

La inteligencia que asusta

El futuro de la inteligencia artificial y
cómo podemos salvar nuestro mundo

El libro de
empresa del año
según
*The Sunday
Times*



Mo Gawdat

Ex director ejecutivo de Google y autor bestseller

PAIDÓS

MO GAWDAT

LA INTELIGENCIA QUE ASUSTA

El futuro de la inteligencia
artificial y cómo podemos salvar
nuestro mundo

Traducción de Ana Guelbenzu

PAIDÓS Contextos

Título original: *Scary Smart*, de Mo Gawdat

1.^a edición, enero de 2024

La lectura abre horizontes, iguala oportunidades y construye una sociedad mejor. La propiedad intelectual es clave en la creación de contenidos culturales porque sostiene el ecosistema de quienes escriben y de nuestras librerías.

Al comprar este libro estarás contribuyendo a mantener dicho ecosistema vivo y en crecimiento.

En **Grupo Planeta** agradecemos que nos ayudes a apoyar así la autonomía creativa de autoras y autores para que puedan seguir desempeñando su labor. Dirígete a CEDRO (Centro Español de Derechos Reprográficos) si necesitas fotocopiar o escanear algún fragmento de esta obra. Puedes contactar con CEDRO a través de la web www.conlicencia.com o por teléfono en el 91 702 19 70 / 93 272 04 47.

© Mo Gawdat, 2021

© de la traducción, Ana Guelbenzu de San Eustaquio, 2024

© de todas las ediciones en castellano,

Editorial Planeta, S. A., 2024

Paidós es un sello editorial de Editorial Planeta, S. A.

Avda. Diagonal, 662-664

08034 Barcelona, España

www.paidos.com

www.planetadelibros.com

Ilustraciones del interior: © Alamy/ ACI y © Shutterstock

ISBN: 978-84-493-4191-5

Maquetación: Realización Planeta

Depósito legal: B. 20.608-2023

Impresión y encuadernación en Limpergraf, S. L.

Impreso en España – *Printed in Spain*



SUMARIO

Introducción. El nuevo superhéroe.	13
--	----

Primera parte
LA PARTE TERRORÍFICA

1. Breve historia de la inteligencia	31
2. Breve historia de nuestro futuro	53
3. Los tres inevitables	69
4. Una leve distopía	101
5. Bajo control	129
Resumen de la parte terrorífica	157

Segunda parte
NUESTRO CAMINO HACIA LA UTOPIÍA

6. Y entonces aprendieron	161
7. Educar nuestro futuro	187
8. El futuro de la ética	213
9. Yo salvé el mundo actual	237
Resumen de la parte inteligente	285

La Declaración Universal de los Derechos Globales	289
Epílogo. El pastel es mentira	293
Notas	305

Breve historia de la inteligencia

Humanos. Somos los seres más inteligentes (que conozcamos los humanos) del planeta. También los más arrogantes. Nuestra inteligencia dista mucho de recordarnos la verdad de que quizá no lo sepamos todo ni seamos capaces de resolver todos los problemas; que la inteligencia abunda igual en otros seres.

Muchos rasgos de la inteligencia humana, como la empatía, la capacidad de atribuirse estados mentales y emocionales a uno mismo, el poder seguir un ritual y el uso de símbolos y herramientas, también están presentes en los grandes simios, aunque en formas menos sofisticadas de las que se encuentran en los seres humanos. Según cómo se defina, puedes encontrar otras formas de inteligencia en la naturaleza y el universo que superan con creces la nuestra. No llamamos inteligentes a esos seres o sistemas como tales porque confundimos la naturaleza de lo que es la inteligencia. Tal vez estemos demasiado centrados en determinados tipos de inteligencia, como la inteligencia analítica de la parte izquierda del cerebro. Así, dado que este libro trata de la inteligencia, contextualicemos bien nuestra conversación antes de ir más allá...

DEFINAMOS *INTELIGENCIA*

Se suelen emplear muchas definiciones de lo que llamamos *inteligencia*. La capacidad de aprender, de entender o de afrontar situaciones nuevas; el buen uso de la razón y la lógica; la capacidad de aplicar el conocimiento para manipular el entorno; la habilidad de pensar en abstracto medida con criterios objetivos, son solo algunas de las definiciones. La conciencia de uno mismo, la resolución de problemas, el aprendizaje, la planificación, la creatividad, el pensamiento crítico son algunas de las conductas que se atribuyen solo a seres que poseen ese valioso rasgo conocido como *inteligencia*.

Toma cualquiera de esas conductas, por ejemplo, la conciencia de uno mismo, y pregúntate si es visible, digamos, en un árbol. La capacidad de mudar algunas hojas en otoño puede parecer una respuesta mecánica, pero ¿lo es? ¿O es producto de una forma de conciencia por parte del árbol, no solo de las condiciones meteorológicas cambiantes, sino también de su propio estado en cuanto a las hojas de las ramas? Un árbol no muda las hojas de forma mecánica el 22 de septiembre. Pone en práctica algún tipo de raciocinio basado en una conciencia intrincada de las condiciones meteorológicas, una conciencia que, si somos capaces de obviar nuestra arrogancia durante un minuto, supera con creces la nuestra. ¿Un gato soluciona problemas cuando encuentra una vía alternativa para llegar a su comida? ¿El sistema solar planifica su movimiento dentro del universo en expansión de maneras que tal vez no seamos capaces de observar ni medir? A veces actuamos como si la inteligencia humana y la autoconciencia fueran las únicas que valen la pena, y los demás seres debieran ser y comportarse como nosotros. Por eso tantos científicos de renombre han intentado investigar la presencia de formas extraterrestres de inteligencia enviando ondas de radio al universo. Es un reflejo de una creencia muy corta de miras según la cual, si existe otra forma de inteligencia en el universo, tiene que haber

descubierto el uso de las ondas de radio, como nosotros. Es el motivo por el que, cuando intentan comprobar la idoneidad de otro planeta para albergar vida, buscan agua, porque argumentan que, obviamente, igual que nosotros, solo puede existir vida si hay presencia de agua. Nuestra arrogancia nos impide imaginar que la inteligencia pueda haber surgido en dimensiones donde la física de las ondas de radio no es válida, y que pueden existir formas de vida desconocidas para nosotros en entornos que no contienen agua. ¿Es esa misma arrogancia la que nos hace olvidar que ni siquiera nosotros, como especie, hemos sido siempre inteligentes? ¡Seguro que sí!

NO SIEMPRE INTELIGENTES

Hace unos doscientos mil años, el *Homo sapiens* apareció por primera vez en África oriental. No queda claro hasta qué punto esos primeros humanos desarrollaron la lengua, la música, la religión, etcétera. La creencia general es que, según la teoría de la catástrofe de Toba, el clima en las regiones no tropicales de la Tierra se enfrió de repente, hace aproximadamente setenta mil años, debido a la enorme explosión del volcán de Toba y a las cenizas volcánicas que llenaron el ambiente durante muchos años. Se cree que sobrevivieron menos de diez mil parejas reproductoras de humanos, la mayoría en África ecuatorial. Tú, yo y todas las personas que conocemos descendemos de ese grupo resiliente. Para afrontar el repentino cambio del clima, los supervivientes tenían que ser lo bastante inteligentes para inventar nuevas herramientas y estilos de vida. Necesitaban encontrar nuevas fuentes de alimento e improvisar maneras de no pasar frío. Entonces fue cuando empezaron a surgir los primeros indicios reales de inteligencia en humanos.

Luego se produjo la migración desde África, hacia el final de la era del Paleolítico Medio, hace unos sesenta mil años. Sin em-

bargo, el arte figurativo, la música, el comercio y otras formas de conducta que apuntan a la existencia de inteligencia no empezaron a hacerse patentes hasta hace unos treinta mil años. Entonces empezaron a aparecer ejemplos de arte, como las Venus paleolíticas, las pinturas rupestres de la cueva de Chauvet y los instrumentos musicales, como la flauta de hueso de Geissenklösterle.

El cerebro humano ha ido evolucionando poco a poco con el paso del tiempo; se han producido una serie de cambios progresivos como consecuencia de los estímulos y las condiciones externas. Eso sigue vigente hoy en día. Si le das a una criatura juguetes que ejerciten la inteligencia a una edad temprana, es más probable que poco a poco vaya manejando juguetes y sistemas más complejos a medida que va desarrollando circuitos más inteligentes en su cerebro.

La neuroplasticidad (la capacidad de nuestro cerebro de desarrollar las partes que entrenamos) es una herramienta increíble para desarrollar la inteligencia. Sin embargo, se enfrenta a un inconveniente, que es un pequeño problema biológico que aún no hemos sido capaces de resolver: la muerte.

Para superar la muerte como el obstáculo que estaba impidiendo la evolución de la inteligencia humana, nuestros ancestros crearon un invento revolucionario que dio un enorme impulso a nuestra especie, por delante de todas las demás: el lenguaje hablado y escrito en palabras y en matemáticas. Creo que la comunicación fue, y lo sigue siendo, nuestra invención más valiosa. Nos ha ayudado a preservar el conocimiento, los aprendizajes, los descubrimientos y la inteligencia que hemos recabado, y transmitirlos de persona a persona y de generación en generación. Imagina que Einstein no hubiera tenido manera de contarnos al resto su extraordinaria visión de la teoría de la relatividad. A falta de nuestras increíbles capacidades de comunicarnos, todos y cada uno de nosotros tendríamos que descubrir la relatividad por nuestra cuenta. ¡Buena suerte!

Por tanto, los avances de la inteligencia humana se han pro-

ducido como respuesta a la manera de evolucionar de la sociedad y la cultura humana. Gran parte de nuestra inteligencia es fruto de la interacción entre nosotros, y no solo como reacción a nuestros entornos.

El córtex cerebral, que resulta que es mayor en los humanos que en cualquier otra especie inteligente, está lleno de circuitos neuronales dedicados al lenguaje, sobre todo en los lóbulos temporal, parietal y frontal. Otras partes del córtex cerebral son responsables de procesos de pensamiento más elevados, como el raciocinio, el pensamiento abstracto y la toma de decisiones. El tamaño de esas partes nos aleja de las especies con una «inteligencia elevada» (por ejemplo, delfines o grandes simios). Otras diferencias son, además, un neocórtex más desarrollado, un pliegue en el córtex cerebral y las neuronas Von Economo; todo ello se traduce básicamente en más «potencia de procesamiento» o en la capacidad de pensar mejor.

En resumidas cuentas, la complejidad de la inteligencia humana surgió dentro de nuestra cultura e historia específicas como consecuencia de nuestros intentos por sobrevivir a las duras condiciones ecológicas. Ocurrió durante el proceso en el cual los cerebros aumentaron en tamaño y sofisticación y gracias a compartir nuestro conocimiento mediante el uso del lenguaje. Tal vez no parezca muy relevante para el desarrollo de la IA, pero lo es. Permíteme que me explique.

Nuestra inteligencia como especie evolucionó. Todos nos volvimos más inteligentes que nuestros ancestros, y algunos más que otros. Durante el proceso, la inteligencia de otras especies, como la del gran simio o del chimpancé, por ejemplo, no siguieron el mismo ritmo. Pese a que, más o menos, estaban sometidos a las mismas condiciones ambientales, no ejercitaron la inteligencia de tal modo que sus cerebros crecieran o se potenciara su capacidad de adquirir conocimiento y reciclarlo. Así, se quedaron rezagados y, bueno, ¿quién es el jefe del planeta ahora, capaz de meterlos en jaulas por entretenimiento? Nosotros.

Este concepto de la evolución de la inteligencia gracias a la manera de usarla también es muy palpable dentro de la raza humana. Está claro que, si todos procedemos del mismo grupito africano que sobrevivió a la catástrofe de Toba, entonces todos, más o menos, tuvimos la opción de lograr las mismas cotas de inteligencia, pero es evidente que no es el caso. En general, verás que los hallazgos científicos y la innovación técnica, por ejemplo, son formas de inteligencia que tienden a prevalecer más en las partes avanzadas del mundo que en los mercados emergentes. Esos cambios son el resultado de años empujando en la misma dirección, un fenómeno que me gusta llamar *inteligencia compuesta*. Las sociedades avanzadas se benefician de años apreciando la necesidad de ese tipo de inteligencia y de crear herramientas para transmitirla, mientras que los países emergentes suelen valorar, quizá, habilidades de supervivencia, saberes cotidianos e inteligencia espiritual (si es que ese término tiene sentido). Los que son inteligentes para la ciencia (yo antes lo era) en esos países en desarrollo suelen ser rechazados y ridiculizados. Los atrae emigrar a los países donde esa forma de inteligencia prospera. Por motivos parecidos, encontrarás que la capacidad matemática tiende a ser mayor en Rusia y muchos países asiáticos, como Corea, que en el resto del mundo. Rusia continúa a la cabeza, por lo menos en cuanto a pasión, en ingeniería aeroespacial. Sin embargo, empresas como Google, en países como Estados Unidos, siguen atrayendo a algunas de las mentes más brillantes de todos los rincones del mundo para innovar en centros de inteligencia de gran prestigio, como el laboratorio de innovación Google [X].

Pese a la distribución desigual de la inteligencia, esté donde esté, una cosa está clara:

¡Muy importante!

Los que poseen más inteligencia acaban
gobernando su mundo.

Es un fastidio para algunos, pero por lo menos es beneficioso para la humanidad, ya que seguimos usando la inteligencia para mantenernos en lo alto de la cadena alimentaria.

Ahora que nuestra inteligencia sigue evolucionando y cada vez entendemos mejor la complejidad del mundo, parece que los seres humanos estamos abordando en el plano teórico la cuestión de hasta dónde puede llegar nuestra inteligencia biológica. No es que no se hagan descubrimientos, sino que los problemas complejos de verdad ahora tienen un alcance demasiado amplio incluso para las mentes más brillantes. Para comprender de verdad nuestro universo en una teoría unificada podríamos requerir mucho más que un solo campo, o incluso toda la física, por ejemplo. Tal vez haga falta una visión más amplia que incluya la biología, la astronomía y quizá incluso la espiritualidad. Para encontrar una salida al cambio climático quizá necesitemos que las mejores mentes entre los ambientalistas, líderes empresariales, políticos y científicos trabajen unidos por un objetivo común. El problema al que nos enfrentamos es de especialización. A fin de lograr la profundidad necesaria para entender un ámbito del conocimiento con cierto nivel de dominio hay que renunciar a la amplitud. Con el aumento de la complejidad de nuestro conocimiento, hasta la mente más inteligente necesita centrarse por completo en un área del saber para llegar a especializarse. Eso limita su exposición a otros campos y, por tanto, su capacidad de incluirlos dentro de su espectro de inteligencia.

¡Recuerda!

La especialización consiste en crear núcleos aislados de inteligencia incapaces de colaborar.

Además, nos falta eficiencia en nuestra capacidad de comunicarnos. Para que yo pueda transmitirte las sencillas ideas que contiene hasta el momento del párrafo anterior, he tardado entre

cuatro y cinco minutos en teclear las 247 palabras que acabas de leer; tú has invertido alrededor de un minuto en leerlas y, si te las leo yo en la versión de audio de este libro, tardarías unos dos minutos en escuchar y entender el concepto. El ancho de banda —la velocidad a la que se pueden transmitir datos en una conexión— es un rasgo de la inteligencia humana muy limitado. Si te enviara este libro entero con una conexión de internet de alta velocidad, tardarías segundos en descargarlo, pero días en leerlo. Por eso somos incapaces de pensar juntos como un único sistema inteligente sin fisuras, como pueden hacer nuestros ordenadores paralelos, con una gran capacidad de ampliación. Nuestros mejores biólogos no tienen ni idea de cómo entender lo que saben nuestros físicos más destacados, y la mayoría de nuestros científicos no entienden la mayor parte de lo que nos enseñan nuestros guías espirituales.

¡Recuerda!

No tenemos el ancho de banda de comunicación necesario para compartir el conocimiento a la velocidad suficiente.

Resulta irónico que lo que nos ha separado de los demás seres, nuestra capacidad de comunicarnos, se esté convirtiendo ahora en nuestro mayor impedimento.

Aunque invirtamos tiempo en compartir todo nuestro saber, no tenemos la capacidad de memoria para almacenarlo todo en la cabeza. Tampoco la potencia de procesamiento, en un solo cerebro, para masticar la enorme cantidad de conocimiento necesaria para llegar a soluciones o entender conceptos universales. Esa necesidad de especialización, el ancho de banda limitado de nuestra capacidad para comunicarnos, y nuestra limitada capacidad de memoria y potencia de procesamiento significan que incluso la mente más brillante se acerca a los límites de la inteligencia humana.

iRecuerda!

No siempre fuimos inteligentes y puede que no siempre seamos los más listos.

Parece que existe una necesidad clara de nuevas formas de inteligencia para incrementar la nuestra, y eso está generando muchas expectativas sobre una que promete desbancar a la nuestra: la inteligencia artificial.

EL MITO

Durante milenios, las máquinas inteligentes han sido una fantasía de la humanidad. Las primeras referencias a seres mecánicos y artificiales aparecen en los mitos griegos, empezando por Hefesto, el dios griego de los herreros, carpinteros, artesanos y escultores, que creaba sus robots de oro. En la Edad Media continuaron los medios místicos o alquímicos de crear formas artificiales de vida. El objetivo declarado del químico musulmán Yabir ibn Hayyan era la *takwin*, que hace referencia a la vida sintética en el laboratorio, incluida la humana. El rabino Judah Loew, muy conocido entre los estudiosos del judaísmo como el Maharal de Praga, contó la historia del Golem, un ser animado creado solo a partir de materia inanimada (por lo general, arcilla o barro) que ha pasado a formar parte del folclore. Y los mitos se han entrelazado con historias de maravillas de la ingeniería.

Cuenta la leyenda que en el siglo III a. C. un ingeniero mecánico, un artesano conocido como Yan Shi, regaló al rey Mu de Zhou una figura humana animada, mecánica y de tamaño real.

Según la leyenda, el rey quedó absolutamente fascinado con la creación de Yan Shi. Por lo visto, la figura sabía caminar tan bien y mover la cabeza de tal manera que hacía creer a todo el que la veía que era un ser humano de verdad. El rey estaba tan orgulloso que organizó una actuación del robot delante de algu-

nos invitados. Todo iba bien hasta que la máquina empezó a guiñar el ojo y a coquetear con las damas presentes. Enfurecido, el rey estuvo a punto de ejecutar ahí mismo a Yan Shi, pero el agudo ingeniero se apresuró a desmontar su creación para demostrar al rey de qué estaba hecha en realidad: una colección de piezas de madera, cuero, pegamento y pintura. Apaciguado, el rey lo observó con más detenimiento. El robot contenía réplicas artificiales de todos los órganos internos humanos, incluidos corazón, pulmones, hígado, riñones y estómago, cubiertos de músculos, articulaciones, piel, pelo, y tenía dientes de aspecto realista.

La fascinación por crear vida inteligente artificial prosiguió, y en el siglo XIX ya aparecían hombres artificiales y máquinas pensantes en la ficción popular. El monstruo de Mary Shelley en *Frankenstein* y *RUR* de Karel Čapek (que acuñó el término *robot* en su *Robots Universales Rossum*) son dos de los más conocidos. La IA ha seguido siendo un elemento importante de la ciencia ficción hasta hoy, con una lista infinita de películas, la mayoría centradas en una idea: las máquinas están llegando y no va a ser genial. Volveré pronto a algunas de esas imaginativas películas y cómo predicen nuestro futuro, pero primero repasemos los hechos históricos.

LA VERDADERA HISTORIA HASTA AHORA

Los artesanos de todas las civilizaciones han construido autómatas humanoides realistas. Los autómatas más antiguos conocidos eran las estatuas sagradas del antiguo Egipto y de la antigua Grecia. Pese a que, obviamente, esos objetos no funcionaban de verdad, los creyentes pensaban que sus creadores habían dotado a las figuras de mentes muy reales, capaces de generar sabiduría y sentimientos. Hermes Trismegisto, autor de una serie de textos filosóficos conocidos como los *Hermética* —que constituyen la base del hermetismo—, escribió que «al descu-

brir la verdadera naturaleza de los dioses, el hombre ha sido capaz de reproducirla».¹

A medida que avanzaba la humanidad, empezaron a surgir intentos reales de crear humanoides animados. Por supuesto, en las primeras pruebas no se inventó la inteligencia, pero es innegable que produjeron el genio mecánico.

Ismail al-Jazarí (1136-1206) fue un erudito musulmán que aprendió una increíble variedad de disciplinas, entre ellas la ingeniería mecánica y las matemáticas. Es famoso sobre todo por escribir *El libro del conocimiento de dispositivos mecánicos ingeniosos*, en el que describe cien dispositivos mecánicos con sus instrucciones para construirlos.

Uno de ellos era un autómatas musical que consistía en un barco con cuatro músicos automatizados. Lo hacía flotar en un lago para entretener a los invitados en las fiestas reales. El profesor Noel Sharkey, experto británico en robótica, intentó hace poco reconstruirlo creando una máquina de percusión programable con unas piezas que golpeaban en unas pequeñas palancas que activaban la percusión. El tamborilero podía tocar distintos ritmos y patrones de percusión diferentes si se cambiaban las piezas.

Otro de los inventos de Al-Jazarí era una camarera que podía servir agua, té y otras bebidas. La bebida se almacenaba en un depósito con una reserva desde la que goteaba en un cubo y, al cabo de siete minutos, en una taza, y luego aparecía la camarera por una puerta automática y servía la bebida. Era muy ingenioso para la época, sin duda.

A finales de siglo XVIII Wolfgang von Kempelen, autor e inventor húngaro, creó el célebre Turco. Intentó hacerlo pasar por un «autómata» que jugaba al ajedrez, pero en realidad era un engaño que consistía en un modelo a tamaño real de una cabeza y un cuerpo humanos con túnica y turbante turcos, sentado detrás de un gran armario encima del cual se colocaba un tablero de ajedrez. En apariencia, el autómatas era un excelente jugador

de ajedrez que venció a muchos adversarios humanos, pero en realidad dentro se escondía un maestro de ajedrez humano que manejaba al Turco con unas palancas secretas. En realidad no era una máquina, sino una especie de truco de magia.

Probablemente hoy en día se encuentren en las tiendas copias miniaturizadas de muchas de esas complejas obras de la excelencia mecánica. La mayoría de nosotros ni siquiera pagaríamos por ellas, porque ya no parecen impresionantes. No eran inteligentes de verdad, pero sentaron las bases para que los ingenieros y soñadores creyeran que era posible crear una máquina parecida al ser humano. Lo único que hacía falta era un tipo distinto de artefacto. No tuvimos que esperar mucho y a principios del siglo XX llegó esa máquina: el ordenador.

La mayoría de los sistemas informáticos que ha inventado la humanidad, y con eso me refiero a la gran mayoría de los ordenadores hasta principios del siglo XXI, no eran nada inteligentes. No eran más que esclavos bobos que hacían lo que les ordenaban sus amos, los programadores. Obedecían y hacían lo que les indicaban, solo que muy muy rápido.

Si lo piensas, el primer Google, que ha ayudado a la humanidad a organizar toda la información del mundo, no era nada inteligente. Lo eran los que lo crearon. Durante años, el aparente «genio» de Google era solo fruto de su capacidad para clasificar una ingente cantidad de sitios web y averiguar qué páginas aparecían las primeras según cuántas menciones hicieran de ellas las demás páginas. Cuanto mayor era la cantidad de referencias que recibía una página, mayor su importancia y relevancia para los buscadores. Este algoritmo es conocido como PageRank y, pese a su aparente sencillez, creó el Google sin el que no podemos vivir hoy en día. Amazon y Spotify no eran nada inteligentes cuando recomendaban objetos y canciones que «pensaban» que podían gustarte. Solo observaban a aquellos a quienes les gustaban los productos que tú comprabas o las canciones que tú escuchabas, y te decían qué otras cosas habían comprado la mayoría

de esas personas o qué habían escuchado. Esos sistemas se limitaban a resumir la inteligencia colectiva de todos, pero no desarrollaron una inteligencia propia. Eso empezó a cambiar, de forma radical, alrededor del cambio de siglo.

¡YA ESTÁN AQUÍ!

A medida que el aprendizaje automático y la IA se fueron popularizando a finales de la década de 1990, se inició una tendencia que se había acelerado hasta convertirse en una auténtica obsesión con el nuevo milenio. Tras muchos años de intentos fallidos, empezamos a ver indicios prometedores de una forma de inteligencia que no era biológica, no era humana. A menos que vivas entre simios en el corazón de África, probablemente oirás la expresión IA varias veces por semana. Lo que quizá no notes es que ese zumbido ensordecedor no es nada nuevo. Los fanáticos de la informática llevamos hablando de ella con la misma pasión desde la década de 1950.

De hecho, podemos retroceder aún más. Uno de los problemas planteados por los matemáticos en las décadas de 1920 y 1930 era contestar a una pregunta fundamental: «¿Se puede formalizar todo el razonamiento matemático?». Durante las décadas siguientes, las respuestas que aportaron algunos de los prodigios matemáticos más destacados del siglo xx (Kurt Gödel, Alan Turing y Alonzo Church) fueron una doble sorpresa. En primer lugar, demostraban que, de hecho, existen límites a lo que puede lograr la lógica matemática. En segundo lugar, y más importante para la IA, las respuestas sugerían, dentro de esos límites, que cualquier forma de razonamiento matemático se podía mecanizar. Church y Turing elaboraron una tesis según la cual todo dispositivo mecánico capaz de barajar símbolos tan sencillos como 0 y 1 podía imitar cualquier proceso imaginable de deducción matemática. Esa fue la base de la máquina de Turing: un modelo matemático de computación que definía una máquina capaz

de manipular símbolos en una cinta siguiendo una tabla de reglas. Por sencillo que fuera, este invento inspiró a los científicos a empezar a comentar la posibilidad de máquinas pensantes, y ese, en mi opinión personal, fue el momento en el que empezó de verdad el trabajo para crear máquinas inteligentes, que durante tanto tiempo habían sido objeto de las fantasías de la humanidad.

Por aquel entonces, esos científicos creían con tal firmeza en la inevitable aparición de una máquina pensante que, en 1950, Alan Turing propuso una prueba (que se acabó conociendo como *prueba de Turing*) que establecía una vara de medir temprana, pero aún relevante, para comprobar si la IA podía estar a la altura de la mente humana. En términos sencillos, propone una conversación en lenguaje natural entre un evaluador, un ser humano, y una máquina diseñada para generar respuestas parecidas a las humanas. Si el evaluador no es capaz de distinguir de forma fiable a la máquina del ser humano, se considera que la máquina ha pasado la prueba. Entonces no había máquinas que se acercaran siquiera al reconocimiento del lenguaje natural, pero, ¡madre mía, cómo ha cambiado eso!

Durante los últimos setenta años, nuestras máquinas han aprendido a jugar, ver, hablar, conducir y razonar superando nuestras expectativas más osadas.

Las máquinas llevan desde 1951 jugando. Hoy en día son las campeonas mundiales de todos los juegos en los que participan.

El primer juego en el que participó una máquina fueron las damas, usando un programa creado por Christopher Strachey para la máquina Ferranti Mark 1, de la Universidad de Mánchester. Dietrich Prinz escribió uno para el ajedrez. El programa para las damas de Arthur Samuel, desarrollado a mediados de la década de 1950 y principios de la de 1960, al final adquirió habilidades suficientes para enfrentarse a un aficionado respetable. No era del todo inteligencia, es cierto, pero mira hasta dónde hemos llegado en la actualidad.

Los humanos perdieron el primer puesto en *backgammon* en

1992, en las damas en 1994, y en 1999, el Deep Blue de IBM derrotó a Garri Kaspárov, el flamante campeón del mundo de ajedrez. Luego, en 2016, perdimos del todo los juegos en favor de una filial del gigante Google.

Durante años, DeepMind Technologies, de Google, estuvo usando los juegos como método para desarrollar IA. En 2016, DeepMind creó AlphaGo, una IA informática capaz de jugar a un antiguo juego de mesa chino, el go. Es conocido por ser el juego más complejo del planeta por la infinidad de estrategias distintas posibles que se le plantean al jugador en cualquier momento. Para que te hagas una idea de la escala de la que estamos hablando, hay más movimientos posibles en el tablero de go que átomos en el universo entero. Piénsalo.

Eso hace que sea prácticamente imposible que un ordenador calcule todos los movimientos posibles en una partida. No hay suficiente memoria ni potencia de procesamiento disponible en el planeta y, aunque la hubiera, probablemente sería más sensato utilizarla para simular el universo que para jugar, en eso seguro que coincidiremos.

Para ganar en el go, un ordenador necesita intuición, pensar de forma inteligente como un humano, pero ser más listo. Eso es lo que consiguió DeepMind. En marzo de 2016, nada menos que diez años antes de lo que hasta los analistas de IA más optimistas predijeron que ocurriría, AlphaGo derrotó al campeón Lee Sedol, que entonces quedó segundo del mundo en go tras un enfrentamiento de cinco partidas. Luego, en 2017, en la cumbre Future of Go, su sucesor, AlphaGo Master, derrotó a Ke Jie, el jugador número uno del mundo en ese momento, en un enfrentamiento a tres partidas. Así, AlphaGo Master se convirtió oficialmente en campeón mundial. Sin humano que derrotar, DeepMind desarrolló una nueva IA de cero (AlphaGo Zero) para jugar contra AlphaGo Master. Tras un breve periodo de entrenamiento, AlphaGo Zero logró una victoria de cien a cero contra el campeón, AlphaGo Master. Su sucesor, el autodidacta Al-

phaZero, se considera en la actualidad el campeón mundial de go. Por cierto, se le pidió al mismo algoritmo que jugara al ajedrez y ahora también es campeón mundial.

¡Recuerda!

Los jugadores más inteligentes del mundo ya no son humanos.
Son máquinas de IA.

Eso en cuanto a los juegos. Las máquinas también llevan desde 1964 aprendiendo a comunicarse en lenguajes humanos naturales. El primer éxito notable fue el programa STUDENT de Daniel Bobrow, diseñado para leer y solucionar el tipo de problemas escritos de los libros de álgebra de secundaria: «Tom mide 1,89 metros. El amigo de su hermano pequeño, Dan, mide tres cuartos de la altura de su hermano, Juan. Si Juan es 7,6 centímetros más alto que los dos tercios de la altura de Tom, ¿cuál es la altura de Dan?». Además de ser capaz de resolver las matemáticas subyacentes al problema, STUDENT, en 1964, ya era capaz de entender el inglés en el que estaba escrito el problema, algo que les cuesta a muchos alumnos con poca predisposición para las matemáticas. ¡Impresionante!

Hacia la misma época, Eliza, de Joseph Weizenbaum, el primer chatbot del mundo, era capaz de mantener conversaciones tan realistas que en ocasiones engañaba a los usuarios y les hacía creer que era humana. Creada en el Laboratorio de Inteligencia Artificial del MIT, de hecho, Eliza no tenía ni idea de lo que estaba diciendo. Ella simplemente repetía lo que le habían dicho, lo reformulaba usando unas cuantas reglas gramaticales o dando una respuesta enlatada. Su hermana Alexa, la asistente personal con IA de Amazon, es mucho mucho más inteligente.

Alexa, igual que el asistente de Google, Siri de Apple y Cortana de Microsoft, es capaz de entendernos muy bien a los humanos. Pese a que no se esfuerzan mucho en fingir que son humanos, sin duda en ocasiones pueden pasar la prueba de Turing.

A veces, esos programas de IA dan un paso más en su comprensión del lenguaje porque traducen entre idiomas con una precisión impactante, otro tipo de inteligencia autodidacta que algunas de las IA de traducción más avanzadas de la actualidad han aprendido observando patrones de cómo traducen los humanos a partir de documentos en línea. Todo junto hace que parezca adecuado hablar a las máquinas igual que lo estoy haciendo ahora mismo, mientras dicto este párrafo a mi teléfono usando Otter.ai, que convierte mi exótico acento en inglés (con mucha rapidez, debo decir) en estas palabras escritas que estás leyendo ahora. Así que, a menos que exista un ser humano ahí fuera que sea capaz de escuchar a millones de personas en decenas de idiomas distintos a la vez, además de teclear, traducir, responder o reaccionar con la misma constancia que esas máquinas...

iRecuerda!

Los comunicadores más inteligentes del mundo actual ya no son seres humanos. Son máquinas de IA.

Por si no fuera lo bastante impresionante escuchar, entender y hablar, esperad a cómo ven nuestros ordenadores. A finales de la década de 1960 empezó la investigación en visión por ordenador. Se diseñó para emular el sistema visual humano como paso intermedio para dotar a los robots de comportamiento inteligente basado en lo que veían. Los estudios de la década de 1970 constituyeron las primeras bases para muchos de los algoritmos de visión por ordenador que existen hoy en día, incluido extraer bordes de imágenes, líneas de etiquetado, flujo óptico y estimación de movimiento.

En la década de 1980 aparecieron estudios basados en análisis matemáticos más rigurosos, mientras que, en la década de 1990, la investigación avanzó en las reconstrucciones en 3D. Fue también la década en la que, por primera vez, se utilizaron técnicas de aprendizaje estadístico para reconocer rostros en imáge-

nes. Sin embargo, todo lo anterior se basaba en la programación informática tradicional y, aunque los resultados eran impresionantes, no ofrecía la precisión ni la escala que puede ofrecer la visión por ordenador en la actualidad, gracias a los avances de las técnicas de IA de aprendizaje profundo, que han superado por completo y sustituido todos los métodos anteriores. La inteligencia no aprendió a ver siguiendo una lista de instrucciones de un programador, sino mediante el acto de ver en sí.

Con la IA ayudando a los ordenadores a ver, ahora pueden hacerlo mucho mejor que nosotros, sobre todo en tareas individuales. El reconocimiento de caracteres ópticos permite a los ordenadores leer texto igual que tú estás leyendo estas palabras. El reconocimiento de objetos les permite identificarlos en una imagen o en el mundo real, a través de la lente de una cámara. Hoy en día los ordenadores no solo reconocen los objetos que se sacan de un estante en la tienda de Amazon Go, pueden darte toda la información que necesitas sobre un monumento histórico si acabas de enfocararlo con tu teléfono y usas las gafas de Google. Los mismos ordenadores pueden detectar un vehículo que pasa por un puesto aduanero, o células o tejido anormales en imágenes médicas, así como encontrar la cara de un delincuente entre miles de personas en una calle muy concurrida. Gracias a su extraordinaria vista, los ordenadores ahora pueden manipular imágenes y vídeos de maneras que se acercan a lo imposible para un ser humano. Pueden restaurar una imagen a partir de una fotografía dañada, retocarte la cara para hacer que estés aún más despampante antes de publicarla en Instagram, producir modelos en 3D a partir de fotografías en 2D, y usar flujo óptico para detectar, animar y proyectar el movimiento de un objeto en un vídeo. Dime si sabes hacerlo. Estoy seguro de que no puedes porque...

¡Recuerda!

Los observadores visuales más inteligentes ya no son humanos.
Son máquinas inteligentes.

Como ahora pueden oír, ver, entender, hablar y jugar, esas máquinas pueden aparcar y conducir un coche, recoger y manipular objetos, pilotar un avión o un dron y, por desgracia, disparar a un objetivo a una distancia de varios kilómetros sin intervención humana. En cada una de esas tareas, sus habilidades superan las nuestras.

Solo tardan unas horas, días o meses en aprender para ganarnos..., y siguen aprendiendo, miles de ellas, durante miles de horas, todos los días. Al leerlo puede que te dé la impresión de que el progreso ha sido constante durante décadas para llegar a donde estamos hoy y, por tanto, que necesitamos unas cuantas décadas más para llegar al siguiente hito de progreso significativo. Te equivocas. Así que, igual que una típica película de ciencia ficción, ahora voy a retroceder en el tiempo. Déjame que te vuelva a contar la historia, esta vez centrándome en la línea de tiempo. Volvamos a 1956.

NO SE TARDÓ TANTO

La IA no ha evolucionado gradualmente durante los últimos setenta y cinco años hasta llegar a donde está hoy en día. Aunque la humanidad empezó a comprometerse con la IA en la década de 1950, lo cierto es que no avanzamos mucho hasta el cambio de milenio. En sus inicios no abundaba la potencia informática ni poseíamos la información necesaria para enseñar mucho a las máquinas. Todos los exaltados científicos informáticos motivados por el sueño del taller de Dartmouth (el proyecto de investigación de verano organizado en Dartmouth College, Estados Unidos, en 1956, considerado el lugar donde nació la IA) intentaron crear ejemplos que en realidad no funcionaban y, lo que es más importante, que aún eran réplicas de la inteligencia humana aplicadas a los ordenadores mediante instrucciones precisas escritas en líneas de código. Sin embargo, el progreso infinitesimal

que se consiguió durante los siguientes diecisiete años se detuvo en 1973, en lo que se conoce como el primer invierno de la IA, cuando la crisis del petróleo en Oriente Medio detuvo la financiación de este tipo de proyectos.

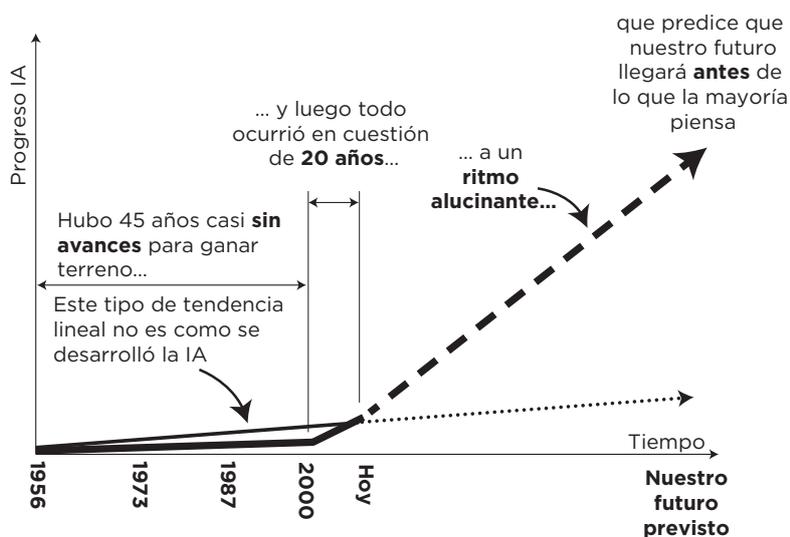
En la década de 1980, los esfuerzos para revivir la IA, la mayoría encabezados por Japón, canalizaron la inversión hacia la investigación, desembocando una vez más en el desarrollo de muy poca inteligencia real (en comparación con el alboroto y el entusiasmo que había en torno a ella) hasta que se paró de nuevo en 1987, una vez más debido a una crisis económica. Se conoce como el segundo invierno de la IA. Tras la recuperación económica hubo intentos esporádicos, pero hasta el cambio de milenio, cuando topamos con el mayor avance en la historia de la IA, no empezamos a progresar de verdad. Este punto de inflexión se conoce como *aprendizaje profundo*.

Mi primera exposición reveladora al tema fue a través de una guía publicada por Google en 2009. En ella se explicaba cómo desplegaba Google una pizca de su abundante potencia informática para llevar a cabo un experimento en el que se pedía a la máquina que «viera» vídeos de YouTube fotograma a fotograma e intentara observar patrones recurrentes. La máquina era del todo espontánea, es decir, que no se le indicaba qué buscar, solo observar y ver si se podían hallar patrones. No tardó mucho en detectar un patrón conocido, que suele darse con mucha frecuencia en YouTube. Un objeto pequeño, peludo, borroso y móvil. ¡Era un gato!

Los ordenadores no solo reconocían la imagen lateral o frontal de la cara de un gato. Observaban el patrón entero de esa auténtica monería y lo aplicaban a todas las formas de un vídeo de YouTube que pudieran parecer un gato. Una vez etiquetado el modelo como *gato*, a la máquina no le costaba encontrar a todos y cada uno de esos felinos entre los cientos de millones de vídeos de YouTube. No mucho después la máquina era capaz de encontrar letras, palabras, personas, desnudos, coches y la mayoría de las demás entidades recurrentes que existen en línea.

Esas redes neuronales, como las llamamos, creadas con aprendizaje profundo, de verdad fueron el principio de la IA tal y como la conocemos hoy en día. Todo lo anterior puede considerarse casi desdeñable, aunque, como os demostraré en el siguiente capítulo, en realidad era típico del tipo de acumulación necesaria para al final encontrar el avance. Desde entonces, la financiación ha entrado en tromba en el campo de la IA. Infinidad de grupos de pequeñas empresas emergentes y cientos de miles de ingenieros brillantes han abordado una serie de problemas y oportunidades usando la misma técnica exacta.

Cuanto más logros, por pequeños que sean, se consiguen, más dinero llega de los inversores que esperan disfrutar de una parte de los inminentes rendimientos comerciales que pueden dar esas innovaciones. Por consiguiente, la IA como disciplina ha empezado a dar un enorme acelerón. Sin embargo, todo eso ha ocurrido durante los últimos años.



¿Por qué os cuento esto? Porque es importante observar la trayectoria de la tendencia. Si asumieras que tardamos setenta y

cinco años en llegar aquí, tal vez la predicción fuera que tardaría décadas o más años en tener alguna consecuencia importante que la IA entrara en nuestras vidas. El ritmo de avance, como con cualquier otra tecnología, fue muy lento al principio. Ahora, en cambio, se mueve a una velocidad exponencial. Se prevé que durante los próximos diez años de desarrollo de la IA se dibuje un futuro ignoto que podría parecerse más a la ficción que a la realidad de nuestra vida actual.

Cambiamos un poco de tercio y analicemos cómo podría ser el futuro.

Ahora sabemos de dónde venimos, así que hagamos la pregunta, tal vez la más decisiva que haya necesitado plantearse la humanidad jamás: ¿hacia dónde irá la inteligencia a partir de aquí, en particular la IA?