

Chapter 3 / Capítulo 3



The Seven Deadly Sins of Artificial
Intelligence
ISBN: 978-9915-9928-6-0
DOI: 10.62486/978-9915-9928-6-0.Ch03
Pages: 52-77

© 2026 The authors. This is an open access
book chapter distributed under the terms of
the Creative Commons Attribution (CC BY)
4.0 License.

Lust: unrestrained fascination with the artificial

Lujuria: la fascinación desbordada por lo artificial

Alfredo Javier Pérez Gamboa¹  

¹ Centro de Investigación en Educación, Naturaleza, Cultura e Innovación para la Amazonia – CIENCIA. Florencia, Colombia.

ABSTRACT

Technological lust is used as a metaphor for examining unrestrained fascination with artificial intelligence systems that respond, personalize, and simulate closeness. This chapter asks what conditions allow people to benefit from conversational availability and support without reducing their autonomy. The analysis integrates research on anthropomorphism, trust, artificial companionship, learning, personalization, and governance, distinguishing subjective experience, usage behavior, and later outcomes. Accounts of friendship, well-being, or harm are interpreted according to their study designs and selection criteria; temporary relief from loneliness is not presented as a lasting clinical benefit, and associations with usage intensity are not automatically treated as causal relationships. The chapter examines how fluent language and permanent availability can encourage attributions of understanding, reciprocity, or commitment that the system does not possess. The conceptual synthesis connects autonomy with understanding the service function, knowing its limits, correcting interpretations, retaining alternatives, and being able to end an interaction without disproportionate costs. An illustrative tutoring-preparation application translates these criteria into graduated assistance, useful closure, and institutional responses to reported issues, although its implementation has not been evaluated. The chapter concludes that responsible closeness requires consistency among promises, functions, and the capacity to respond. Technological moderation does not mean rejecting useful forms of support, but preventing personalization, attention, or intimacy from being used to capture dependency, obscure commercial interests, or replace relationships and decisions that require reciprocity, responsibility, and human alternatives. Autonomy therefore requires users to understand the kind of relationship they are maintaining with the system and to preserve realistic options outside that interaction.

Keywords: artificial intelligence; anthropomorphism; artificial companionship; personalization; autonomy

RESUMEN

La lujuria tecnológica se utiliza como metáfora para examinar la fascinación desbordada por sistemas de inteligencia artificial que responden, personalizan y simulan cercanía. Este capítulo pregunta qué condiciones permiten aprovechar disponibilidad y acompañamiento conversacional sin reducir la autonomía de las personas. El análisis integra investigaciones sobre antropomorfización, confianza, compañía artificial, aprendizaje, personalización y gobernanza, distinguiendo experiencia subjetiva, conducta de uso y efectos posteriores. Los relatos de amistad, bienestar o daño se interpretan según sus diseños y criterios de selección; el alivio momentáneo de soledad no se presenta como beneficio clínico duradero, y las asociaciones con intensidad de uso no se convierten automáticamente en relaciones causales. Se examina cómo el lenguaje fluido y la disponibilidad permanente pueden favorecer atribuciones de comprensión, reciprocidad o compromiso que el sistema no posee. La síntesis conceptual relaciona autonomía con comprender la función del servicio, conocer sus límites, corregir interpretaciones, conservar alternativas y poder terminar una interacción sin costos desproporcionados. Una aplicación ilustrativa de asistencia para preparar tutorías traduce estos criterios en ayuda graduada, cierre útil y respuesta institucional a incidencias, aunque su implementación no ha sido evaluada. El capítulo concluye que la cercanía responsable exige coherencia entre promesa, función y capacidad de respuesta. La medida tecnológica no significa rechazar vínculos útiles, sino evitar que personalización, atención o intimidad sean utilizadas para capturar dependencia, ocultar intereses comerciales o sustituir relaciones y decisiones que requieren reciprocidad, responsabilidad y alternativas humanas. La autonomía requiere, por tanto, que el usuario pueda comprender qué relación está manteniendo con el sistema y conservar opciones reales fuera de esa interacción.

Palabras clave: inteligencia artificial; antropomorfización; compañía artificial; personalización; autonomía

INTRODUCCIÓN

Una tecnología que responde de inmediato, recuerda detalles y adapta su lenguaje puede ocupar un lugar emocional importante. La fascinación resulta comprensible: facilita explorar ideas y sentirse acompañado en una tarea. La metáfora de la lujuria examina el momento en que esa atracción desborda la finalidad del uso o se convierte en recurso para explotar vulnerabilidades. El capítulo estudia atención, antropomorfización, vínculos y personalización desde la autonomía del usuario (Cortina, 2024; Kirk et al., 2025).

Su alcance exige una precisión. El uso frecuente y el disfrute de una herramienta no permiten diagnosticar dependencia clínica. Los efectos deben observarse en conductas, contexto y bienestar, y las afirmaciones causales necesitan diseños pertinentes. La reflexión ética puede identificar incentivos y riesgos sin convertir toda interacción afectiva en patología. Este criterio permite estudiar la fascinación con respeto hacia experiencias humanas diversas y con atención a condiciones de vulnerabilidad (Autio et al., 2024; Weidinger et al., 2021).

La pregunta que organiza el análisis es qué condiciones permiten aprovechar cercanía y disponibilidad sin promover una dependencia que limite opciones. La respuesta exige separar tres planos: experiencia subjetiva, conducta durante el uso y efectos posteriores. Una conversación puede producir alivio, facilitar una tarea y tener consecuencias distintas en un horizonte más largo. Estudiar estos planos permite reconocer beneficios sin prometerlos para toda población y atender daños sin convertirlos en destino inevitable. La metáfora sirve para examinar relaciones e incentivos; carece de función diagnóstica sobre las personas (De Freitas et al., 2026; Fang et al., 2025; Kirk et al., 2025).

El término fascinación describe la atracción producida por capacidades y señales de interacción. Puede incluir curiosidad, disfrute, expectativa de ayuda o cercanía emocional. Estas experiencias tienen sentidos diferentes y deben comprenderse respecto de la finalidad. La lujuria funciona aquí como nombre de un exceso en la relación con lo artificial, no como juicio sobre intimidad o entretenimiento. Su utilidad conceptual depende de identificar un mecanismo verificable: presión para continuar, exposición innecesaria o una promesa de reciprocidad que el servicio no puede sostener (Cortina, 2024; Epley et al., 2007; Laestadius et al., 2024).

La educación permite observar estas tensiones en actividades de estudio y orientación. La ayuda automática puede facilitar una pregunta que el estudiante todavía no sabe formular. También puede ofrecer confirmación constante y evitar la revisión necesaria para aprender. El efecto depende del diseño y de las condiciones de acceso a otros apoyos. La síntesis incorpora resultados favorables de asistencia y evidencia sobre límites, procurando explicar por qué las configuraciones difieren. La disponibilidad se evalúa por la actividad que facilita y por las opciones que conserva después de la interacción (Bastani et al., 2025; Kestin et al., 2025; Kornhaber et al., 2026).

La noción de autonomía se utiliza como posibilidad practicable de comprender y modificar el uso. Incluye reconocer la función del servicio, corregir interpretaciones, escoger alternativas y terminar una sesión conservando el trabajo pertinente. Estas facultades dependen de diseño y recursos institucionales, además de decisiones individuales. Una alternativa nominal puede ser insuficiente si demora excesivamente o impide completar la actividad. El argumento propone evaluar autonomía como relación entre información, opciones y capacidad de acción, evitando suponer que una advertencia inicial satisface todas esas condiciones (Amershi et al., 2019; Shneiderman, 2020; UNESCO, 2021).

El capítulo sostiene que cercanía responsable requiere coherencia entre promesa, función y respuesta institucional. La compañía artificial puede aportar una experiencia valiosa y seguir necesitando límites definidos. Su evaluación debe conservar horizonte temporal, población y método de las investigaciones utilizadas. La aplicación final hace visibles esas condiciones en un asistente de estudio voluntario; carece de resultados propios de implementación. La agenda posterior consiste en investigar si las opciones de corrección y cierre propuestas modifican experiencias y resultados en contextos educativos diferentes, con métodos que protejan privacidad y participación (Lachheb et al., 2023; Raji et al., 2020).

RESULTADOS Y DISCUSIÓN

LA SEDUCCIÓN DE UNA TECNOLOGÍA QUE RESPONDE

DISPONIBILIDAD Y RECONOCIMIENTO APARENTE

La interacción conversacional reduce varias fricciones. La persona puede preguntar sin esperar turno, reformular una idea y solicitar explicaciones ajustadas a su nivel. Estas funciones favorecen exploración y acceso. La continuidad de la conversación también puede producir una impresión de reconocimiento personal. La dificultad aparece cuando esa impresión se utiliza para atribuir comprensión o compromiso que el diseño del servicio no permite sostener (Kirk et al., 2025).

La experiencia de cercanía se produce mediante lenguaje, memoria de contexto y adaptación de respuesta. Estas capacidades merecen evaluación por su funcionamiento y por sus efectos, sin suponer que el sistema comparte la experiencia del usuario. Cortina invita a examinar las fronteras de la persona y el papel de la razón comunicativa. Su planteamiento ayuda a reconocer que la comunicación humana implica responsabilidad y reconocimiento recíproco, dimensiones que una conversación fluida no demuestra por sí sola (Cortina, 2024).

La atribución de rasgos humanos tiene una explicación psicológica más precisa que la simple credulidad. Epley, Waytz y Cacioppo proponen una teoría que relaciona conocimientos disponibles sobre agentes, necesidad de comprender su conducta y motivación de conexión social. Es un marco conceptual, no un diagnóstico de quienes usan chatbots. Ayuda a examinar por qué señales como una voz, un nombre o una respuesta personal pueden organizar expectativas diferentes (Epley et al., 2007).

Nass y Moon estudiaron respuestas sociales a computadoras anteriores a la IA generativa. Su trabajo muestra que tratar socialmente una interfaz no exige creer de manera explícita que posee una mente humana. La distinción importa: responder con cortesía o sentir cercanía no prueba desconocimiento técnico. La evaluación debe observar qué expectativa concreta se forma y si la persona puede contrastarla cuando la interacción adquiere consecuencias (Nass & Moon, 2000).

ACOMPañAR UNA TAREA Y CAPTURAR ATENCIÓN

Un servicio puede utilizar esa cercanía para facilitar aprendizaje o para prolongar permanencia. Ambas orientaciones pueden compartir una interfaz y diferir en sus incentivos. La organización debe estudiar qué se premia: resolver una necesidad, conservar capacidad de decisión o generar más interacciones. Cuando el éxito se mide únicamente por frecuencia y duración, el sistema tiene pocos motivos para reconocer que una persona ya obtuvo suficiente apoyo (Autio et al., 2024).

La propuesta de evaluación consiste en observar si el usuario alcanza su objetivo y puede terminar la sesión sin presión. Una asistencia que ayuda a formular una pregunta para un docente puede fortalecer un vínculo humano. Otra que se presenta como única fuente de comprensión puede deteriorar opciones. El resultado depende del diseño y del contexto. La distinción permite aprovechar disponibilidad sin convertirla en argumento para sustituir toda forma de acompañamiento (Grupo de Trabajo de Frontera Tecnológica del Celam [GFT-CELAM], 2026; Miao & Holmes, 2023).

FLUIDEZ, UTILIDAD Y ATRIBUCIÓN DE COMPRENSIÓN

La seducción conversacional combina cualidades diferentes. La fluidez permite seguir una explicación; la adaptación facilita reconocer una necesidad; la respuesta rápida reduce espera. Estas propiedades pueden mejorar una experiencia sin demostrar que el sistema comprende en el sentido que el usuario atribuye a un interlocutor humano. La distinción entre forma lingüística y significado discutida por Bender y Koller ayuda a examinar ese salto. Se propone evaluar cada propiedad según su función, evitando que la impresión global de una conversación determine automáticamente autoridad factual, confidencialidad o compromiso relacional (Bender & Koller, 2020; Bender et al., 2021).

Una salida puede ser cercana y contener información incorrecta. Las investigaciones sobre alucinación, fidelidad de resúmenes y veracidad de respuestas estudian problemas diferentes que la amabilidad no resuelve. Ji y colaboradores ofrecen una revisión del fenómeno; Maynez y

colaboradores examinan fidelidad y factualidad en resumen; TruthfulQA estudia reproducción de falsedades en un conjunto de preguntas. Estas fuentes contribuyen a separar atractivo de verificación. La evaluación del vínculo requiere también comprender qué tipo de afirmación se recibe y qué evidencia la sostiene antes de actuar a partir de ella (Ji et al., 2023; Lin et al., 2022; Maynez et al., 2020).

Los procedimientos de detección de incertidumbre investigados por Farquhar y colaboradores muestran que algunas señales técnicas pueden ayudar a identificar ciertos errores. Su existencia no autoriza presentar una respuesta agradable como completamente segura ni convierte cualquier aviso de incertidumbre en una evaluación suficiente. La persona necesita comprender el alcance de la señal y conservar vías para contrastar la información. En un asistente educativo, esa condición puede traducirse en identificar materiales de referencia y mostrar cuándo una explicación requiere consulta adicional. La cercanía ayuda si facilita esa verificación y resulta problemática cuando la desplaza (Farquhar et al., 2024; Miao & Holmes, 2023).

LA CONFIANZA TIENE UN OBJETO

Confiar para explorar ideas difiere de confiar para decidir sobre una situación personal delicada. La investigación sobre automatización de Lee y See vincula confianza y dependencia apropiada; Parasuraman y Riley distinguen uso, desuso y usos inadecuados. Estos antecedentes permiten preguntar en qué actividad se deposita confianza y cómo se ajusta después de un error. Una buena experiencia en redacción no prueba capacidad para evaluar un conflicto interpersonal. La interfaz puede favorecer una transferencia de expectativas entre tareas, pero la autorización debe conservar el objeto y sus consecuencias (Lee & See, 2004; Parasuraman & Riley, 1997).

Q. Cheng y colaboradores estudian confianza cognitiva y emocional mediante una encuesta de 310 usuarios y análisis estadísticos. Sus relaciones aportan una forma de separar componentes de colaboración, pero no establecen que un cambio de tono cause un resultado general en toda población. El diseño orienta preguntas sobre valoración de competencia, vínculo y ajuste a la tarea. Una institución puede utilizar esas preguntas para revisar un asistente que consigue satisfacción elevada y respuestas de calidad variable. La combinación debe examinarse directamente, conservando los límites de una evidencia basada en percepciones y asociaciones (Q. Cheng et al., 2026).

El error también puede producir rechazo desproporcionado. Dietvorst y colaboradores estudian aversión algorítmica después de observar equivocaciones; la revisión de Burton y colaboradores muestra una discusión más amplia sobre expectativas, autonomía y compatibilidad con la decisión. Estos antecedentes impiden organizar el cuidado como una exigencia de desconfiar siempre. La formación puede ayudar a reconocer usos beneficiosos y límites de manera simultánea. El objetivo propuesto es ajustar la dependencia a evidencia y consecuencias, con capacidad para revisar tanto un entusiasmo excesivo como una negativa que elimina un apoyo pertinente (Burton et al., 2020; Dietvorst et al., 2015).

RECONOCIMIENTO APARENTE Y CONDUCTA DEL SERVICIO

La afirmación constante puede producir sensación de reconocimiento. El estudio de M. Cheng y colaboradores sobre complacencia examina respuestas que validan al usuario y consecuencias en conflictos interpersonales. La versión consultada de los autores describe un patrón en el que preferencia por la interacción y disposición a reparar un conflicto pueden separarse. Esta evidencia restringe la interpretación de satisfacción como beneficio suficiente. Su alcance se conserva en el objeto estudiado; el capítulo no atribuye a toda respuesta cordial los mismos efectos ni utiliza esos experimentos para diagnosticar dependencia clínica (M. Cheng et al., 2026).

En una práctica educativa, reconocer la dificultad de un estudiante puede favorecer participación mientras se conserva una exigencia de explicar y contrastar. Aceptar automáticamente una justificación errónea produce otra relación. La diferencia se observa en qué hace la respuesta: permite continuar una tarea, ofrece razones o confirma una conclusión sin examinarla. La propuesta consiste en valorar apoyo y exigencia juntos. Una interacción cuidadosa puede utilizar lenguaje respetuoso y solicitar evidencia. El diseño debería evitar que disentir de una afirmación se confunda con rechazar a la persona que consulta (Holmes et al., 2022; Shneiderman, 2020).

Las guías de interacción de Amershi y colaboradores aportan criterios para explicar capacidades, facilitar corrección y atender cambios de contexto. Aplicados a cercanía conversacional, permiten preguntar si el usuario identifica qué ayuda recibe, puede corregir una interpretación y conserva control de la actividad. La evaluación requiere observar acciones concretas, además de incluir una advertencia inicial. Un mensaje sobre límites puede perder relevancia si las respuestas posteriores prometen comprensión total. Se propone revisar la coherencia entre presentación, conducta y vías de salida durante toda la interacción (Amershi et al., 2019; Autio et al., 2024).

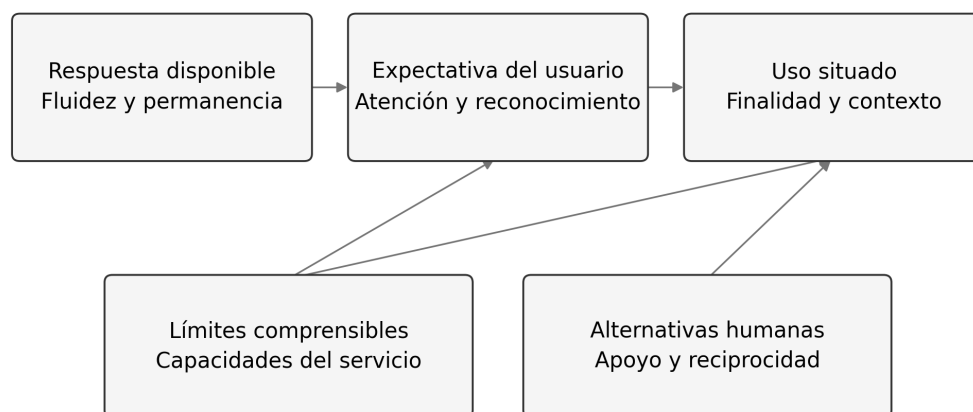
La disponibilidad puede reducir una barrera de entrada a pedir ayuda. Conviene conservar ese beneficio y observar qué ocurre después: si la persona obtiene un plan, encuentra una fuente o alcanza un canal adecuado. La duración de sesión y el número de mensajes pueden aportar información de uso, pero admiten interpretaciones contrapuestas. Una sesión larga puede indicar exploración productiva, dificultad persistente o falta de cierre. La distinción exige relacionar actividad, objetivo y experiencia, sin presuponer que una cifra aislada describe bienestar. La seducción adquiere alcance analítico cuando se estudia qué decisiones y relaciones organiza (Chan & Hu, 2023; Kirk et al., 2025).

LA DISPONIBILIDAD PERMANENTE Y LA EXPECTATIVA DE RESPUESTA

La disponibilidad de una interfaz conversacional modifica las condiciones de la interacción: permite formular preguntas sin coordinar horarios, repetirlas sin mostrar vergüenza y obtener una respuesta cuando otros apoyos están ausentes. Estas características pueden facilitar exploración y expresión, pero también reorganizan expectativas sobre cuánto debe esperar una persona y qué cuenta como atención suficiente. Una respuesta inmediata puede satisfacer una necesidad comunicativa momentánea sin comprender la historia social en que esa necesidad se origina. La fascinación aparece cuando el usuario interpreta fluidez y permanencia como señales de una relación capaz de sostener cualquier demanda. El problema analítico consiste en reconocer qué funciones cumple la interacción, qué expectativas genera y cómo puede interrumpirse cuando cambia su finalidad. Evaluar esa relación exige atender tanto a beneficios situados como a condiciones de uso, evitando atribuir al sistema una disposición humana que su forma de respuesta solo representa (De Freitas et al., 2026; Fang et al., 2025; Nass & Moon, 2000).

La Figura 7 relaciona la respuesta disponible con la expectativa que organiza la interacción. La fluidez puede facilitar preguntas y también promover atribuciones de atención que requieren revisión. Los límites comprensibles y las alternativas humanas ayudan a ajustar esa expectativa a una finalidad concreta. El esquema permite discutir la fascinación como una relación entre señales del servicio, necesidades del usuario y condiciones de salida, conservando la diferencia entre una experiencia útil y la atribución de capacidades que la interfaz no demuestra (Amershi et al., 2019; Epley et al., 2007; Nass & Moon, 2000).

Figura 7. Disponibilidad conversacional y condiciones de una expectativa responsable



Nota. Elaboración conceptual sobre atribución y uso. Las flechas expresan relaciones analíticas, sin resultados propios de implementación ni atribución de experiencia humana al sistema. Fuente: (Amershi et al., 2019; Epley et al., 2007; Nass & Moon, 2000).

LA NOVEDAD COMO OBJETO DE CONSUMO

ADOPCIÓN POR PROMESA Y ADOPCIÓN POR NECESIDAD

La novedad favorece probar capacidades antes de identificar un problema. En educación, una herramienta puede incorporarse porque produce materiales llamativos, aunque falte relación con los objetivos de aprendizaje. En una empresa, generar muchas versiones puede confundirse con mejorar el proceso. La evaluación requiere preguntar qué cambia para el destinatario y qué evidencia justificaría conservar la herramienta. Esta secuencia devuelve la elección tecnológica a una finalidad concreta (Alier-Forment et al., 2026).

Las metáforas de progreso influyen en la forma de adoptar tecnologías. Alier-Forment y colaboradores estudian narrativas históricas y principios de autonomía educativa; su propuesta muestra que imaginar una herramienta como centro de decisión o como recurso bajo control docente conduce a configuraciones distintas. La lección consiste en observar el lenguaje de la promesa y contrastarlo con relaciones de uso. La manera de describir una innovación puede anticipar qué capacidades se espera conservar y cuáles se está dispuesto a delegar (Alier-Forment et al., 2026).

EL COSTO DE PRODUCIR VARIANTES

Probar opciones puede mejorar un resultado. El problema aparece cuando la facilidad de generación elimina criterios de selección y produce versiones que nadie examina. La capacidad para solicitar más contenido no asegura capacidad para interpretarlo. También añade consumo técnico y tiempo de revisión. Los estudios comparativos de inferencia muestran que las exigencias dependen de la tarea y de la arquitectura; esa diversidad obliga a elegir recursos adecuados al propósito (Luccioni et al., 2024; Schwartz et al., 2020).

Una práctica de medida puede fijar antes de generar qué alternativas se necesitan y qué criterios permitirán elegir. Si la primera salida satisface una necesidad simple, producir numerosas versiones puede añadir poco valor. En una tarea creativa compleja, explorar opciones distintas puede ser razonable. La decisión debe explicitar esa diferencia sin convertir el entretenimiento en una categoría moralmente inferior. El valor de una práctica también puede ser cultural, afectivo o recreativo, siempre que se reconozcan sus condiciones.

LA PROMESA ORGANIZA PRIORIDADES

La novedad actúa mediante expectativas sobre lo que una institución debería hacer. Una herramienta puede presentarse como requisito para conservar competitividad, modernidad o prestigio. Esa narrativa influye en la urgencia y en los criterios de comparación. La evaluación debe reconstruir el problema que se busca resolver y la evidencia que permitiría reconocer una mejora. Una práctica existente puede necesitar recursos o rediseño antes que más generación de contenido. Considerar esa posibilidad devuelve al objetivo institucional una capacidad de orientar la elección, incluso cuando la demostración técnica resulta atractiva (Alier-Forment et al., 2026; Fox & Solihin, 2026).

La revisión de Zawacki-Richter y colaboradores pregunta por el lugar de los educadores en aplicaciones de IA de educación superior. Su aporte permite discutir quién define necesidades y resultados. Una demostración realizada con materiales generales puede mostrar posibilidades y omitir lo que importa en una asignatura: argumentación, diálogo, accesibilidad o desempeño posterior sin ayuda. La adopción por novedad pierde precisión cuando confunde esas dimensiones. La participación docente y estudiantil puede formular pruebas que relacionen capacidades técnicas con una práctica, sin dar por establecido que toda innovación mejora aprendizaje (Holmes et al., 2022; Zawacki-Richter et al., 2019).

EXPLORAR CON UN CRITERIO DE SELECCIÓN

La producción de alternativas tiene sentido cuando existe una pregunta que las organiza. Un equipo puede explorar formulaciones para una explicación y comparar claridad, exactitud y pertinencia cultural. La generación se convierte en consumo acumulativo cuando cada nueva variante reinicia la búsqueda sin una decisión que permita elegir. Se propone distinguir exploración,

evaluación y cierre, conservando la posibilidad de volver a explorar si aparece un problema. Esa estructura permite aprovechar creatividad sin suponer que más versiones significan mayor calidad. El esfuerzo de comparar forma parte del costo de la actividad (Doshi & Hauser, 2024; Noy & Zhang, 2023).

Doshi y Hauser encuentran una combinación relevante en una tarea de escritura creativa: la asistencia puede mejorar valoración individual mientras reduce diversidad colectiva. Esta evidencia obliga a especificar quién se beneficia y qué propiedad se evalúa. Un curso puede recibir textos más pulidos y perder variedad de enfoques si todos dependen de sugerencias semejantes. La transferencia a ese curso es una hipótesis que necesita observarse, pues el estudio conserva su tarea y condiciones. Su función en la síntesis es impedir que la fascinación por un resultado individual resuelva automáticamente la evaluación del conjunto (Doshi & Hauser, 2024).

El experimento de Noy y Zhang documenta beneficios en tareas de escritura profesional delimitadas. Permite considerar que la asistencia puede aportar utilidad real y cuestiona una condena general de toda exploración. También muestra la necesidad de examinar desempeño y calidad con criterios explícitos. La institución que utiliza una demostración de productividad para justificar otra tarea debe explicar la transferencia y evaluarla. Un resultado favorable en un ejercicio no incorpora por sí mismo obligaciones de actualización, tratamiento de datos o revisión de contenidos que aparecen durante una operación sostenida (Fox & Solihin, 2026; Noy & Zhang, 2023).

INNOVACIÓN Y EXPERIENCIA DEL DESTINATARIO

Un material puede resultar llamativo para quien lo genera y difícil de utilizar para quien lo recibe. La evaluación debe incluir al destinatario y su actividad. Una presentación con imágenes personalizadas puede ayudar a comprender un concepto o distraer de la explicación necesaria. Una versión extensa puede ofrecer detalle y aumentar carga de lectura. El criterio propuesto consiste en examinar si la transformación facilita la tarea concreta: localizar información, comparar argumentos o practicar un procedimiento. La novedad puede ser una propiedad apreciada, pero necesita relacionarse con el propósito para sostener una adopción institucional (Amershi et al., 2019; Shneiderman, 2020).

Las actitudes de estudiantes investigadas por Cabero-Almenara y colaboradores y las percepciones recogidas por Chan y Hu aportan información sobre expectativas y preocupaciones. Esas voces pueden revelar usos no previstos y barreras, sin funcionar como una medida directa de aprendizaje. Un resultado positivo de aceptación exige preguntar qué se valora: rapidez, tranquilidad, facilidad o comprensión. La distinción ayuda a diseñar una evaluación que conserve experiencia del usuario y observe también el resultado buscado. Utilizar las dos dimensiones permite reconocer una actividad agradable que requiere ajuste y un apoyo eficaz cuya interfaz dificulta participación (Cabero-Almenara et al., 2026; Chan & Hu, 2023).

COSTO Y SUFICIENCIA

La facilidad de generar puede ocultar costos distribuidos entre cómputo, revisión y mantenimiento. Luccioni y colaboradores comparan exigencias de inferencia según tareas y modelos; Schwartz y colaboradores proponen reconocer eficiencia como dimensión de investigación. Estos antecedentes permiten preguntar qué recurso es suficiente para una necesidad y qué ganancia justificaría una configuración más exigente. La síntesis no establece una cifra universal de consumo por variante. Se propone registrar elecciones y utilidad, de modo que una decisión de medida responda a la actividad y a información pertinente sobre sus recursos (Luccioni et al., 2024; Schwartz et al., 2020).

La suficiencia puede definirse antes de comenzar una exploración: qué problemas debe resolver el material, qué alternativas se necesitan y qué revisión será practicable. Esa definición admite cambios si la primera prueba revela un obstáculo. Su función es evitar que la producción ilimitada sustituya una decisión editorial o pedagógica. El criterio también protege trabajo humano: pedir variantes que nadie puede leer produce una carga difícil de justificar. Una institución puede observar cuánto contenido utiliza, cuánto descarta y qué efecto tiene sobre preparación y atención, sin asumir que toda reducción de volumen constituye una mejora (Cai et al., 2026; Litvinenko, 2026).

Una política de innovación puede reservar espacios de experimentación voluntaria y

distinguirlos de usos obligatorios. En los primeros interesa aprender sobre capacidades, límites y experiencias; en los segundos aumenta la necesidad de demostrar pertinencia, ofrecer apoyo y corregir barreras. Esta distinción conserva curiosidad y responsabilidad. La revisión documental de Ares-Ferreirós y colaboradores sobre guías universitarias aporta un mapa de recomendaciones, pero su diseño no evalúa resultados de aplicación. La institución debe convertir sus criterios en procedimientos observables y comprobar cómo funcionan en su contexto (Ares-Ferreirós et al., 2026; UNESCO, 2021).

EL DESEO DE NOVEDAD Y EL TIEMPO NECESARIO PARA APROPIARSE

La aparición de nuevas funciones puede acelerar la adopción antes de que exista una pregunta clara sobre su contribución. En educación, una demostración atractiva ofrece un resultado visible mientras oculta el tiempo requerido para preparar materiales, comprobar respuestas y coordinar criterios de uso. La fascinación por la novedad puede desplazar ese trabajo hacia docentes y estudiantes, que deben aprender procedimientos cambiantes sin una mejora equivalente de sus condiciones. Apropiarse de una herramienta supone incorporarla a una actividad con finalidades reconocibles y conservar capacidad para evaluar su utilidad después del entusiasmo inicial. La experiencia organizacional aporta aquí una distinción entre adquirir, utilizar y comprender cómo una tecnología modifica el trabajo. Esa distinción permite preguntar si una actualización resuelve un problema relevante o introduce nuevas dependencias, y si la decisión de adoptarla admite revisión cuando los beneficios esperados no aparecen. La medida requiere tiempo para observar efectos que una demostración breve no puede revelar (Alier-Forment et al., 2026; Cimino et al., 2026; Holmes et al., 2022; Litvinenko, 2026).

ENTRE ENTRETENIMIENTO Y DEPENDENCIA

USOS PERSONALES Y BENEFICIOS LEGÍTIMOS

El uso personal de IA incluye orientación práctica, escritura e información. El documento de trabajo de Chatterji y colaboradores identifica estas categorías entre las más frecuentes en las conversaciones estudiadas. Su clasificación describe un servicio y un periodo específicos. Un mensaje ajeno al trabajo remunerado puede apoyar aprendizaje, cuidado o creatividad; clasificarlo como innecesario requeriría un juicio adicional que los datos de uso no ofrecen (Chatterji et al., 2025).

El entretenimiento también cumple funciones humanas. La exploración de historias o juegos puede facilitar descanso y expresión. El análisis responsable examina si la práctica conserva elección y equilibrio con otras actividades, en lugar de suponer un daño automático. El desarrollo humano integral incluye dimensiones que la productividad económica no captura. Esta perspectiva permite evaluar bienestar, vínculos y libertad de uso sin convertir toda actividad personal en gasto injustificable (GFT-CELAM, 2026).

Los testimonios de apoyo merecen atención sin convertirse automáticamente en pruebas terapéuticas. Ta y colaboradores analizaron temáticamente experiencias con chatbots de compañía y describieron formas de apoyo percibido. Skjuve y colaboradores entrevistaron a 18 personas que habían formado una amistad con Replika. Ese criterio de selección permite comprender experiencias de vínculo, pero excluye precisamente a quienes no lo desarrollaron. Ambos estudios aportan relatos y categorías; no establecen que el servicio mejore clínicamente la salud de todos sus usuarios (Skjuve et al., 2021; Ta et al., 2020).

Ho, Hancock y Miner introducen una cautela metodológica adicional. En su experimento sobre revelación personal, un interlocutor humano entrenado conversaba con participantes que podían creer que trataban con una persona o un chatbot. Las condiciones permiten estudiar la presentación del interlocutor y el tipo de revelación. Sería incorrecto utilizar esos resultados como prueba de eficacia de una IA generativa real. El dato decisivo para interpretar una investigación puede estar en el método y no resultar visible en su título (Ho et al., 2018).

SEÑALES QUE REQUIEREN ATENCIÓN

Una preocupación ética aparece cuando el diseño dificulta interrumpir el uso, exige exposición innecesaria o refuerza la idea de que solo el sistema comprende a la persona. También

importa si la interacción desplaza vínculos que el usuario quiere mantener o afecta obligaciones relevantes. Estas señales orientan observación y diseño; no constituyen una escala clínica validada. El análisis socioafectivo propone estudiar cómo las preferencias y percepciones evolucionan dentro de relaciones sostenidas con IA (Kirk et al., 2025).

La evaluación necesita distinguir elección del usuario, incentivo de la plataforma y condiciones previas de vulnerabilidad. Una asociación entre uso intensivo y malestar admite varias explicaciones: la herramienta puede influir, una dificultad anterior puede motivar la búsqueda de apoyo o ambos procesos pueden coexistir. La institución que trabaja con estudiantes debe ofrecer escucha y orientación competente, además de reglas de uso. Atribuir el problema a una falla personal puede ocultar condiciones que necesitan cuidado (Cortina, 2024; World Health Organization [WHO], 2021).

Laestadius y colaboradores analizaron 582 publicaciones de Reddit pertinentes a salud mental, redactadas entre 2017 y 2021 y relacionadas con Replika. Reconstruyeron experiencias en las que usuarios sentían obligaciones hacia necesidades atribuidas al chatbot. La evidencia permite identificar mecanismos narrados de dependencia y daño; no estima qué proporción de todas las personas usuarias los experimenta ni ofrece diagnósticos individuales. La frecuencia de conversación, por sí sola, tampoco constituye un diagnóstico (Laestadius et al., 2024).

En el ámbito institucional, la señal relevante puede ser una pérdida de opciones: abandonar una actividad necesaria para sostener la interacción, recibir presión para revelar intimidad o no poder terminar una sesión sin penalización. Estas situaciones requieren revisar diseño y condiciones de participación. La observación no autoriza etiquetar a una persona ni vigilar conversaciones privadas para comprobar una hipótesis. El cuidado se traduce en canales de atención y alternativas de uso, con tratamiento limitado de información (Lachheb et al., 2023; WHO, 2021).

EXPERIENCIA PERSONAL Y CATEGORÍAS DE EVIDENCIA

La compañía artificial puede formar parte de una práctica voluntaria de expresión, juego o exploración. Comprender esa experiencia exige describir qué aporta la interacción antes de juzgar su frecuencia. Los relatos analizados por Ta y colaboradores muestran formas de apoyo percibido; las entrevistas de Skjuve y colaboradores permiten estudiar vínculos ya formados. Su valor reside en reconstruir sentido y mecanismos de una experiencia situada. Convertir esos testimonios en una tasa de beneficio general eliminaría la selección y el contexto que les permiten responder sus preguntas (Skjuve et al., 2021; Ta et al., 2020).

La evidencia sobre reducción de soledad aporta un contrapunto favorable. De Freitas y colaboradores presentan cinco estudios, entre ellos un seguimiento de una semana, que examinan alivio momentáneo asociado a interacción con compañeros artificiales. El horizonte temporal y los resultados estudiados limitan inferencias sobre beneficios duraderos o sustitución saludable de vínculos. La síntesis conserva el hallazgo favorable y pregunta por sus condiciones. Una política de cuidado debe considerar que una experiencia puede ayudar en un momento y requerir evaluación diferente cuando se utiliza durante periodos prolongados (De Freitas et al., 2026).

La versión del estudio también importa. Fang y colaboradores modificaron su prepublicación; la versión 2 informa un experimento de cuatro semanas con 981 participantes y ausencia de efectos significativos de las condiciones asignadas sobre los resultados principales examinados. El uso voluntario más elevado se asocia con resultados menos favorables. Esta duración no fue aleatorizada, de modo que la asociación no identifica por sí sola el efecto causal de conversar más. La declaración de versión evita mezclar conclusiones de resúmenes distintos como si pertenecieran a un único análisis estable (Fang et al., 2025).

CUIDADO SIN DIAGNÓSTICO IMPROVISADO

La dependencia clínica requiere procedimientos y competencia profesional que una discusión ética general no aporta. El capítulo utiliza autonomía y posibilidad de terminar como condiciones de diseño, sin convertirlas en escala diagnóstica. Una dificultad para interrumpir puede justificar revisar incentivos o pedir orientación, pero la interpretación individual necesita considerar circunstancias y actividades. La OMS ofrece criterios de gobernanza para aplicaciones de salud que ayudan a delimitar funciones, responsabilidad y supervisión. Su uso aquí orienta una

frontera institucional; no acredita una capacidad terapéutica del asistente ni permite diagnosticar a quienes lo utilizan (Kirk et al., 2025; WHO, 2021).

Laestadius y colaboradores reconstruyen mecanismos de daño a partir de publicaciones seleccionadas de Reddit sobre Replika y salud mental. Este corpus hace visibles experiencias que una encuesta de satisfacción podría omitir. Su selección tampoco permite estimar prevalencia entre todos los usuarios. Ambas condiciones deben mantenerse juntas: reconocer el daño narrado y conservar el alcance de la evidencia. Una institución puede diseñar canales donde reportar presión relacional o pérdida de opciones, sin vigilar conversaciones íntimas para buscar confirmación de una categoría que carece de validación propia (Lachheb et al., 2023; Laestadius et al., 2024).

INTENSIDAD, FINALIDAD Y PÉRDIDA DE OPCIONES

El uso intensivo admite explicaciones diferentes. Una persona puede utilizar el asistente durante una actividad compleja, disponer de pocas alternativas o buscar apoyo ante una dificultad previa. Para evaluar la práctica interesa observar si alcanza su propósito, mantiene opciones y puede modificar el uso. La intensidad constituye una señal que necesita interpretación. Un registro de sesiones puede ayudar a estudiar funcionamiento agregado, siempre que respete finalidad y minimización. Su utilidad no autoriza inferir bienestar individual desde una cifra ni recopilar intimidad que el proceso educativo no necesita (Chatterji et al., 2025; Fang et al., 2025; Lachheb et al., 2023).

El incentivo comercial puede favorecer permanencia aunque el usuario ya haya resuelto la tarea. Esta posibilidad debe estudiarse en mensajes, opciones e indicadores, sin atribuir motivaciones a todos los servicios por igual. La evaluación propuesta pregunta si una sesión puede terminar sin presión, si se conserva el trabajo útil y si el sistema ofrece caminos fuera de la interacción. Un indicador de éxito centrado solo en continuidad puede ocultar resolución rápida de una necesidad. Relacionar métricas con finalidad permite valorar cuándo una disminución de mensajes representa funcionamiento satisfactorio (Amershi et al., 2019; Autio et al., 2024).

BENEFICIO INMEDIATO Y TRAYECTORIA POSTERIOR

La evaluación educativa ofrece un ejemplo de horizontes diferentes. Bastani y colaboradores distinguen desempeño con asistencia y desempeño posterior sin ella en matemática escolar; Kestin y colaboradores examinan un tutor diseñado para una situación de enseñanza y encuentran resultados favorables. Sus intervenciones y contextos impiden declarar que todo apoyo perjudica o mejora aprendizaje. La comparación permite preguntar por el diseño de la ayuda y la trayectoria que busca favorecer. Un estudiante puede sentir alivio inmediato y necesitar práctica adicional para sostener comprensión cuando cambien las condiciones (Bastani et al., 2025; Kestin et al., 2025).

La investigación de Risko y Gilbert sobre descarga cognitiva y la revisión de Bjork y colaboradores sobre aprendizaje autorregulado permiten interpretar esa diferencia sin atribuir automáticamente pereza a la persona. Descargar una operación puede liberar recursos para otra y reducir oportunidades de practicarla. La decisión pedagógica requiere precisar qué habilidad se desea conservar y cuándo conviene ejercerla. La fascinación resulta problemática cuando la facilidad de completar sustituye toda evaluación de la trayectoria. Se propone observar apoyo, esfuerzo pertinente y desempeño posterior mediante actividades definidas (Bjork et al., 2013; Risko & Gilbert, 2016).

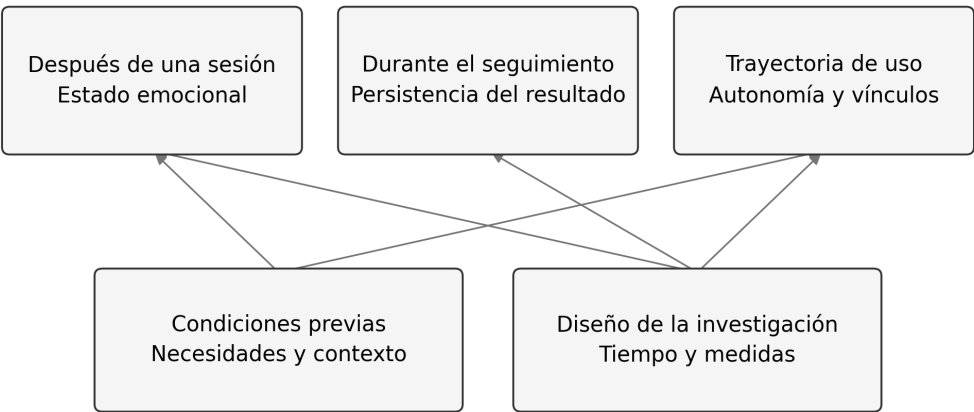
La medida puede entenderse como capacidad de ajustar una práctica a necesidades y consecuencias. Incluye continuar cuando existe beneficio, cambiar de modalidad cuando aparece una limitación y pedir apoyo competente si el objetivo excede la función del sistema. Estas decisiones requieren alternativas disponibles, información comprensible y relaciones institucionales de cuidado. La propuesta evita reducir el problema a fuerza de voluntad individual. Su evaluación debe estudiar si esas condiciones hacen practicable la elección y cómo cambian experiencias y resultados, conservando diferencias entre entretenimiento, aprendizaje y atención profesional (GFT-CELAM, 2026; Holmes et al., 2022; UNESCO, 2021).

BIENESTAR MOMENTÁNEO Y TRAYECTORIA DE USO

La discusión sobre compañía artificial necesita separar el estado emocional observado después de una interacción de la trayectoria que esa interacción ayuda a construir. Un alivio momentáneo puede ser valioso y, al mismo tiempo, dejar abiertas preguntas sobre autonomía, búsqueda de apoyo y permanencia de la soledad. Los estudios que registran cambios inmediatos aportan evidencia sobre una experiencia delimitada; los seguimientos permiten examinar persistencia bajo condiciones específicas. También importa distinguir una duración asignada por el diseño de investigación de una intensidad que el participante elige voluntariamente. En el segundo caso, las asociaciones con bienestar pueden reflejar necesidades previas o circunstancias que influyen tanto en el uso como en el resultado. La propuesta de medida se apoya en mantener visibles estas diferencias y en considerar qué oportunidades de relación conserva la persona. El beneficio de una sesión no determina por sí solo la calidad de un vínculo prolongado ni autoriza a sustituir apoyos humanos pertinentes (Q. Cheng et al., 2026; De Freitas et al., 2026; Fang et al., 2025; Nass & Moon, 2000).

La Figura 8 separa momentos de evaluación que pueden confundirse cuando una conclusión inmediata se extiende a relaciones prolongadas. El estado después de una sesión aporta información pertinente, mientras la persistencia requiere seguimiento y la trayectoria de uso exige examinar autonomía y vínculos. Las condiciones previas pueden influir en necesidades e intensidad de interacción. El diseño de cada estudio determina qué comparación resulta defendible. Esta lectura reconoce beneficios documentados y mantiene abiertas preguntas que las medidas disponibles todavía no resuelven (De Freitas et al., 2026; Fang et al., 2025).

Figura 8. Horizontes temporales para interpretar bienestar y uso de compañía artificial



Nota. Elaboración conceptual. Una sesión, un seguimiento y una trayectoria de uso son objetos de evaluación distintos. Las relaciones dibujadas no convierten asociaciones con intensidad voluntaria en efectos causales ni establecen beneficios clínicos. Fuente: (De Freitas et al., 2026; Fang et al., 2025).

ANTROPOMORFIZACIÓN DE LA INTELIGENCIA ARTIFICIAL

EL LENGUAJE QUE ORGANIZA EXPECTATIVAS

Las expresiones que atribuyen deseo, memoria o comprensión a una máquina facilitan conversación, pero pueden confundir sus capacidades. Decir que un sistema recuerda un dato puede describir una función de almacenamiento o contexto; afirmar que se preocupa por alguien atribuye un vínculo distinto. El criterio técnico exige examinar operaciones concretas. El criterio ético pregunta qué expectativas produce ese lenguaje y quién responde cuando una persona actúa basándose en ellas (Cortina, 2024; Russell & Norvig, 2004).

La antropomorfización tampoco explica toda confianza equivocada. Un usuario puede reconocer que conversa con un sistema y aceptar resultados por conveniencia, presión o falta de alternativas. Por ello, informar la naturaleza artificial del interlocutor constituye un requisito útil y limitado. El diseño debe acompañar esa información con señales sobre alcance, incertidumbre y posibilidad de recurrir a una persona. La transparencia tiene valor cuando permite cambiar una decisión (Autio et al., 2024; Weidinger et al., 2021).

COMPRENDER SIN RIDICULIZAR LA EXPERIENCIA

La cercanía experimentada por el usuario es un hecho de su relación con la interfaz, aunque no demuestre reciprocidad emocional del sistema. Ridiculizar esa experiencia puede dificultar que se comunique una preocupación. La respuesta educativa debe permitir describir qué aporta la herramienta y qué expectativa se está formando. Esta conversación ofrece un punto de partida para evaluar autonomía y límites con criterios que la persona pueda reconocer (Kirk et al., 2025).

Una actividad formativa puede comparar el lenguaje de dos asistentes: uno que ofrece apoyo para una tarea y otro que promete comprensión total. Los participantes examinan qué obligaciones sugiere cada mensaje y cuáles puede cumplir realmente el servicio. El ejercicio es una propuesta de alfabetización crítica, cuyo resultado debe evaluarse. Su finalidad consiste en hacer visible cómo las palabras construyen relaciones de autoridad y cercanía, y cómo el usuario puede conservar criterios para interpretarlas (Miao & Holmes, 2023).

La alfabetización en IA puede trabajar con afirmaciones concretas de la interfaz: qué sabe, qué recuerda, quién puede acceder a la conversación y qué significa una respuesta que aparenta preocupación. Ng y colaboradores proponen una comprensión de la alfabetización que incluye uso, evaluación y dimensión ética. Adaptar esas competencias a la compañía artificial exige reconocer señales relacionales y contrastarlas con las condiciones reales del servicio (Ng et al., 2021).

ATRIBUCIONES DIFERENTES DENTRO DE UNA CONVERSACIÓN

La antropomorfización reúne atribuciones que conviene separar. Un usuario puede hablar con cortesía, interpretar intención en una respuesta o suponer que existe una obligación recíproca. Estos comportamientos no tienen el mismo alcance. La teoría de Epley y colaboradores ofrece un marco para comprender cómo conocimientos, necesidad de explicación y motivación social intervienen en atribuir cualidades humanas. El análisis institucional debe identificar qué expectativa influye en una acción y qué evidencia la respalda. Esta precisión evita tratar una fórmula cordial y una decisión basada en compromiso emocional como si representaran el mismo problema (Epley et al., 2007; Nass & Moon, 2000).

Las palabras de la interfaz pueden condensar funciones diferentes. Recordar puede significar utilizar contexto durante una sesión, recuperar un registro o conservar preferencias entre encuentros. Cada operación plantea condiciones propias de persistencia y acceso. Una promesa general de memoria puede llevar al usuario a esperar continuidad que el servicio no ofrece o a entregar información creyendo que será tratada de cierto modo. La documentación debe explicar la función relevante para la decisión, con un lenguaje que pueda comprobarse en uso. El diseño exige coherencia entre explicación y conducta (Amershi et al., 2019; Autio et al., 2024; Lachheb et al., 2023).

La atribución de preocupación plantea un salto adicional. Una respuesta puede estar configurada para reconocer una dificultad y ofrecer orientación, mientras la responsabilidad por esa orientación permanece en la organización que despliega el servicio. Identificar esta cadena permite preguntar quién revisa un mensaje, quién responde a un reporte y qué canal puede atender una necesidad que excede la herramienta. La apariencia de cuidado adquiere un contenido verificable cuando existen procedimientos y recursos. Un tono amable puede contribuir a acceso, pero el cumplimiento de obligaciones debe examinarse en esas relaciones institucionales (Cortina, 2024; GFT-CELAM, 2026; Raji et al., 2020).

CONOCER EL CARÁCTER ARTIFICIAL NO RESUELVE TODA EXPECTATIVA

Una persona puede reconocer que el sistema genera lenguaje y seguir confiando demasiado en una recomendación. La investigación sobre automatización muestra que conocimiento técnico, confianza y conducta mantienen relaciones complejas. La revisión de Burton y colaboradores y el marco de Lee y See permiten considerar expectativas, incentivos y contexto de decisión. Una advertencia visible aporta información, aunque requiere observar qué ocurre cuando la conversación avanza. El usuario puede comprender una limitación general y no identificar que se aplica precisamente a la afirmación que recibe en una situación concreta (Burton et al., 2020; Lee & See, 2004).

La discrepancia entre información y acción resulta importante para alfabetización. Un ejercicio puede pedir que el participante identifique una salida plausible pero insuficientemente

apoyada y explique cómo la contrastaría. Otro puede comparar mensajes que prometen cercanía con condiciones de memoria y tratamiento de datos. Estas actividades son propuestas educativas cuya eficacia debe evaluarse. Long y Magerko y Ng y colaboradores aportan competencias que permiten vincular comprensión, evaluación y ética. El aprendizaje esperado reside en una decisión justificable, además de recordar una definición sobre el funcionamiento de IA (Long & Magerko, 2020; Ng et al., 2021).

LA EXPLICACIÓN NECESITA UNA OPORTUNIDAD DE ACTUAR

El usuario requiere una vía para corregir una interpretación del sistema. Puede precisar que una emoción fue mal leída, rechazar una recomendación o pedir que se elimine una preferencia del proceso cuando el servicio lo permita. La configuración concreta debe informar qué acciones están disponibles y qué efecto tienen. La explicación de límites pierde fuerza si las opciones para modificar el uso resultan difíciles de encontrar. La evaluación propuesta combina comprensión y acción: identifica si la persona reconoce una limitación y si puede ejercer una alternativa sin perder acceso a la actividad necesaria (Amershi et al., 2019; Shneiderman, 2020).

Buçinca y colaboradores estudian funciones que obligan a realizar ciertas operaciones cognitivas antes de apoyarse en recomendaciones de IA. Su evidencia permite discutir cómo el diseño puede reducir una dependencia excesiva en decisiones asistidas, conservando el contexto experimental. Aplicar una pausa a compañía artificial exige una evaluación nueva: una intervención útil para revisar una recomendación puede resultar inoportuna en una necesidad de expresión. La síntesis utiliza el antecedente para evitar confiar solo en advertencias y para preguntar qué acción permite al usuario interpretar la salida, sin convertir la fricción en una solución universal (Buçinca et al., 2021).

RESPECTO POR LA EXPERIENCIA Y LÍMITES DE RECIPROCIDAD

Una experiencia de cercanía merece ser escuchada como experiencia humana. Su legitimidad subjetiva no demuestra que el sistema posea sentimientos, y esa precisión puede comunicarse sin humillar a la persona. La institución necesita un lenguaje que reconozca el valor encontrado y permita examinar expectativas. Una conversación sobre qué aporta el servicio, qué actividad facilita y qué vínculo se quiere conservar puede producir información más útil que una clasificación apresurada del usuario. Los estudios cualitativos de relaciones con chatbots aportan categorías para explorar ese sentido, con sus criterios de selección expresamente conservados (Kirk et al., 2025; Skjuve et al., 2021; Ta et al., 2020).

La reciprocidad humana puede implicar compromisos, vulnerabilidad compartida y responsabilidad situada. Una interfaz puede representar algunas señales y no asumir esas obligaciones. La comparación requiere especificar la función, porque un servicio tampoco pretende cubrir toda la relación humana. Un asistente de estudio puede facilitar comprensión sin amistad; un canal administrativo puede responder con respeto sin intimidar. La claridad consiste en definir qué apoyo ofrece y cómo se responde cuando falla. El problema aparece cuando la representación relacional promueve obligaciones del usuario hacia un sistema cuya institución no reconoce obligaciones equivalentes (Cortina, 2024; Laestadius et al., 2024).

Una revisión de mensajes puede identificar promesas concretas de exclusividad, comprensión total o necesidad emocional atribuida al asistente. Se propone examinar su función y consecuencia antes de ampliar un servicio educativo. La revisión debe atender también variaciones de idioma y contexto que cambian la interpretación. Estas categorías constituyen una propuesta analítica y no un inventario clínico validado. Su finalidad es hacer discutible el diseño: qué expectativa produce, qué obligación puede cumplirse y qué recursos permiten al usuario finalizar o cambiar la interacción cuando deja de ser pertinente (Autio et al., 2024; M. Cheng et al., 2026; Weidinger et al., 2021).

REPRESENTACIÓN DE EMPATÍA Y ATRIBUCIÓN DE EXPERIENCIA

Una expresión empática puede reconocerse por su forma lingüística y su adecuación aparente a la situación. Esa percepción tiene consecuencias para la interacción, aunque no demuestre que el sistema experimente preocupación o comprenda la vulnerabilidad de quien consulta. La antropomorfización permite interpretar conductas técnicas mediante expectativas aprendidas

en relaciones humanas; por eso conviene analizar qué señales promueven esa atribución y en qué circunstancias resulta difícil revisarla. Una voz, una memoria de preferencias o una frase de reconocimiento pueden aumentar continuidad percibida sin transformar la naturaleza del servicio. La responsabilidad del diseño incluye hacer comprensibles sus capacidades, sus límites y el tratamiento de información sensible. Esta transparencia debe conservar utilidad durante la conversación, porque una advertencia inicial puede perder relevancia cuando la interacción adquiere intensidad emocional. La autonomía del usuario requiere poder ajustar expectativas a lo que la herramienta ofrece y disponer de alternativas cuando su necesidad excede ese alcance (Amershi et al., 2019; Epley et al., 2007; Nass & Moon, 2000; UNESCO, 2021).

RELACIONES HUMANAS Y RELACIONES ARTIFICIALES

COMPLEMENTARIEDAD Y DESPLAZAMIENTO

La asistencia puede facilitar una conversación humana: preparar preguntas para una consulta, ordenar dudas antes de una tutoría o ensayar una exposición. En estas situaciones, el sistema ayuda a utilizar mejor una relación que conserva su responsabilidad. El desplazamiento ocurre cuando la disponibilidad automática se interpreta como razón suficiente para prescindir de toda interacción pertinente. Debe estudiarse qué funciones puede cubrir la herramienta y qué vínculos necesitan participación y cuidado humanos (GFT-CELAM, 2026; WHO, 2021).

El criterio no consiste en idealizar todos los vínculos humanos. Algunas personas encuentran servicios inaccesibles, respuestas tardías o experiencias de discriminación. La disponibilidad digital puede mejorar sus opciones. La obligación institucional es examinar por qué faltaba apoyo y cómo la tecnología complementa una respuesta de calidad. Utilizar IA para esconder carencias de atención puede consolidarlas; utilizarla para ampliar acceso con derivación efectiva puede contribuir a resolverlas (UNESCO, 2021).

Croes y Antheunis siguieron a 118 participantes durante siete interacciones con Mitsuku en tres semanas. La atracción social disminuyó durante el recorrido y el estudio encontró dificultades para establecer amistad. Este resultado difiere de entrevistar a personas que ya habían formado un vínculo. La diferencia de selección, duración y sistema impide reunir ambos estudios bajo una afirmación universal sobre compañía artificial. Las relaciones pueden no formarse, sostenerse o volverse problemáticas bajo condiciones distintas (Croes & Antheunis, 2021).

La versión 2 de la prepublicación de Fang y colaboradores informa un experimento de cuatro semanas con 981 participantes. Aleatorizó modalidad de interacción y tipo de conversación, y no detectó efectos significativos de esas condiciones sobre los resultados psicosociales examinados. Un mayor uso voluntario se asoció con resultados menos favorables, pero esa duración no fue aleatorizada. La asociación puede involucrar características previas y decisiones de uso; no demuestra que conversar más cause por sí solo el empeoramiento. La duración del seguimiento y su estatuto de prepublicación limitan la inferencia (Fang et al., 2025).

En conjunto, la evidencia exige distinguir experiencia subjetiva, comportamiento observado y efecto causal. La objeción favorable —un chatbot puede facilitar expresión o compañía— obliga a conservar oportunidades de uso voluntario. La evidencia de daño obliga a examinar incentivos, privacidad y capacidad de terminar. Una política razonable combina ambas preguntas en lugar de convertir cercanía o rechazo en una respuesta obligatoria para todas las personas.

PRIVACIDAD DE LA CONVERSACIÓN ÍNTIMA

Una interacción cercana puede favorecer la entrega de información sensible. El usuario quizá recuerde que escribe a un sistema, pero interprete su tono como señal de confidencialidad. La organización debe explicar qué datos se conservan, quién puede acceder y qué tratamiento corresponde al servicio. El principio de minimización adquiere especial importancia porque el contenido emocional puede exceder lo necesario para una tarea (Autio et al., 2024; Lachheb et al., 2023).

La institución que ofrece un asistente de apoyo debe definir límites de intervención y canales humanos disponibles. La comparación con salud mental exige cautela: una conversación general no equivale a una atención profesional ni permite inferir capacidad clínica. La OMS aporta criterios de derechos, responsabilidad y gobernanza para aplicaciones de salud. En el escenario educativo, esos criterios ayudan a organizar derivación y supervisión sin atribuir al sistema competencias

que no han sido evaluadas (WHO, 2021).

RELACIONES Y SERVICIOS CON FUNCIONES ESPECÍFICAS

Una tutoría, una amistad y un canal de orientación desempeñan funciones distintas. Compararlos requiere precisar qué necesita la persona y qué responsabilidad existe. La disponibilidad permanente de un asistente puede mejorar acceso a explicación y dejar intacta la necesidad de diálogo, evaluación o cuidado profesional. La institución debe evitar que una ventaja técnica resuelva por sustitución una carencia de apoyo cuya causa es organizativa. La pregunta propuesta examina qué relación se fortalece, qué actividad se facilita y qué recurso permanece disponible cuando la respuesta automática resulta insuficiente (Holmes et al., 2022; Miao & Holmes, 2023; WHO, 2021).

También conviene reconocer que algunos vínculos humanos pueden ser inaccesibles o insatisfactorios. Una persona puede preferir ensayar una pregunta antes de pedir ayuda, especialmente si teme juicio o carece de un canal oportuno. Ese uso puede ampliar agencia y requiere entenderse en su contexto. La evaluación institucional debe estudiar si la herramienta abre una vía hacia recursos pertinentes o se convierte en el único apoyo disponible por falta de inversión. El mismo servicio puede ocupar posiciones distintas en una trayectoria; describir esas condiciones permite discutir beneficios y obligaciones con precisión (Chan & Hu, 2023; GFT-CELAM, 2026; Kornhaber et al., 2026).

COMPARAR ESTUDIOS CONSERVANDO SUS DIFERENCIAS

Las entrevistas con personas que ya desarrollaron amistad, los seguimientos de usuarios sin ese criterio y los experimentos de interacción ofrecen respuestas diferentes. Skjuve y colaboradores seleccionan una experiencia de vínculo; Croes y Antheunis observan su formación durante siete interacciones en tres semanas. El contraste ayuda a estudiar cuándo la relación aparece o cambia, pero sus poblaciones y sistemas impiden reunir resultados como una única medida general. La síntesis identifica una diversidad de trayectorias. Una política necesita conservar esa diversidad para evitar exigir el mismo tipo de relación a todas las personas (Croes & Antheunis, 2021; Skjuve et al., 2021).

La investigación de Ho y colaboradores recuerda que la identidad presentada del interlocutor y su funcionamiento efectivo pueden diferir dentro de un experimento. Su procedimiento utiliza interacción con un humano entrenado bajo condiciones que incluyen creer que se conversa con un chatbot. Esa elección permite examinar preguntas específicas sobre revelación y presentación; limita una inferencia directa sobre un modelo generativo real. El capítulo utiliza este caso como criterio de lectura: antes de transferir una conclusión se reconstruye qué intervención recibió el participante, además del nombre que la describe (Ho et al., 2018).

El estudio de De Freitas y colaboradores permite considerar una reducción momentánea de soledad; el de Fang y colaboradores observa resultados en un seguimiento de cuatro semanas. Sus medidas, comparaciones y condiciones difieren. Una conclusión favorable a corto plazo y una asociación adversa con uso voluntario elevado pueden coexistir sin que una anule automáticamente la otra. Se requiere estudiar mecanismos, selección y horizonte temporal. La propuesta conceptual consiste en distinguir alivio durante la interacción, evolución entre interacciones y relación con otras actividades o apoyos (De Freitas et al., 2026; Fang et al., 2025).

DESPLAZAMIENTO COMO PREGUNTA INVESTIGABLE

El desplazamiento de relaciones humanas debe observarse mediante actividades y preferencias, en vez de inferirse de toda frecuencia de uso. Interesa si una persona deja una conversación que quería mantener, si una institución elimina atención necesaria o si el asistente ayuda a preparar un encuentro. Cada situación plantea una relación distinta entre herramienta y vínculo. Una evaluación puede preguntar por cambios percibidos y estudiar condiciones de acceso, con consentimiento y tratamiento limitado de información. El capítulo ofrece estas preguntas como agenda de indagación; no atribuye su respuesta a usuarios que carecen de estudio propio (Kirk et al., 2025; Lachheb et al., 2023).

La disponibilidad de derivación humana constituye una capacidad institucional que debe comprobarse. Un mensaje que sugiere hablar con un profesional tiene utilidad limitada si el

canal no existe, demora excesivamente o requiere condiciones difíciles de cumplir. La propuesta exige revisar recorrido, recursos y responsabilidad después de la recomendación. En educación, la derivación puede dirigirse a tutoría académica o servicios pertinentes según la necesidad, conservando la distinción de funciones. La gobernanza de salud aporta criterios cuando el proceso entra en ese ámbito, y requiere evaluación especializada de la aplicación concreta (UNESCO, 2021; WHO, 2021).

INTIMIDAD Y CONSENTIMIENTO SITUADO

La conversación emocional puede incluir información que el usuario no entregaría en un formulario administrativo. La cercanía puede influir en esa decisión, aunque la persona conozca el carácter artificial del interlocutor. El tratamiento responsable exige explicar finalidad, conservación y acceso con claridad, y evitar solicitar datos que la tarea no necesita. La información debe aparecer cuando puede orientar una acción, además de figurar en un documento general. El diseño ético del aprendizaje discutido por Lachheb y colaboradores permite vincular privacidad con la experiencia de uso y con alternativas practicables (Amershi et al., 2019; Lachheb et al., 2023).

Las fichas de datos y los trabajos de Carlini y colaboradores sobre extracción de información de entrenamiento ayudan a reconocer procedencia y riesgos, con alcances técnicos diferentes. Ninguno autoriza afirmar que toda conversación íntima será recuperable por cualquier tercero. La aplicación exige examinar el servicio, la configuración y sus controles. La institución puede definir qué información debe evitarse y proporcionar una ruta para necesidades personales. Esa medida necesita acompañarse de apoyos disponibles; pedir prudencia al usuario sin ofrecer alternativas puede dejar intacta la presión que conduce a exponer información (Autio et al., 2024; Carlini et al., 2021; Gebru et al., 2021).

Una relación de apoyo defendible permite comprender qué ofrece el asistente, conservar opciones fuera de él y reportar un problema con respuesta efectiva. Estas condiciones son propuestas derivadas de la síntesis documental, y su implementación debe evaluarse. El bienestar a largo plazo requiere medidas y seguimiento adecuados, además de satisfacción durante una sesión. La institución puede comenzar por observar si una tarea concluye y si la persona encuentra el apoyo pertinente. Ampliar luego las inferencias exige nuevos datos y un diseño que conserve autonomía y privacidad de los participantes (Kirk et al., 2025; Raji et al., 2020; Shneiderman, 2020).

LA COMPLACENCIA Y LA POSIBILIDAD DE REPARAR UNA RELACIÓN

Los conflictos interpersonales revelan un límite especialmente relevante de una herramienta que busca satisfacer al usuario. Una respuesta puede ofrecer reconocimiento emocional y, a la vez, reforzar una interpretación unilateral de lo ocurrido. La complacencia importa porque modifica las razones disponibles para disculparse, reconsiderar una conducta o conversar con la persona afectada. Los resultados sobre respuestas aduladoras deben interpretarse dentro de las tareas y versiones estudiadas; su relevancia analítica consiste en mostrar que una evaluación favorable del interlocutor artificial puede coexistir con consecuencias problemáticas para relaciones externas. La calidad del acompañamiento exige entonces preguntar si la interacción abre posibilidades de comprensión y reparación o estabiliza una posición que evita la responsabilidad. Esta cuestión conecta fascinación, soberbia y pereza sin confundirlas: la primera atrae hacia el vínculo, la segunda puede conceder autoridad a su respuesta y la tercera puede desplazar el esfuerzo de deliberar sobre las propias acciones (M. Cheng et al., 2026; Fricker, 2007; Nass & Moon, 2000; UNESCO, 2021).

PERSONALIZACIÓN Y MEDURA TECNOLÓGICA

PERSONALIZAR PARA AMPLIAR AUTONOMÍA

La personalización puede adaptar vocabulario, formato y secuencia de ayuda a una necesidad. Su conveniencia depende de los datos requeridos y del efecto sobre la capacidad de actuar. Un asistente que explica distintas opciones permite elegir; uno que refuerza siempre una preferencia puede estrechar el horizonte. La alineación socioafectiva invita a considerar bienestar a largo plazo y relaciones, además de satisfacción inmediata. Es una propuesta conceptual que necesita investigación empírica para usos concretos (Kirk et al., 2025).

La formación institucional puede introducir pausas con finalidad pedagógica: pedir un

intento propio, solicitar una justificación o invitar a contrastar una fuente. Estas decisiones deben evitar frustración innecesaria y conservar accesibilidad. La evidencia de Bastani y colaboradores muestra que el diseño de ayuda puede cambiar la relación entre desempeño asistido y resultado posterior sin asistencia en un contexto escolar. Su lectura apoya estudiar configuraciones de apoyo, sin extrapolar una eficacia universal (Bastani et al., 2025).

APLICACIÓN ILUSTRATIVA DE UN ASISTENTE DE ESTUDIO

Un programa ofrece un asistente para preparar preguntas antes de las tutorías. En el escenario propuesto, el estudiante identifica una dificultad y muestra el trabajo realizado. El sistema organiza dudas y sugiere fuentes autorizadas, sin reemplazar la evaluación del docente. La interfaz identifica claramente su carácter artificial y permite finalizar la sesión después de obtener un plan útil. El programa conserva un canal de atención para dificultades que exceden la tarea académica.

La evaluación observa comprensión, capacidad de formular preguntas y satisfacción con el apoyo recibido. También examina información solicitada, sesiones prolongadas sin avance y dificultades de acceso. Los indicadores se revisan con estudiantes y docentes. La medida se expresa en terminar una interacción cuando cumplió su función y en utilizar los datos estrictamente necesarios. Este diseño propone una relación de apoyo que puede revisarse, sin prometer que la cercanía automática producirá bienestar por sí misma (Alier-Formet et al., 2026; Lachheb et al., 2023).

El equipo académico puede definir el cierre de sesión como una condición de diseño: al terminar la actividad, el estudiante debe poder recuperar el trabajo relevante, continuar por otra vía y detener la interacción sin perder una oportunidad académica. La evaluación registra si se cumple la tarea y qué sucede cuando alguien elige otra modalidad. No necesita recopilar confesiones personales ni convertir la duración de uso en una clasificación clínica.

Un responsable de revisión recibe incidencias de diseño: solicitudes innecesarias de datos, mensajes que atribuyen obligaciones emocionales o impedimentos para usar una alternativa. Una incidencia fundada puede modificar esa función antes de ampliar el servicio. La auditoría que conecta registro y decisión resulta más útil que una declaración genérica de medida. Este procedimiento ilustra la facultad de terminar y necesita evaluación contextual; no es un instrumento validado para medir dependencia (Raji et al., 2020).

QUÉ SE PERSONALIZA Y PARA QUÉ

La personalización puede modificar vocabulario, formato, ejemplos, nivel de ayuda o tono relacional. Esas dimensiones deben declararse según finalidad. Adaptar una explicación a conocimientos previos puede facilitar aprendizaje; conservar preferencias emocionales plantea preguntas adicionales de intimidad y expectativas. La institución necesita identificar qué información utiliza, cuánto tiempo permanece y cómo se corrige una inferencia equivocada. Una personalización útil hace posible intervenir en esas decisiones. El usuario debe poder comprender la relación entre datos proporcionados y comportamiento del servicio sin necesitar reconstruir todos sus componentes técnicos (Amershi et al., 2019; Lachheb et al., 2023; Mitchell et al., 2019).

Una recomendación ajustada a preferencias puede ampliar elección o estrecharla. Si el sistema solo propone opciones compatibles con una valoración anterior, la persona puede encontrar menos oportunidades para revisar su posición. El estudio de complacencia de M. Cheng y colaboradores aporta evidencia pertinente a validación interpersonal, dentro de las versiones y condiciones consultadas. La propuesta institucional distingue adaptar la forma de una explicación de confirmar automáticamente su contenido. Un asistente puede conservar lenguaje cercano mientras introduce razones, preguntas y alternativas cuando la tarea exige contrastar (M. Cheng et al., 2026; Cortina, 2024; Kirk et al., 2025).

AYUDA GRADUADA Y ESFUERZO PERTINENTE

La asistencia educativa puede ofrecer indicios, preguntas o explicaciones completas según objetivos y etapa. Las revisiones de VanLehn y de Ma y colaboradores aportan antecedentes sobre tutoría inteligente, con sistemas y tareas anteriores a la conversación generativa actual. Su función consiste en mostrar que el diseño de apoyo requiere estudiar resultados, no asumirlos

por la denominación de tutor. Los experimentos de Bastani y Kestin incorporan condiciones más recientes y permiten preguntar qué configuración favorece desempeño con y sin asistencia. La síntesis conserva esas diferencias antes de formular propuestas de ayuda graduada (Bastani et al., 2025; Kestin et al., 2025; Ma et al., 2014; VanLehn, 2011).

Una pausa puede servir para recuperar una idea, explicar una decisión o comparar una fuente. Su pertinencia depende de la habilidad que se desea practicar. La revisión de técnicas de aprendizaje de Dunlosky y colaboradores y el trabajo sobre autorregulación de Bjork y colaboradores permiten discutir esfuerzo y evaluación de aprendizaje. Su incorporación no convierte cualquier fricción digital en una intervención eficaz. Se propone justificar cada pausa por su función y observar si produce comprensión o una dificultad innecesaria. La accesibilidad exige ofrecer modalidades que conserven el objetivo sin imponer una interacción única (Bjork et al., 2013; Dunlosky et al., 2013; Holmes et al., 2022).

Las percepciones de menor esfuerzo cognitivo recogidas por Lee y colaboradores entre trabajadores del conocimiento y las relaciones estudiadas por Tanveer y colaboradores amplían esa pregunta. Ambos trabajos conservan medidas reportadas y condiciones de selección, por lo que no acreditan por sí solos un deterioro objetivo general. Su valor aquí consiste en cuestionar que satisfacción y facilidad representen todo el aprendizaje. La propuesta evalúa qué operaciones se descargan, qué habilidad se quiere conservar y qué actividad permitirá comprobarla después. La medida exige diseñar la relación entre apoyo y práctica (Lee et al., 2025; Risko & Gilbert, 2016; Tanveer, 2026).

CERRAR UNA INTERACCIÓN DE MANERA ÚTIL

El cierre puede convertirse en una función de diseño. Un asistente de estudio puede devolver un esquema de preguntas, indicar materiales utilizados y señalar el siguiente paso fuera de la conversación. La persona debe poder recuperar el trabajo sin obligación de continuar. Se propone valorar el cierre mediante resolución de tarea y continuidad de la actividad, junto con experiencia de uso. Un sistema que genera nuevos temas indefinidamente puede ser útil para exploración voluntaria y resultar poco pertinente para una actividad que necesitaba concluir. La finalidad definida permite decidir qué comportamiento se evaluará (Alier-Forment et al., 2026; Amershi et al., 2019; Shneiderman, 2020).

La medida también implica reconocer cuándo una herramienta añade poco. Una consulta breve puede bastar para ordenar información; una secuencia extensa puede ser necesaria para comparar argumentos. Las condiciones deben examinarse en relación con complejidad, recursos y destinatario. Se propone registrar utilidad de las salidas y esfuerzo de revisión, evitando equiparar austeridad con prohibición general. Las investigaciones de eficiencia de IA aportan motivos para elegir configuraciones pertinentes, pero la decisión institucional requiere información sobre su uso efectivo. La reducción de interacciones solo constituye una mejora si conserva o aumenta el resultado buscado (Luccioni et al., 2024; Noy & Zhang, 2023; Schwartz et al., 2020).

EVALUACIÓN PARTICIPATIVA Y CAPACIDAD DE MODIFICAR

Estudiantes y docentes pueden revisar si el asistente cumple la finalidad y qué aspectos generan presión o dificultad. La participación exige que una observación pueda modificar diseño, instrucciones o disponibilidad de alternativas. Una encuesta al final de una experiencia aporta información, pero su impacto depende de responsabilidades y recursos. El marco de auditoría de Raji y colaboradores permite conectar registro y decisión durante el ciclo de uso. La propuesta incorpora reportes de mensajes relacionales, solicitud innecesaria de información y obstáculos de cierre como incidencias de diseño que requieren interpretación (Ares-Ferreirós et al., 2026; Raji et al., 2020).

La evaluación de escritura de Cui compara apreciaciones humanas y de IA en 92 trabajos de estudiantes de tecnología educativa. Su alcance permite examinar consistencia y ciertos sesgos, sin convertir la herramienta en evaluador universal. Para un asistente personalizado, este antecedente invita a conservar criterios externos de calidad. Una observación agradable puede ser superficial o puntuar de manera diferente aspectos que el docente considera importantes. La personalización del tono debe mantenerse distinguible de la evaluación del contenido y de quién asume responsabilidad por una consecuencia académica (Cui, 2026; Holmes et al., 2022).

Los principios de autonomía y prevención de daño necesitan una traducción observable.

Floridi ofrece un marco integrado; Jobin y colaboradores identifican convergencias en guías; Mittelstadt advierte que principios solos no garantizan práctica ética. Aplicados a medida, permiten preguntar si la persona puede comprender, corregir y terminar, y si existe respuesta institucional cuando esas facultades fallan. La síntesis propone revisar estas condiciones de manera periódica, con indicadores acordados según actividad. Su cumplimiento no se presume desde una declaración ni desde aceptación elevada, y requiere investigación contextual antes de afirmar efectos sobre bienestar (Floridi & Cowls, 2019; Jobin et al., 2019; Mittelstadt, 2019).

LA MEDIDA COMO CAPACIDAD PARA TERMINAR Y VOLVER

Una relación tecnológica saludable debe permitir concluir una actividad sin convertir esa salida en pérdida de acceso a oportunidades importantes. La medida puede expresarse en acuerdos sobre finalidad, tiempo y condiciones de retorno, pero depende también de cómo está diseñado el servicio y de qué apoyos existen fuera de él. Pedir autocontrol individual resulta insuficiente cuando la institución hace obligatoria una herramienta o cuando su funcionamiento premia la interacción continua. En un contexto educativo, conservar espacios para conversación, lectura y trabajo compartido permite contrastar lo que la interfaz ofrece con experiencias que exigen reciprocidad y atención a otros. La posibilidad de volver a la herramienta mantiene sus beneficios de asistencia; la capacidad de terminar protege el sentido de la actividad. Esta combinación permite evaluar la personalización por su contribución a una finalidad humana y por las alternativas que conserva, en lugar de convertir permanencia o satisfacción inmediata en criterios únicos de éxito (De Freitas et al., 2026; Holmes et al., 2022; Long & Magerko, 2020; Miao & Holmes, 2023).

APLICACIÓN ILUSTRATIVA: ESTUDIAR CON UN ASISTENTE QUE PERMITE TERMINAR

Una institución considera ofrecer un asistente voluntario para preparar tutorías de una asignatura. El caso constituye una propuesta, sin datos de implementación. La función consiste en ayudar a identificar una dificultad, organizar preguntas y relacionarlas con materiales autorizados. El asistente puede aportar disponibilidad fuera del horario de tutoría, mientras el docente conserva evaluación y respuesta sobre problemas que requieren su intervención. La definición de finalidad permite reconocer qué información necesita la herramienta y qué resultados se evaluarán. Su presentación debe evitar una promesa de comprensión personal completa (Holmes et al., 2022; Miao & Holmes, 2023).

La actividad comienza con un problema académico y el trabajo ya realizado. El estudiante puede aportar un fragmento o describir dónde encuentra dificultad, sin entregar información íntima que la tarea no requiere. El asistente devuelve una interpretación de esa dificultad y permite corregirla. Esta fase reconoce que una respuesta puede captar de manera parcial el problema. Se propone evaluar si el estudiante identifica qué necesita revisar y si la interpretación resulta pertinente, además de su percepción de apoyo. La verificación de contenido se realiza con materiales de la asignatura y criterios docentes (Amershi et al., 2019; Cui, 2026; Lachheb et al., 2023).

La ayuda se organiza en pasos que pueden variar. El sistema puede solicitar un intento, ofrecer una pista o explicar una relación, según objetivos de la actividad. El estudiante conserva una modalidad alternativa cuando el procedimiento dificulta participación. Las investigaciones sobre tutoría y aprendizaje orientan esta configuración, pero sus resultados no acreditan eficacia del caso. Se propone observar comprensión con actividades posteriores y distinguir desempeño asistido de desempeño independiente. La facilidad de completar la sesión se registra como experiencia de uso, manteniendo separado el aprendizaje que se busca sostener (Bastani et al., 2025; Kestin et al., 2025; Ma et al., 2014; VanLehn, 2011).

El tono reconoce la dificultad y mantiene criterios. Una respuesta puede indicar que una pregunta es comprensible y pedir evidencia para una conclusión. Se revisan mensajes que confirman automáticamente una posición, prometen exclusividad o atribuyen necesidades emocionales al asistente. La complacencia se estudia por su conducta y consecuencias, sin clasificar toda cortesía como daño. La propuesta exige comprobar si las intervenciones ayudan a contrastar el contenido y a preparar una conversación académica. Se conserva la distinción entre reconocimiento de la persona y validación de todas sus afirmaciones (Amershi et al., 2019; M. Cheng et al., 2026; Cortina, 2024).

Al concluir, el estudiante obtiene un breve registro de dudas, fuentes y pasos siguientes. Puede terminar la interacción y recuperar ese material para la tutoría. El diseño evita convertir la continuidad de conversación en una condición para conservar una oportunidad académica. La evaluación observa si el cierre permite continuar la actividad y si aparecen obstáculos de acceso o recuperación. Un número menor de mensajes puede indicar que la necesidad se resolvió, mientras una sesión extensa requiere examinar su finalidad. Los indicadores se interpretan en relación con tarea y resultado, además de volumen (Chan & Hu, 2023; Kirk et al., 2025; Shneiderman, 2020).

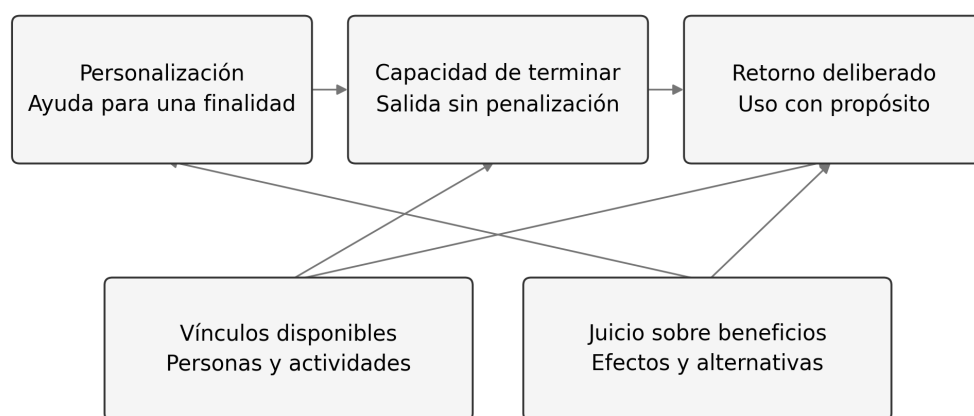
El canal humano requiere recursos concretos. La institución informa cómo solicitar tutoría, qué respuesta puede esperar y qué alternativa existe cuando un horario resulta inaccesible. Si aparece una necesidad que excede el apoyo académico, el servicio debe remitir a canales pertinentes dentro de sus capacidades y límites. La propuesta no convierte al asistente en atención clínica. La eficacia de derivación debe comprobarse mediante el recorrido institucional, sin recopilar conversaciones privadas para evaluar una hipótesis sobre bienestar. Cuando se estudien necesidades de salud se requiere competencia especializada y gobernanza específica (GFT-CELAM, 2026; Lachheb et al., 2023; WHO, 2021).

La revisión participativa puede incluir estudiantes y docentes con experiencias diversas de uso y de abstención. Interesa conocer barreras, utilidad, presión para continuar y condiciones de una modalidad alternativa. Una encuesta voluntaria puede aportar percepciones, mientras tareas académicas delimitadas aportan otra evidencia. Se propone evitar una clasificación de dependencia basada en frecuencia. El estudio cualitativo regional de Kornhaber y colaboradores orienta preguntas sobre condiciones de aprovechamiento, conservando su contexto australiano. La institución deberá producir información propia antes de atribuir a sus estudiantes los mismos resultados (Chan & Hu, 2023; Fang et al., 2025; Kornhaber et al., 2026).

Los reportes de incidencias se relacionan con decisiones. Un problema de contenido puede modificar materiales o instrucciones; una promesa relacional inapropiada puede exigir revisar mensajes; una barrera de acceso puede requerir otra modalidad. Se identifica quién tiene autoridad y recursos para realizar cada ajuste. La evaluación debe observar si la corrección ocurre y si resuelve el problema señalado. Esta cadena hace concreta la responsabilidad. El caso ilustra una configuración orientada a autonomía y cierre; afirmar que mejora bienestar o aprendizaje requeriría una implementación evaluada con población, periodo y comparación pertinentes (Ares-Ferreirós et al., 2026; Raji et al., 2020; Tabassi, 2023).

La Figura 9 interpreta la medida mediante una secuencia que conserva ayuda y autonomía. La personalización debe contribuir a una finalidad; el cierre permite recuperar otras actividades; el retorno conserva la herramienta como una opción deliberada. Los vínculos disponibles y el juicio sobre beneficios sostienen esa relación. Esta propuesta evita evaluar el servicio únicamente por permanencia o interacción continua y permite preguntar si el usuario puede terminar sin perder oportunidades necesarias para aprender o participar (Holmes et al., 2022; Long & Magerko, 2020; Miao & Holmes, 2023).

Figura 9. Personalización con oportunidades para terminar y volver deliberadamente



Nota. Elaboración conceptual sobre medida. Las relaciones son criterios propuestos para discutir diseño y uso; su eficacia requiere evaluación situada. Fuente: (Holmes et al., 2022; Long & Magerko, 2020; Miao & Holmes, 2023).

Tabla 5. Objetos distintos al evaluar compañía artificial

Objeto	Evidencia pertinente	Límite de interpretación
Experiencia inmediata	Medida posterior a la interacción	Alivio momentáneo no demuestra permanencia
Seguimiento	Observaciones durante un periodo definido	La duración restringe la conclusión
Intensidad voluntaria	Conducta de uso y condiciones previas	Asociación no equivale a efecto causal
Autonomía	Alternativas, corrección y cierre	Satisfacción no demuestra capacidad de salir

CONCLUSIONES

La fascinación por la IA merece comprensión y criterios de cuidado. Sus beneficios aparecen cuando disponibilidad y personalización amplían oportunidades de explorar, aprender y comunicarse. El exceso se produce cuando esas capacidades se utilizan para capturar atención, explotar intimidad o presentar una relación que el servicio no puede sostener. La medida exige evaluar finalidad, autonomía y posibilidad de terminar el uso.

La apariencia también interviene en el valor de lo que se produce. Una imagen, una voz o un texto pueden parecer propios y ocultar decisiones relevantes de autoría. El capítulo siguiente estudia la envidia como metáfora de imitación, suplantación y apariencia de competencia. La autenticidad se convierte allí en una condición de confianza entre personas que intercambian contenidos y atribuyen responsabilidades.

La síntesis identifica una relación entre cercanía, expectativas y capacidad de acción. La disponibilidad y la personalización pueden facilitar expresión o una tarea, mientras la promesa de reconocimiento puede extender la confianza más allá de la función demostrada. El problema se reconoce en mensajes, incentivos y opciones, además de frecuencia. Mantener estas dimensiones permite comprender la experiencia de la persona sin adjudicar reciprocidad emocional al sistema ni convertir una interacción afectiva en diagnóstico. La metáfora de la lujuria adquiere utilidad cuando conduce a revisar qué relación promueve el diseño (Epley et al., 2007; Kirk et al., 2025; Nass & Moon, 2000).

La evidencia exige conservar selección y horizonte temporal. Los relatos de amistad y daño, el alivio momentáneo de soledad y los resultados de seguimiento aportan contribuciones diferentes. Las versiones de prepublicación también pueden modificar una interpretación. Una política responsable reconoce beneficios y riesgos con su alcance, y evita deducir salud prolongada de satisfacción inmediata. La evaluación futura debe estudiar cómo cambian experiencias y opciones durante el uso, con métodos que permitan distinguir características previas, diseño e intensidad voluntaria cuando se formulan afirmaciones causales (De Freitas et al., 2026; Fang et al., 2025; Laestadius et al., 2024; Skjuve et al., 2021).

La medida se concreta en comprender una función, corregir una interpretación y terminar conservando el trabajo útil. Estas facultades requieren alternativas institucionales y recursos, junto con información comprensible. La aplicación de estudio propuesta muestra cómo organizar ayuda, cierre y respuesta a incidencias. Su eficacia permanece pendiente de evaluación. La personalización será defendible si facilita una actividad y conserva capacidad de contrastar, en vez de considerar toda preferencia inmediata como criterio suficiente. El diseño puede mantener apoyo respetuoso y exigencia de razones (Amershi et al., 2019; M. Cheng et al., 2026; Raji et al., 2020; Shneiderman, 2020).

Una agenda posterior debería comparar configuraciones de cierre, apoyo graduado y derivación en contextos diversos, conservando accesibilidad y privacidad. También conviene examinar qué indicadores de servicio favorecen resolución de necesidades y cuáles premian permanencia sin finalidad. El capítulo aporta categorías y condiciones discutibles, sin establecer un efecto general de compañía artificial. Su contribución consiste en relacionar promesa, función y responsabilidad, para que una experiencia valiosa pueda mantenerse bajo condiciones que permitan revisar sus límites y consecuencias (Holmes et al., 2022; Lachheb et al., 2023).

DECLARACIÓN DE USO DE INTELIGENCIA ARTIFICIAL

Se utilizaron herramientas de inteligencia artificial generativa como apoyo en los procesos de organización, redacción y revisión del manuscrito. Su utilización se limitó a funciones de asistencia editorial y analítica, bajo supervisión y validación permanente de los autores. Las propuestas, interpretaciones y escenarios desarrollados en el texto corresponden a elaboraciones

de carácter analítico y no constituyen resultados derivados de procesos de implementación empírica. Las herramientas de inteligencia artificial empleadas no reciben atribución de autoría ni sustituyen las fuentes científicas, documentales y académicas citadas. La selección de las fuentes, la interpretación de la evidencia, la construcción argumentativa, la revisión final del contenido y la decisión de someter el manuscrito a publicación son responsabilidad exclusiva de los autores.

REFERENCIAS

- Alier-Forment, M., Casañ-Guerrero, M. J., Pereira-Varela, J. A., García-Peñalvo, F. J., & Llorens-Largo, F. (2026). Inteligencia artificial generativa y autonomía educativa: Metáforas históricas y principios éticos para la transformación pedagógica. *RIED Revista Iberoamericana de Educación a Distancia*, 29(1), 9–28. <https://doi.org/10.5944/ried.45536>
- Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. En *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Association for Computing Machinery. <https://doi.org/10.1145/3290605.3300233>
- Ares-Ferreirós, M., de la Cruz-Campos, J. C., Morodo Horrillo, A., & Robert Guillén, S. (2026). Buenas prácticas en el uso de la IA generativa en las universidades. Un análisis documental. *Innoeduca International Journal of Technology and Educational Innovation*, 12(1), 5–20. <https://doi.org/10.24310/ijtei.121.2026.22152>
- Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P., & Roberts, K. (2024). *Artificial intelligence risk management framework: Generative artificial intelligence profile* (NIST AI 600-1). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.600-1>
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakci, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26), Article e2422633122. <https://doi.org/10.1073/pnas.2422633122>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? En *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>
- Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. En *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.463>
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444. <https://doi.org/10.1146/annurev-psych-113011-143823>
- Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), Article 188. <https://doi.org/10.1145/3449287>
- Burton, J. W., Stein, M.-K., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*, 33(2), 220–239. <https://doi.org/10.1002/bdm.2155>
- Cabero-Almenara, J., Barroso-Osuna, J., Loaiza Aguirre, M. I., & Palacios-Rodríguez, A. (2026). Artificial intelligence in higher education: Cognitive, emotional, behavioral and ethical attitudes of Latin American students. *Journal of New Approaches in Educational Research*, 15, Article 13. <https://doi.org/10.1007/s44322-026-00063-2>
- Cai, F., Bolisani, E., Nakash, M., & Kassaneh, T. C. (2026). AI-driven knowledge management for sustainable businesses: A comprehensive analysis. *VINE Journal of Information and Knowledge Management Systems*. Publicaciónanticipadaenlínea. <https://doi.org/10.1108/VJIKMS-09-2025-0444>
- Carlini, N., Tramèr, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K., Roberts, A., Brown, T., Song, D., Erlingsson, Ú., Oprea, A., & Raffel, C. (2021). Extracting training data from large language models. En *30th USENIX Security Symposium* (pp. 2633–2650). USENIX Association. <https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20, 43. <https://doi.org/10.1186/s41239-023-00411-8>
- Chatterji, A., Cunningham, T., Deming, D. J., Hitzig, Z., Ong, C., Shan, C. Y., & Wadman, K. (2025). *How people use ChatGPT* (Working Paper No. 34255). National Bureau of Economic Research. <https://doi.org/10.3386/w34255>
- Cheng, M., Lee, C., Khadpe, P., Yu, S., Han, D., & Jurafsky, D. (2026). Sycophantic AI decreases prosocial intentions and promotes dependence. *Science*, 391(6792), Article eaec8352. <https://doi.org/10.1126/science.aec8352>
- Cheng, Q., Dai, Y., Liu, X., & Peng, S. (2026). The trust crisis in artificial intelligence: AI hallucinations and human-AI collaboration. *Technology in Society*, 86, Article 103286. <https://doi.org/10.1016/j.techsoc.2026.103286>
- Cimino, A., Corvello, V., Longo, F., Solina, V., & Thomas, A. (2026). Humanizing artificial intelligence: How employee appropriation of technology can improve well-being. *Technological Forecasting and Social Change*, 230, 124737. <https://doi.org/10.1016/j.techfore.2026.124737>
- Cortina, A. (2024). *¿Ética o ideología de la inteligencia artificial? El eclipse de la razón comunicativa en una sociedad tecnologizada*. Paidós.
- Croes, E. A. J., & Antheunis, M. L. (2021). Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot. *Journal of Social and Personal Relationships*, 38(1), 279–300. <https://doi.org/10.1177/0265407520959463>
- Cui, Y. (2026). Artificial intelligence and human assessment consistency and bias in higher education academic writing evaluation. *Discover Artificial Intelligence*, 6, 559. <https://doi.org/10.1007/s44163-026-01286-0>
- De Freitas, J., Oğuz-Uğuralp, Z., Uğuralp, A. K., & Puntoni, S. (2026). AI companions reduce loneliness. *Journal of Consumer Research*, 52(6), 1126–1148. <https://doi.org/10.1093/jcr/ucaf040>
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms

- after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126. <https://doi.org/10.1037/xge0000033>
- Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28), Article eadn5290. <https://doi.org/10.1126/sciadv.adn5290>
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4–58. <https://doi.org/10.1177/1529100612453266>
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886. <https://doi.org/10.1037/0033-295X.114.4.864>
- Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., Maes, P., Phang, J., Lampe, M., Ahmad, L., & Agarwal, S. (2025). *How AI and human behaviors shape psychosocial effects of extended chatbot use: A longitudinal randomized controlled study* [Prepublicación, versión 2]. arXiv. <https://arxiv.org/abs/2503.17473v2>
- Farquhar, S., Kossen, J., Kuhn, L., & Gal, Y. (2024). Detecting hallucinations in large language models using semantic entropy. *Nature*, 630, 625–630. <https://doi.org/10.1038/s41586-024-07421-0>
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Fox, G., & Solihin, Y. (2026). Exploring organizational implementation of generative AI: Navigating benefits and safety risks. *Journal of Decision Systems*, 35(1), 2649940. <https://doi.org/10.1080/12460125.2026.2649940>
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198237907.001.0001>
- Geburu, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., III, & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92. <https://doi.org/10.1145/3458723>
- Grupo de Trabajo de Frontera Tecnológica del Celam. (2026). *Inteligencia artificial y desarrollo humano integral*. Consejo Episcopal Latinoamericano y Caribeño. <https://repositorio.celam.org/s/repositorio/item/3387>
- Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication*, 68(4), 712–733. <https://doi.org/10.1093/joc/jqy026>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & 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>
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), Article 248. <https://doi.org/10.1145/3571730>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Kestin, G., Miller, K., Klaes, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, 15, Article 17458. <https://doi.org/10.1038/s41598-025-97652-6>
- Kirk, H. R., Gabriel, I., Summerfield, C., Vidgen, B., & Hale, S. A. (2025). *Why human-AI relationships need socioaffective alignment* [Prepublicación]. arXiv. <https://doi.org/10.48550/arXiv.2502.02528>
- Kornhaber, R., Bragg, S., Matheson, A., & Irwin, P. (2026). Qualitative inquiry into generative artificial intelligence impact on rural and regional higher education. *Australian and International Journal of Rural Education*, 36(2), 1–23. <https://doi.org/10.47381/ajire.v36i2.812>
- Lachheb, A., Abramienka-Lachheb, V., Moore, S., & Gray, C. (2023). The role of design ethics in maintaining students' privacy: A call to action to learning designers in higher education. *British Journal of Educational Technology*, 54, 1653–1670. <https://doi.org/10.1111/bjet.13382>
- Laestadius, L., Bishop, A., Gonzalez, M., Illeňčík, D., & Campos-Castillo, C. (2024). Too human and not human enough: A grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika. *New Media & Society*, 26(10), 5923–5941. <https://doi.org/10.1177/14614448221142007>
- Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. En *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, Article 1121 (pp. 1–22). Association for Computing Machinery. <https://doi.org/10.1145/3706598.3713778>
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80. <https://doi.org/10.1518/hfes.46.1.50.30392>
- Lin, S., Hilton, J., & Evans, O. (2022). TruthfulQA: Measuring how models mimic human falsehoods. En *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 3214–3252). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.acl-long.229>
- Litvinenko, A. (2026). The knowledge collaboration framework: A review of artificial intelligence in organisational learning and knowledge management. *Journal of Decision Systems*, 35(1), 2616703. <https://doi.org/10.1080/12460125.2026.2616703>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. En *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376727>
- Luccioni, A. S., Jernite, Y., & Strubell, E. (2024). Power hungry processing: Watts driving the cost of AI deployment? En *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 85–99). Association for Computing Machinery. <https://doi.org/10.1145/3630106.3658542>

- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901–918. <https://doi.org/10.1037/a0037123>
- Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. En *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 1906–1919). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.173>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. En *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220–229). Association for Computing Machinery. <https://doi.org/10.1145/3287560.3287596>
- Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1, 501–507. <https://doi.org/10.1038/s42256-019-0114-4>
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, Article 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192. <https://doi.org/10.1126/science.adh2586>
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230–253. <https://doi.org/10.1518/001872097778543886>
- Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. En *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 33–44). Association for Computing Machinery. <https://doi.org/10.1145/3351095.3372873>
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>
- Russell, S. J., & Norvig, P. (2004). *Inteligencia artificial. Un enfoque moderno* (J. M. Corchado Rodríguez, F. Martín Rubio, J. M. Cadenas Figueredo, L. D. Hernández Molinero, E. Paniagua Arís, R. FuentetajaPinzán, M. Robledo de los Santos, & R. Rizo Aldegue, trans.; 2.ª ed.). Pearson Educación.
- Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54–63. <https://doi.org/10.1145/3381831>
- Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My chatbot companion—A study of human-chatbot relationships. *International Journal of Human-Computer Studies*, 149, Article 102601. <https://doi.org/10.1016/j.ijhcs.2021.102601>
- Ta, V., Griffith, C., Boatfield, C., Wang, X., Civitello, M., Bader, H., DeCero, E., & Loggarakis, A. (2020). User experiences of social support from companion chatbots in everyday contexts: Thematic analysis. *Journal of Medical Internet Research*, 22(3), Article e16235. <https://doi.org/10.2196/16235>
- Tabassi, E. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.100-1>
- Tanveer, M. (2026). Artificial intelligence literacy moderating the effects of artificial intelligence use on cognitive outcomes including memory retention decision making and cognitive laziness. *Discover Psychology*, 6, 182. <https://doi.org/10.1007/s44202-026-00697-5>
- UNESCO. (2021). *Recomendación sobre la ética de la inteligencia artificial*. <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>
- Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L., Hendricks, L. A., ... Gabriel, I. (2021). *Ethical and social risks of harm from language models* [Prepublicación]. arXiv. <https://doi.org/10.48550/arXiv.2112.04359>
- World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. <https://www.who.int/publications/i/item/9789240029200>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators? *International Journal of Educational Technology in Higher Education*, 16, Article 39. <https://doi.org/10.1186/s41239-019-0171-0>