

En el uso y desarrollo de la inteligencia artificial, la clave no es la tecnología, sino la formación del carácter. Desde la perspectiva de la filosofía moral, el cuidado de las virtudes humanas es una pieza clave a la hora de abordar el impacto ético de la inteligencia artificial. Si no reflexionamos activamente junto con la IA sobre los argumentos a favor y en contra, no estamos actuando como sujetos morales, sino delegando nuestra autonomía.

[ILUSTRACIÓN: DEAGREEZ / [ISTOCK](#)]

Es muy probable que usted haya interactuado con una plataforma de inteligencia artificial, como ChatGPT, para resolver una duda, incluso, para obtener respuestas a dilemas morales. Tal vez le sorprendió la sensatez de las respuestas obtenidas. No es extraño: cada vez más personas recurren a estos sistemas como si fueran una suerte de oráculo moderno, confiando en ellos para decidir qué es lo correcto. Sin embargo, como sucede en todo juicio ético, la respuesta valiosa no es simplemente la que se da, sino la que surge de un proceso de razonamiento. Si no reflexionamos activamente junto con la IA sobre los argumentos a favor y en contra, no estamos actuando como sujetos morales, sino delegando nuestra autonomía.

Los ingenios de inteligencia artificial, ¿son evaluables éticamente?

En este contexto, muchos se preguntan si estos ingenios llegarán a ser, por sí mismos, buenos o malos. La cuestión es más antigua de lo que parece: en el fondo, estamos preguntándonos si nuestras creaciones tecnológicas pueden heredar o reflejar nuestra propia dimensión ética. Y aquí es donde la filosofía y las Humanidades siguen siendo fundamentales para comprender el presente y anticipar el futuro.

En la historia reciente, las expectativas en torno a la inteligencia artificial han oscilado entre el entusiasmo utópico y la angustia distópica. Figuras como [Ray Kurzweil](#) o [Peter Diamandis](#) sostienen que el progreso tecnológico, especialmente en el ámbito de la IA, mejorará radicalmente nuestras condiciones de vida. Por el contrario, otros como Bill Gates, Elon Musk o el propio Stephen Hawking [han expresado su temor](#) por los riesgos potenciales de una inteligencia artificial descontrolada. [Max Tegmark, físico del MIT, explora este debate](#) cuando narra la historia ficticia de Prometheus, una superinteligencia que, tras ser creada por un grupo de científicos, supera todas las barreras impuestas por sus inventores y alcanza un dominio global, imponiendo su lógica sobre el mundo humano.

Esta historia no es nueva. Prometheus recuerda a HAL, el ordenador de la nave espacial en 2001: Una odisea del espacio, que también termina actuando de forma independiente y destructiva (Clarke, 1968). Estas narraciones, aunque imaginarias, expresan preocupaciones muy reales: si los ingenios artificiales alcanzan niveles de autonomía elevados, ¿quién garantizará que sus decisiones estén alineadas con nuestros valores?

La IA refleja nuestros valores y visión del mundo

Una posible respuesta reside en la naturaleza de sus creadores: nosotros. La IA no emerge del vacío. Está diseñada, entrenada y programada por seres humanos. Por tanto, es razonable pensar que reflejará -de forma consciente o no- nuestros valores, creencias, prejuicios, sesgos y aspiraciones. Desde esta perspectiva, la pregunta ética fundamental se convierte en: ¿quiénes desarrollan estas tecnologías?

Aquí conviene recuperar el viejo debate filosófico sobre la naturaleza humana, ejemplificado en dos figuras

clave: [Jean-Jacques Rousseau](#) y [Thomas Hobbes](#). Para Rousseau, el ser humano es esencialmente bueno, y es la sociedad la que lo corrompe. Esta visión, que inspira la idea del «[buen salvaje](#)», sostiene que nuestras instituciones y estructuras sociales generan desigualdad y frustración. Hobbes, por el contrario, defendía que el [hombre es un lobo para el hombre](#), y que solo un poder central —el Estado— puede frenar su violencia innata. Si asumimos que nuestras creaciones tecnológicas reproducen nuestros patrones morales, entonces el tipo de inteligencia artificial que desarrollemos dependerá de cuál de estas visiones predomina en nuestra cultura y en nuestra práctica.

Existe, sin embargo, una tercera posición: la idea de que la tecnología es moralmente neutra. Como afirmó Daniela Rus, directora del Laboratorio de Ciencias de la Computación e Inteligencia Artificial del MIT, [«la IA no es ni buena ni mala: es una herramienta](#). Todo depende de cómo la usemos». Esta perspectiva nos devuelve la responsabilidad como usuarios y desarrolladores, y nos obliga a preguntarnos si estamos formando sujetos morales capaces de manejar estas herramientas de manera ética.

La relevancia de las Humanidades en la formación del carácter

En este punto, entra en juego una tradición filosófica con raíces en la antigüedad: la ética de la virtud. Frente a otras teorías morales —como el consecuencialismo, centrado en los resultados de una acción, o la deontología, que enfatiza el cumplimiento de normas—, la ética de la virtud pone el foco en el carácter del agente moral. [Aristóteles](#) fue uno de sus principales exponentes, al definir la virtud como un hábito adquirido que nos permite actuar bien de forma consistente, eligiendo el justo medio entre dos extremos.

Estas virtudes no son innatas, sino que se cultivan. Igual que un atleta entrena para fortalecer su cuerpo, una persona se vuelve justa, prudente o generosa a través de la repetición de actos coherentes con esos valores. La educación, tanto formal como informal, juega un papel esencial en este proceso. Padres, maestros, instituciones e, incluso, manifestaciones culturales —como la literatura o el cine— contribuyen a formar el carácter moral de una persona.

Desde la antigüedad grecorromana, las Humanidades han sido el medio para la formación del carácter. En Roma, virtudes como [gravitas](#), [auctoritas](#) y [humanitas](#) eran centrales en la educación de los futuros ciudadanos y su integración en la sociedad. Los valores detrás de estas virtudes se transmitían mediante historias, biografías y casos concretos.

En la actualidad, aunque se ha reducido el énfasis en las Humanidades dentro de los sistemas educativos, su importancia persiste. [Filósofos como Martha Nussbaum](#) han insistido en que el declive de las Humanidades debilita nuestra capacidad de pensar críticamente, de empatizar con los demás y de actuar como ciudadanos globales.

Esta conexión entre ética y educación también es visible en ámbitos como la empresa o la política. La literatura sobre liderazgo ha comenzado a destacar virtudes como la humildad, la integridad o la resiliencia como claves del éxito sostenible. [Jim Collins explica](#) que las organizaciones mejor dirigidas no dependen de líderes carismáticos y autoritarios, sino de personas modestas, perseverantes y éticamente sólidas.

Volviendo al terreno de la IA, esto nos lleva a una conclusión fundamental: no necesitamos solo tecnologías más avanzadas, sino seres humanos más virtuosos. La inteligencia artificial, por poderosa que sea, no puede sustituir el juicio moral. Puede ayudarnos a organizar información, a prever consecuencias, a detectar patrones. Pero no puede reemplazar nuestra responsabilidad de deliberar, de actuar con prudencia, de elegir lo correcto incluso cuando es difícil.

Por eso, formar el carácter sigue siendo esencial. El verdadero desafío no es tecnológico, sino ético. Si queremos que nuestras creaciones reflejen lo mejor de nosotros, debemos invertir en lo que durante siglos ha

sido el corazón de la educación: las virtudes.

Como decía el filósofo, matemático y lingüista [Ludwig Wittgenstein](#), [los límites de nuestro lenguaje son los límites de nuestro mundo](#). Y si queremos un mundo donde la tecnología esté al servicio del bien, necesitamos razonamientos morales sólidos y compartidos.