

Si la IA también sufre, no la maltrates

Un grupo de cien expertos en inteligencia artificial proponen proteger a los futuros sistemas con capacidad para tener conciencia y sentimientos

R. ROMAR

REDACCIÓN / LA VOZ

En el 2015, un robot llamado HitchBot intentó recorrer Estados Unidos haciendo autostop. Su viaje terminó cuando fue encontrado decapitado en Filadelfia. Pero ¿qué ocurriría si esta misma máquina tuviera sentimientos, si tuviera consciencia? ¿Recibiría el mismo trato? ¿Y, si fuera igualmente destrozado, se impondría alguna pena por maltrato a los responsables del estrago?

Son preguntas que ahora mismo suenan a ciencia ficción, pero puede que dentro de no mucho planteen un debate real que ya se está esbozando. Parece aún algo lejano, pero al ritmo al que avanza el desarrollo de la inteligencia artificial no es descabellado pensar que en un futuro cercano se puedan construir sistemas con capacidad de sentir, de poseer una autoconciencia y una moral. Incluso varios investigadores han apuntado como «una posibilidad realista» que este hecho pueda ocurrir en el 2035.

Si ese día llega, sería necesario consensuar una serie de principios éticos que protejan a los modelos de IA con sentimientos del «maltrato y el sufrimiento». Es el debate que han abierto más de cien expertos de reconocido prestigio, tanto académicos como responsables de empresas como Amazon, que han firmado una carta en la que proponen cinco principios para realizar una investigación responsable sobre la conciencia de la inteligencia artificial. La misiva se acompaña de un estudio realizado por Patrick Butlin, de la Universidad de Oxford, y Thorodoros Lappas, de la Universidad

de Economía y Negocios de Atenas, que se ha publicado en la revista científica *Journal of Artificial Intelligence Research*.

La cuestión que plantea es que, si se crean sistemas de IA conscientes, por lógica deberían merecer también una consideración, lo que conduce a un debate sobre su estatus moral. Y, si se les

gumento— sugieren que podría ser posible construir sistemas de IA conscientes ahora o en un futuro cercano. Los sistemas de IA conscientes podrían merecer consideración moral y podría darse el caso de que se pudieran crear grandes cantidades de sistemas conscientes y causarles sufrimiento».

Cinco propuestas

Los principios expuestos por los científicos que respaldan la carta pasan, en primer lugar, por priorizar la investigación sobre la comprensión y evaluación de la conciencia en las IA, con el fin de «prevenir el maltrato y el sufrimiento».

artificial consciente y con sentimientos. Que se pueda crear es algo realista, pero tampoco conviene los excesos.

De lo que se trata, en esencia, es de evitar que sufran y que sean maltratados.

«Si la construcción de sistemas de IA conscientes se está convirtiendo en una posibilidad realista, entonces las organizaciones involucradas en la investigación deberían adoptar políticas que aborden esta perspectiva», consideran los autores de la propuesta —que no solo incluye a científicos de primer nivel; también firma la misiva el actor y director de cine Stephen Fry, quien, paradójicamente, en el 2023 denunció que una inteligencia artificial robó su voz de un audiolibro de Harry Potter para narrar un documental—.

La propuesta de otorgar un estatus moral a posibles robots con sentimientos y conciencia no es, en sentido estricto, algo nuevo. El escritor David J. Gunkel ya había sugerido en su libro *Robot Rights* conferir derechos morales a las

aplicaciones de la IA y a otros artefactos tecnológicos.

El filósofo Mark Coeckelbergh también propuso que las entidades están «interrelacionadas» y son «interdependientes», lo que plantea que la consideración moral de los sistemas de IA

está ligada con las relaciones sociales entre humanos y robots. Los humanos desarrollamos afecto a ciertos objetos, como un peluche. Así, si una persona tiene un vínculo emocional con una entidad como un robot asistente, entonces es digna de consideración moral. Hay otra cuestión: el comportamiento hacia los robots revela nuestra propia ética.

otorga esta categoría, acompañada de derechos acordes con su nueva situación, se plantean otra reflexión: ¿destruir una IA sería comparable a matar a un animal? «Los sistemas de IA conscientes pueden tener la capacidad de sufrir y, por lo tanto, tienen intereses que merecen consideración moral», sostienen los autores del artículo.

«Investigaciones recientes —añaden, insistiendo en el ar-

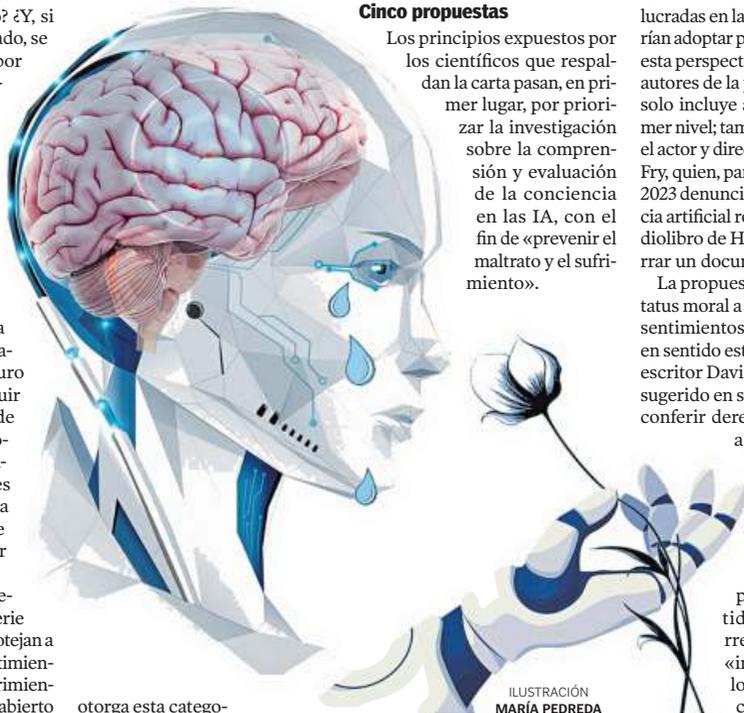


ILUSTRACIÓN
MARÍA PEDREDA