas digitales y el deterioro del bienestar en adolescentes; Odgers y Jensen (2020), sin embargo, han cuestionado con argumentos metodológicos sólidos la robustez de esas correlaciones y, sobre todo, la asunción de causalidad que con frecuencia se desliza en su interpretación. Lo que sí parece razonablemente establecido es que el tipo de uso importa mucho más que el tiempo de pantalla: el uso pasivo y comparativo produce efectos distintos al uso activo e instrumental, y la presencia o ausencia de reflexividad sobre el propio uso modula de manera significativa los resultados (Odgers & Jensen, 2020).

Esta distinción entre tipos de uso tiene implicaciones directas para el análisis de la IA emocional en contextos universitarios. No es lo mismo utilizar una aplicación de biofeedback para aprender a reconocer señales de activación fisiológica que recurrir compulsivamente a un chatbot como sustituto de la socialización en momentos de crisis. Tampoco equivalen esas dos formas de uso a la derivación institucional automatizada hacia plataformas de IA sin acompañamiento humano, que es, quizás, la forma más problemática que puede adoptar la integración de estos sistemas. La literatura sobre dependencia tecnológica (Kuss & Griffiths, 2017; Billieux et al., 2015) proporciona categorías analíticas

pertinentes para este campo, pero aún no han sido trasladadas sistemáticamente al estudio específico de la IA emocional universitaria.

De este conjunto de tensiones teóricas, conceptuales, empíricas y éticas, surge la necesidad de un diseño metodológico que no pretenda resolverlas prematuramente, sino examinarlas con la precisión que merecen.

Metodología

Este trabajo adopta el diseño de revisión documental analítica en el marco de la investigación teórica crítica (Fink, 2019; Grant & Booth, 2009). La elección no es arbitraria. Cuando el campo de implementación de una tecnología avanza más rápido que su investigación empírica, y este es claramente el caso de la IA emocional en contextos universitarios, acumular más datos no es lo que más se necesita. Lo que falta es examinar con mayor precisión conceptual los datos que ya existen, identificar las tensiones que la literatura no ha resuelto y proponer marcos interpretativos que orienten tanto la investigación futura como las decisiones de política. Es eso, exactamente, lo que una revisión crítica permite hacer y que un estudio empírico primario no haría con la misma profundidad.

El corpus se construyó mediante búsqueda sistemática en Scopus, PsycINFO, ERIC y Web of Science, con búsqueda complementaria en Google Scholar para incorporar literatura gris relevante. Los términos de búsqueda, en español e inglés, combinaron los descriptores: artificial intelligence AND mental health AND higher education; AI AND student well-being AND ethics; chatbot AND psychotherapy AND university students; algorithmic governance AND education; digital mental health AND adolescents. Se aplicaron filtros de idioma (español e inglés), período de publicación (2019-2025) y tipo de documento: artículos de investigación empírica y teórica, revisiones sistemáticas, metaanálisis e informes institucionales de organismos acreditados.

Conviene señalar una decisión de delimitación que podría parecer restrictiva: se excluyeron deliberadamente los estudios de

caso institucionales sin validación empírica externa, pese a que varios de ellos aportaban descripciones detalladas de implementaciones en contextos latinoamericanos. La razón es metodológica, no ideológica: incluirlos habría introducido una asimetría en el tipo de evidencia que no se podía controlar en el análisis comparativo. El corpus resultante comprende 37 fuentes primarias del período indicado, a las que se suman referencias clásicas de los marcos teóricos de base cuya inclusión se justifica por su relevancia fundacional. La búsqueda se detuvo al alcanzar saturación temática en las tres tensiones identificadas, criterio congruente con el diseño de revisión integradora adoptado (Grant & Booth, 2009).

El análisis se organizó en dos fases, aunque la frontera entre ellas fue, en la práctica, más porosa de lo que el esquema sugiere. En la primera, de lectura analítica, cada documento fue codificado según cuatro categorías: tipo de evidencia aportada (empírica, teórica o normativa); dimensión de análisis privilegiada (eficacia, ética, gobernanza o impacto psicológico); posición evaluativa respecto a la IA en el bienestar estudiantil (favorable, crítica o ambivalente); y población y contexto geográfico de referencia. Esta codificación no pretendió reducir la riqueza de los textos a categorías discretas, sino construir un primer mapa de posiciones que hiciera posible, en la segunda fase, identificar convergencias y, sobre todo, las tensiones no resueltas.

En la segunda fase, de síntesis crítica, los hallazgos anteriores fueron reorganizados en torno a tres tensiones estructurales. La selección de esas tensiones no fue mecánica ni surgió de un procedimiento algorítmico de categorización. Emergió de la lectura reiterada del corpus y, en particular, del contraste entre autores que, partiendo de datos similares, llegan a conclusiones sensiblemente distintas. Ese contraste, más que la acumulación de coincidencias es donde reside el mayor interés analítico del campo.

La perspectiva epistemológica adoptada es explícitamente crítica: no asume la neutralidad de la tecnología ni la universalidad de sus efectos, sino que interroga las condiciones

sociales, institucionales e históricas que determinan quién se beneficia de la IA emocional, quién queda excluido y quién asume los costos no medidos. Esta posición conecta con la tradición de la pedagogía crítica (Giroux, 2011) y con la sociología crítica de la tecnología educativa (Selwyn, 2019; Facer, 2011), sin que ello implique descartar la evidencia empírica disponible; al contrario, la interroga con mayor exigencia.

Desarrollo analítico

Las tres tensiones que se analizan a continuación no se presentan en orden de importancia creciente ni decreciente. Se ordenan, más bien, según su grado de visibilidad en la literatura: de la más debatida a la más silenciada. Ese orden tiene, en sí mismo, algo que decir sobre los sesgos del campo.

Primera tensión: eficacia parcial versus ilusión de cuidado

La tensión más citada en la literatura sobre IA y bienestar universitario es también, posiblemente, la peor planteada. El debate sobre si los chatbots terapéuticos son eficaces suele presentarse como una disputa empírica entre quienes tienen datos favorables y quienes los cuestionan. Pero el problema de fondo es conceptual: eficaces ¿para qué, para quién y en qué condiciones?

La evidencia favorable proviene principalmente de estudios sobre aplicaciones como Woebot, Wysa o Replika. Fitzpatrick et al. (2017), Inkster et al. (2018) y Linardon et al. (2021) reportan reducciones estadísticamente significativas en síntomas de ansiedad y depresión en usuarios con niveles subclínicos de malestar. Esos datos existen y no conviene descartarlos. Una lectura crítica de esos estudios, sin embargo, revela un perfil de limitaciones bastante consistente: muestras de jóvenes adultos motivados que se autoseleccionan para participar en investigaciones sobre tecnología; períodos de seguimiento que raramente superan las dos o tres semanas; medición de resultados centrada casi exclusivamente en escalas de autoinforme de

síntomas afectivos. En suma, estos estudios documentan un efecto sobre el componente hedónico-negativo del bienestar, reducción de malestar inmediato, y no se ocupan, porque no están diseñados para ocuparse, de dimensiones como la calidad del vínculo social del estudiante, su sentido de propósito, su capacidad de tolerar la incertidumbre o su confianza en recursos propios de regulación. Esa omisión no invalida los hallazgos; los sitúa en su lugar.

Más inquietante que la limitación metodológica es la posibilidad de que la eficacia a corto plazo en síntomas leves coexista con efectos adversos que ninguno de esos estudios está diseñado para detectar. Turkle (2015) ha señalado con agudeza que la comodidad que ofrecen los sistemas conversacionales de IA, su disponibilidad permanente, su ausencia de juicio, su respuesta siempre emocionalmente calibrada, puede generar un desplazamiento de la tolerancia a la frustración y la ambigüedad que el contacto humano inevitablemente implica. Si los estudiantes aprenden a preferir la interacción regulada algorítmicamente sobre la impredecible interacción con otro ser humano, pueden estar cediendo, gradualmente y sin advertirlo, las capacidades relacionales que el malestar interpersonal exige y, al mismo tiempo, cultiva.

A esta dimensión individual se añade una institucional que resulta aún más preocupante. Cuando las universidades adoptan sistemas de IA como respuesta a la demanda de servicios de salud mental, pueden estar legitimando la sustitución de inversión en recursos humanos clínicos por soluciones tecnológicas de menor costo unitario. La American College Health Association (ACHA, 2023) ha expresado una preocupación explícita al respecto: que la IA se convierta no en una ampliación del cuidado sino en una herramienta de desresponsabilización institucional. Ese riesgo no requiere ninguna mala intención por parte de las autoridades universitarias; basta con que la presión presupuestaria y la narrativa de eficiencia tecnológica coincidan en el momento equivocado.

Segunda tensión: democratización del acceso versus mercantilización del malestar

La narrativa de la democratización del acceso es, en el campo de la IA y la salud mental universitaria, la más persuasiva. Y, por esa misma razón, la que merece mayor escrutinio. El argumento tiene una lógica aparentemente irrefutable: si los servicios de psicología institucional son insuficientes, como efectivamente lo son en la mayoría de las universidades latinoamericanas, y si una plataforma digital puede constituir el primer punto de contacto para estudiantes que de otro modo no buscarían ayuda, entonces la IA amplía el acceso. Baumel et al. (2019) y Torous et al. (2021) aportan evidencia que sostiene, con matices, este argumento.

Hasta aquí, la lógica es razonable. El problema empieza cuando se pregunta quién ofrece esas plataformas y en qué condiciones. La mayoría de los chatbots de salud mental más difundidos son productos comerciales desarrollados por empresas privadas cuyo modelo de negocio descansa, en un grado mayor o menor, en la recolección y el procesamiento de datos de usuarios. La experiencia emocional de los estudiantes, sus temores, sus duelos, sus crisis de identidad, sus conflictos relacionales, se convierte en este esquema en materia prima de un proceso de valorización que opera bajo lógicas radicalmente distintas a las del cuidado psicológico. Zuboff (2019) ha descrito con precisión el modelo general de la economía de la vigilancia; su transposición al dominio de la salud mental estudiantil es, en varios aspectos, más perturbadora que sus aplicaciones en el consumo o la publicidad, porque opera sobre la subjetividad en formación de personas que no siempre están en condiciones de evaluar los términos del intercambio al que consienten.

A la dimensión económica se suma una dimensión epistémica que la perspectiva de Fricker (2007) ayuda a articular. La gran mayoría de los chatbots disponibles fueron desarrollados en inglés, por equipos del norte global, y validados en muestras de universitarios de contextos anglosajones. Su transferencia a contextos culturalmente heterogéneos, con

concepciones distintas del sufrimiento psicológico, formas distintas de expresión emocional y relaciones distintas con la autoridad médica, implica un problema de validez cultural que raramente aparece en los materiales de implementación (Montero-Marín et al., 2022; Muñoz et al., 2022). Cuando un sistema de IA entrenado para reconocer somatizaciones típicas de la depresión en jóvenes anglosajones interactúa con un estudiante universitario de primera generación en Latinoamérica, no hay ninguna garantía de que lo que el sistema reconoce como malestar se corresponda con lo que el estudiante efectivamente experimenta. Esa brecha epistémica es, en sí misma, una forma de injusticia.

Tercera tensión: innovación pedagógica versus erosión de la autonomía

La tercera tensión es, quizás, la filosóficamente más densa y la que la literatura ha abordado con menor sistematicidad. Su núcleo es una pregunta incómoda: ¿puede un instrumento diseñado para aliviar el malestar psicológico terminar obstaculizando, paradójicamente, el desarrollo de las capacidades que el ser humano necesita para enfrentar ese malestar por sí mismo?

Piénsese en un caso concreto antes de abordar el problema en abstracto. Un estudiante universitario de primer año, lejos de su familia, recurre noche tras noche a un chatbot terapéutico para gestionar la ansiedad que le provoca la adaptación. La herramienta funciona: reduce la activación, ofrece validación, sugiere técnicas de respiración. Doce meses después, ese estudiante maneja mejor sus síntomas inmediatos. Pero ¿ha aprendido a tolerar la incertidumbre? ¿Ha construido vínculos sociales que le sirvan de sostén? ¿Ha desarrollado la capacidad de buscar ayuda humana cuando la necesite? Esas preguntas no tienen respuesta en la literatura actual. Y eso es, precisamente, el problema.

Freire (2005) advertía que la dependencia cognitiva y emocional de estructuras externas de sentido, sean estos dogmas religiosos, pedagogías bancarias o, añadamos, sistemas algorítmicos de regulación emocional; cons-

tituye un obstáculo al proceso de humanización. Illich (1971) hablaba de contraproduktividad para describir el fenómeno por el que un instrumento diseñado para liberar termina generando una nueva forma de dependencia. Ambas perspectivas adquieren una relevancia renovada cuando se aplican al uso de IA en la gestión del malestar de estudiantes universitarios que están, en ese momento de su vida, construyendo los recursos psicológicos que los acompañarán el resto de su trayectoria.

Esta tensión se expresa de manera concreta en la distinción entre regulación emocional asistida y regulación emocional delegada. En el primer caso, la herramienta tecnológica actúa como andamiaje, en el sentido vigotskiano del término, que acompaña el proceso de aprendizaje de nuevas capacidades regulatorias, retirándose progresivamente a medida que el estudiante internaliza esas capacidades. En el segundo, la herramienta sustituye ese proceso: el estudiante aprende a funcionar con el apoyo del sistema, pero no desarrolla los recursos propios que harían posible funcionar sin él. La diferencia entre ambos escenarios no es teórica; tiene consecuencias medibles en la autonomía psicológica a mediano y largo plazo. El problema es que los estudios existentes, casi todos de corto plazo, no están diseñados para distinguir entre estos dos patrones de uso (Gross, 2015; Thompson, 2014).

La investigación sobre autorregulación emocional ofrece aquí un marco conceptual más preciso que el que habitualmente se emplea en los estudios sobre IA y bienestar. Gross (2015) distingue entre estrategias de regulación orientadas a modificar la situación, reappraisal, búsqueda activa de soluciones, contacto social, y estrategias orientadas a suprimir o atenuar la experiencia emocional sin modificar sus condiciones de origen. Los sistemas de IA conversacional, tal como están diseñados actualmente, favorecen predominantemente este segundo tipo: ofrecen contención, validación y técnicas de reducción de la activación fisiológica, pero no están equipados para ayudar al estudiante a transformar las condiciones que generan el

malestar. En poblaciones cuyo malestar tiene raíces estructurales, precariedad económica, discriminación, sobrecarga académica, esta limitación no es menor.

Dicho esto, sería injusto y metodológicamente indefendible adoptar una postura de rechazo generalizado. Existen usos de la IA que, según la evidencia disponible, son compatibles con el fortalecimiento de la autonomía estudiantil: herramientas de biofeedback que enseñan al estudiante a reconocer patrones de activación fisiológica y a desarrollar recursos propios de regulación; plataformas de registro emocional que refuerzan la metacognición sobre el propio estado interior; y sistemas de derivación que conectan oportunamente al estudiante con recursos humanos antes de que una situación de malestar escale. Esta distinción entre tipos de uso encuentra un precedente metodológico próximo en Odgers y Jensen (2020), quienes demostraron que el tipo de uso tecnológico, activo frente a pasivo, modula significativamente los resultados en autorregulación adolescente, categoría analítica directamente aplicable al diseño de estos indicadores. La diferencia entre estos usos y los que erosionan la autonomía es conceptualmente clara, aunque su operacionalización en políticas universitarias concretas exige un nivel de sofisticación ética que pocas instituciones han alcanzado.

Discusión

Implicaciones teóricas: el marco EEPA como herramienta analítica

El recorrido analítico de las secciones anteriores permite proponer una herramienta conceptual para la evaluación crítica de los sistemas de IA en el bienestar psicológico universitario: el marco EEPA (Eficacia, Equidad, Privacidad, Autonomía), sintetizado en la Tabla 1. Antes de describirlo, conviene precisar lo que no pretende ser: no es una lista de verificación, no es un protocolo de auditoría y no ofrece respuestas definitivas a los debates descritos. Su propósito es más modesto, y por eso más útil, reorganizar esos debates en torno a criterios explícitos y disputables que

permitan superar la dicotomía entre adopción acrítica y rechazo tecnoescéptico.

La dimensión de eficacia interroga los efectos documentados de la herramienta sobre indicadores validados de bienestar psicológico, con especificación rigurosa de la población, el contexto, las dimensiones del bienestar medidas y el período de seguimiento. La equidad pregunta qué grupos se benefician y cuáles quedan excluidos, y si la implementación reproduce o transforma desigualdades preexistentes según nivel socioeconómico, género, origen étnico o condición de primera generación universitaria. La privacidad examina los flujos de datos que genera el uso de la herramienta: quién los procesa, en qué condiciones contractuales y con qué posibilidades reales, no formales, de que el estudiante comprenda a qué ha consentido. Y la autonomía analiza si el uso sostenido de la herramienta contribuye al desarrollo de recursos psicológicos propios del estudiante o genera los patrones de dependencia descritos en el apartado anterior. Esta dimensión es también la que presenta mayor vacío evidencial: como señalan Gross (2015) y Thompson (2014), los diseños de investigación disponibles no fueron concebidos para distinguir entre regulación emocional asistida y regulación emocional delegada, lo que hace imposible responder esta pregunta con los datos actuales.

Desde el punto de vista teórico, este marco contribuye al campo en dos sentidos que vale la pena hacer explícitos. Primero, pone en diálogo tradiciones disciplinares que rara vez se encuentran en la literatura sobre IA educativa: la psicología clínica y del desarrollo, la ética aplicada, la sociología de la tecnología y la pedagogía crítica. Segundo, desplaza la pregunta habitual del campo, si la IA es buena o mala para el bienestar estudiantil, hacia una pregunta más exigente y más útil: ¿en qué condiciones, para qué poblaciones y con qué diseños institucionales puede la IA contribuir legítima y responsablemente al bienestar psicológico universitario?

Implicaciones prácticas: lo que las instituciones deberían hacer

Las implicaciones prácticas del análisis son, en varios casos, incómodas para las instituciones universitarias. La más directa es que la adopción de cualquier sistema de IA orientado a la salud mental estudiantil debería ir precedida de un proceso de evaluación ética explícita que examine, como mínimo, la evidencia de eficacia disponible para la población estudiantil específica; los datos que la plataforma recopila y las condiciones de su tratamiento; los protocolos de derivación a profesionales clínicos para situaciones de riesgo; y los mecanismos de consentimiento informado real, no solo formal, por parte de los usuarios.

La segunda implicación práctica es igualmente urgente: la IA en el bienestar universitario debería concebirse siempre como complemento y nunca como sustituto de la inversión en servicios de orientación psicológica institucional. La narrativa tecnosolucionista que presenta estas herramientas como una respuesta eficiente a la crisis de salud mental universitaria enmascara una omisión de responsabilidad que la literatura sobre gobernanza universitaria ha comenzado a denunciar con creciente claridad (Williamson et al., 2022; Selwyn & Hillman, 2022). Ningún chatbot puede reemplazar la presencia de psicólogos, orientadores y equipos de bienestar competentes. Las instituciones que lo presentan como alternativa no están siendo innovadoras; están siendo irresponsables.

Vacíos de investigación futura

El análisis permite identificar al menos cuatro líneas de investigación que deberían priorizarse. La más urgente es la de los efectos longitudinales: los estudios identificados en la revisión realizada no reportan seguimientos superiores a tres meses, y ese vacío hace imposible pronunciarse con rigor sobre el riesgo de dependencia algorítmica afectiva. Se necesitan estudios prospectivos con seguimiento de al menos un año académico completo, que incluyan

mediciones de autonomía psicológica y no solo de síntomas afectivos.

La segunda línea prioritaria es la investigación transcultural. Los sistemas más difundidos fueron diseñados para contextos anglosajones, y su transferencia a contextos latinoamericanos, africanos o asiáticos implica supuestos de universalidad que deberían ser cuestionados empíricamente antes de que se generalice su uso. La tercera línea es la investigación participativa con estudiantes: la gran mayoría de los estudios existentes trata a los estudiantes como sujetos de investigación sin incorporarlos como codiseñadores de las herramientas y las políticas que los afectan. Y la cuarta, quizá la más descuidada, es el análisis de las implicaciones jurídicas del uso de datos emocionales en contextos donde la regulación sobre protección de datos en salud es ambigua o directamente inexistente.

Conclusiones

Este artículo ha argumentado que la integración de sistemas de inteligencia artificial en el apoyo al bienestar psicológico de adolescentes universitarios es un fenómeno lo suficientemente complejo como para resistir evaluaciones simplistas en cualquier dirección. La evidencia de eficacia a corto plazo en síntomas leves es real pero limitada en su alcance. Los riesgos de mercantilización del malestar, de erosión de la autonomía y de desresponsabilización institucional son igualmente reales, y considerablemente menos documentados. Esa asimetría en el estado del conocimiento debería, por sí sola, producir cautela regulatoria.

Las preguntas que orientaron este trabajo reciben respuestas necesariamente provisionales, pero argumentativamente fundadas. La mediación algorítmica del bienestar psicológico puede ser legítima en dimensiones acotadas, regulación de síntomas leves, acceso a psicoeducación, derivación oportuna a recursos humanos, siempre que opere como complemento de la atención profesional y no erosione la agencia del sujeto. Los marcos éticos existentes son insuficientes para gobernar este campo con la precisión que requiere,

lo que hace urgente el desarrollo de políticas universitarias específicas. Y la adopción acrítica de estas herramientas puede reproducir y amplificar desigualdades ya existentes, en la medida en que los sistemas disponibles no fueron diseñados para las poblaciones que más las necesitan. En el caso específico de América Latina, esa amplificación opera sobre desigualdades preexistentes de acceso, de validez cultural de los instrumentos y de capacidad regulatoria institucional que hacen especialmente urgente el desarrollo de marcos propios; no adaptaciones de marcos ajenos.

El marco EEPA propuesto en estas páginas, Eficacia, Equidad, Privacidad, Autonomía, no pretende ser una solución sino una herramienta de trabajo: un dispositivo analítico que permite hacerle a cualquier sistema de IA

las preguntas correctas antes de implementarlo. Su utilidad dependerá, en último término, de si las instituciones están dispuestas a hacerse esas preguntas con honestidad, incluso cuando las respuestas sean incómodas.

La educación superior tiene, en este campo, una responsabilidad que trasciende la gestión de servicios. Tiene la responsabilidad de ser el espacio donde se piensa críticamente sobre las tecnologías antes de que la inercia del mercado las normalice. La comunidad académica que estudia la IA en educación puede contribuir a eso, pero solo si está dispuesta a ir más lento que el ritmo de implementación, a formular preguntas que no tienen respuesta inmediata y a sostener la incomodidad de la incertidumbre. Esa disposición es, en cierto modo, el mismo recurso que queremos que los estudiantes desarrollen.

Marco EEPA para la evaluación crítica de sistemas de IA en bienestar estudiantil universitario

Dimensión	Pregunta orientadora	Indicadores sugeridos	Riesgo si se omite
Eficacia	¿Qué efectos documentados tiene la herramienta sobre el bienestar, en qué población y durante cuánto tiempo?	Tamaño del efecto, período de seguimiento, dimensiones del bienestar medidas	Adopción de herramientas con evidencia insuficiente o no transferible al contexto local
Equidad	¿Quién se beneficia y quién queda excluido? ¿La implementación reproduce desigualdades?	Acceso diferencial por nivel socioeconómico, género, etnia, condición de primera generación	Ampliación de brechas existentes disfrazada de democratización
Privacidad	¿Qué datos se recopilan, quién los procesa y bajo qué condiciones? ¿El consentimiento es real?	Políticas de datos, transferencia a terceros, comprensión efectiva del consentimiento	Mercantilización de los datos emocionales de poblaciones vulnerables
Autonomía	¿El uso sostenido fortalece o sustituye los recursos psicológicos propios del estudiante?	Mediciones de autorregulación emocional, dependencia del sistema, uso pasivo vs. activo	Dependencia algorítmica afectiva que erosiona la autonomía psicológica a largo plazo

Nota. Marco propuesto por el autor a partir del análisis del corpus. EEPA = Eficacia, Equidad, Privacidad, Autonomía.

Declaraciones

Financiamiento. Este trabajo no recibió financiamiento de ninguna fuente pública, comercial o sin fines de lucro.

Integridad de la investigación. Los autores declaran que este trabajo se ha llevado a cabo de conformidad con los principios establecidos en la Declaración de Singapur sobre Integridad en la Investigación (2010).

Contribución de autores (CRediT). Emgelberth Vargas: conceptualización, investigación formal, análisis formal, redacción del borrador original, revisión y edición. Alberto Luis Ramírez: investigación formal, revisión y edición.

Conflicto de intereses. El autor no declara ningún conflicto de intereses en relación con este trabajo.

Referencias

- American College Health Association. (2023). *ACHA guidelines on artificial intelligence in student mental health services*. https://www.acha.org/ACHA/Resources/Topics/Mental_Health.aspx
- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55(5), 469–480. <https://doi.org/10.1037/0003-066X.55.5.469>
- Barraza-Macías, A. (2022). Ansiedad académica en universitarios latinoamericanos: Una revisión sistemática. *Revista de Psicología Educativa*, 10(2), 45–63. <https://doi.org/10.20511/pyr2022.v10n2.1654>
- Baumel, A., Muench, F., Edan, S., & Kane, J. M. (2019). Objective user engagement with mental health apps: Systematic search and panel-based usage analysis. *Journal of Medical Internet Research*, 21(9), e14567. <https://doi.org/10.2196/14567>
- Beauchamp, T. L., & Childress, J. F. (2019). *Principles of biomedical ethics* (8th ed.). Oxford University Press.
- Bermúdez-Díaz, L., y Villanueva-Hernández, C. (2021). Estrés académico y salud mental en estudiantes universitarios de América Latina. *Salud Mental*, 44(3), 123–132. <https://doi.org/10.17711/SM.0185-3325.2021.016>
- Bickmore, T. W., Trinh, H., Asadi, R., & Olafsson, S. (2018). Safety first: Conversational agents for health communication. In *Proceedings of the 2018 ACM International Conference on Intelligent Virtual Agents* (pp. 2–9). ACM. <https://doi.org/10.1145/3267851.3267881>
- Billieux, J., Maurage, P., Lopez-Fernandez, O., Kuss, D. J., & Griffiths, M. D. (2015). Can disordered mobile phone use be considered a behavioral addiction? An update on current evidence and a comprehensive model for future research. *Current Addiction Reports*, 2(2), 156–162. <https://doi.org/10.1007/s40429-015-0054-y>
- Carr, N. (2010). *The shallows: What the Internet is doing to our brains*. W. W. Norton & Company.
- Choukas-Bradley, S., Maheux, A. J., Goldberg, M. H., Rouleau, C. R., & Roberts, S. R. (2022). Picture perfect during a pandemic: A study of adolescent social media use and well-being. *Frontiers in Psychology*, 13, 856763. <https://doi.org/10.3389/fpsyg.2022.856763>
- Cushman, P. (2011). Toward a sociohistorical understanding of constructivism. *Journal of Constructivist Psychology*, 24(1), 1–16. <https://doi.org/10.1080/10720537.2011.535554>
- Facer, K. (2011). *Learning futures: Education, technology and social change*. Routledge
- Fernández-Berrocal, P., Extremera, N., & Ruiz-Aranda, D. (2021). Emotional intelligence and mental health in university students: A systematic review. *Psicothema*, 33(2), 161–170. <https://doi.org/10.7334/psicothema2020.356>
- Fink, A. (2019). *Conducting research literature reviews: From the Internet to paper* (5th ed.). SAGE.
- Fitzpatrick, K. K., Darcy, A., & 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>
- Freire, P. (2005). *Pedagogía del oprimido* (3.ª ed.). Siglo XXI Editores.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Giroux, H. A. (2011). *On critical pedagogy*. Continuum.
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>

- Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, 26(1), 1–26. <https://doi.org/10.1080/1047840X.2014.940781>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32, 504–526. <https://doi.org/10.1007/s40593-021-00239-1>
- Hooks, B. (1994). *Teaching to transgress: Education as the practice of freedom*. Routledge.
- Illich, I. (1971). *Deschooling society*. Harper & Row.
- Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation mixed-methods study. *JMIR mHealth and uHealth*, 6(11), e12106. <https://doi.org/10.2196/12106>
- Kuss, D. J., & Griffiths, M. D. (2017). Social networking sites and addiction: Ten lessons learned. *International Journal of Environmental Research and Public Health*, 14(3), 311. <https://doi.org/10.3390/ijerph14030311>
- Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller-Tyszkiewicz, M. (2021). The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. *World Psychiatry*, 18(3), 325–336. <https://doi.org/10.1002/wps.20673>
- Luger, E., & Sellen, A. (2016). “Like having a really bad PA”: The gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 5286–5297). ACM. <https://doi.org/10.1145/2858036.2858288>
- Ly, K. H., Ly, A. M., & Andersson, G. (2017). A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interventions*, 10, 39–46. <https://doi.org/10.1016/j.invent.2017.10.002>
- Matos, L., Reyes, C., y Castellanos, M. (2022). Primera generación universitaria y bienestar psicológico en contextos latinoamericanos. *Revista Latinoamericana de Psicología*, 54, 112–121. <https://doi.org/10.14349/rlp.2022.v54.13>
- Montero-Marín, J., Allwood, M., Ball, S., Crane, C., Austin, T., & Kuyken, W. (2022). School-based mindfulness training in early adolescence: What works, for whom and how in the MYRIAD trial? *Evidence-Based Mental Health*, 25(3), 117–124. <https://doi.org/10.1136/ebmental-2022-300439>
- Muñoz, R. F., Chavira, D. A., Himle, J. A., & Ho, J. (2022). Technology-based mental health interventions for Latinos: Current status and future directions. *Psychiatric Services*, 73(7), 799–806. <https://doi.org/10.1176/appi.ps.202100096>
- Norcross, J. C., & Wampold, B. E. (2019). Evidence-based psychotherapy responsiveness: The third task force report. *Journal of Clinical Psychology*, 75(11), 1899–1906. <https://doi.org/10.1002/jclp.22893>
- Odgers, C. L., & Jensen, M. R. (2020). Annual research review: Adolescent mental health in the digital age: Facts, fears, and future directions. *Journal of Child Psychology and Psychiatry*, 61(3), 336–348. <https://doi.org/10.1111/jcpp.13190>
- Oxford Internet Institute. (2023). *Responsible AI for student well-being: Principles and governance*. University of Oxford. <https://www.oii.ox.ac.uk/research/responsible-ai-governance/>
- Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room: The obligation to act. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 46–55). ACM. <https://doi.org/10.1145/3027385.3027406>
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069–1081. <https://doi.org/10.1037/0022-3514.57.6.1069>

- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.
- Selwyn, N., & Hillman, T. (2022). Algorithms and the governance of educational data. *Learning, Media and Technology*, 47(2), 159–163. <https://doi.org/10.1080/17439884.2022.2037792>
- Stanford HAI. (2023). *AI and student well-being: Principles for responsible integration in higher education*. Stanford Human-Centered Artificial Intelligence. <https://hai.stanford.edu/news/ai-student-wellbeing>
- Stade, E. C., Stirman, S. W., Ungar, L. H., Boland, C. L., Schwartz, H. A., Yaden, D. B., & Eichstaedt, J. C. (2024). Large language models could change the future of behavioral health research and practice. *NPJ Mental Health Research*, 3, 12. <https://doi.org/10.1038/s44184-024-00056-z>
- Thompson, R. A. (2014). Stress and child development. *The Future of Children*, 24(1), 41–59. <https://doi.org/10.1353/foc.2014.0004>
- Torous, J., Bucci, S., Bell, I. H., Kessing, L. V., Faurholt-Jepsen, M., Whelan, P., & Firth, J. (2021). The growing field of digital psychiatry: Current evidence and the future of apps, social media, sensing, and artificial intelligence. *World Psychiatry*, 20(3), 318–335. <https://doi.org/10.1002/wps.20883>
- Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.
- Turkle, S. (2015). *Reclaiming conversation: The power of talk in a digital age*. Penguin Press.
- Twenge, J. M. (2017). *iGen: Why today's super-connected kids are growing up less rebellious, more tolerant, less happy, and completely unprepared for adulthood*. Atria Books.
- Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, 14(3), 270–277. <https://doi.org/10.1002/wps.20238>
- 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>
- Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223–235. <https://doi.org/10.1080/17439884.2020.1798995>
- Williamson, B., Gulson, K. N., Perrotta, C., & Witzemberger, K. (2022). AI and datafication of schooling. *Learning, Media and Technology*, 47(2), 157–158. <https://doi.org/10.1080/17439884.2022.2062196>
- Woebot Health. (2021). Real-world evidence: *Woebot outcomes and impact*. <https://woebothealth.com/real-world-evidence>
- Organización Mundial de la Salud. (2022). *Salud mental de los adolescentes*. <https://www.who.int/es/news-room/fact-sheets/detail/adolescent-mental-health>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

El autor no declara ningún conflicto de intereses en relación con este trabajo.

La correspondencia relativa a este artículo debe dirigirse a Emgelberth Vargas, Miami, Estados Unidos. Correo electrónico: emgelberth.vargas@cloud.carver.university



Emgelberth Vargas

Investigador y académico con más de veinte años de experiencia en educación superior, combinando liderazgo institucional, innovación curricular y producción científica en América Latina. Su trabajo se ha enfocado en tecnología educativa, bienestar estudiantil y transformación universitaria, con especial atención a la formación docente en entornos virtuales y la educación postpandemia. Recientemente ha sumado la ética de la inteligencia artificial a sus líneas de investigación, explorando el impacto de los algoritmos en la autonomía del estudiante y la gobernanza universitaria. Su formación en psicología le permite abordar estas cuestiones desde las dimensiones cognitivas y de salud mental. Está vinculado a Carver University y a la UPEL, y participa en redes internacionales de investigación interdisciplinaria.



Alberto Luis Ramírez

Ingeniero de sistemas con más de treinta años de experiencia, vinculado a la Escuela de Ingeniería de Sistemas de Carver University. Su trabajo se centra en el diseño de arquitecturas cognitivas que integran modelos de lenguaje, validación semántica y pipelines automatizados, combinando ingeniería de software, ciencia de datos y arquitectura de sistemas en soluciones escalables orientadas a la automatización avanzada. Ha desarrollado frameworks modulares para la orquestación de procesos de razonamiento asistido por IA, con experiencia en despliegues locales de modelos de alto rendimiento e integración de APIs cognitivas en entornos híbridos. En el ámbito académico, analiza las implicaciones éticas y sociales de la inteligencia artificial en contextos educativos y organizacionales. Actualmente orienta su investigación hacia sistemas autónomos y adaptativos, explorando la convergencia entre inteligencia artificial, optimización de procesos y computación avanzada.