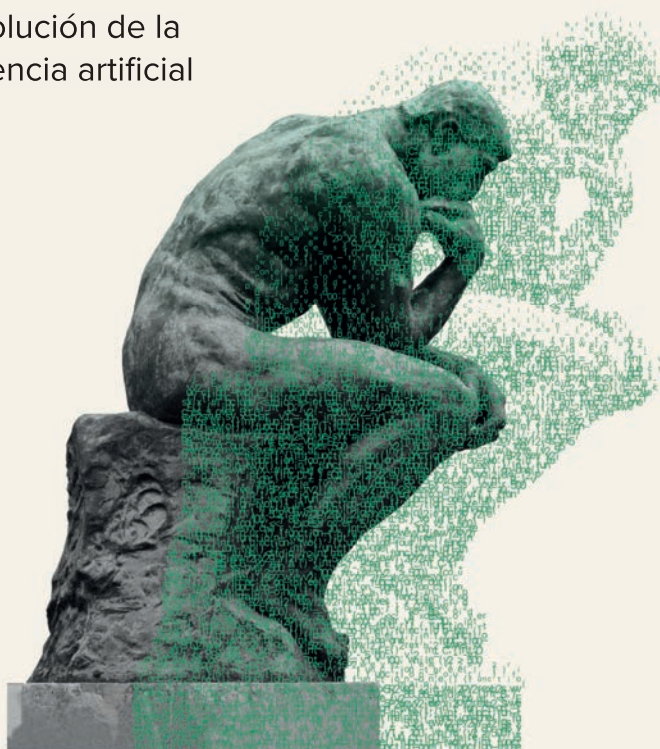


Antonio L. Flores Galea

Una mente infinita

La revolución de la
inteligencia artificial



Antonio L. Flores Galea

UNA MENTE INFINITA

La revolución de la inteligencia artificial

TUSQUETS
EDITORES

1.ª edición: marzo de 2024

© Antonio Luis Flores Galea, 2024

Reservados todos los derechos de esta edición para
Tusquets Editores, S.A. – Avda. Diagonal, 662-664 – 08034 Barcelona
www.tusquetseditores.com
ISBN: 978-84-1107-431-5
Depósito legal: B. 2.083-2024
Fotocomposición: David Pablo
Impresión y encuadernación: Liberdúplex, S.L.
Impreso en España

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.



| | |
|--|-----|
| Prólogo..... | 11 |
| Primera parte: Introducción a la inteligencia artificial | |
| 1. Qué es la inteligencia artificial..... | 17 |
| 2. Un poco de historia..... | 24 |
| 3. Inteligencia artificial genérica y específica..... | 46 |
| 4. Capacidades de la inteligencia artificial..... | 52 |
| Segunda parte: Redes neuronales y <i>Machine Learning</i> | |
| 5. Qué es <i>Machine Learning</i> | 61 |
| 6. Redes neuronales artificiales. El perceptrón..... | 70 |
| 7. Arquitecturas de las redes neuronales artificiales ... | 83 |
| 8. Tipos de aprendizaje de las redes neuronales artificiales | 91 |
| Tercera parte: Diseño de los sistemas de inteligencia artificial | |
| 9. Algoritmos de aprendizaje supervisado..... | 121 |
| 10. Algoritmos de aprendizaje no supervisado..... | 144 |
| 11. Diseño de una red neuronal artificial..... | 153 |
| 12. Optimización de una red neuronal artificial..... | 170 |
| Cuarta parte: Inteligencia artificial avanzada | |
| 13. Qué es <i>Deep Learning</i> y retos que presenta..... | 181 |
| 14. Visión artificial..... | 195 |

| | |
|--|-----|
| 15. Procesamiento de lenguaje natural (NLP) | 205 |
| 16. Inteligencia artificial generativa. El motor GPT | 211 |

Quinta parte: Implementación de una solución
de inteligencia artificial

| | |
|---|-----|
| 17. Planificación de un proyecto de inteligencia artificial | 225 |
| 18. Herramientas de desarrollo | 242 |
| 19. Retos de la inteligencia artificial. Ajuste y sesgo. . . . | 249 |
| 20. Riesgos de la inteligencia artificial. Inteligencia artificial confiable | 258 |
| 21. Control de la inteligencia artificial. Inteligencia artificial transparente y explicable | 270 |

Sexta parte: Inteligencia artificial en la empresa
y la sociedad

| | |
|--|-----|
| 22. Beneficios de la inteligencia artificial para la empresa | 285 |
| 23. Habilitadores y barreras para su adopción. | 297 |
| 24. Ética y regulación de la inteligencia artificial | 309 |
| 25. El futuro de la inteligencia artificial | 334 |

Apéndices

| | |
|---|-----|
| Bibliografía | 353 |
| Fuentes de las figuras | 361 |
| Abreviaturas empleadas en esta obra | 363 |
| Agradecimientos | 365 |

Capítulo 1

Qué es la inteligencia artificial

La inteligencia artificial, también conocida por sus siglas IA o AI (en inglés, *artificial intelligence*), es la habilidad de un ordenador o sistema digital para realizar tareas que normalmente desarrollan los seres inteligentes. La inteligencia se puede describir como la capacidad de realizar determinadas actividades mentales, como el aprendizaje, el razonamiento y la comprensión. Así, teniendo en cuenta que el adjetivo «artificial» hace referencia a algo creado que no ocurre de manera «natural», se puede concluir que la inteligencia artificial es la *capacidad que el hombre ha dado a determinadas máquinas de aprender, razonar y comprender conceptos, el entorno, o cualquier otra cosa*.

Sin embargo, en la literatura no existe una definición exacta de qué es y qué no es inteligencia artificial. Mientras, por ejemplo, hay autores que defienden una definición antropocéntrica —sistemas que piensan o actúan «como las personas»—, otros prefieren definir la inteligencia artificial como la capacidad de los sistemas para «razonar». En consecuencia, los límites de la inteligencia artificial siguen siendo confusos.

Lo que sí se acepta de manera general es que determinadas tareas propias de seres inteligentes y basadas en el razonamiento, como es la resolución de operaciones matemáticas *per se* (realizar una suma, una división o una raíz cuadrada, entre otras), no pueden ser consideradas actividades de la IA. Por ejemplo, una calculadora no es un sistema de IA. Pero, por el contrario, sí lo es

un programa de ajedrez que permite a una máquina jugar contra otro jugador, ya sea humano u otro ordenador con un programa similar. ¿Por qué? Pues parece que simplemente porque la comunidad científica considera actualmente que un programa de ajedrez aplica reglas «complejas», mientras que una calculadora implementa reglas de cálculo «sencillas». La ambigüedad y subjetividad de estos dos conceptos —complejo y sencillo— podría llevarnos a pensar que la IA habría tenido una definición cambiante a lo largo de los años, conforme la tecnología ha ido avanzando y que esta podría continuar evolucionando a lo largo de los próximos años. Desde el punto de vista del autor, así ha ocurrido y es más que probable que continúe ocurriendo, al menos hasta que la IA alcance un nivel de madurez suficiente para que su definición pueda mantenerse invariable y concreta por un periodo largo de tiempo.

Cuando en siglos pasados se comenzaron a fabricar autómatas, bien podrían haber pertenecido estos al campo de la IA, aunque en tal época el concepto no existiera. Para las personas de principios del siglo xx, una calculadora sí habría sido considerada como IA, al igual que la máquina de encriptación Enigma durante la Segunda Guerra Mundial. No obstante, el concepto IA surgió a mitad del siglo xx, precisamente tras la guerra, vinculado al emergente desarrollo del ordenador. Es curioso que, mientras en el mundo anglosajón denominaron a estas máquinas «computadoras» (*computers*), en alusión a su capacidad para realizar operaciones matemáticas de manera increíblemente rápida, en España preferimos llamarlas «ordenadores» (del francés *ordinateur*), enfatizando su capacidad para ejecutar una secuencia de órdenes de manera ordenada. Así, una computadora sugiere ser una «calculadora avanzada», mientras que un ordenador sugiere ser una especie de «robot que recibe órdenes». En ninguno de los dos casos, no obstante, se intuye el concepto «pensar», propio de la IA.

Sin embargo, la IA nació bajo la forma de algunos programas informáticos que eran capaces de ejecutar órdenes y realizar cálculos, con dos salvedades:

- Estas órdenes eran más avanzadas que las primitivas básicas de un ordenador: mientras las secuencias típicas de un ordenador son del tipo «sumar los números de los registros A y B», estos programas permitían órdenes más abstractas como, por ejemplo, «aumentar un 10 % el control de temperatura del depósito de agua».
- La ordenación de las secuencias también era bastante más avanzada: cuando los programas tradicionales de ordenador ejecutan una secuencia tras otra y, como mucho, forman «bucles», repitiendo varias veces un conjunto de secuencias, estos programas permitían escenarios condicionales —denominados «*if... then...*» en inglés, o «si... entonces...» en español— dependientes de algunas variables del entorno. Por ejemplo, una orden de este tipo podría ser: «si el tanque está lleno por encima del 50 % de su capacidad, la temperatura del líquido que contiene está por debajo de 50 °C y lleva calentándose menos de una hora, abrir el grifo para añadir más agua»).

Es decir, la IA representa una serie de reglas de razonamiento que consideramos complejas, que provocan determinadas acciones o resultados, basadas en una comprensión específica del entorno y que tratan de reproducir la forma de razonar de los humanos. Tanto es así que las modernas técnicas de IA tratan de emular incluso cómo aprendemos y procesamos la información para ser capaces de razonar, llegar a unas conclusiones y, si es necesario, tomar decisiones al respecto. Existe una expresión acuñada para describir las tecnologías orientadas a conseguir esto con mayor eficacia o efi-

ciencia:¹ «tecnologías cognitivas». Así, las tecnologías cognitivas se asemejan a la acción de la mente humana, lo que significa que proporcionan a los ordenadores las funciones de pensar y actuar como seres humanos.

Sin embargo, como ya se ha indicado, el concepto de IA está vivo y sigue evolucionando. La conocida como «prueba de Turing», ideada por el matemático e informático Alan Turing en 1950, también denominada en inglés «*the imitation game*» («el juego de la imitación»), consiste en utilizar la técnica de exponer la resolución de un problema realizada por un humano y por un ordenador a un observador. Aunque el observador sabe *a priori* que una de las dos soluciones (o representaciones; imaginemos, por ejemplo, una imagen) está hecha por ordenador, se determinará que tal ordenador está ejecutando un algoritmo de IA si dicho observador es incapaz de identificar cuál de las dos es. Esta definición, aunque es muy gráfica y sencilla, y está generalmente aceptada en la actualidad, no tiene en cuenta que las capacidades y el entrenamiento del propio observador para identificar lo que es humano y lo que es artificial ha ido evolucionando a lo largo del tiempo.

De todo lo anterior se puede extraer a modo de resumen que, aunque hay varias definiciones para la IA, muchas de ellas vagas o cambiantes a lo largo del tiempo, existe un consenso en la comunidad científica que acepta que el núcleo de la IA son las teorías, métodos, tecnologías y aplicaciones de investigación para simular, complementar y expandir la inteligencia humana. Sin embargo, como veremos más adelante, esta idea abre senderos aún inexplorados y que, con gran probabilidad, podrían transfor-

1. Existe una confusión generalizada en el uso de estos dos términos: mientras «eficacia» hace referencia a alcanzar el máximo grado de consecución de un objetivo, «eficiencia» indica alcanzar el objetivo de una manera aceptable utilizando el mínimo de recursos posibles.

mar radicalmente la IA, tal como la conocemos actualmente. La raíz de la cuestión está en la gran diferencia que existe entre los conceptos «simular», «complementar» y «expandir» antes mencionados. En la actualidad, las técnicas y soluciones de IA están más centradas en los dos primeros, pero cuando comience la siguiente revolución centrada en expandir la inteligencia humana a nuevas fronteras, probablemente todo cambiará como nunca antes ha ocurrido en la historia.

Como se indica en [art 26],¹ los campos de investigación de la IA incluyen ingeniería y sistemas, ciencias del cerebro, psicología, ciencias cognitivas, matemáticas, informática y muchos otros. Por ejemplo, hay soluciones de IA para reconocimiento de voz, procesamiento de imágenes, procesamiento del lenguaje natural (o NLP, por las siglas de «*natural language processing*», en inglés), robots avanzados inteligentes, vehículos autónomos, sistemas de distribución de la energía, atención sanitaria y telemedicina, soluciones financieras avanzadas, etcétera.

De hecho, en algunas soluciones concretas, la IA ha superado tanto en eficacia como en eficiencia a los humanos y constantemente se leen noticias que sugieren que la IA realiza el trabajo de determinadas personas con mayor fiabilidad, en menos tiempo y con un coste menor. Aunque esto pueda generar, *a priori*, cierta preocupación, conviene resaltar que todas estas soluciones, a pesar de lo espectaculares que puedan parecer o atractivas para abrir titulares en los medios, no dejan de pertenecer al grupo de la IA «estrecha» o «restringida» (ANI, por las siglas de «*artificial narrow intelligence*», en inglés). Este concepto se asocia a la capacidad de una solución de IA para superar el promedio de las capacidades humanas en una determinada tarea. Aunque, lógicamente, transformarán determinados procesos y formas de acometer ciertos trabajos, no suponen un mayor riesgo, al igual que te-

1. Véase la bibliografía al final de este libro.

ner calculadoras, la máquina de vapor o el avión, o incluso la pólvora o la bomba atómica no ha supuesto una amenaza para la especie humana, al menos hasta el momento.

Sin embargo, si una solución de IA fuera capaz de resolver *todas* las tareas propias de un humano, pertenecería al grupo de la IA «general» (AGI, por las siglas de «*artificial general intelligence*», en inglés) y esto sí podría suponer un serio problema. ¿Por qué? Es fácil deducir que una solución de AGI no solo sería capaz de reemplazar por completo a un humano en cualquier campo, sino que también sería capaz de sustituirlo para *pensar* en cómo mejorar las propias soluciones de IA. El momento en que esto ocurra en el tiempo se conoce como la «singularidad» y constituye el punto en el que la IA evolucionará por sí misma, con independencia de la especie humana. El riesgo inherente que percibimos es que, quizás, la IA llegue a convertirse en la *especie* dominante sobre el planeta y, al igual que hemos hecho los humanos con el resto del ecosistema, someta a todos los seres «inferiores», incluidos nosotros. Es más, aunque desconocemos los motivos, la realidad es que el *Homo sapiens* es el único homínido que ha quedado vivo sobre el planeta. ¿Exterminó nuestra especie a quienes eran muy parecidos a nosotros para evitar una amenaza? ¿Tuvimos ese punto extra de inteligencia y de crueldad para hacerlo? ¿Ocurrió de manera natural? Sea como fuere, la situación entre los humanos y la AGI será muy similar y, por lo tanto, hasta cierto punto es razonable que exista el miedo a resultar exterminados como especie por esta.

Sin embargo, aunque algunas voces pronostican la llegada inminente de la AGI, para alcanzar la singularidad la IA no debería necesitar ser programada explícitamente para realizar una tarea inteligente. Debería poder sentir, interpretar, aprender, planificar, comprender y actuar por sí misma, lo que significa que la IA debería poder interpretar correctamente datos externos, aprender de estos datos y utilizar este aprendizaje para lograr objetivos

y tareas específicas a través de una adaptación flexible y sin la intervención humana. Además, esto debería lograrse sin seguir unas reglas predeterminadas ni que existiera un observador evaluando los resultados de las acciones durante todo el proceso. Así, aunque en algunas aplicaciones experimentales se sigue avanzando en esta línea, las soluciones que vemos en el mercado, por muy avanzadas y complejas que nos parezcan, distan mucho aún de este punto.

A lo largo de este libro, se hará un recorrido sobre las técnicas y soluciones de IA, comenzando por las originales y más simples para continuar con las más complejas y acabar planteando los posibles escenarios a los que cualquier persona o empresa se enfrenta tanto en la actualidad como en un futuro más o menos próximo.

Comencemos, por tanto, viendo un poco de historia.