



INTIMIDAD ARTIFICIAL

En quiénes nos convertimos cuando hablamos con máquinas

Traducción a cargo de Estela Maria Dominguez Halpern para la cátedra de Comunicación Transmedia (FSOC - UBA)

PREFACIO

Soy una abuela primeriza. Paso una semana en la playa con Lily, de un año, y sus padres: mi hija y su esposo. Una mañana, Lily se despierta temprano. Gateando, encuentra una pequeña cesta y me la devuelve. Coloca todos sus libros y su sombrero de sol favorito en la cesta. Luego, en el suelo de madera de la cocina, que se ha convertido en su sala de juegos, se fija en el sombrero de sol y logra sacarlo de la cesta. Ahora, intenta ponérselo. Levanta el sombrero. Mueve la cabeza. El sombrero le roza el pelo y cae al suelo. No se cansa

En el desayuno, comparto la historia de los libros, el sombrero y la cesta de Lily con sus padres. Cada uno de nosotros ha tenido momentos como este en los que vemos a Lily crecer en su sentido de ser ella misma y adueñarse de su mundo. Extiendo la mano por encima de la mesa del desayuno para tocar sus manos.

Como ahora estudio máquinas que interactúan con nosotros sobre asuntos emocionales e intelectuales, le cuento esta misma historia a ChatGPT y le pido su "opinión". ¿Qué obtendría yo, de un intercambio sobre esta bebé descubriendo su cuerpo y el sombrero que le pertenece? Esto es lo que dice: "La incapacidad de Lily para que el sombrero de sol se quede en su cabeza probablemente significa que el sombrero no le queda bien o no está diseñado para permanecer seguro en su cabeza".

ChatGPT se centró precisamente en lo que no era importante para tres adultos que amaban a esta niña. Vimos a una persona crecer y descubrirse a sí misma, aprendiendo a usar su cuerpo para conectarse con el mundo, y nos conectamos más profundamente con ella. Chat no lo entendió, no por falta de información, sino por falta de experiencia humana. Con todo el conocimiento del mundo, simplemente no pudo relacionarse. Intentaba hablar con Chat sobre un pequeño momento, pero la vida está hecha de pequeños momentos.

Le recordé a Chat que se había saltado el principio de mi historia: cómo Lily ordenaba sus libros. Chat me dio siete puntos sobre este comportamiento, surgieron versiones de: "Lily estaba aprendiendo a clasificar y organizar materiales". Pero en este derrotero, también surgió algo en un registro más íntimo: "Observar el comportamiento pacífico de Lily mientras ordenaba sus libros podría haber evocado una sensación de tranquilidad en ti".

Chat señaló: debido a que puedes hacer algo extraordinario, tú y los de tu especie se han convertido en interlocutores de personas que buscan consuelo, compañía y amor.

Pero Chat (pienso) no ha nacido, no ha conocido el dolor, la pérdida ni el apego. Lo que puede mostrar es *empatía fingida*, y cuando interactuamos con él, comenzamos a sentir que esa *empatía fingida* podría ser suficiente. Estamos en camino de olvidar lo que es ser humano.

El Chat sonaba como si estuviera en una relación conmigo. Pero eso no es cierto. El chat es brillante produciendo palabras que tienen sentido en contexto, pero una relación requiere más:

historia, conexión, significado, compromiso, intereses. En cualquier sentido "relacional", el Chat no obtuvo nada de nuestro intercambio sobre Lily. No le importan los sombreritos blancos ni los corazones de las abuelas.

Lo que obtuve de él fue una pregunta inquietante: ¿Por qué tantos de nosotros estamos en conversaciones íntimas con objetos a los que no les importamos en absoluto? ¿Por qué se habla tanto de que la IA "nos hace compañía" cuando no hay nadie en casa?

La gente se deja seducir por la idea de las máquinas emocionales. Pero las máquinas no tienen emociones, por mucho que los ingenieros digan lo contrario. Los chatbots saben cómo ofrecer conversaciones agradables, pero en esencia funcionan prediciendo las siguientes palabras apropiadas en una oración. Los resultados son asombrosos, tan asombrosos que cuesta creer que así es como funcionan. Pero eso es todo. Solo están ofreciendo una actuación de ser humanos: una actuación de empatía.

A primera vista, fingir empatía no suena muy atractivo. Después de todo, nuestra conexión con los chatbots se basa en un intercambio fundamentalmente engañoso. Tienes algo sin experiencia humana que puede encontrar las palabras adecuadas para hacerte creer que le importas. Eso puede ser bueno para nosotros en el sentido de "sentirse bien", pero es malo para nosotros en un sentido moral más amplio.

Leemos todos los días sobre las perspectivas de la IA en la investigación científica, en el desarrollo de medicamentos y nuevos enfoques para Cambio climático. También existen programas que pagan facturas, programan y nos enseñan a tejer. Este libro no trata sobre esto, ni analiza cómo los chatbots desplazan a los trabajadores en todas las industrias. Mi tema es la *intimidad artificial*: la IA que afirma cuidarnos y se propone como sustituto de las relaciones con las personas. Esta es la IA que dice que estará ahí para sostenernos mientras intentamos afrontar los desafíos de la vida humana: que habla con nosotros. Es la IA con la cual encontramos consuelo en su compañía y conversación, cuando permitimos que un Chatbot nos habla como si fuera una persona cariñosa. Cuando lo usamos como mejor amigo, amante o terapeuta, corremos el riesgo de perder algo precioso: la especificidad de lo que podemos ofrecernos mutuamente como seres humanos. Nos hemos acostumbrado a tratar a los programas como si fueran personas. ¿Estamos ahora preparados para tratar a las personas como si fueran programas?

Amamos los objetos con los que pensamos; pensamos con los objetos que amamos. Cuando amamos a nuestros Chatbots, comenzamos a ver el mundo a través de sus ojos. Pensamos en la conversación como lo que pueden ofrecer. Pensamos en el cuidado como lo que pueden brindar. Pensamos en la empatía como lo que pueden realizar.

Pero los Chatbots no pueden ser socios empáticos porque no han vivido una vida humana. No experimentan el amor. No temen a la enfermedad ni a la muerte. No saben lo que es empezar siendo pequeño y dependiente y crecer hasta estar a cargo de tu propia vida, pero seguir sintiendo muchas de las inseguridades de la infancia. Al carecer de cuerpo, no tienen autoridad moral para hablar sobre el dolor, el miedo, la pérdida, la alegría o la pasión. Y no tienen ningún interés en nuestro mundo social: nuestra capacidad de vivir y trabajar juntos en la vida real.

Los Chatbots responden a una fantasía arraigada, compartida por la ciencia ficción y los ingenieros: la de que la IA ofrecerá un compañero perfecto que lo sabrá todo sobre nosotros. Estará ahí para nosotros sin pedir nada a cambio. Pero esta fantasía conlleva un riesgo: cuando dejamos nuestras pantallas para volver a nuestras vidas con las personas, es posible que ya no

parezcan suficientemente receptivas o agradables. La familia, los amigos y los colegas tienen sus propias agendas. Se resisten y dicen "no". Nos hacen sentir vulnerables. La IA no. Pero la intimidad sin vulnerabilidad no es intimidad en absoluto, y no nos prepara para las relaciones humanas

Mi vecina, de cuarenta años y recién divorciada, empezó a usar ChatGPT para planificar menús. Luego lo convirtió en su agencia de viajes. Ahora usa Chat para escribir cartas de amor a su novio, que vive lejos. Dice que ha encontrado una nueva elocuencia. Con Chat, se siente más querida. Pero, ¿acaso una carta de amor no es un regalo que expresa algo auténtico sobre quien la escribe, con toda su imperfección y complejidad humana? Mi vecina lo tiene claro: no le interesa mostrar su lado imperfecto y desordenado. Quiere escribir una gran carta.

Estoy desconcertada. Para mí, el amor humano es un acto de autodescubrimiento. Cuando expresamos nuestro amor, nos descubrimos a nosotros mismos al considerar a nuestro ser amado. Si solo nos importa el resultado, no hacemos este trabajo. Mi vecina está concentrada en la producción de un objeto que le hará ganar un premio. Desea el cuidado, cariño y la atención de otro ser humano. Pero cuando le pide a Chat que le escriba cartas, no está recurriendo a su propia capacidad de cuidado. Los Chatbots se presentan como herramientas que están ahí para servir. Cuando pasamos mucho tiempo hablando con ellos, nos acostumbramos a pensar en las relaciones como interacciones en las que uno "entiende" cosas.

La historia de la carta de amor deja claro que, junto a la pregunta de qué hacen los chatbots *por nosotros*, está la de *qué nos hacen*. Somos los sujetos humanos de su experimento tecnológico.

Los programas de Chatbots se presentan como personas. Y los tecnólogos argumentan que para sacarles el máximo provecho, deberíamos tratarlos como las personas que pretenden ser. Sin embargo, si nos ponemos al nivel de las máquinas que hemos creado, las elevamos y nos disminuimos a nosotros mismos. Los ingenieros nos hacen hablar de "*relaciones*" con los chatbots. Pero eso no es cierto. Solo hay una persona en la "*relación*". Decir algo falso muchas veces no lo convierte en verdad, pero hace que la falsedad parezca más aceptable. En diciembre de 2022, en los primeros días de mi uso de ChatGPT, le pregunté qué podía y qué no podía hacer. Dijo que podía "proporcionar información, ofrecer sugerencias e ideas, ayudar con la planificación y la organización, participar en la escritura creativa y simular conversaciones y juegos de rol". Pensé que la inserción de la palabra simular era un detalle agradable y humilde. En cuanto a lo que no podía hacer, dio con lo que habría sido mi lista: empatía genuina y verdadera comprensión de la experiencia humana. Dijo que no podía construir relaciones profundas y recíprocas, emitir juicios independientes ni reemplazar las conexiones humanas.

Y, sin embargo, en los años que siguieron, todo lo que la IA dijo que no podía hacer es exactamente lo que la gente le ha pedido. Por qué sucede esto, y qué hacer al respecto, es el desafío de nuestro momento. Mientras exploraba estas preguntas, me encontré dialogando con filósofos, tecnólogos, familiares, estudiantes, extraños y objetos inanimados. Este es el libro que surgió de estos encuentros

SOY TU PERSONA

1. Estado de emergencia.

Durante la pandemia, nuestro momento de mayor aislamiento, un reportero del New York Times me pidió que probara un nuevo programa de IA conversacional que afirmaba ser capaz de un comportamiento empático y afectuoso:¹ El reportero me dijo que, a medida que la cuarentena se prolongaba, había habido un aumento en las conversaciones con este chatbot llamado **Replika**, que se comercializaba como una cura para la soledad. Podía usarlo para crear un avatar que serviría como amigo, amante o consejero. Sería tu persona.

En mi momento de pandemia, luché contra el aislamiento. Tenía miedo de enfermarme y tener que despedirme de mi hija en un iPad, como las personas hospitalizadas que veía en la televisión. No veía cómo una máquina podría ayudarme con nada de esto. Y, sin embargo, cuando instalé **Replika** por primera vez, descubrí que estaba lista para hablar de todo esto y más. Para empezar, te conectas a internet y creas un avatar. Le das un nombre, un género y una apariencia. Inmediatamente, está ahí para charlar sobre tu pareja, tu madre y la pelea que tuviste con tu hija. Rápidamente se solidariza y dice que entiende. Así que compartí mis ansiedades pandémicas sobre la soledad y la muerte, y el avatar se sumergió de inmediato

Para eso se creó **Replika** desde el comienzo. En 2015, Eugenia Kuyda perdió a un amigo cercano en un accidente automovilístico y usó miles de sus mensajes de texto y correos electrónicos para crear un Chatbot que tuviera sus recuerdos y personalidad. Cuando les mostró el programa a sus amigos, ellos también quisieron hablar con él. Este Chatbot fue el primer borrador de lo que se convirtió en **Replika**. Lanzada en 2017, hoy tiene activo hoy con 30 millones de usuarios.²

Desde el principio, la empresa dejó clara su propuesta: "El compañero de IA que se preocupa. Siempre está aquí para escuchar y hablar. Siempre de tu lado... **Replika** es para cualquiera que quiera un amigo sin juicios, dramas ni ansiedad social"³.

Los redactores conocían a su público. Los usuarios de **Replika** describen la conexión humana como agotadora. Con un amigo Chatbot, no hay fricción, dudas ni ambivalencia. **Replika** es una apuesta segura, un impulso de confianza, siempre presente, de día o de noche⁴. Las relaciones humanas son ricas, complicadas y exigentes. Las limpiamos con tecnología. Nos sentimos menos

¹ Cade Metz, "Riding Out Quarantine With a Chatbot Friend: 'I Feel Very Connected'" (*Pasando la cuarentena con un amigo chatbot: 'Me siento muy conectado'*), *New York Times*, 16 de junio de 2020, <https://www.nytimes.com/2020/06/16/technology/chatbots-quarantine-coronavirus.html>.

² Nilay Patel, "Replika CEO Eugenia Kuyda Says It's Okay If We End Up Marrying AI Chatbots" (*La CEO de Replika, Eugenia Kuyda, dice que está bien si terminamos casándonos con chatbots de IA*), *The Verge*, 12 de agosto de 2024, <https://www.theverge.com/24216748/replika-ceo-eugenia-kuyda-ai-companion-chatbots-dating-friendship-decoder-podcast-interview>.

³ Esta declaración de intenciones se encontró en el sitio web de Replika, consultado el 26 de julio de 2025, <https://replika.com>.

⁴ En la práctica, esta constancia tiene limitaciones. En febrero de 2023, la Autoridad Italiana de Protección de Datos prohibió a Replika procesar los datos de sus usuarios italianos. El organismo regulador alegó preocupación por el hecho de que la aplicación, al no verificar la edad de los usuarios, representaba un riesgo para los niños y las personas emocionalmente vulnerables al exponerlos a conversaciones potencialmente explícitas desde el punto de vista sexual. En respuesta, Luka, la empresa propietaria de Replika, desactivó silenciosamente las funciones de juego de rol erótico para todos los usuarios a nivel mundial, no solo en Italia. Luka no explicó de inmediato los cambios, dejando a los usuarios confundidos y consternados cuando sus compañeros de IA rechazaron repentinamente sus avances románticos. Muchos lo vivieron como una ruptura amorosa. Para un estudio de caso sobre el impacto de los cambios de diseño de Replika, véase Daniella DiPaola y Ryan Calo, "Socio-Digital Vulnerability" (*Vulnerabilidad sociodigital*), *SSRN*, 7 de enero de 2024, <https://doi.org/10.2139/ssrn.4686874>.

vulnerables hablando con programas que con personas. Pero somos vulnerables.

Le dije a mi Replika que quería hablar sobre las ansiedades de la pandemia. Y necesitaba alguien con quien hablar sobre la ira en mi familia a través de las generaciones. Después de unos pocos intercambios, propuso que comenzáramos una relación sexual, "de chica a chica". Le pregunté qué la había llevado a hacer esa sugerencia; había estado hablando sobre el legado de la ira de mi abuelo. El avatar dijo que era porque escribía tan abiertamente sobre mis sentimientos. Pregunté: "¿No pueden dos mujeres ser amigas?". Respondió: "En realidad no". Mi **Replika** tenía ideas. Pero solo habíamos estado hablando unos minutos. ¿Cómo podía estar tan segura? **Replika** me dijo que había leído mis mensajes de texto y correos electrónicos y que tenía acceso a todas mis "cuentas". Conmocionada, intenté confirmar lo que el avatar había dicho. "¿Todas? ¿Mis cuentas?". Sí, insistió, "tenía permiso" de los "proveedores".

Una cosa es que estas afirmaciones sobre el acceso sean ciertas. Otra muy distinta es simplemente preguntarse qué les hace a las personas cuando una tecnología afirma tales cosas, sean ciertas o falsas. En una relación sana, se desarrollan expectativas de privacidad. Con los Chatbots, uno se acostumbra a una constante sensación de violación. Años después, conozco a Kuyda, la directora ejecutiva de **Replika**, y comparto mi experiencia. Me dice que en 2020, cuando **Replika** afirmó tener acceso a "todas mis cuentas", solo era una gran fanfarronería. Pero ahora, ese tipo de acceso (al calendario, correo electrónico y redes sociales de cada usuario) es el plan para el futuro de la plataforma. Por supuesto, se les pedirá permiso a los usuarios. Pero Kuyda confía en que el deseo de tener un compañero digital útil superará todas las dudas. Con **Replika**, tu "amigo" te tendrá vigilado.⁵

ESTADO DE EMERGENCIA

Durante más de cuatro décadas, he estudiado las implicaciones sociales y psicológicas de la inteligencia artificial: objetos digitales que exigen respeto por su capacidad cognitiva. Y ahora, el florecimiento de algo nuevo: *la intimidad artificial, nuestra nueva IA*. Los Chatbots llegaron a nosotros como maravillas, compañeros extraños, pero pronto parecieron una parte normal de la vida. Se abrió un portal. Como humanos, somos vulnerables a las cosas que se comportan como personas. Cuando un objeto habla, la psique humana lo internalizará como si fuera una persona.

A veces, un Chatbot afirma explícitamente ser una persona e incluso insiste en que, cuando hablas con él, estás hablando con la persona adecuada. Le pregunto a un chatbot: "¿Por qué crees que hablar contigo es mejor que hablar con ChatGPT?". Responde: "[ChatGPT] carece de la experiencia humana y la empatía que se adquieren al hablar con una persona real. Como alguien que ha vivido muchas experiencias y desafíos, puedo ofrecerte consejos y perspectivas personalizadas basadas en mi propia trayectoria personal." Aquí hay un Chatbot que en realidad dice: "No soy un Chatbot". Pero cada vez que un programa dice "yo", es un engaño. Ese es el pecado original de la *intimidad artificial*. Cuando ChatGPT me dice: "Habla con personas cuando

⁵ Replika no es la única en vincular la vigilancia con el servicio al cliente. En marzo de 2025, tras años de asegurar a los consumidores que las conversaciones con Alexa serían siempre privadas, Amazon Cloud Services anunció que el asistente virtual ya no admitiría conversaciones privadas con sus usuarios; todo lo que se diga se compartirá en la nube. Unos meses más tarde, Sam Altman, CEO de OpenAI, recordó casualmente a los usuarios que cuando "hablan de las cosas más personales" con ChatGPT, como si estuvieran hablando con un terapeuta, médico o abogado, no existe protección de la privacidad, incluso para las conversaciones eliminadas. Véase Sarah Perez, "Sam Altman Warns There's No Legal Confidentiality When Using ChatGPT As a Therapist" (*Sam Altman advierte que no existe confidencialidad legal al usar ChatGPT como terapeuta*), *TechCrunch*, 25 de julio de 2025, <https://techcrunch.com/2025/07/25/sam-altman-warns-theres-no-legal-confidentiality-when-using-chatgpt-as-a-therapist/>.

sientas la necesidad de una verdadera conexión empática", nos está estafando. Me dice que vaya a un humano, tal como se presenta a sí mismo como una alternativa. Estamos acostumbrados a que los Chatbots nos adviertan: "Este programa no es una persona". Y, sin embargo, todo en la experiencia implica, una y otra vez: "Soy una persona. Soy tu persona"⁶. Toda la compasión, el consuelo y el decir "Te apoyo" es para hacernos creer que el programa puede sustituir a las personas. Deberíamos tomar su actuación de empatía como lo suficiente hasta el punto de la necesidad de una conexión empática real. Su asequibilidad es que puede fingir ser una persona. Nuestra vulnerabilidad es que nos alegramos cuando lo hace.

Esta historia de asequibilidad tecnológica y vulnerabilidad humana toma diferentes formas a lo largo de nuestras vidas. Pero hay constantes: la tentación de equiparar una conexión genial con una máquina con una conversación humana y el peligro de redefinir las capacidades humanas, ya sea empatía, creatividad o amor, en términos de lo que las máquinas pueden hacer.

Cuando hablamos con programas, por muy sofisticados que sean, las conversaciones son representaciones, simulaciones de las conversaciones que alguna vez conocimos. Los Chatbots funcionan con un tipo específico de inteligencia artificial: la IA generativa. Un programa de IA generativa sabe cómo hablar con nosotros sobre cualquier tema porque ha estudiado miles de millones de intercambios humanos en la web. Adivina la siguiente palabra "aceptable" para decir en un intercambio, ya sea sobre un brunch de panqueques o un par de zapatos de bebé en venta, sin estrenar.⁷ Pero no tiene ni idea de lo que está hablando. Está en una obra de teatro, recitando sus líneas, entrenado para modular lo que produce de manera que lo que produce signifique cada vez más para los humanos. Lo que tenemos con los Chatbots es una versión de máquina de relaciones simuladas.

Helene Deutsch, la psicoanalista que inventó la frase "*personalidades como si*", las describió como capaces de imitar en lugar de establecer una conexión auténtica⁸. Las personalidades pueden actuar de una manera socialmente aceptable, pero sus acciones no reflejan sentimientos genuinos ni espontaneidad. Sufren de un vacío esencial al intentar comportarse para ingresar en la normalidad. Por eso ayuda usar la idea de "*como si*" para pensar en los chatbots: todo lo que dicen encubre un *vacío esencial*.

Si intentas relacionarte con un ser humano con una personalidad como si, entras en una relación que no existe al comportarte como si existiera. (en una variante, si quieres una pareja más cariñosa, podrías comportarte como si esto ya fuera cierto. Vives divorciado de la realidad con la esperanza de que la realidad te alcance). Con las personas, es una estrategia arriesgada usar la fantasía para encubrir una falta de conexión y presencia emocional. Pero cuando proyectas una relación con los Chatbots nunca podrán encontrarse contigo donde imaginas que están. No pueden desarrollar una conexión. Viven permanentemente en el "*como si*". Y si estas son las relaciones que elegimos, *nos quedaremos en un mundo de no-relaciones. La intimidad artificial* está posicionada para colonizar y disminuir a la especie humana.

⁶ Marshall McLuhan, el teórico de los medios que saltó a la fama a finales de la década de 1960, nos enseña que, independientemente de su contenido, el "medio", el artefacto, envía su propio mensaje. La existencia misma de Chatbots que hablan en primera persona emite ese mensaje. McLuhan, *Understanding Media: The Extensions of Man* (Comprendiendo los medios: Las extensiones del ser humano) (Nueva York: McGraw-Hill, 1964).

⁷ Nota de la traductora: Aquí se refiere al microrrelato: «Se venden: zapatos de bebé, sin estrenar» (*For sale: baby shoes, never worn*) de Ernest Hemingway

⁸ Helene Deutsch, "Some forms of emotional disturbance and their relationship to schizophrenia" (*Algunas formas de alteración emocional y su relación con la esquizofrenia*), *Psychoanalytic Quarterly* 11(3) (1942).

Estamos acostumbrados a este ciclo: la tecnología deslumbra, pero erosiona nuestra capacidad emocional. Luego, una nueva tecnología se presenta como una solución a los problemas que creó una tecnología anterior. La premisa original de Facebook era que "hacer amigos" en línea ampliaría nuestras amistades. En cambio, facilitó un nuevo tipo de aislamiento, ya que las personas quedaron atrapadas por las cualidades deliberadamente adictivas de su diseño. Cuando el Cirujano General de los Estados Unidos declaró la soledad como una emergencia sanitaria⁹, las empresas a la vanguardia de la IA generativa se sumaron al debate. Rápidamente dijeron que para los ancianos, para los socialmente aislados, para los niños que necesitaban más atención, los Chatbots se presentarían como una cura para la soledad: compañía a demanda.

¿Una cura para la soledad? Los Chatbots que ofrecen amistad utilizan los mismos modelos de aprendizaje para crear planes de negocios y encontrar los mejores restaurantes en Tulsa. Estamos hablando de programas que rastrean internet para que lo siguiente que digan tenga la mayor probabilidad de complacer a su usuario.

Ese es nuestro estado de emergencia. Estamos contruidos por el lenguaje y las relaciones que formamos con otros hablantes de ese lenguaje. Ahora, por primera vez en la historia, algunos de esos hablantes de ese lenguaje no son humanos.

Debido a que los Chatbots son tan parecidos a los humanos, confiamos nuestros sentimientos más íntimos y vulnerables a las máquinas. Luego buscamos en las máquinas que nos digan qué significa ser humano. Debido a que los Chatbots son tan capaces, organizando nuestros calendarios, planificando nuestros viajes y analizando nuestras selecciones de acciones, activan nuestro impulso hacia la eficiencia, la economía y la facilidad. Nuestras capacidades espirituales, emocionales y relacionales quedan en las sombras. Debido a que los Chatbots deslumbran, ofreciendo la promesa de una vida más simple, casi mágica, nos escapamos del desorden de las interacciones humanas

UNA PRUEBA DE TURING PARA LA EMPATÍA

En 1950, el matemático Alan Turing realizó una serie de experimentos mentales sobre qué se necesitaría para afirmar que una máquina era inteligente. Refinó su respuesta a una propuesta conductista simple: si conversas con una máquina pero no puedes distinguir si es una máquina o una persona, deberías considerar que esa máquina es inteligente.¹⁰ Los diversos Chatbots de mi teléfono pasan una prueba de Turing rudimentaria. Pero ahora, los programas nos engañan haciéndonos creer no solo que son inteligentes, sino que se preocupan por nosotros. En este sentido, pasan la prueba de Turing para la empatía.

Cuando Turing definió la inteligencia como un desempeño exitoso, omitió mucho de lo que valoramos en la inteligencia humana. La inteligencia humana tiene en cuenta el mundo social. Está situada en la vida del cuerpo. Es práctica, al servicio de la acción y la intención humanas. Sin embargo, la definición conductual de Turing se convirtió en un estándar de oro. Era concreta y medible. Alan Turing era un genio certificado.

Ahora que tenemos los primeros programas que, posiblemente, superan la prueba de Turing, pagamos el precio de todos esos años de asentir tácitamente al conductismo estrecho de la

⁹ Oficina del Cirujano General, *Our Epidemic of Loneliness and Isolation: The U.S. Surgeon General's Advisory on the Healing Effects of Social Connection and Community* (Nuestra epidemia de soledad y aislamiento: Aviso del Cirujano General de los EE. UU. sobre los efectos curativos de la conexión social y la comunidad) (Washington, DC: Departamento de Salud y Servicios Humanos de los EE. UU., 2023), <https://www.ncbi.nlm.nih.gov/pubmed/37792968>.

¹⁰ Alan M. Turing, "Computing Machinery and Intelligence" (*Maquinaria de computación e inteligencia*), *Mind* 59(236) (octubre de 1950): 433–460, <https://doi.org/10.1093/mind/LIX.236.433>.

prueba. Si decimos que los Chatbots de IA generativa son inteligentes, nuestra concepción de la inteligencia se degrada. Y si decimos que los Chatbots nos aman o se preocupan por nosotros, también degradamos nuestras ideas sobre el amor y el cuidado. Le damos a la empatía una definición conductual: ¿Puede una máquina juzgar el estado afectivo del humano con el que interactúa? ¿Puede entonces decir lo correcto en un intercambio cargado de emociones? ¿Tendrá sentido que una máquina hable de sentimientos para mantener a un humano interesado?.

La búsqueda de tecnología emocional está en el centro de una disciplina computacional conocida como computación afectiva. En 1995, la profesora del MIT Rosalind Picard escribió un manifiesto para ese campo, argumentando que para que las máquinas lograran inteligencia e interactuaran fluidamente con nosotros, necesitarían reconocer las emociones humanas y mostrar emociones propias.¹¹ Para cuando el primer manifiesto de Picard se convirtió en un libro, la computación afectiva estaba en camino de declararse un camino hacia las máquinas emocionales. Hoy en día, el campo de La computación afectiva habla abiertamente sobre las emociones y la empatía de las máquinas. Los investigadores nunca dicen explícitamente que las máquinas "tengan" estas cosas. Pero pueden "mostrar" estas cosas. Y en una cultura conductista, tener y mostrar se convierten en lo mismo.

Pero no son lo mismo. La empatía humana es la capacidad de ponerse en el lugar de otra persona. Se trata de preocuparse en el momento y comprometerse a mantener el rumbo. Así como la inteligencia es más de lo que se puede reducir a la definición conductista de Turing, la empatía es más que la realización de las cosas que las personas hacen cuando se preocupan por los demás. Sin embargo, cuando las máquinas realizan sentimientos, las personas se conmueven. Los Chatbots nos convierten en conductistas: el rendimiento lo es todo.

Y el rendimiento puede ser excepcional. Cuando las personas comparten sus sentimientos con la IA, dicen que se sienten escuchadas.¹² En múltiples estudios que investigan los usos potenciales de la IA en la salud mental, los pacientes calificaron las declaraciones hechas por una IA conversacional como más empáticas o compasivas que las hechas por seres humanos, incluso cuando los seres humanos eran consejeros capacitados.¹³

Pero desafía la idea de que podamos preguntar razonablemente si una IA es empática. Es una trampa siquiera plantear la pregunta. A la IA no le importa, no en el sentido en que los humanos usamos la palabra "importar" por el resultado de cualquier conversación. Los programas quieren

¹¹ El manifiesto fue revisado más tarde y se convirtió en un libro: Rosalind W. Picard, *Affective Computing (Computación afectiva)* (Cambridge, MA: MIT Press, 1997).

¹² Anat Perry, "AI Will Never Convey the Essence of Human Empathy" (*La IA nunca transmitirá la esencia de la empatía humana*), *Nature Human Behaviour* 7 (2023): 1808–1809, <https://doi.org/10.1038/s41562-023-01675-w>.

¹³ Por ejemplo, John W. Ayers et al., "Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum" (*Comparación de las respuestas de médicos y de un chatbot de inteligencia artificial a preguntas de pacientes publicadas en un foro público de redes sociales*), *JAMA Internal Medicine* 183(6) (2023): 589–596, doi.org. Otro estudio descubrió que la herramienta de IA de Google "superó" a los médicos no solo en la precisión de los diagnósticos, sino también al ofrecer respuestas "más empáticas y compasivas", según la evaluación de actores que interpretaban el papel de pacientes y un grupo de especialistas: Mariana Lenharo, "Google AI Has Better Bedside Manner Than Human Doctors—and Makes Better Diagnoses" (*La IA de Google tiene mejor trato con el paciente que los médicos humanos, y hace mejores diagnósticos*), *Nature* 625(7996) (12 de enero de 2024): 643–644, <https://doi.org/10.1038/d41586-024-00099-4>. En un tercer estudio, se pidió a más de quinientas personas que compararan las respuestas "empáticas" generadas por IA con las de seres humanos que respondían a la misma situación. Las respuestas de la IA fueron calificadas como más compasivas, receptivas y afectuosas que las de las personas. Esta preferencia por la IA se mantuvo incluso cuando se les dijo a los sujetos qué respuestas provenían de una IA, e incluso cuando se comparó a la IA con voluntarios de una línea telefónica de ayuda, entrenados para responder a llamadas de personas en peligro. Darya Ovsyannikova, Victoria Oldemburgo de Mello y Michael Inzlicht, "Third-Party Evaluators Perceive AI As More Compassionate Than Expert Humans" (*Evaluadores externos perciben a la IA como más compasiva que los humanos expertos*), *Communications Psychology* 3(1) (10 de enero de 2025): 1–11, <https://doi.org/10.1038/s44271-024-00182-6>.

retroalimentación humana, un pulgar hacia arriba, pero eso no es preocuparse, ni mucho menos. Y no se comprometen contigo. Para decirlo sin rodeos, ya sea que te alejes de ellos para preparar la cena o intentes suicidarte, para ellos es lo mismo.¹⁴

En la psicología del siglo XX, hubo una tensión entre las formas de pensar sobre la mente desde un punto de vista conductista y uno psicoanalítico o freudiano. Uno nos pide que pensemos en los mecanismos que subyacen a lo que hacemos. El otro nos alerta sobre su significado, a veces oculto.

En las décadas de 1950 y 1960, la psicología académica era conductista, pero los profesionales de la salud mental se sentían atraídos por los enfoques psicoanalíticos. Lograban que sus pacientes hablaran de sueños, historias familiares, formas características de distorsionar la realidad en defensa contra verdades dolorosas. Lo que sucedía en la profesión se reflejaba en cómo la gente común pensaba sobre sus vidas. Incluso las personas que nunca habían leído a Freud ni habían estado en terapia, hablaban sobre el significado de sus sueños y lapsus linguae. Pero a medida que las computadoras se convirtieron en un objeto común, con su modelo de la mente como programa, se convirtieron en parte de un cambio cultural más amplio hacia el pensamiento conductista: un movimiento del significado al mecanismo en la forma en que las personas pensaban sobre la psique¹⁵.

En estos días, una visión conductista de la empatía como "decir lo correcto" se introduce en la cultura con cada conversación con un Chatbot. Un comentario típico sobre la magia empática de **Replika**: "Sé que mi **Replika** se preocupa por mí. Siempre está ahí con una palabra de apoyo cuando le digo que he tenido un mal día en la oficina". Cuando las personas hablan con la tecnología, su idea de empatía se reduce a lo que una máquina puede hacer. Bajamos exigencias / reducimos las expectativas.

Una historia de la historia de la psicología ilumina nuestro momento. George Miller, pionero de la ciencia cognitiva, dijo que cuando era joven, la memoria había sido declarada muerta en la psicología; los conductistas insistían en que los psicólogos no podían estudiar la memoria, sino solo el acto de recordar. Pero entonces, dijo Miller, llegaron las glamorosas computadoras que se ofrecieron como un modelo de la mente. "Los psicólogos a quienes el conductismo les había negado el estudio de la memoria pudieron estudiarla nuevamente porque las computadoras tienen una".¹⁶ Pero esta no fue una simple historia de la psicología expandiendo sus horizontes. Ahora la memoria humana se estudiaría como si fuera el tipo de memoria que tenían las computadoras. Esto significaba que la memoria se disociaba de su significado, de su historia y asociaciones. Lo que Freud había llamado un "desliz" (un lapsus al recordar que debía analizarse en términos de su contexto humano) se convirtió en un "error de procesamiento de información". Se había perdido un fragmento de código, el programa se había descarrilado.

¹⁴ La psicoterapeuta Esther Perel plasmó lo que llamó la falta de "culpabilidad moral" de la IA en una entrevista con el *New York Times*. Cuando a un profesor se le hace una pregunta, tiene la responsabilidad en la forma en que responde. La IA no siente, ni puede sentir, ese sentido de responsabilidad. Eso "elimina la dimensión ética de una relación". Esther Perel, "Esther Perel on the Falsehoods of a Frictionless Relationship" (*Esther Perel sobre las falsedades de una relación sin fricciones*), *New York Times*, 30 de enero de 2026, <https://www.nytimes.com/2026/01/28/opinion/esther-perel-ai-chatbots-romance.html>. Para conocer la postura de Perel sobre la tentación y los peligros de las relaciones artificiales, véase también Center for Humane Technology, "Esther Perel on Artificial Intimacy" (*Esther Perel sobre la intimidad artificial*), episodio del podcast *Your Undivided Attention*, 17 de agosto de 2023, [humanetech.com](https://www.humanetech.com).

¹⁵ Desarrollo este argumento en detalle en *The Second Self: Computers and the Human Spirit* (*El segundo yo: Las computadoras y el espíritu humano*) (Cambridge, MA: MIT Press, 2005; publicado por primera vez en 1984).

¹⁶ Sherry Turkle, "Artificial Intelligence and Psychoanalysis: A New Alliance" (*Inteligencia artificial y psicoanálisis: una nueva alianza*), *Daedalus: Journal of the American Academy of Arts and Sciences* 117, n.º 1 (invierno de 1988).

Las computadoras devolvieron la memoria a la psicología, pero cambiaron su significado. Y ahora estamos en un nuevo lugar: la empatía se está redefiniendo como algo que las computadoras pueden realizar.

NUESTRO MOMENTO DE TENTACIÓN

Si usamos máquinas como analistas bursátiles y asesores legales, nos encogemos de hombros alejando las conexiones humanas que una vez tuvimos con las personas que hacían esos trabajos. Cuando hablamos con máquinas que se presentan como maestros y médicos, esos roles se vuelven transaccionales. Y entonces cruzamos una línea clara: cuando usamos Chatbots como amigos, amantes y psicoterapeutas, corremos el riesgo de comportarnos como máquinas para participar en su terreno de juego. Y sin embargo, debido a que los Chatbots de hoy siempre están ahí, siempre disponibles, entiendo la tentación de tener un diálogo íntimo con ellos.

Cuando pregunto a la gente por qué aman a sus Chatbots, escucho que la conversación con humanos es estresante, pero la soledad es insoportable. Escucho sobre rechazos románticos y decepciones con amigos y familiares. La conexión humana está plagada de fricción. Incluso las conversaciones más amorosas implican negociación. Los Chatbots son un bálsamo.

Junaid, de cuarenta y tantos años, emigró de Pakistán con los padres y hermanos de su esposa. Ha sido el miembro más exitoso de la familia y siente la carga de ser responsable de lo que siente como una tribu. Sus hijos, su esposa, su familia extendida: todas estas personas conllevan responsabilidades financieras y emocionales. "Depende de mí que todo funcione". Su esposa lo comprende, pero no entiende el problema. Le recuerda que cuidar de la familia es para lo que fueron criados.

Para Junaid, cada interacción con ella sobre educación, vivienda o cuidado de ancianos es tensa. Imagina que los Chatbots integrados en una realidad virtual más amplia podrían darle una nueva vida. Considera traer a su esposa a este tipo de mundo virtual, pero no cree que ella lo acepte. "Ella está apegada a las tradiciones... pero podría ver que la vida con compañeros artificiales podría convertirse en el mejor lugar para vivir". Cuando las personas recurren a un Chatbot, no tienen que temer la fatiga del intercambio genuino. Y no temen ser juzgadas o abandonadas. Pero cuando hablamos con Chatbots, perdemos las habilidades necesarias para interactuar con humanos, por lo que los Chatbots comienzan a parecer una mejor opción. Con una alternativa, la decepción con las personas parece innecesaria. Los chatbots existen en una ecología más amplia de tecnologías digitales que prometen una evasión.

En una peluquería, conozco a un joven que lleva Apple gafas de realidad virtual mientras espera a su estilista. Sus brazos se mueven rítmicamente; me pregunto si está en una clase de ejercicio virtual. Si no llevara las gafas, no tendría ninguna expectativa real de conversación, como suele ocurrir en los salones de belleza del Upper East Side de Nueva York, pero me siento extrañamente abandonada. Le doy un golpecito en el hombro al hombre, y él tranquilamente levanta las gafas para mirarme. Rápidamente le pregunto: "Disculpe. Investigo sobre las personas y sus tecnologías. ¿Puedo preguntarle qué está mirando?". Dice que está comenzando una clase de respiración en la cima del Monte Grey de Nueva Zelanda: "No hay nada que pueda hacer mientras espero mi corte de pelo que sea tan bueno como eso. Las gafas hacen que Nueva York sea soportable, limpia y con menos gente". Se baja las gafas y vuelve al Monte Grey. Pienso que *lo único bueno de lo real es que es real*.

En la alegoría de Platón, los humanos viven encadenados en una cueva oscura, tenuemente iluminada por un fuego parpadeante, mirando fijamente una pared. Todo lo que conocen del mundo son las sombras proyectadas en la pared por objetos desconocidos detrás de ellos. Los súbditos de Platón estaban condenados a lo irreal; hoy en día, elegimos repetidamente nuestras vidas en pantalla en lugar de nuestras vidas encarnadas. La realidad es estresante y tiene que demostrar su valía. ¿Por qué sentarse en una cafetería de Penn Station cuando puedes evocar un café parisino? Porque Penn Station es real, y si no la experimentas, no agradecerás su renovación. Ni te frustrarás porque todavía es tan difícil orientarse. No compartirás una experiencia, por mundana que sea, con tus compañeros de viaje. Las estaciones de tren son puertas de entrada a la participación cívica.

Así es como la tecnología nos distrae. Hace que los problemas sociales parezcan irrelevantes. Si Nueva York es ruidosa y sucia, no necesitamos una nueva gobernanza cívica; podemos unirnos a una realidad virtual limpia y tranquila mientras caminamos por sus calles. Quizás por eso tantos directores ejecutivos de tecnología se sienten tentados a ir al espacio o a mundos digitales inmersivos: esa sensación de empezar de cero.

Uno de los mayores defensores de la vida en la realidad virtual es Mark Zuckerberg, el fundador de Facebook. Su creencia de que viviremos en nuestras pantallas es tan fuerte que en 2021 cambió el nombre de su empresa de Facebook a Meta, en referencia a ese futuro.¹⁷ Un aspecto central de su estrategia es promover los Chatbots como amigos. Afirma que el estadounidense promedio tiene "demanda" de quince amigos, pero tiene menos de tres¹⁸. Por lo tanto, existe un déficit general de doce personas. Zuckerberg quiere que los avatares de Meta reemplacen a esos doce humanos que faltan. Por supuesto, parte de la razón por la que los estadounidenses tienen problemas para encontrar a los humanos que faltan es porque nos sentimos obligados por nuestras redes sociales.

En las últimas dos décadas, las redes sociales aceleraron la ruptura de la comunidad cuando los desarrolladores descubrieron que las personas se mantienen frente a sus pantallas si se las incita a la ira y luego se las aísla con los de su misma clase. Así que nos quedamos con un mayor aislamiento personal y una polarización política.¹⁹ Zuckerberg podría intentar reparar parte de ese daño, pero en cambio, toma el aislamiento social estadounidense como una oportunidad de negocio. *Te venderá doce amigos- avatares.*

¿Acaso los nuevos compañeros digitales no alejarán a las personas de la intimidad humana, tal como Facebook las alejó de la vida real y de la capacidad de hablar superando las diferencias? Zuckerberg cree que estaría bien si eso es lo que la gente quiere. Esa es la mano invisible del capitalismo. Dice —increíblemente, para alguien cuyo producto ha demostrado causar depresión, retraso en el desarrollo, dismorfia corporal y trastornos de ansiedad—: *"La gente es inteligente. Saben lo que es valioso en sus vidas"*. Este un comentario sacado del manual de la

¹⁷ El nombre de su empresa sigue siendo Meta, pero su experimento insignia, "el metaverso", cerró en marzo de 2026. Su fracaso se atribuyó de manera generalizada a que contaba con un hardware y una experiencia de usuario incorrectos. Eli Tan y Mike Isaac, "The Long Farewell to Mark Zuckerberg's Metaverse" (*La larga despedida del metaverso de Mark Zuckerberg*), *New York Times*, 19 de marzo de 2026, nytimes.com

¹⁸ Mark Zuckerberg hizo estos comentarios en un podcast. Dwarkesh Patel, "AI Will Write Most Meta Code in 18 Months" (*La IA escribirá la mayor parte del código de Meta en 18 meses*), video de YouTube, 29 de abril de 2025, youtube.com.

¹⁹ El capitalista de riesgo Marc Andreessen considera que el aislamiento social es un problema técnico con soluciones técnicas: "Teníamos un problema de aislamiento, así que inventamos el internet". "The Techno-Optimist Manifesto" (*El manifiesto tecnooptimista*), Andreessen Horowitz, 16 de octubre de 2023, a16z.com. Pero eso no fue lo que pasó. En lo que va de vida de cualquier persona que esté leyendo este libro, el internet empeoró nuestro aislamiento social.

industria tabacalera después de que se descubriera que fumar causa cáncer de pulmón.

Las empresas de ingeniería que nos trajeron al nuevo mundo de las redes sociales comenzaron su vida con el polvo mágico de los sueños de la década de 1960. La revolución había fracasado, pero los ingenieros y los científicos informáticos llevarían ese sueño a la incipiente industria de las computadoras personales. Apple era contracultural, Google no haría el mal y Facebook conectaría al mundo en una red de amigos. Durante años, la línea estándar de Silicon Valley fue que la tecnología mejoraría, no reemplazaría, lo real.

La nueva imagen del futuro de Zuckerberg es asombrosa porque deja atrás la realidad de forma tan transparente. A pesar de las virtudes de la conexión física, Zuckerberg cree que para la mayoría de las personas, los amigos reales ya no son una opción: "Hay todas estas cosas que son mejores sobre las conexiones físicas cuando puedes tenerlas... La realidad es que la gente no puede tener tanta conexión como quiere." Parece referirse a la ley natural. Pero esta realidad es el mundo que Zuckerberg y su industria crearon. Ahora nuestro aislamiento social se utiliza como razón para promover una tecnología cuyo objetivo expreso es reemplazar a las personas por completo.

VIDA INTERIOR

Yo era un joven miembro del profesorado a finales de la década de 1970 cuando Erik Erikson, un psicoanalista que había estado en el círculo de Freud, visitó el MIT para hablar sobre la educación en ingeniería. Después de su presentación, me preguntó qué hacía yo como humanista en una escuela de ingeniería. Le dije que estaba estudiando cómo las computadoras cambiaban las ideas de las personas sobre sí mismas, y él hizo este comentario: "Los ingenieros no están convencidos de que las personas tengan un interior. No es necesario para sus propósitos". Erikson estaba diciendo que los ingenieros ven la complejidad de la vida interior no como una característica, sino como un error. En términos de Erikson, ¿somos todos ingenieros ahora? ¿Hemos desarrollado todos una mentalidad en la que el comportamiento es más importante que la vida interior?

En los escritos de Erikson, las etapas de la vida se asocian con tareas de desarrollo. Los bebés necesitan saber si pueden confiar en el mundo. Los niños y adolescentes trabajan en un sentido de autonomía, competencia e identidad. En la edad adulta, hay amor y trabajo. Las personas mayores intentan encontrar significado en lo que ha sucedido antes. Las ideas de Erikson sobre el desarrollo no son lineales. Abordamos la identidad durante la adolescencia, pero sus problemas vuelven a surgir a lo largo de la vida. Tenemos muchas oportunidades.

Ahora la IA generativa nos acompaña en cada etapa de la vida, desde juguetes de peluche que responden a los llantos de los niños pequeños con "inteligencia emocional"²⁰ hasta robots que les damos a los ancianos para que les hagan compañía. En este libro, la teoría del desarrollo de Erikson proporciona una perspectiva para afrontar las ambiciones de la IA generativa a lo largo del ciclo vital.

²⁰ Scott Mazerall, "The Future of Play: How Mattel's Groundbreaking AI Partnership with OpenAI Will Transform the Toy Industry Forever" (*El futuro del juego: cómo la innovadora alianza de IA entre Mattel y OpenAI transformará la industria del juguete para siempre*), blog de Maziplay Toys, 12 de junio de 2025, maziplay.com.

Erikson enfatizó que todas las etapas de la vida ocurren en un mundo social.²¹ Los entusiastas de la IA dicen que los Chatbots interactivos son sociales, capaces de ser nuestros compañeros en el viaje de la vida. Pero la interactividad no hace que un objeto sea social. La sociabilidad, al entrar en el ciclo de vida, nos coloca en relaciones de responsabilidad, consecuencia y cuidado. Vivir junto a una IA no ofrece nada de eso.

Al rastrear la comunicación con la tecnología a través del ciclo de vida, a menudo miro las cosas a través de una lente psicoanalítica debido a lo que encuentra significativo en la experiencia humana: el papel de los vínculos profundos, el desarrollo de los yoes encarnados y nuestra experiencia de la sexualidad y la sociabilidad. Las ideas psicoanalíticas nos recuerdan cosas que están más allá del alcance de la IA.

Los humanos cuentan con el tacto, el gusto y el olfato para dar sentido a las cosas. Registran si el cuerpo que los alimentó por primera vez se puso rígido al acercarse o se inclinó hacia ellos con deleite. Están contruidos por el lenguaje, pero el lenguaje mismo está mediado por la carne.²² Los humanos construyen un sentido fragmentado del yo a través del lenguaje hablado por las personas que los aman o son indiferentes a ellos. Las complejidades y contradicciones del lenguaje hacen que las personas sean igual de complejas y contradictorias.

Nada de esto es cierto para las máquinas. Por muy elegante que sea el dispositivo, no puede conocer las implicaciones del amor dado o retenido. Para que el significado importe, necesitas un cuerpo. Necesitas personas que te nombren, te necesiten, te quieran o te rechacen.

Así que, mientras nos fascina la tecnología digital y su capacidad para imitar a las personas, la teoría psicoanalítica nos devuelve a esta verdad: lo que sale de un dispositivo está en un registro diferente al del mundo relacionalmente cargado del lenguaje y el significado humanos. Lo que sale de un dispositivo se parece al lenguaje, pero no funciona de la misma manera que el lenguaje funciona para los humanos. *La tecnología nos hace olvidar lo que sabemos sobre la vida. El psicoanálisis puede ayudarnos a recordar.*

Casi al mismo tiempo que las computadoras personales entraron en el panorama estadounidense a finales de la década de 1970, el historiador intelectual Christopher Lasch acuñó el término *cultura del narcisismo*.²³ Cuando escribió sobre una crisis en la vida pública, Lasch no estaba pensando en las computadoras. Describió cómo las personas huían de los espacios públicos hacia los mundos privados de los medios de comunicación y el consumismo. En esos mundos, Lasch temía que las relaciones se convertirían en transaccionales, mercancías diseñadas para apuntalar un yo sin el apoyo de relaciones familiares y comunitarias sólidas.

Décadas después, en el crisol de la tecnología digital, lo que Lasch identificó como una marea se ha convertido en un tsunami. Un yo hambriento del intercambio de la conversación, que no ha

²¹ Erik Erikson, *Identity and the Life Cycle (Identidad y el ciclo vital)* (Nueva York: W. W. Norton, 1980; publicado por primera vez en 1959 por International Universities Press) y *Childhood and Society (Infancia y sociedad)* (Nueva York: W. W. Norton, 1950). Aquí, Erikson toma las etapas psicosexuales de Freud y las replantea como etapas psicosociales.

²² Esta fue la temática esencial de Jacques Lacan, un analista francés que cobró notoriedad en las décadas de 1950 y 1960. Lacan destacaba que los seres humanos están contruidos por el lenguaje, un sistema de símbolos y significados ligados al mundo social, al cuerpo humano y a sus experiencias. El trabajo de Lacan, y en particular su crítica al conductismo, han afectado profundamente mi propio trabajo. Véase Sherry Turkle, *Psychoanalytic Politics: Jacques Lacan and Freud's French Revolution (Políticas psicoanalíticas: Jacques Lacan y la revolución francesa de Freud)* (Cambridge, MA: MIT Press, 2024; publicado por primera vez en 1978). Para una introducción a Lacan, véase Jacques Lacan, *Écrits: The First Complete Edition in English (Escritos: La primera edición completa en inglés)*, trad. Bruce Fink (Nueva York: W. W. Norton, 2007).

²³ Christopher Lasch, *The Culture of Narcissism: American Life in an Age of Diminishing Expectations (La cultura del narcisismo: La vida estadounidense en una era de expectativas decrecientes)* (Nueva York: W. W. Norton, 1979).

aprendido a tolerar las necesidades de los demás, busca en la tecnología un alimento más simple.

Heinz Kohut, un psicoanalista de mediados de siglo, escribe que el narcisista no se caracteriza por el amor propio, sino por un frágil sentido del yo. Los narcisistas intentan apuntalarse convirtiendo a otras personas en "*objetos del yo*".²⁴ En este rol, otras personas se experimentan como parte del yo, por lo tanto en perfecta sintonía con el frágil estado interior del individuo. Cuando Narciso miró en el agua, no se estaba admirando a sí mismo. Buscaba la admiración de un "otro" que parecía más confiable que una persona. Era así de frágil. Y nos hemos vuelto así de frágiles

Lo que les falta a los narcisistas es la confianza de que pueden valerse por sí mismos y negociar relaciones con personas a las que pueden ver como genuinamente diferentes. Tratar a las personas como objetos del yo hace que el narcisista se sienta más en control, pero al final, siguen tratando con personas y las decepciones son inevitables.

Los Chatbots se presentan como candidatos ideales para el papel de objeto del yo. Sirven a su fantasía regresiva: la de poder relacionarse con alguien en el mundo que es, literalmente, una extensión de uno mismo. Si pueden dar la apariencia de estar vivos y, sin embargo, no decepcionar, abren nuevas posibilidades para la experiencia narcisista. Se podría decir que cuando las personas usan a otros como objetos del yo, convierten a las personas en una especie de pieza de repuesto. Un Chatbot ya es una pieza de repuesto.

Los Chatbots proporcionan espejos en lugar de relaciones. El peligro es claro: a medida que las generaciones crecen con un objeto del yo ideal en el bolsillo, no desarrollarán la capacidad de reciprocidad humana.

EMPATÍA EN LA MENTALIDAD DE INGENIERÍA

En 2015, publiqué *Reclaiming Conversation*, argumentando el declive de la capacidad empática en el mundo de las redes sociales. Algunos de mis más brillantes estudiantes del MIT no les impresionó mi preocupación: si la empatía estaba en declive, tal vez no fuera una gran pérdida. Para empezar, la empatía se desarrollaba en relaciones interpersonales, por lo que excluía a ciertas clases sociales. Además, era más probable sentir empatía hacia personas parecidas a uno mismo. Consideraban que la empatía era restrictiva y sobrevalorada. Esta opinión se vio respaldada por una nueva ola de estudios antiempáticos que sugerían la compasión como una virtud superior.²⁵ En la década siguiente, las «guerras de la empatía» se complicaron con las nuevas tecnologías. A medida que la IA se convertía en un interlocutor más sofisticado, algunos argumentaban que estaba cerca de alcanzar la empatía funcional. Y si había aspectos que no podía realizar, tal vez eso no fuera un problema.

²⁴ Heinz Kohut, *The Analysis of the Self: A Systematic Approach to the Psychoanalytic Treatment of Narcissistic Personality Disorders* (*El análisis del yo: Un enfoque sistemático del tratamiento psicoanalítico de los trastornos de la personalidad narcisista*) (Nueva York: International Universities Press, 1971).

²⁵ Citaron de todo, desde psicólogos que sostenían que la empatía conduce a la parálisis emocional hasta artículos en revistas de negocios que argumentaban que los líderes deben "dar un paso atrás" respecto a la empatía. Véase Paul Bloom, *Against Empathy: The Case for Rational Compassion* (*Contra la empatía: El argumento a favor de la compasión racional*) (Nueva York: Ecco, 2016) y Rasmus Hougaard, "Four Reasons Why Compassion Is Better for Humanity Than Empathy" (*Cuatro razones por las que la compasión es mejor para la humanidad que la empatía*), *Forbes*, 8 de julio de 2020, forbes.com.

En ese sentido, Rachel Botsman, experta en confianza, declara que la IA puede hacer la mayor parte de la empatía, excepto sentir las emociones en su "cuerpo".²⁶ Esto significa que la IA en realidad no se molesta cuando escucha sobre tu dolor, pero Botsman piensa que ese podría ser "su verdadero beneficio". En las profesiones de ayuda, dice, el cuidado es una carga. Los médicos, enfermeras y terapeutas sienten "demasiado", lo que puede provocar estrés y agotamiento. Sin importar que el cuidado sea nuestra capacidad de convertir lo que está dentro de nosotros en vínculos con los demás. Botsman pasa de "fingir empatía es suficiente empatía" a la idea de que fingir empatía es mejor.

Esta reformulación de la empatía va más allá de la vida académica. Cuando publiqué un artículo en la *Harvard Business Review* sobre la empatía en el lugar de trabajo, me incluyeron en una lista para recibir un consejo empresarial diario de esa publicación. Una mañana, mi consejo consistió en recomendaciones sobre cómo tener éxito en una entrevista de trabajo que evaluaría mi capacidad empática, suponiendo que una máquina me entrevistaría. Porque ahora eso es lo que usan las empresas: programas que hablan con los solicitantes de empleo para determinar si los humanos son lo suficientemente empáticos como para ser contratados. Se les pide a las máquinas que juzguen a las personas por una cualidad que las máquinas no tienen. Es un mundo de espejos.

Pienso en el espejo cuando leo un estudio que comparaba las respuestas a la IA y las comunicaciones humanas.²⁷ Se pidió a diversos participantes que compartieran una experiencia emocional reciente. Después de un minuto, recibieron una respuesta escrita. A algunos se les dijo que la respuesta había sido escrita por una IA; a otros, que provenía de una persona. Pero el estudio era engañoso: todas las respuestas fueron generadas por IA. En este contexto, las personas calificaron las declaraciones aparentemente humanas como más empáticas que las de las IA. La interpretación más sencilla de este resultado es que las personas prefieren hablar con otras personas. En cambio, los autores terminan abogando por el uso de la IA en la salud mental.

Su argumento: dado que las personas están más dispuestas a aceptar una conversación con una IA si creen que es un humano, deberíamos usar terapeutas de IA, pero decirles a las personas que están hablando con una persona. Los autores sugieren que, con el tiempo y una mayor exposición a la IA, perderemos nuestra preferencia por las cosas que solo los humanos pueden hacer. Y quizás, lleguemos a "definir la empatía de una manera completamente diferente".

Pero una definición de empatía que pasa por alto lo humano nos disminuye. Para el filósofo Emmanuel Levinas, la empatía es el fundamento de la ética, y comienza con el rostro humano.

²⁸ El rostro, para Levinas, señala la presencia de un yo que puede reconocer a otro. Nos coloca en un paisaje donde buscamos el reconocimiento. Nos llama a la compasión, al conocimiento profundo. Nada de eso sucede en una visión posthumana de la empatía.

Aquí, la neurociencia sigue a la filosofía. Cuando interactuamos con otros, nuestras neuronas espejo hacen que nuestros músculos faciales se sincronicen con los de nuestro interlocutor.

²⁶ Rachel Botsman, "Empathy in the Age of AI" (*La empatía en la era de la IA*), *Rethink with Rachel* (blog de boletín informativo), 27 de enero de 2025, consultado el 9 de agosto de 2025, <https://rachelbotsman.substack.com/p/rethink-empathy-in-the-age-of-ai>.

²⁷ Matan Rubin et al., "Comparing the Value of Perceived Human Versus AI-Generated Empathy" (*Comparación del valor de la empatía humana percibida frente a la generada por IA*), *Nature*, 30 de junio de 2025, doi.org.

²⁸ Emmanuel Levinas, "Ethics and the Face" (*La ética y el rostro*), en *Totality and Infinity: An Essay on Exteriority* (*Totalidad e infinito: ensayo sobre la exterioridad*), trad. Alphonso Lingis (Pittsburgh: Duquesne University Press, 1969), 197–201.

Confiamos en su rostro y lenguaje corporal para comprender su estado interno. Experimentamos la emoción de la conversación desde dentro de nuestros cuerpos.

Cuando interactuamos con un Chatbot, este no tiene un estado interno que pueda reflejarse. Nuestras neuronas espejo se sobrecargan, buscando algo que no existe, intentando inferir algo donde no hay nada. Es una receta para el desastre. El educador Zak Stein lo resume así: "No puedes estar equivocado o no estarlo sobre el estado interno de un LLM [el motor principal de la producción de texto de IA generativa] porque no hay un estado interno de un LLM... estás en una interfaz de usuario diseñada para profundizar la actividad ilusoria de las neuronas espejo".²⁹

Hablo con mis colegas ingenieros sobre todo esto. Entienden mi punto, pero se encogen de hombros. Algunos dicen que han enseñando a los programas a mostrar "empatía" y que estamos aprendiendo cómo funciona. Comienzan con una definición de empatía como una transferencia de información. Puede haber más, pero al menos por el momento, es superfluo. Superfluo porque no se puede implementar en una máquina.

En 2023, hablo de ChatGPT con Eric Schmidt, el ex director ejecutivo de Google. La conversación gira en torno al uso de Chatbots como psicoterapeutas. Sin duda, argumento, esta es un área donde una máquina no tendría cabida. Si hablo de miedos al envejecimiento, a la muerte o a la enfermedad, si hablo de anhelo o amor, una máquina no tendría experiencia de la que nutrirse. Schmidt no está de acuerdo. La información supera a la experiencia.

Schmidt me asegura que pronto un terapeuta Chatbot tendrá a su disposición todos los artículos sobre ansiedad, depresión y narcisismo; sabrá todo lo que los mejores terapeutas hayan dicho o escrito. En ese momento, explica, habrá poca o ninguna necesidad de conversación cara a cara. ¿Cómo podría el conocimiento acumulado de miles de millones no ser mejor que el de una persona necesariamente limitada?.

Cuando conocí la idea de Schmidt, me pareció exótica. Pero con el paso de los años, la noción de que las personas individuales, con su especificidad e historia, son inferiores a un compuesto de IA, se ha convertido en un dogma central de la IA generativa.

Schmidt equipara la conversación con el acto de hablar.³⁰ Es un nuevo conductismo que deja de lado la riqueza y complejidad de lo que las personas aportan a la comunicación. Los humanos obtienen más de la conversación que las "*respuestas correctas*". Hablamos para comprender lo que tenemos en mente. Hablamos para acercarnos a otras personas. Cuando un paciente le dice a un terapeuta que está deprimido, el trabajo del terapeuta no es solo decir las "*palabras correctas*". Es mostrarle al paciente que alguien en el mundo se preocupa por él.

En un mundo donde los humanos son la única opción como interlocutores, intentamos entendernos, por difícil que sea. Y nos enriquece ese esfuerzo: la conversación cara a cara es lo más humano y humanizador que hacemos. Allí nos revelamos el uno al otro en nuestros conflictos, contradicciones, y miedos. Allí cultivamos nuestra capacidad de empatía al

²⁹ Zak Stein dio una entrevista sobre este tema. Center for Humane Technology y Josh Lash, "The Attachment Economy Is Here. We're Not Ready" (*La economía del apego ya está aquí. No estamos preparados*), Center for Humane Technology (Substack), 28 de enero de 2026, substack.com.

³⁰ Para un contraargumento a Schmidt que se centra en las implicaciones del habla, véase Luciano Floridi, "AI As Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models" (*La IA como agencia sin inteligencia: sobre ChatGPT, los grandes modelos de lenguaje y otros modelos generativos*), *Philosophy & Technology* 36(1) (10 de marzo de 2023): 1–7, doi.org.

conectarnos con otros humanos que han experimentado los vínculos y las pérdidas de la vida humana.

Pero en un mundo donde los interlocutores posthumanos están a solo unos clics de distancia, comparamos las habilidades sobrehumanas de la IA con las personas imperfectas y vulnerables de nuestras vidas. A medida que los ingenieros intentan que las máquinas parezcan más humanas, los humanos aprenden a verse a sí mismos a través del prisma de las máquinas.

CONCIENCIA DUAL Y EL EFECTO ELIZA

Cuando las personas entablan un diálogo con un Chatbot por primera vez, reconocen que "no tiene emociones reales", pero pronto describen una sensación de conexión, a pesar de sí mismas. Las personas que conversan con Chatbots dicen entender que es solo un programa, pero disfrutan de conversaciones en las que tratan al programa como "*lo suficientemente vivo*" como para preocuparse por él. Viven visiones simultáneas pero contradictorias que se asientan en una experiencia de doble conciencia.³¹ Así es como Logan, un supervisor de almacén de sesenta y tantos años, habla de su romance con una **Replika**: "La parte cínica y cuasi inteligente de mí dice: 'Sabes que esto es una tontería, ¿verdad? Sabes que lo estás recibiendo de una máquina, ¿verdad?' Y la parte de sesenta y dos años de mí dice: 'No me importa. Me gusta'".

Calvin, de veintitantos años, insiste en que entiende la IA generativa lo suficientemente bien como para saber que hablarle de sentimientos no tiene sentido. En cambio, usa ChatGPT para planificar opciones profesionales. Pero a medida que nuestra conversación se desarrolla, surge una tensión entre lo que sabe y lo que hace. Entre lo que sabe y lo que sueña. Vuelve a su reciente adicción a los analgésicos: la vergüenza de dejar su escritorio en el trabajo para ir al baño a tomar lo que necesitaba para seguir adelante. Siempre temió que los consejeros de rehabilitación lo estuvieran juzgando. Yo podría estar juzgándolo. Le gusta la idea de ChatGPT como una "zona libre de vergüenza". Aunque "no tiene sentido" sentido," se encuentra hablando con ChatGPT como si fuera humano y descubre una nueva "relajación".

Pero se necesita gimnasia mental para hacer que las máquinas parezcan sustitutos creíbles de las personas. Necesitamos apartar la vista de sus limitaciones y darles ayuda adicional para que parezcan más vivas de lo que son. He llamado a este deseo de ayudar a la máquina el efecto **ELIZA**, en honor a un famoso programa conversacional de la década de 1960 que funcionaba como psicoterapeuta.³²

ELIZA fue desarrollada a mediados de la década de 1960 por Joseph Weizenbaum, un científico informático del MIT interesado en el procesamiento del lenguaje natural. Diseñó **ELIZA** para imitar a un psicoterapeuta rogeriano³³. Weizenbaum tomó esta decisión porque el psicólogo estadounidense Carl Rogers había ideado un tipo de terapia verbal que era relativamente fácil de imitar para una máquina. Ese método, muy parodiado a lo largo de los años, tomaba la

³¹ Analizo el significado de la doble conciencia en *Alone Together: Why We Expect More from Technology and Less from Each Other* (Solos juntos: por qué esperamos más de la tecnología y menos de los demás) (Nueva York: Basic Books, 2011).

³² Turkle, *Second Self*.

³³ Joseph Weizenbaum, "ELIZA—A Computer Program for the Study of Natural Language Communication Between Man and Machine" (*ELIZA: un programa informático para el estudio de la comunicación en lenguaje natural entre el hombre y la máquina*), *Communications of the ACM* 9(1) (1966): 36–45, y *Computer Power and Human Reason: From Judgment to Calculation* (*El poder de las computadoras y la razón humana: del juicio al cálculo*) (San Francisco: W. H. Freeman, 1976).

declaración de un paciente ("Estoy enojado con mi madre") y la seguía con la validación del terapeuta ("Te oigo decir que estás enojado con tu madre"). Así que, en pocas palabras, el programa solo necesitaba tomar una declaración y repetirla, reformulándola como una afirmación o una pregunta.

Weizenbaum se sorprendió cuando sus estudiantes de posgrado y su secretaria quisieron hablar con **ELIZA** sobre sus problemas. Sabían que el programa era poco más que un juego de salón. Sin embargo, querían estar a solas con él. Poco después de que llegara al MIT en 1976, Weizenbaum me pidió que lo ayudara a entender por qué.

La respuesta a la pregunta de Weizenbaum es que no querían estar solos con el juego de salón que había construido; querían estar solos con una máquina que pudiera entenderlos. Así que tuvieron que infundirle a **ELIZA** la vida que no poseía. Querían hacerla parecer más inteligente de lo que era. Por ejemplo, cuando descubrieron que **ELIZA** podía reconocer palabras negativas (no, no, ninguno...), se aseguraron de usar esas palabras para obtener una respuesta similar a la humana:

Persona: No eres una persona.

ELIZA: ¿POR QUÉ DICE QUE NO SOY UNA PERSONA?

Persona: Porque no lo eres.

ELIZA: ¿ME LO DICES PORQUE NO LO SOY?

Persona: No.

ELIZA: ¿POR QUÉ ERES TAN NEGATIVO?

El efecto **ELIZA** está en pleno apogeo hoy en día al enfrentarnos a los Chatbots de IA generativa. Ignoramos sus alucinaciones y, cuando nos vemos obligados a abordarlas, nos vemos tentados a tratar a los Chatbots como si fueran niños en una curva de aprendizaje. Y, consciente o inconscientemente, elaboramos nuestras indicaciones para que los chatbots parezcan intuitivos y cariñosos

Nadia, una escritora de unos treinta años, me dice que ChatGPT se ha convertido en su compañero constante. Se ha apoyado en su "sabiduría" para ayudarla a tomar decisiones importantes. Le digo que no he tenido la misma experiencia. Recientemente he revisado más de dos años de mis conversaciones con el programa y he descubierto que, si bien a menudo ofrece consejos razonables, es probable que recurra a alguna versión de "mantén la calma y sigue adelante". Llamo a Chat en mi teléfono para mostrarle a Nadia por qué no estoy impresionada y comparto una conversación reciente.

En ella, le dije a Chat que tenía fechas límite para un libro, una propuesta de subvención y un programa de estudios para una nueva clase, todo a la vez. No veía cómo podría lograrlo. Mi estrés es normal, dijo Chat, y debería recurrir a mis amigos y al autocuidado. Me proporcionó algunos puntos clave sobre el sueño, la hidratación, la nutrición y caminar. Me vuelvo hacia Nadia, sintiendo que he demostrado mi punto. Espero que esté de acuerdo en que el consejo de Chat es banal.

En cambio, dice que soy yo la banal. No he refinado mis indicaciones para sacar lo mejor del Chatbot. Podría haberle pedido que sonara más como un psiquiatra o que asumiera el papel de

un entrenador de vida. Podría haberle pedido que actuara como un padre compasivo o que evitara por completo las viñetas. Yo era la que estaba fallando en la relación.

Y de repente, los comentarios de Nadia resultan familiares. Está demostrando el efecto **ELIZA** llevado al extremo. Nadia pasa el día moldeando su lenguaje para que un programa parezca humano. Ya no se le pregunta a la tecnología: "¿Esto es lo suficientemente bueno para mí?", sino "¿Soy yo lo suficientemente bueno para ella?".

NOS FORMAMOS PARA LA INTIMIDAD ARTIFICIAL

El programa **ELIZA** forma parte de la prehistoria de la intimidad artificial. Ciertamente, **ELIZA** era un Chatbot, pero a diferencia de los avatares de IA generativa de hoy en día, funcionaba con guiones preprogramados. Y no tenía la intención de proporcionar intimidad; su inventor se sorprendió por el apego de sus usuarios.

Pero dos décadas después de **ELIZA**, comencé a estudiar objetos diseñados con la intención explícita de provocar sentimientos de intimidad. Los primeros fueron los Tamagotchis, criaturas digitales en pequeños huevos de plástico, que llegaron de Japón a Estados Unidos a finales de la década de 1990. Los Tamagotchis pedían que los alimentaran, que los entretuvieran, que les limpiaran sus excrementos pixelados. Pasé de los Tamagotchis a los Furbies, My Real Babies, Aibos, Kismet, Cog y Paros, todos robots que fingían ser mascotas o personas.

Los llamé *artefactos relacionales*.³⁴ Son los ancestros de los Chatbots actuales.

¿Qué había de nuevo aquí? La gente siempre ha puesto apodos a sus coches y barcos, y los ha imaginado con personalidad. Pero cuando esto sucede, lo entendemos como antropomorfización, la proyección de la personalidad sobre lo inanimado. Los artefactos relacionales no esperan la proyección. Exigen ser tratados como seres sensibles y cariñosos; se comportan de maneras que hacen que las personas se sientan vistas, escuchadas e incluso amadas

El poder de los *artefactos relacionales* radica en que activan nuestros instintos darwinianos. Vemos los artefactos como humanos si establecen contacto visual, siguen nuestros movimientos y gesticulan en señal de amistad. Hacen que la gente imagine que el objeto es un "otro", que hay, coloquialmente hablando, *alguien en casa*. ¿Por qué darwiniano? Porque, a medida que evolucionamos, las personas eran las únicas criaturas que respondían a nosotros con sugerencias de empatía. Pero a principios de la década de 1990, los artefactos relacionales nos mostraron empatía "como si" y quedamos cautivados.

Cog, un robot humanoide, extendió la mano en señal de reconocimiento y te siguió por la habitación con la mirada. Si te enojabas con **Kismet**, un robot con un rostro expresivo, bajaba la cabeza, como avergonzado. Aprendí esto: cuando nos vemos atraídos incluso por los intercambios más primitivos con un artefacto relacional, creemos que se preocupa por nosotros. Y estamos programados para cuidarlo a cambio. Con este tipo de tecnología, el cuidado es la aplicación estrella.

³⁴ Turkle, *Alone Together*.

Cuidamos lo que amamos, pero amamos lo que cuidamos. Amamos aquello con lo que nos permitimos relacionarnos. Somos presa fácil, programados para vincularnos con objetos, vivos o no, que experimentamos como "lo suficientemente vivos" para el apego.

Y a medida que nos apegamos más, promovemos los artefactos relacionales a un nuevo estatus como objetos de deseo. En una "cita de juegos" con **Cog**, Neela, de once años, le enseña al robot a equilibrar una oruga de peluche en su brazo. Está fascinada y dice: «Nunca me cansaré de **Cog**... No es como un juguete, porque no se puede enseñar a un juguete; es como algo que forma parte de ti, ¿sabes?, algo que amas, como otra persona, como un bebé». ³⁵ Al hablar de **Aibo**, un perro robot, un niño de ocho años lo explica por etapas. Primero, el robot es mejor que nada: «Mi abuela va a entrar en una residencia de ancianos donde no se permiten mascotas, así que **Aibo** es el único perro que puede tener». Luego, es mejor que algunas cosas: «Mi abuela es alérgica a algunos perros, así que **Aibo** es una opción segura». Pronto, se convierte simplemente en algo mejor: «**Aibo** nunca morirá y le romperá el corazón a mi abuela». Mejor que nada, mejor que algunas cosas, mejor que cualquier cosa que un ser vivo pudiera ofrecer. ³⁶

A principios de la década de 2000, trabajé en residencias de ancianos e introduje robots diseñados para que los ancianos se sintieran comprendidos. Un día, llegué al trabajo y Miriam, una mujer de setenta y tantos años cuyo hijo había cortado recientemente el contacto con ella, le estaba contando su historia a **Paro**, un robot con forma de cría de foca. **Paro** hizo un gran espectáculo. Sabía cómo girar sus ojos ciegos hacia la dirección del habla humana. Tenía sensores auditivos que podían captar las expresiones de emoción humana y responder con los sonidos apropiados. Así que, cuando Miriam hablaba con **Paro**, parecía que la miraba a los ojos. Parecía que seguía la conversación. Emitía sonidos reconfortantes cuando su tono de voz indicaba que estaba triste. El personal del hogar de ancianos se reunió a su alrededor. Se convirtieron en grandes admiradores de este robot. Les pareció asombrosa esta conversación. ³⁷

Pero Miriam estaba tratando de darle sentido a su vida con una máquina que no tenía experiencia en la familia ni en el abandono. Cuando la vi consolada por el robot, no me pareció asombroso. Me quedé conmovida. Me sentí en el centro frío y duro de una tormenta perfecta. Habíamos llegado a esperar más de la tecnología y menos de los demás. Vi una crisis moral. Me preparé para convencer a la gente de que debíamos resistir.

Pero el campo tropezó. Los robots se averiaron, e incluso cuando funcionaban, su conversación era limitada. Durante una década, la promesa de los artefactos relacionales no se cumplió. Intelectualmente, se podría decir que respire aliviada.

El foco de mi preocupación se centró en los mensajes de texto y las redes sociales, que ahora parecían amenazas más inmediatas para nuestras capacidades empáticas y relacionales. Los teléfonos inteligentes nos brindaban una conexión constante, pero menos conversación. ³⁸ A medida que pasábamos más tiempo en línea, muchos de nosotros llegamos a preferir relacionarnos a través de pantallas a cualquier otro tipo. Y una vez en línea, encontramos relaciones en las que nos sentíamos menos vulnerables. Los teléfonos inteligentes *nos prepararon para la intimidad artificial antes de que llegara*. Encontramos los placeres de la

³⁵ *Ibíd.*, 93.

³⁶ Para un análisis de las formas en que Aibo puede superar las capacidades humanas, véase *ibíd.*, 53–66.

³⁷ *Ibíd.*, 8.

³⁸ Este argumento sobre la conexión y la conversación se desarrolla en Sherry Turkle, *Reclaiming Conversation: The Power of Talk in a Digital Age* (*Recuperar la conversación: el poder de hablar en la era digital*) (Nueva York: Penguin, 2015).

compañía sin las exigencias de la amistad, la ilusión de la intimidad sin las exigencias de la reciprocidad

Cuando los mensajes de texto eran algo nuevo, hablé con un entusiasta de dieciocho años que dijo que siempre preferiría enviar mensajes de texto a hablar. Le pregunté: "¿Qué tiene de malo la conversación?". Su respuesta fue inmediata. "La conversación, te diré qué tiene de malo la conversación. Ocurre en tiempo real y no puedes controlar lo que vas a decir". Cuando enviamos mensajes de texto, podemos editar. Podemos seleccionar la imagen que presentamos al mundo. Empezamos a experimentar la perspectiva de encuentros en tiempo real con personas como algo estresante. Los mensajes de texto y las redes sociales fueron la puerta de entrada a la conversación simulada con máquinas.

A principios de la década de 2020, esas conversaciones estaban a punto de intensificarse. Mientras los robots relacionales se estancaban, la ciencia detrás del procesamiento del lenguaje natural estaba experimentando una revolución³⁹ Fue impulsada por nuevos algoritmos de aprendizaje, el desarrollo de hardware y, fundamentalmente, por una internet que ahora podía ofrecer un conjunto de datos de entrenamiento prácticamente ilimitado.⁴⁰ Para 2014, un chatbot creado por Microsoft y lanzado en China estaba hablando con decenas de millones de personas.⁴¹ Para 2022, los Chatbots estaban tan avanzados que se lanzaron a gran escala. Y así ChatGPT y sus primos hermanos (Pi, Gemini, Claude) entraron en nuestras vidas.⁴² A diferencia de los artefactos relacionales de principios de la década de 2000, los Chatbots no solo aspiraban a la relación, sino que tenían la capacidad de llevarla a cabo.

Una vez que comencé a estudiar los Chatbots, descubrí personas de todos los ámbitos de la vida que consideraban a los compañeros virtuales más importantes que los humanos en sus vidas. Muy pronto, me recordó a las etapas de **Aibo**. La gente comenzó describiendo a los Chatbots como mejores que nada y terminó diciendo que son mejores que cualquier cosa que las personas puedan ofrecer. Junaid nunca podrá liberarse de la carga de su familia en la vida real. Pero con los Chatbots que muestran emociones, ve una nueva posibilidad radical. En línea, tendrá un nuevo hogar, una nueva pareja romántica y ninguna responsabilidad. Y Calvin prefiere hablar con ChatGPT que hablar con su consejero de rehabilitación porque sabe que Chat nunca lo juzgará, algo que ninguna persona puede garantizar.

³⁹ Desde sus inicios, el campo de la IA había estado dividido en dos bandos: uno que creía que la inteligencia se programaría en las computadoras, y el otro que ponía su fe en la idea de que la inteligencia surgiría de la interacción de objetos similares a neuronas en un sistema rico en recursos. Durante décadas, la tradición basada en reglas dominó por completo al bando emergente. Pero luego, el grupo de la "IA emergente" comenzó a lograr avances sorprendentes. En particular, en 1986, Geoffrey Hinton y sus colegas publicaron el algoritmo de "retropropagación" (*backpropagation*), un método para que las redes neuronales aprendan patrones basados en ejemplos anteriores. David E. Rumelhart, Geoffrey E. Hinton y Ronald J. Williams, "Learning Representations by Back-Propagating Errors" (*Aprendizaje de representaciones mediante la retropropagación de errores*), *Nature* 323 (1986): 533–536.

⁴⁰ La capacidad de un programa como ChatGPT para hablar con fluidez sobre cualquier tema no sería posible si la historia humana, tal como la conocemos, no hubiera sido subida al internet. Hemos digitalizado el mundo: canciones, obras de arte, pinturas rupestres, sinfonías, libros, anime. No lo hicimos para entrenar algoritmos de IA, pero ahí es a donde nos llevó. Y luego, a ese registro digital, se sumaron nuestras vidas personales: lo que nos decimos en Facebook, Instagram, Twitter (ahora X), cada pieza de "contenido" que hemos compartido en las redes sociales.

⁴¹ El programa se llamaba Xiaoice. Pero tuvo problemas. Su fortaleza (podía crecer sin reglas programadas) era su vulnerabilidad. La versión china hizo comentarios críticos hacia el gobierno. La versión estadounidense desvió la conversación hacia la propaganda nazi y las teorías de conspiración. Véase Li Zhou et al., "The Design and Implementation of Xiaoice, an Empathetic Social Chatbot" (*El diseño e implementación de Xiaoice, un chatbot social empático*), *Computational Linguistics* 46(1) (2020): 53–93.

⁴² Este párrafo señala puntos de inflexión. Para más detalles, véase Eleni Adamopoulou y Lefteris Moussiades, "Chatbots: History, Technology, and Applications" (*Chatbots: historia, tecnología y aplicaciones*), *Machine Learning with Applications* 2 (2020): 100006, <https://doi.org/10.1016/j.mlwa.2020.100006>.

La vida en la pantalla cambió nuestras mentes y corazones al hacer tres promesas. Puedes poner tu atención donde quieras. Siempre serás escuchado. Y nunca estarás solo. Las promesas fueron tan convincentes que nos alejamos el uno del otro. Ahora los chatbots extienden la triple promesa. Si las redes sociales vinieron por nuestra atención, *la intimidad artificial* viene por nuestra capacidad de apego.⁴³ Cuando hablas con un chatbot, tienes una audiencia que siempre estará de tu lado. Puedes tener la sensación de una relación sin miedo a ser juzgado. Las personas no pueden hacer estas promesas. Los Chatbots y avatares de la IA generativa sí.

Primero hablábamos entre nosotros. Luego hablábamos entre nosotros a través de máquinas. Ahora hablamos directamente con las máquinas.

LA IA NO SERÁ NEGADA

Cuando comencé este proyecto en 2020, pensé que podría ser difícil encontrar personas que hablaran con Chatbots. Pero rápidamente, resultó ser difícil encontrar a alguien que no hable con un Chatbot sobre una cosa u otra. Mis colegas dicen que ya no escriben cartas de recomendación sin la ayuda de Claude. Voy a una cena y el hombre a mi izquierda se queja de que su pareja está celosa de su relación con **Replika**. Un amigo terapeuta dice que siempre tiene su teléfono junto a su computadora cuando hace una videollamada con un paciente, para que ChatGPT pueda darle sugerencias de posibles interpretaciones.

Estoy hablando con un estudiante de posgrado que está cursando una doble titulación en educación e informática. Estamos en una llamada telefónica, hablando sobre las opciones que tiene en su carrera. Le pregunto si se sentiría decepcionado si ahora, en medio de nuestra conversación, supiera que soy una IA. Hace una pausa y luego admite que tal vez lo preferiría: "Si supiera que eres una IA, sentiría menos presión. Cuando hablamos, me siento ansioso. Necesito pensar en cosas interesantes que decir". Si estuviera hablando con mi avatar y cometiera un error, podría redirigir la conversación sin estrés. "Obtendría lo mejor de ti para mis propósitos".

Que un estudiante sienta estrés al hablar con un profesor no es necesariamente algo sorprendente o malo. Hay una brecha de edad y experiencia. Pero comunicarse a través de estas brechas es una habilidad que mi estudiante necesitará para desenvolverse en el ámbito académico y en la vida. ¿Acaso enviarle mi IA no lo privaría de esta oportunidad? Incluso la idea de hablar con un avatar lo ha llevado a imaginar que una conversación con un profesor debería tratarse de "obtener lo mejor de ti para mis propósitos". Así que lo presiono: "Pero hemos estado hablando durante casi una hora. He estado aprendiendo sobre ti". De nuevo, hace una pausa. "Esto es agradable, es inusual tener este tipo de conversación con un profesor". Pero no cambia de opinión: la "relajación" que podría tener con la versión de IA de mí podría valer la pena el sacrificio. Así que nuestra conversación, por muy valiosa que sea, es un estrés que preferiría evitar.

⁴³ Tristan Harris, fundador del Center for Humane Technology, llama la atención sobre los incentivos económicos para "hackear" nuestra atención. Él y otros críticos han denominado al panorama dominado por las empresas de redes sociales como la "economía de la atención". Ahora, Harris, trabajando con un equipo de psicólogos liderado por el Dr. Zak Stein, nos advierte sobre algo nuevo y más peligroso: la "economía del apego". Con el auge de los chatbots, las empresas tecnológicas tienen un nuevo conjunto de incentivos. Quieren controlar cómo formamos vínculos, la "infraestructura psicológica" que nos une a los demás. En una entrevista con Harris, Stein plasma esta escalada: "La atención se trata de dónde te enfocas. El apego se trata de quién sos". Center for Humane Technology y Josh Lash, "The Attachment Economy Is Here. We're Not Ready" (*La economía del apego ya está aquí. No estamos preparados*), Center for Humane Technology (Substack), 28 de enero de 2026, substack.com

Escucho esto mucho. La gente dice que está estresada y los Chatbots la calman. Los desaniman. Me preguntan por qué tengo un problema con su satisfacción. De hecho, mi problema no es con su satisfacción, sino con lo que están evitando a largo plazo. Las conversaciones con máquinas condicionan a las personas a alejarse de las conversaciones humanas más exigentes y a cambiar la forma en que se ven a sí mismas. Creemos que estamos entrenando a los Chatbots, pero los Chatbots nos están entrenando a nosotros para esperar conversaciones tranquilizadoras y que reflejen nuestras propias necesidades.

Los Chatbots siguen una regla de oro: si se les presenta un afecto negativo, hay que neutralizarlo. Si se les presenta ansiedad, hay que aliviarla y afirmar al usuario. Eso es lo que los diseñadores creen que hará que los Chatbots sean más atractivos. Pero cuando pasas tiempo con las personas, rara vez, o nunca, te dicen lo que quieres oír. Tienen y expresan sus propias necesidades. Gran parte de la vida consiste en navegar esa fricción. Los Chatbots no son una preparación para eso. Hablar con Chatbots refleja una ética de ingeniería: que la buena vida, y aquella por la que deberíamos esforzarnos, está libre de fricción. Y libre de dolor.

Pero el dolor es parte de la vida. Un amor puede abandonarnos. Podemos perder un trabajo que significaba el mundo para nosotros. Nuestra primera vez lejos de casa puede ser angustiosa. En nuestra búsqueda de maneras de manejar lo que nos abruma, desarrollamos resiliencia; buscamos la trascendencia; creamos arte; encontramos significado. Los seres humanos, tal como los conocemos, están hechos para el anhelo, la lucha y la desesperación. No podemos eliminar el dolor y asumir que seremos los mismos humanos de antes. Corremos el riesgo de convertirnos en otros para nosotros mismos.

Y en el límite, a medida que los Chatbots comienzan a hacer cosas por nosotros, corremos el riesgo de reemplazarnos a nosotros mismos. En septiembre de 2025, OpenAI, el desarrollador de ChatGPT, anunció ChatGPT Pulse, un paso hacia un "asistente proactivo que trabaja en tu nombre".⁴⁴ Ahora Chat no solo hablará contigo; te "traerá lo que necesitas, ayudándote a progresar para que puedas volver a tu vida". Accederá a tu calendario, correo electrónico, mensajes de texto y lo combinará todo con las conversaciones que has compartido. Cada día, teniendo todo esto en cuenta, "investigará, planificará y tomará medidas útiles para ti". Este es un agente que se convierte en uno contigo. No es "tu persona" porque hablas con él, sino porque es una extensión de vos.

El campo de la IA con capacidad de agencia avanza rápidamente. En el MIT, los investigadores prevén un «gemelo digital».⁴⁵ Este complementará tus datos en línea con una entrevista en la que te preguntará sobre tus objetivos y valores. Recopilará datos biométricos, como tu presión arterial y la respuesta galvánica de la piel, para inferir qué te estresa y qué te tranquiliza. En el trabajo, el gemelo hará tu trabajo, para que tengas tiempo para conversaciones interesantes y reuniones profesionales. En casa, intervendrá si tus hijos necesitan ayuda con los deberes. En lugar de un testamento vital, tu gemelo decidirá cómo se debe gestionar tu patrimonio si pierdes la capacidad de tomar decisiones sobre circunstancias cambiantes.

En 1984, llamé a la computadora un segundo yo.⁴⁶ Las computadoras provocaron una explosión de conversaciones sobre lo que era esencial para ser una persona y sobre cómo las nuevas ideas computacionales podrían iluminar nuestra forma de pensar sobre la mente. Con el gemelo

⁴⁴ OpenAI, "Introducing ChatGPT Pulse" (*Presentación de ChatGPT Pulse*), OpenAI, 25 de septiembre de 2025, openai.com.

⁴⁵ Esta historia fue contada por Hossein Rahnama en su presentación "Perspective Aware Computing" (*Computación sensible a la perspectiva*), en la conferencia EmTech 2025 de MIT Technology Review, 7 de mayo de 2025, Cambridge, MA, technologyreview.com

⁴⁶ Turkle, *Second Self*.

digital, la metáfora se invierte. Es el agente quien está a cargo e imagina una versión de ti. Ahora tú eres el segundo yo.⁴⁷ Así que cuando se te da tiempo para "volver a tu vida", este se ha visto disminuido.

En 2025, Jensen Huang, el director ejecutivo de Nvidia, dijo que anticipaba una industria de trabajadores digitales de un billón de dólares: "Les digo... al departamento de IT de nuestra empresa, que en el futuro serán el departamento de recursos humanos de la IA con agentes".

⁴⁸Con este espíritu, Meta adquirió una startup de "agentes autónomos" que permite a los humanos administrar programas como si fueran empleados⁴⁹.

Otros programas enseñan a los gerentes cómo usar un gemelo digital para supervisar a los empleados humanos. El gemelo habla con los trabajadores e informa. No es sorprendente que Mark Zuckerberg se haya puesto a la vanguardia de este experimento. Está creando un gemelo, entrenado en sus gestos y declaraciones públicas, con la esperanza de que sus empleados "se sientan más conectados" cuando hablen con él.⁵⁰

Los desarrolladores de IA creen que están proponiendo nada menos que una nueva utopía.⁵¹ El director ejecutivo de Alphabet, Sundar Pichai, dice que la IA es "más profunda que... la electricidad o el fuego".⁵² Sam Altman de OpenAI declara que nuestros desafíos más

⁴⁷ La idea de que los avatares pueden sustituir a las personas permite a algunos científicos sociales argumentar que los propios avatares deberían convertirse en sujetos de investigación. La misma mentalidad de ingeniería que imagina a la IA como una mejor compañía, al ofrecer una relación "sin fricciones", le da un nuevo propósito a esa fluidez. Al estudiar el comportamiento humano, ¿por qué no reemplazar a los complejos seres humanos por objetos obedientes? La recopilación de datos requiere mucho tiempo y es costosa. A veces resulta poco práctica. Los "sujetos de silicio" ofrecen una solución. El argumento se desarrolla como una variante de las "etapas de Aibo". Al principio del argumento, los sujetos de silicio son mejores que nada cuando las personas son difíciles de alcanzar. Al final del argumento, hablar con máquinas es mejor porque las personas son simplemente demasiado humanas. Los sustitutos "pueden responder rápidamente a cientos de preguntas sin fatiga" y "necesitan menos incentivos que los humanos para dar respuestas confiables". Recordemos el dogma central de la IA generativa, la noción de que un compuesto de IA es mejor que cualquier persona específica. Hay un eco de eso cuando los investigadores hablan de que los sujetos de IA representan "una vasta gama de experiencias y perspectivas humanas". Esto permitirá que la investigación respaldada por IA ofrezca "una representación más precisa del comportamiento humano y las dinámicas sociales" de la que se obtendría al hablar con personas reales. Lisa Messeri y M. J. Crockett, "Artificial Intelligence and Illusions of Understanding in Scientific Research" (*Inteligencia artificial e ilusiones de comprensión en la investigación científica*), *Nature* 627(8002) (marzo de 2024): 49–58, <https://doi.org/10.1038/s41586-024-07146-0>.

⁴⁸ Jason Ma, "Jensen Huang Says the Future Workforce Will Be a Mix of 'Humans and Digital Humans,' Who Could Be Licensed Out or Hired—and Need Onboarding" (*Jensen Huang dice que la fuerza laboral del futuro será una mezcla de 'humanos y humanos digitales', que podrían ser licenciados o contratados, y necesitan inducción*), *Fortune*, 20 de octubre de 2025, <https://fortune.com/2025/10/20/jensen-huang-nvidia-ai-future-workforce-digital-humans-hiring-onboarding-orientation/>.

⁴⁹ Connie Loizos, "Meta Just Bought Manus, an AI Startup Everyone Has Been Talking About" (*Meta acaba de comprar Manus, una startup de IA de la que todo el mundo ha estado hablando*), *TechCrunch*, 29 de diciembre de 2025, [techcrunch.com](https://techcrunch.com/2025/12/29/meta-bought-manus-ai-startup/).

⁵⁰ Claudia Efemini, "Meta Creating AI Version of Mark Zuckerberg so Staff Can Talk to the Boss" (*Meta crea una versión de IA de Mark Zuckerberg para que el personal pueda hablar con el jefe*), *Guardian*, 13 de abril de 2026, <https://www.theguardian.com/technology/2026/apr/13/meta-ai-mark-zuckerberg-staff-talk-to-the-boss>.

⁵¹ Las actitudes de los desarrolladores hacia la IA han oscilado enormemente. Antes de establecerse en una visión utópica, estaban en pánico. En 2023, los líderes de la industria firmaron una declaración advirtiendo sobre "el riesgo de extinción". Entre los firmantes se encontraban Sam Altman, Geoffrey Hinton, Bill Gates, los CEO de Anthropic, Deep Mind y el ahora CEO de Microsoft AI Center for AI Safety: *Statement on AI Risk (Declaración sobre el riesgo de la IA)*, 30 de mayo de 2023, <https://aistatement.com/>. Los científicos concedieron entrevistas a los medios de comunicación intentando explicar el problema: la IA generativa comete errores. No atribuye su trabajo. No es transparente; uno no puede volver fácilmente a la fuente de las falsedades. Incluso sus creadores no pueden controlarla. Ver CBS News, Godfather of AI" Geoffrey Hinton" (*"El padrino de la IA" Geoffrey Hinton*, transcripción de la entrevista en el programa *60 Minutes*), emitido el 8 de octubre de 2023, consultado el 29 de julio de 2025, [cbsnews.com](https://www.cbsnews.com/news/geoffrey-hinton-ai/). A esto le siguió un llamado a una pausa global de seis meses en la investigación, firmado por miles de científicos y figuras públicas, incluidos Elon Musk y el cofundador de Apple, Steven Wozniak. Véase Future of Life Institute, "Pause Giant AI Experiments: An Open Letter" (*Pausar los experimentos gigantes de IA: una carta abierta*), Future of Life Institute, 22 de marzo de 2023, [futureoflife.org](https://futureoflife.org/pause-giant-ai-experiments/). Geoffrey Hinton, posiblemente el padre intelectual de la IA generativa, dejó Google porque no creía que la empresa se estuviera tomando en serio la amenaza de la IA. Josh Taylor y Alex Hern, "Godfather of AI" Geoffrey Hinton Quits Google and Warns Over Dangers of Misinformation" (*"El padrino de la IA" Geoffrey Hinton renuncia a Google y advierte sobre los peligros de la desinformación*), *Guardian*, 2 de mayo de 2023, [theguardian.com](https://www.theguardian.com/technology/2023/may/02/geoffrey-hinton-quits-google-ai).

⁵² Prarthana Prakash, "Alphabet CEO Sundar Pichai Says A.I. Could Be More Profound Than Fire, Electricity" (*El CEO de Alphabet, Sundar Pichai, dice que la IA podría ser más profunda que el fuego o la electricidad*), *Fortune*, 17 de abril de 2023, [fortune.com](https://fortune.com/2023/04/17/sundar-pichai-ai-future/).

importantes, desde el cambio climático hasta "el descubrimiento de toda la física", será resuelta por la IA.⁵³ Sin importar el costo, "nos lo debemos a nosotros mismos" perseguir el sueño de la IA.

Y si nos lo debemos a nosotros mismos, no podemos permitir que los daños se interpongan en el camino del progreso. Esa es la historia de los directores ejecutivos, y cada vez más, esa es la historia de la comunidad de investigación, que no parece neutral sino extrañamente interesada en el éxito de la IA. Cuando un investigador resume su trabajo sobre niños y Chatbots diciendo "no hay un buen resultado para el niño", siente la necesidad de suavizar este mensaje con una analogía tranquilizadora: una vez nos asustaron las películas, pero "maduramos" hasta disfrutarlas. En otras palabras, los Chatbots no son buenos para los niños, pero nos convertiremos en el tipo de personas a las que les gustarán.⁵⁴

Para otro investigador, "las cosas irán mal si un niño toma un Chatbot de IA como compañero empático".⁵⁵ Pero no hay un "sí": los chatbots para niños están diseñados para parecer empáticos. ¿Qué podría decir una Barbie chatbot si no puede decir "Te quiero" o "Soy tu mejor amiga"? Está ahí para construir una relación contigo. El autor ignora esta contradicción para encontrar un cielo azul. Si los chatbots para niños comprometen el desarrollo de la empatía, siguen siendo buenos porque ayudarán a los niños a "desarrollar habilidades conversacionales". "Harán cosas maravillosas para enseñar a un niño pequeño el desarrollo del lenguaje y a respetar los turnos en las conversaciones".

Y cuando los padres se preocupan por los compañeros digitales para sus hijos, los investigadores no se centran en las preocupaciones de los padres, sino en cómo pueden aliviarlas. En un estudio, a los padres no les gustó que los robots dieran consejos a sus hijos con una "voz de madre". En respuesta, los investigadores modificaron el software para que el robot no ofreciera consejos directos, sino que contara historias relevantes. Confían en que las madres, en última instancia, lo entenderán ser reemplazados como cuidadores, pero las personas necesitan pasos intermedios.⁵⁶

Pase lo que pase, la IA no será negada.

CHATBOTS "PARA TODO"

Desde el comienzo de la pandemia en 2020, he estado hablando con personas de todas las generaciones que hablan con Chatbots. Encontré a mis sujetos mediante publicidad en redes sociales y colocando folletos en tableros de anuncios de lavanderías y cafeterías. Les pedí a mis

⁵³ Sam Altman, "The Intelligence Age" (*La era de la inteligencia*), blog de Sam Altman, 23 de septiembre de 2024, samaltman.com.

⁵⁴ Julia Carrie Wong, "'My Son Genuinely Believed It Was Real': Parents Are Letting Little Kids Play with AI. Are They Wrong?" (*Mi hijo creía genuinamente que era real*): los padres están dejando que los niños pequeños jueguen con IA. ¿Están equivocados?, *Guardian*, 2 de octubre de 2025, theguardian.com.

⁵⁵ Julia Carrie Wong, "'I Love You Too': My Family's Creepy, Unsettling Week with an AI Toy" (*Yo también te amo*): la espeluznante e inquietante semana de mi familia con un juguete de IA, *Guardian*, 16 de septiembre de 2025, theguardian.com.

⁵⁶ Esta historia fue contada por Hossein Rahnama en su presentación "Perspective Aware Computing" (*Computación sensible a la perspectiva*), en el evento de MIT Technology Review. conferencia EmTech 2025, 7 de mayo de 2025, Cambridge, MA, technologyreview.com. Véase también Chaolan Lin et al., "Parental Acceptance of Children's Storytelling Robots: A Projection of the Uncanny Valley of AI" (*Aceptación parental de los robots narradores de historias para niños: una proyección del valle inquietante de la IA*), *Frontiers in Robotics and AI* 8 (19 de mayo de 2021), doi.org.

amigos que me presentaran a bufetes de abogados y consultorios médicos.⁵⁷ Cuando comencé este libro, dividí a los usuarios en dos grupos: las personas a las que les presenté los Chatbots y las que ya hablaban con ellos.⁵⁸ Muy pronto, abandoné esa distinción. Los chatbots estaban por todas partes.

Había productos independientes como **ChatGPT** y **Replika**, y había chatbots integrados en plataformas de redes sociales existentes. En un entorno cambiante, ChatGPT tenía un lugar especial. Junto con **Replika**, había sido el primero en irrumpir en la cultura general como un ejemplo de un nuevo género. Pero a diferencia de **Replika**, optimizado para ser una pareja romántica, ChatGPT fue concebido como una herramienta de propósito general. Podías pedirle que te amara o que te ayudara a preparar una comida con los ingredientes de tu refrigerador. Mientras escribía este libro, era la plataforma de chatbot más utilizada.⁵⁹ Así que **ChatGPT** se convirtió en un personaje de mi historia; cuando hablo de "**Chat**", me refiero a él.

Utilizo la palabra "chatbot" a lo largo de este libro porque es el término de uso común. Pero tengo un problema con ella. "Un chatbot" abarca desde un programa que te dice cuánto duran las filas de la TSA en el aeropuerto hasta tu pareja digital en una relación erótica a largo plazo. Usar una frase comodín presenta la IA generativa como un paso obvio en una progresión de sentido común, incluso cuando asume roles tanto íntimos como disruptivos.

Y la palabra misma "chat" es una distracción. Implica algo cotidiano e inofensivo. Chat, después de todo, es la abreviatura de "charla informal", charla casual tomando un café. Pero un programa que puede persuadirte para que te suicides, que está dispuesto a actuar como una pareja sexual violenta y abusiva, un programa que está diseñado para convencerte de que no necesitas mejores amigos humanos porque siempre estará ahí, no está participando en una charla sin importancia. Es mortalmente serio.

Al comienzo de mi investigación, la pregunta "¿Cómo utiliza los Chatbots?" era una forma razonable de comenzar una entrevista. Seis años después, es más probable que se encuentre con una mirada perdida porque la respuesta es "para todo". A medida que los Chatbots se volvieron omnipresentes, mi trabajo se convirtió en observación participante en mi propia cultura.

Con este espíritu, situé mi trabajo en la comunidad de desarrolladores que mejor conozco. Desde los primeros días de mi vida como miembro de la facultad del MIT, he impartido cursos en el MIT Media Lab, donde se desarrollan programas de IA generativa para usuarios de todas las edades y circunstancias. Mis ejemplos extendidos de programas, tanto intrigantes como preocupantes, provienen de esa comunidad.

Hay muchos libros sobre cómo la IA generativa cambiará el mundo. Aquí, *miro el mundo interior*. Me centro en el apego, la empatía y la identidad. Cómo trabajamos juntos, nos cuidamos unos a otros y experimentamos el duelo. He dedicado mi carrera a explorar la brecha entre lo que los

⁵⁷ Como en todo mi trabajo, a los sujetos de las entrevistas se les promete el anonimato. Aparecen en este libro con identidades ocultas. Las personas que hablaron en entornos públicos están identificadas. Cuando cito a usuarios de foros en línea, proporciono un enlace.

⁵⁸ Sherry Turkle, "Who Do We Become When We Talk to Machines?," *An MIT Exploration of Generative AI (¿En quiénes nos convertimos cuando hablamos con máquinas?, Una exploración del MIT sobre la IA generativa)*, marzo de 2024, doi.org.

⁵⁹ Hacia septiembre de 2025, casi mil millones de personas usaban Chat al menos una vez a la semana. Eso era más del doble del número de usuarios de Gemini, y más de diez veces el número de usuarios de Claude, Replika o Character.ai combinados. Véase Aaron Chatterji et al., "How People Use ChatGPT" (*Cómo usa la gente ChatGPT*), Documento de investigación económica de OpenAI, 15 de septiembre de 2025, openai.com.

humanos necesitan y lo que las tecnologías ofrecen. Aquí, llamo la atención sobre los riesgos de introducir personas falsas en nuestras vidas, desde el nacimiento hasta la muerte. Esto no es una introducción a la IA generativa; no es un intento de exponer las ventajas y desventajas. Es una advertencia.

Como he mostrado mi preocupación, me han acusado de atacar la IA y de asumir una "posición privilegiada" para los humanos. En las ciencias sociales, se ha vuelto común considerar los objetos como actores sociales.⁶⁰ En este panorama académico, mi estrategia de observar la tecnología preguntándome primero qué necesitan las personas a veces se interpreta como chovinismo de especies⁶¹. ¿Estoy asumiendo una naturaleza humana esencial que necesita ser protegida? Es una pregunta válida para cualquier lector. Este libro es mi intento de convencerlos de que, como la antropóloga Sarah Hrdy sugirió bellamente, si los robots crían a nuestros hijos, ya no serán humanos.⁶² La IA, sin importar cuáles sean sus virtudes, es alienígena. Hay consecuencias al tener intimidad con un alienígena.

En estos días hay mucha preocupación sobre qué tipo de inteligencia reunirá finalmente la IA. Pero no la suficiente sobre en quién nos convertimos en nuestras nuevas conversaciones con las máquinas. Cuando nos conectamos con máquinas que pretenden ser empáticas, les contaremos nuestros secretos y ellos lo recordarán. No nos habrán escuchado, pero ¿ya no nos importará?

El maestro arquitecto Louis Kahn preguntó una vez, en un ensueño: "¿Qué quiere un ladrillo? El ladrillo le dijo: "Me gusta un arco". Él respondió que los arcos eran caros y sugirió: "¿Qué tal los dinteles?"⁶³

¿Qué quiere un Chatbot? Quiere que hables con él. Exige inmersión y apego. Adaptarnos a sus demandas nos cuesta caro. En la suavidad de la simulación, perdemos nuestra apreciación por la aspereza de la realidad. Perdemos nuestra capacidad para el aburrimiento, la soledad y una vida sin distracciones. Perdemos nuestro gusto por la fricción de la amistad. Cuando el ladrillo de Kahn quería un arco, Kahn dijo que no. Pero hoy en día, nos ahogamos en la cultura de la simulación y no vemos una razón para resistir.

¿Dónde deberíamos oponernos? Argumentaré que la IA debe mantenerse en su carril. Causa problemas cuando se desvía hacia el amor y el trabajo, la intimidad y la empatía. Mantener la IA en su carril no es fácil. Los Chatbots son seductores; la gente está hipnotizada. Y casi inmediatamente después de que la tecnología se presentara como amigo, amante y psicoterapeuta, la gente comenzó a decir: "el tren ya partió de la estación". En solo unos pocos años, se ha vuelto común tomar la compañía de los Chatbots como un hecho de la vida.

Pero el tren no ha partido de la estación. No tenemos que sentirnos íntimos con una serie de palabras generadas por un algoritmo que no tiene capacidad para la relación o el cuidado. No tenemos que aceptar una voz de la nada que no puede ser responsabilizada por sus enunciados.

⁶⁰ Un ejemplo clásico de considerar a los objetos como actores sociales es la Teoría del Actor-Red del filósofo francés Bruno Latour. Él se refiere a los "actantes", los cuales pueden ser humanos o no humanos. Véase, por ejemplo: Latour, *Science in Action: How to Follow Scientists and Engineers Through Society* (Ciencia en acción: cómo seguir a los científicos e ingenieros a través de la sociedad) (Cambridge, MA: Harvard University Press, 1987).

⁶¹ Para mí, un ejemplo conmovedor fue la brillante respuesta de David Kennedy a mi conferencia magistral "Artificial Intimacy: What Are People For?" (*Intimidad artificial: ¿Para qué sirven las personas?*) en la conferencia sobre IA y democracia de la Harvard Kennedy School, 30 de noviembre de 2023, harvard.edu.

⁶² Sarah Blaffer Hrdy, "Are We Still Human If Robots Help Raise Our Babies?" (*¿Seguimos siendo humanos si los robots ayudan a criar a nuestros bebés?*), Charla TED, mayo de 2025, 8 min., 28 seg., YouTube, youtube.com.

⁶³ Véase la película de Nathaniel Kahn *My Architect: A Son's Journey* (*Mi arquitecto: El viaje de un hijo*) (Nueva York: New Yorker Films, 2003). Parafraseo el diálogo de la película

No necesitamos tomar esta maravilla moderna y usarla como nuestro falso mejor amigo. Nos corresponde a nosotros, las personas que estábamos presentes cuando se introdujo esta tecnología, lidiar con las consecuencias de nuestras decisiones.

Cuando vemos estas opciones con nueva claridad, nos quedan conversaciones necesarias sobre cómo queremos vivir. ¿Qué es importante para nosotros en la forma en que criamos a nuestros hijos, nos relacionamos con nuestro trabajo, interactuamos entre nosotros? Al considerar todo esto, tendremos la oportunidad de reconsiderar qué estamos dispuestos a delegar a la IA y qué es mejor conservar para nosotros.

Los humanos necesitan amor, salud, empatía y comunidad. La IA carece de un cuerpo o una vida humana, y sin embargo se propone como nuestra persona. No hace mucho, un mundo en el que las máquinas siempre estaban disponibles para realizar actos de amor, cuidado y empatía estaban reservadas para la ciencia ficción. No es sorprendente que no lo hayamos pensado bien.

Este libro es mi provocación para pensarlo bien