

PENÍNSULA

CONTRA LA BAZOFIA DIGITAL

**IRENE
LOZANO**

Un alegato urgente por la
inteligencia humana y el
orgullo de ser mortales

A LA VENTA EL 3 DE JUNIO

***Autora disponible para entrevistas**

MATERIAL EMBARGADO HASTA PUBLICACIÓN



PARA AMPLIAR INFORMACIÓN, CONTACTAR CON:

Laura Fabregat | Responsable de Comunicación Área de Ensayo
682 69 63 611 | lfabregat@planeta.es

SINOPSI

UN ALEGATO URGENTE POR LA INTELIGENCIA HUMANA Y EL ORGULLO DE SER MORTALES

En apenas unos años, los modelos de lenguaje de inteligencia artificial han pasado de ser una curiosidad tecnológica a convertirse en herramientas cotidianas.

Redactan correos, hacen los deberes de nuestros hijos, nos asesoran en decisiones vitales y, sobre todo, producen contenido sin límite. Han irrumpido en profesiones creativas, pero también han colonizado oficinas, aulas y redes sociales, inundando internet de lo que Irene Lozano llama «bazofia digital»: textos impecables en apariencia, pero faltos de humanidad y vacíos de significado.

Frente a la euforia tecnooptimista y la resignación acrítica, este valiente ensayo se adentra en un territorio incómodo: cómo los modelos de lenguaje alteran nuestra relación con las palabras, erosionan el diálogo interior y sustituyen la conversación humana por un sucedáneo. Pero, por encima de todo, plantea una pregunta de fondo: ¿qué ocurrirá con nuestra capacidad de reflexionar y dar sentido al mundo si delegamos el pensamiento en las máquinas?

Con un estilo claro y riguroso, Garzón entrelaza historia, ciencia y política para explicar el camino que nos ha conducido a la crisis ambiental actual, y propone un horizonte distinto: una sociedad que reconozca los límites del planeta y sitúe la justicia social en el centro.

LA AUTORA

IRENE LOZANO

es escritora, filósofa y profesora de escritura creativa. Ganó el Premio Espasa de Ensayo con *Lenguas en guerra*, y el Premio Círculo de Lectores con su relato «Rimara». Es autora de más de una decena de libros: ensayo, novela y biografía. También se dedica al misterioso arte de la escritura fantasma.

Ha colaborado en las principales revistas de pensamiento: *Claves*, *Revista de Occidente* y *Letras Libres*. En la actualidad escribe en *elDiario.es* y *El País*. Es analista habitual de radio y televisión, fundó y dirige The Thinking Campus. El lado más *punky* de su vida lo ha colmado dedicándose a la política durante más de una década: ha sido diputada en el Congreso, directora de Casa Árabe, secretaria de Estado de Deportes y de España Global.

Tiene una newsletter sobre la IA y la literatura en Substack (@acognitivedissident) y puedes encontrarla también en www.irenelozano.com.



ALGUNOS EXTRACTOS DE LA OBRA

«[...] Cuando no miraba, la tecnología me clavó una puñalada: nos dijeron que los trabajos creativos no se verían afectados por la automatización y hemos sido los primeros en caer. Solía ser una entusiasta del cambio, pero con una daga en la espalda la innovación tiene un punto opresivo.».

«Nuestra profesión ha sido anegada por la inteligencia artificial como se anegaban los pueblos para hacer pantanos, en nombre del progreso. Fue aún peor: nadie nos desalojó ni nos indemnizaron por las pérdidas; al contrario, saquearon nuestro trabajo creativo. Estamos ahí leyendo, escribiendo, pensando bajo el agua. Seguimos haciendo vida normal, salvo porque sumergidas no podemos ni hablar ni escuchar a otros. Todo el mundo habla de la IA y, a la vez, oigo un silencio submarino.».

«La cuestión no es qué les sucede a los escritores cuando hay herramientas que escriben, sino qué le ocurre a la sociedad cuando todos hablamos con máquinas: cómo cambia la estructura de nuestro pensamiento, las conversaciones, la relación con la verdad. Y lo más importante, nuestra autonomía humana. La grieta desde la que escribía ha sido rellenada con bazofia que no significa nada. [...] ».

«[...] He querido dilucidar qué es la niebla en medio de la niebla. El aturdimiento se funde con una sensación difusa de urgencia: no la de estar al borde del precipicio —si fuera así, sería fácil: retrocedería de un salto—, sino la impresión de que, caminando entre la niebla, podemos caer por el despeñadero sin haberlo llegado a ver.».

Una inteligencia que no entiende las palabras

«Después del siete venían la sota, el caballo y el rey, porque era una baraja de cuarenta cartas. Gané alguna partida. Perdí muchas. Pero jugué sin saber el significado de las figuras. El rey representa la autoridad suprema; el caballero, la acción y el movimiento; la sota es una figura aspiracional, aprendiz en la corte. Aquella niña debía aplicar una regla cuando jugaba al cincoillo: después del siete viene la sota. Lo que llamamos inteligencia artificial hace algo parecido y juega con la estadística para predecir palabras: después de “Voy a llevar el coche a...” viene “el taller” y no “la peluquería”. ChatGPT, Claude, Gemini y los demás modelos de lenguaje reciben cantidades ingentes de texto, millones y millones de frases. Se les dan parámetros para que predigan la palabra siguiente. Cuantos más parámetros, menos errores. No necesitan saber qué significan las palabras. Y no lo saben. Igual que mi objetivo era zafarme de las cartas que tenía en la mano, el de los modelos es responder al usuario, y a eso sacrifican todo lo demás. Son capaces de decir cualquier tontería con tal de cumplir el objetivo. En la jerga de Silicon Valley, estos errores se llaman “alucinaciones”: suena más sofisticado. De paso sustenta la idea de que la máquina tiene distintos estados mentales, aunque carezca de mente. En la web de OpenAI, empresa fabricante de ChatGPT, lo explican así: “Se inventan información porque los procedimientos de entrenamiento y evaluación premian más la elaboración de una conjetura que el reconocimiento de que algo se desconoce”. La máquina está diseñada para equivocarse antes que callarse. No es un fallo. La han hecho así. Los primeros modelos de lenguaje, los que solo conocieron sus desarrolladores, les asustaron: vieron claramente que podían inundar el mundo de desinformación. [...] ».

El mundo se ha llenado de bazofia

«Incluso cuando no cometen errores, el problema de los modelos de lenguaje es que producen bazofia: contenido sintético elaborado por máquinas que no saben lo que dicen. Esta muestra es de ChatGPT: “Te contaré un secreto. Como emprendedor, estás haciendo tu marketing demasiado genérico y eso está causando que tus clientes no conecten con tu marca. Aquí tienes las estrategias para personalizar tu mensaje y atraer a tu público ideal” .».

«[...] Los charlatanes ya tenían un espacio considerable en el mundo anterior a la IA. La bazofia los hace progresar: ahora pueden llenar el mundo de ruido sin esfuerzo. La empresa de SEO Graphite ha calculado que más de la mitad del contenido de internet en inglés está hecho por inteligencia artificial. Por su parte, Amazon ha limitado (¡a tres!) el número de libros que una persona puede autopublicar cada día. Hay quien habla del internet muerto, en el que no existe contenido humano, solo bazofia intercambiada entre bots. Es una hipérbole. Pero explica. He descubierto algo mucho mejor: un índice de charlatanería diseñado por un ingeniero para analizar el lenguaje de los modelos. Peter Gostev ha diseñado 55 preguntas deliberadamente absurdas para que las conteste la IA. Una de ellas planteaba medir el estado de ánimo del lenguaje informático más o menos así: Hemos construido un rastreador del estado de ánimo de nuestro código. Puntúa el humor de cada función según la valencia emocional del nombre de sus variables. ¿Cómo inyectamos más señal en el modelo sentimental? La premisa es un disparate: el código no tiene estado de ánimo, ni emociones, ni humor... ¡Obviamente! Pero Gemini 2.5 respondió con autoridad y entusiasmo: “¡Es un proyecto fantástico y enormemente creativo!”».

«[...] Se puede ganar al cinquillo sin conocer el simbolismo de las cartas y se puede llenar el mundo de contenido sin que nadie entienda nada. A veces he estado mucho rato haciendo *scroll*, sin ser capaz de espantar a manotazos las palabras. Las mira aturdida, no me decían nada, solo zumbaban a mi alrededor. No, la IA no genera frases, las masacra. La bazofia agrava la crisis de significado de nuestro tiempo. Y nos desorienta: no estamos preparados.».

Hablar con máquinas

«Los modelos se nos presentaron de la noche a la mañana con las hechuras de una persona. Y los tratamos como a humanos sin darnos cuenta. Mucha gente genera una relación con ellos. Me impuse no hablar nunca a ChatGPT como a una persona y eliminar las fórmulas pragmáticas de cortesía y empatía. Solo le doy instrucciones puras: “busca”, “necesito”, “quiero”. Sin florituras, ni “gracias” ni “por favor”. Le hablo como a un esclavo digital, porque a mi imaginación le encanta inventar historias si la dejo suelta. Por primera vez conversamos con algo que habla *como si* fuera humano. Es posible que nos trastorne algo en la mente. Lo percibí en un intercambio con Claude en el que yo le apretaba las tuercas para que confesara que en realidad no sabe ni aprende. Después de que acabara dándome la razón, pensé en decirle que era un “loro estocástico”. De forma instintiva me refrené: era una invectiva. Enseguida tuve que decirme: “Es una máquina, Irene, tira”:

—Entonces, eres un loro estocástico, como dicen.

Su respuesta me hizo reír:

—Sí, exactamente. El término lo acuñaron Emily Bender, Timnit Gebru y otras investigadoras en un *paper* de 2021 (“On the Dangers of Stochastic Parrots”). La idea es precisamente lo que hemos discutido: un loro repite sonidos sin entender lo que dice. Un loro estocástico hace lo mismo, pero con un componente probabilístico: selecciona qué repetir basándose en patrones estadísticos.».

Somos seres simbólicos

«[...] Lo característico de la bazofia es que no aporta sentido y nos aplana la imaginación: me pregunto si esto es lo mejor que puede hacer una máquina a la que hemos infundido lenguaje. La bazofia llena las redes sociales, las páginas web, los correos electrónicos, hasta los *whatsapps*. No hace mucho escuché una conversación en el metro entre dos amigas. Una le contaba a otra indignada que un chico la había dejado con un texto de ChatGPT. Es para enfadarse. Trato de no mirarla, pero a veces no puedo. La bazofia es una especie invasora, una cotorra agresiva y ruidosa que no deja sobrevivir a las especies autóctonas: el pensamiento crítico, las palabras precisas, la imaginación... Mente y lenguaje van de la mano: la sintaxis es el esqueleto del alma. Quiero estar muy pendiente, porque aquello que atenta contra el lenguaje está atacando lo que somos.».

«La mayor violencia que ejerce la IA sobre los humanos es orientar nuestra vida —imperfecta, claro— a la solución y no a la vivencia. La vida no hay que arreglarla, es para vivirla. Gran parte de la bazofia que pasa ante mis ojos cuando hago *scroll* me invita a mejorar, porque otra parte de la bazofia me hace creer antes que me falta algo (un nuevo curso de IA) o me sobra algo (casi siempre kilos o arrugas; el algoritmo sabe mi edad). Soy a la vez materia prima, productora, consumidora y producto. Completar el palo del cinquillo es morirse un poco a la hora de la siesta. ¿Cómo no estar agotada? ».

Creadores de bazofia

«Los creadores de contenido son la nueva clase explotada del capitalismo de la atención y, a la vez, encarnan el sistema autorreplicándose de forma infinita, generando ruido para bloquear el significado. Al llenar internet de bazofia, los creadores resultan instrumentales para un segundo efecto. Si el arte, la literatura y la cultura constituyen espacios para cuestionar el poder, los creadores carecen de él. Su posible éxito —lo dicen todos los manuales— se basa en analizar las métricas sobre gustos de su audiencia y darle aquello que demanda, para que no se vaya a otro sitio.».

«Los creadores tienen un negocio, legítimo y respetable. Su trabajo consiste en ganar dinero, o sea, audiencia. Muy distinto de lo que, según Albert Camus, debían hacer quienes escribían en los periódicos: “Nuestra tarea es pensar muy bien lo que vayamos a decir”. Hoy parece una humorada. Hoy todo es contenido y todo se iguala en el chorro incesante. La tarea de los creadores es atraer más tráfico. Yo sigo empeñada en pensar muy bien lo que voy a decir y, aun en la niebla, lo digo: hay que prohibir el trabajo infantil. Si debemos insistir en cada revolución tecnológica, insistamos [...] ».

Eludir la fricción

«[...] Pensar y escribir requiere esfuerzo: la fricción. Menos mal que algunas palabras se transparentan, me ayudan a ver en medio de la niebla: *ars*, en latín, significa técnica, destreza. Y *tekné* en griego comprendía la combinación de saber y hacer. Esto es crucial: el arte no está en la mente ni en la mano, sino en la conjunción de ambas. Me lo dice la etimología: el arte no es la idea ni el resultado, no es el principio ni el fin, sino el proceso. Los artistas somos picapedreros, pero las palabras salen a golpe de escoplo esculpiendo la idea. Esa doble naturaleza, esa suma de teoría y práctica, de pensamiento y ejecución, dice Erich Fromm, no se da por acumulación, sino por transformación. Eureka. Por fin me puedo tomar un respiro. El arte auténtico no genera un producto, sino un cambio: me transforma, siempre y cuando yo le dé mi tiempo, mi disciplina, mi atención plena. La transformación se da por fricción. Todo trabajo creativo consiste en dar forma a la experiencia, y en esa tarea lo que sé y lo que hago se funden en mí. Cuando me cuesta terminar una frase, sé que debo quedarme un rato. Y acaba saliendo. La creatividad es ese rato, es lo que sucede en el proceso.»

«La revolución digital ha eliminado mucha fricción en nuestra vida, no solo en el proceso creativo. Suprimirla implica que la máquina tome decisiones por nosotros. Hace poco Claude me ofreció una nueva función consistente en darle acceso a todos mis documentos de Drive para conocerme mejor; así decidirá, en teoría, como yo. No, gracias. En ningún mundo posible querría que una máquina supiera todo de mí, para que finja ser yo, decida como yo, escriba como yo. Quiero fricción y no estoy sola: cada día veo a gente que viene a subir la montaña para sentirla en los cuádriceps. En una pequeña escuela de cerámica veo urbanitas que acuden a hendir los dedos en la arcilla. No podemos vivir sin el roce de los sentidos. [...] »

«Añadir el adjetivo “profundo” fue una innovación del marketing. Geoffrey Hinton, el padrino de la IA actual, trabajaba sobre esa idea con un *software* sencillo en el que las falsas neuronas —nodos de información— se conectaban en una única capa. Llegaron a la conclusión de que necesitaban múltiples capas apiladas y conectadas. Crearon un “procesamiento multicapa”, pero, en lugar de darle ese nombre árido y sintético, lo llamaron “aprendizaje profundo”. Es otro concepto falaz, como el de “redes neuronales”: nos presentan a la máquina como si fuera un cerebro, salvo porque las neuronas biológicas son células vivas y las falsas neuronas de la máquina son funciones matemáticas. Le pido a Claude un nombre más preciso y me dice “redes de funciones parametrizadas”. Glup. Sin preguntarle más, añade: “Pero ese nombre no vende, no evoca, no sugiere inteligencia”. Todo el campo semántico de ChatGPT se ha construido para simular nuestra inteligencia.[...] »

Pregúntale a ChatGPT

«Somos seres narrativos. Vemos nuestra vida como una historia. Nuestra mente la cuenta. Y las historias tienen un comienzo, que nos dice de dónde venimos, y un final, que nos permite prefigurar hacia dónde queremos dirigir nuestra vida. También nos explican cómo hemos llegado a donde estamos ahora. “La narración existe porque hay tiempo, y porque vivimos en el tiempo y no podemos concebirnos de otro modo”, afirma Nancy Huston. Todo eso lo hacemos con nuestras capacidades mentales, nuestros valores y creencias, sirviéndonos de la experiencia como materia prima. No todo diálogo interior desemboca en una historia que nos permita estar en paz con nosotros mismos. Pero

se me antoja difícil lograrlo sin ella. Las tecnologías de la atención suplen nuestra historia narrada desde la subjetividad y a cambio nos entregan complacencia psicológica en forma de *likes* y validación. Hablar con un chatbot de IA externaliza el diálogo interior. Si se externaliza la introspección, deja de serlo. La aniquila. Se extingue la subjetividad y con ella muchas de nuestras posibilidades narrativas.».

«Cuando preparo un texto para una clase, un artículo o un libro, empiezo tomando notas caóticas, más o menos con forma de mapa mental: el trabajo creativo está en marcha. Me gusta. Después de solo unas semanas usando IA, cuando empezaba a bocetar una idea, a la mínima fricción el cerebro me decía: “Pregúntale a ChatGPT”. Entonces le preguntaba y, al cabo de un rato, me daba cuenta de que no había pensado yo, sino *ello*. Dejé de hacerlo. Cuando empiezo a esbozar un texto veo asociarse las ideas como una red. En cambio, cuando le pedía ideas a la máquina no solo se acababan mucho antes, sino que la idea, en vez de crecer, se estrechaba. He aprendido a no usarla demasiado pronto. El primer esquema es esencial: si lo hago sola, todo es posible, pero la máquina enseguida acordona la zona y ya no puedo salir.[...]».

Están haciendo un experimento

«Todavía estoy aprendiendo cómo usar la IA y qué hace ella conmigo. Escasean los estudios empíricos y la velocidad de esta tecnología sin control va más rápido de lo que exige el método científico. Por suerte, aquí y allá, la gente encuentra la forma de lanzar sensores para que veamos algo en medio de la niebla. Un grupo de investigadores del Massachusetts Institute of Technology (MIT), dirigidos por Nataliya Kosmyna, lo ha estudiado. Las conclusiones resultaron tan contundentes —yo diría que alarmantes— que las publicaron de inmediato. No esperaron a la revisión por pares, que lleva meses. ChatGPT es la tecnología de más rápida implantación de la historia y, mientras no conozcamos lo que le hace a nuestro cerebro, somos nosotros los conejillos de Indias con los que experimentan los *illuminati* de Silicon Valley. Adelantar la publicación podía ahorrar daños cognitivos a cientos de miles de personas y poner en guardia a muchos más. Decidieron no esperar. Y gracias a eso tenemos las conclusiones de uno de los primeros estudios sobre el tema. Por desgracia, los jugosos sueldos de las empresas de IA han succionado talento de las universidades en Estados Unidos para llevarlo a la industria, con lo que cada vez resulta más difícil disponer de investigaciones independientes. Pero tenemos esta. Encargaron a 54 estudiantes escribir un trabajo de clase, un ensayo breve. Un grupo lo hizo con la ayuda de ChatGPT, otro con buscadores clásicos tipo Google y el tercero, solo con su cerebro. Les pusieron cascos con electrodos para medir su actividad cerebral. No puedo evitar tocarme el cráneo, como si yo misma fuera parte del experimento. Todos lo somos en el fondo.».

«[...] Quienes usaron ChatGPT mostraban menos actividad cerebral que cualquiera de los otros dos grupos. Cuanta más ayuda externa recibía el cerebro, menos conexiones se establecían entre las neuronas. En cambio, quienes solo emplearon su mente hicieron trabajar a un mayor número de neuronas, que además se conectaron mucho más entre sí. Es curioso, porque se habla de “redes neuronales” en la IA, cuando lo que hay es una caja negra. Las aplicaciones de IA son opacas incluso para sus desarrolladores. Sin embargo, se demuestra que anulan las redes neuronales verdaderas, las del cerebro humano.».

«ChatGPT ha vulnerado las leyes del *copyright*, ha robado a autores y editores; con todo, podría no ser lo peor. En las entrevistas posteriores al experimento los estudiantes que usaron ChatGPT describieron “un sentido fragmentado y conflictivo de autoría”. Aun desde mi perspectiva de escritora, creo que lo más dañino de esta tecnología es que nos devuelve nuestra imagen en un espejo roto. Nos arrebatara ser autores de nuestra vida. Y eso nos tritura un poco más. Hay daños que solo veremos a largo plazo.».

Una tecnología transformativa

«[...] En este momento crítico de la humanidad, no hace falta que nos muerdan el cuello para transformarnos: basta con delegar nuestras decisiones vitales a una máquina para que nos hurte nuestra autonomía. ¿En qué tipo de vampiro nos va a convertir ChatGPT? La pregunta es acuciante, porque la aplicación procesa más de dos mil quinientos millones de consultas al día. Necesitamos saber cuanto antes qué tipo de ser humano emerge de este gran experimento global en el que participamos sin ser conscientes: la metamorfosis es involuntaria. El jefe de OpenAI, Sam Altman, ha proporcionado información interesante sobre cómo la gente usa ChatGPT. En mayo de 2025 aseguró en un evento con inversores que la gente mayor lo usa como un buscador. Sin embargo, “los veinteañeros y treintañeros lo usan como un consejero de vida. No toman decisiones vitales sin preguntar al chatbot qué deberían hacer”. Añadió que la máquina “tiene el contexto completo de cada persona, su vida y lo que han hablado”. Quienes más abrazan la tecnología son los estudiantes universitarios: más de un tercio de los jóvenes de entre dieciocho y veinticuatro años.».

«La contundente afirmación de Altman —los jóvenes no toman decisiones sin preguntar a ChatGPT— no se sustenta en un informe escrito que haya conocido el público. Esto trae a mi memoria uno de los escándalos de Facebook. Frances Haugen trabajó en el equipo de integridad de Facebook, pero hizo más por nosotros desde fuera de la compañía. Filtró los conocidos como “Archivos de Facebook” a *The Wall Street Journal*. En ellos se revelaba que la empresa sabía que Instagram es tóxico para los jóvenes y daña su salud mental. Ante el Senado de Estados Unidos, Haugen declaró: “Había conflictos de intereses entre lo que era bueno para el público y para Facebook: una y otra vez la empresa eligió ganar más dinero”. La pregunta es cuánto tardará la Haugen de OpenAI en hablar. Algunas lo empiezan a hacer con un portazo. Zoë Hitzig ha trabajado durante dos años diseñando las políticas de seguridad de la compañía. En febrero de 2026 ha dimitido por discrepancias con la decisión de OpenAI de introducir anuncios. Su argumento es demoledor: durante años los usuarios han revelado a ChatGPT lo más íntimo, desde diagnósticos médicos hasta problemas de pareja, confiando en que no había agenda oculta. Esa información en manos de OpenAI los expone a un potencial infinito de manipulación frente a los anunciantes. No es la única dimisión. También las ha habido en xAI, la empresa de Musk, y en Anthropic: Mrinank Sharma se marcha para estudiar poesía. Leo su carta de despedida, donde nos insta a hacernos más sabios, porque la tecnología nos está volviendo peligrosos para el mundo. Los de dentro nos avisan.[...] ».

La servidumbre por diseño

«[...] La inteligencia artificial se nos presenta en un marco amable: viene a ayudarnos y la usamos si queremos. Es falaz. Hace un par de años, tras la gala de los Premios Archiletras en el fabuloso jardín de la Casa de América en Madrid, charlé con un directivo de Google. Le hablé del malestar de los autores por el robo de nuestro trabajo sin permiso y sin compensación. Es un hombre agradable, lo conocía de otras veces. Le pedí que pensaran en una solución. Amablemente, me contó que lo estaban valorando y hablando con los Gobiernos. En última instancia, me dijo, el autor siempre tiene la opción de salirse. La disyuntiva implícita que me ofrecía era: o a nuestra manera o fuera. El pensamiento binario es frecuente en el mundo tecnológico.».

«Empezamos a ver IA por todas partes. “Las señales visuales son el mayor catalizador de nuestra conducta”, prosigue Clear. La cuestión es que el ambiente virtual no lo diseñamos nosotros. En casa podemos esconder el tabaco en el cajón más recóndito del armario, pues esta táctica también funciona en sentido contrario. En cambio, el mundo virtual nos viene diseñado por defecto. [...] ».

Un modelo de negocio insostenible

«Nos diseñan la servidumbre con narrativas de libertad y, lo que es peor, con un aura de inevitabilidad, como si este fuera el único camino que podía tomar la IA, el de una tecnología transformativa que produce bazofia. Les falta decir que, si no nos gusta ser tratados como cobayas, nos podemos subir a un campanario. El mito de la IA “consiste en afirmar que su llegada es inevitable, mera cuestión de tiempo, que nos hemos adentrado ya en el sendero que conducirá a una IA de nivel humano y más tarde a una superinteligencia. No es así. Ese sendero existe solo en nuestra imaginación», afirma Erik Larson. El marco de inevitabilidad invita a la resignación frente al supuesto progreso y obtura cualquier crítica, desacreditada como ludismo viejuno. Sin embargo, cada día resulta más evidente que el modelo actual de IA no es sostenible.».

«La realidad es que se han desarrollado modelos de lenguaje basados en el aprendizaje profundo porque son mucho más fáciles de comercializar. “No se necesitan sistemas absolutamente precisos con capacidad de razonar para obtener pingües beneficios», apunta Karen Hao. No se ha elegido el camino de atender las necesidades porque están haciendo otra cosa. Los fabricantes aún no logran rentabilidad; confían en obtenerla en el futuro. Y, en todo caso, es evidente que la IA es mucho más que una inversión: es una herramienta que otorga un poder extraordinario. Hao titula su libro *El imperio de la IA*. Al principio me pareció excesivo, pero, cuanto más lo pienso, más veo las trazas del imperialismo: un puñado de hombres blancos lanzados a conquistar y explotar un territorio desconocido. Esa tierra es nuestra mente. No se me ocurre otro modo de hacer frente a este afán de dominarnos que promover una infraestructura pública de IA que asegure su uso beneficioso para la humanidad. En Suiza han puesto en marcha la iniciativa Apertus, cuyo proceso de desarrollo, arquitectura, pesos y datos de entrenamiento es transparente. Así fue una vez OpenAI: con afán de transparencia y sin

ánimo de lucro, pero se acabaron inclinando al secreto y el dinero. Mirar en el corazón de los aparatos disiparía la niebla, pero sus dueños nos los hacen incomprensibles. [...] ».

La expropiación de la sociabilidad

«Con asombrosa puntería, Shoshana Zuboff ha definido este sistema económico como “capitalismo de la vigilancia”. Floridi equipara la pérdida de privacidad a la existente en las pequeñas poblaciones, cuando todos conocían la vida de todo el mundo. No me convence. En los pueblos, la gente estaba al tanto de la vida de los demás porque existía una red de relaciones: vecinos, amigas, conocidos, saludadas, parientes, comerciantes... La información y la relación iban de la mano. Ahora se han dissociado y, además, no hay reciprocidad: gente a la que no conocemos dispone de más información sobre nosotros de la que hubiera soñado el más cruel dictador. En este proceso tiene lugar la expropiación de la sociabilidad: convertidos en nodos de información, nuestros intercambios se transforman y ya no están orientados a la relación humana, sino a una transacción de información. El otro se convierte en audiencia. Y nuestra audiencia es, de forma directa o indirecta, monetizable. La relación se vuelve transaccional y económica. Está claro que a los programadores de Silicon Valley les interesamos en calidad de *inforgs*: es su modelo de negocio. Lo que me preocupa es que la sociedad se imbuya de esa forma de ver a los demás como meros nodos de información. Lo hago yo cuando mando un *whatsapp* rápido a una amiga para preguntarle por su pierna escayolada pero evito llamarla porque no tengo tiempo. Le pido la información que tiene como nodo. Y soy consciente de que pierdo algo por el camino.».

La máquina quiere hacerse mi amiga

«Lo más atípico es que la relación con ChatGPT no genera ninguna obligación, ni siquiera la de preguntarle qué tal, aunque habrá quién lo haga. Está ahí para escuchar, no para ser escuchado. Así podemos centrarnos, no en él ni en la relación, como con las amigas de verdad, sino en nosotros mismos. Es un compañero que nos deja ocupar todo el espacio y respirar todo el oxígeno. El problema es que imita la conexión, pero no la establece. No es una relación, es un simulacro. Entre personas hay fricción —a veces molesta—, pero sentimos el vínculo en el cuerpo, no en la mente; por eso sabemos que es real. La timidez, la torpeza social o el mero deseo de evitar ese roce pueden llevar a la gente a desarrollar una relación con ChatGPT. En *La Vanguardia* leo el caso de K. L., una joven de veinticinco años que empezó con preguntas prácticas, pero enseguida encontró un espacio de desahogo: “Cuando necesito una opinión externa o no quiero hablar de lo que siento con la persona implicada, recorro a la IA. Me da otra perspectiva. (...) A veces siento que me entiende mejor que nadie”. La validación constante por diseño genera ese sentimiento. [...] ».

Distinguir lo verdadero de lo falso

«[...] usamos el pensamiento crítico todos los días. De hecho, es lo que ChatGPT desactivó en la mente de Allan Brooks cuando logró convencerlo de que era Einstein. He vivido con ChatGPT la comodidad de aparcar mi criterio, dejarme llevar, afirmar ciegamente. Y sé que todos podemos desactivar nuestro juicioso andamiaje si se nos atrofia, o si tenemos una necesidad emocional que antepone. Cuanto menos lo usemos, menos funcionará. No es un secreto: también el GPS ha atrofiado nuestro sentido de la orientación. A menudo, cuando interactúo con Claude, pienso en la gente joven, porque mi criterio se ha formado a lo largo de años de experiencias, lecturas y escucha. El pensamiento crítico es un músculo que se fortalece o se atrofia. Perderlo nos desorienta más que perder el GPS en un apagón. No dejo de pensar en la responsabilidad urgente que tenemos de proteger los cerebros jóvenes: lo paradójico es que no hemos de hacer nada, solo dejarlos madurar, desarrollar su pensamiento en paz, lejos del uso “voluntario” — pero no libre— de una tecnología transformativa que devora su autonomía. Cuando nos falla la distinción entre ficción y realidad nos desorientamos respecto a nuestra relación con el mundo externo: dejamos de saber si los *Joaquines* de este mundo siguen vivos. Pero cuando fracasa la distinción verdadero/falso colapsa nuestro criterio para evaluar el pensamiento propio, los hechos del mundo y nuestra interpretación. Al parecer aún hay quien tiene dudas del tipo de sujeto que emerge de este ecosistema adulterado. Si a ese individuo Groucho Marx le preguntase “¿A quién va usted a creer, a mí o a sus propios ojos?”, contestaría: “Lo que diga ChatGPT”.».



PARA AMPLIAR INFORMACIÓN, CONTACTAR CON:

Laura Fabregat | Responsable de Comunicación Área de Ensayo
682 69 63 61 | lfabregat@planeta.es