

Los robots están saliendo de las fábricas y comienzan a convivir en sociedad. Sin embargo, los que colaboran en hospitales y escuelas, los futuros coches sin conductor y el resto de las máquinas semiautónomas surgidas de la inteligencia artificial se enfrentan al reto de demostrar que poseen la ética indispensable para confiar en ellas.

Manual de ética para robots



Texto de **Antonio Orti**
y fotos de **Mané Espinosa**

Se llama Reem, mide 1,70 metros de estatura y trabaja las 24 horas. No duerme, no come, no fuma y no pierde la compostura cuando las cosas se ponen feas. En principio, su destino es Dubái, donde ya operan otros robots policía como él, capaces de desplazarse a 4 km/h, abrir puertas, identificar a sospechosos (comparando sus caras con las de una base de datos), leer matrículas e incluso gestionar denuncias a través de la pantalla táctil que lucen en el pecho.

La falta de cabrioleo de sus ojos (dos cámaras de videovigilancia) no ha impedido que el Gobierno de Dubái haya manifestado su propósito de que, si la experiencia es exitosa, para el año 2030, un 25% de sus fuerzas de seguridad sean robóticas. Eso sí, este androide no podrá arrestar a personas ni disparar armas. “Pero Reem también podría ser recepcionista de hotel, azafato de congresos (o punto de información dinámico), incluso ocuparse de atender a los clientes”, indica Francesco Ferro, director general de Pal Robotics, la empresa que lo fabrica.

Desde una sala acristalada, Ferro observa el ballet imposible que ejecutan otros tres autómatas. El más cercano, por ejemplo, se llama Tiago y es capaz de monitorizar las constantes vitales de un ser humano, así como de proponerle ejercicios físicos, “*trackeando* la silueta de la persona para comprobar si los ejecuta bien”, explica Judith Vilado-

mat, responsable de marketing de la empresa. También avisa a las personas enfermas o dependientes sobre la hora a la que han de tomar la pastilla y les recomienda platos saludables.

La familia de robots de esta empresa (en las imágenes) es más numerosa e incluye a Stockbot (su especialidad es inventariar tiendas y almacenes) y Talos, el único humanoide, que dispone de dos piernas y cuyo punto fuerte es trabajar en entornos peligrosos y fatigosos, como las cadenas de montaje de aviones.

La tecnología avanza rápido y los ingenieros trabajan duro para dotar a estos ingenios mecánicos surgidos de la inteligencia artificial (IA) con algo parecido a la ética, en vista de que los robots ya están saliendo de las cadenas de montaje de las fábricas y comienzan a entrar en los hogares y a circular por las carreteras, como ya sucede en Arizona, donde es posible cruzarse con camiones sin nadie al volante, pilotados por algoritmos.

Para garantizar que los robots no se comporten imprevisiblemente, Karolina Zawieska, una investigadora especializada en ética y aprendizaje cultural de la robótica que trabaja en la Universidad de Montfort, en Leicester (Reino Unido), explica que desde el 2004 ha surgido una nueva disciplina, la roboética, que se centra en aplicar la ética humana a la robótica y que está propiciando que los programadores colaboren cada vez más estrechamente con lingüistas, historiadores, antropólogos, filósofos y psicólogos.

“Los desarrolladores de ro- →

Si los robots van a realizar según qué tareas, que hasta ahora hacían humanos, deberán tener en cuenta las preocupaciones, preferencias y necesidades humanas



Uno de los dilemas entre los ingenieros, pero también entre los expertos de otras disciplinas, es hasta qué punto se deben humanizar los robots

¿DEBERÍA EL COCHE SIN CONDUCTOR PODER ELEGIR A SUS VÍCTIMAS?

¿Debería desviarse de su trayectoria para matar a una persona en lugar de a cinco? ¿Sería mejor que atropellara a una anciana en vez de a una embarazada? ¿Y si tuviera que elegir entre un indigente y un ejecutivo? Hasta la fecha, más de dos millones de personas han respondido a estas preguntas, una encuesta impulsada por el Instituto de Tecnología de Massachusetts (MIT) que busca determinar si se debería programar el comportamiento del coche autónomo ante un posible atropello. Los resultados sugieren que sería complicado implantar un único conjunto de normas en todos los países: mientras los occidentales tienden a salvar en mayor medida a los jóvenes, en Asia protegen más a las personas mayores.

O, según las respuestas, los estadounidenses se inclinan por salvaguardar más a las personas de compleción atlética que a las obesas, pero los países del sur amparan en mayor medida a las mujeres, sean como sean. Los expertos creen que los coches autónomos se deberían limitar a cumplir las normas de circulación sin preestablecer preferencias en función de la edad, la constitución física o la clase social de una posible víctima, ya que todas las vidas valen lo mismo. Y ponen este ejemplo: si por una acera caminan cinco personas necesitadas de un trasplante (de riñón, hígado, etcétera) y por la acera contraria anda un peatón sano... ¿sería éticamente admisible sacrificar al sano para salvar a las otras cinco?

→ bots –explica Zawieska– son expertos en la construcción de robots, pero no necesariamente en abordar las preocupaciones de la sociedad y las preferencias y necesidades humanas”. Se trata de garantizar, por ejemplo, que cuando un robot haga el trabajo de un enfermero sepa comportarse como es debido, de cara a garantizar no solamente la seguridad del paciente, sino también su bienestar mental. Una forma de conseguirlo es incrustar en el código de programación del autómatas las normas éticas de Buda, Aristóteles, Platón, Descartes y otros grandes filósofos antiguos y actuales.

minante, indica Zawieska, es que “la roboética no es la ética de los robots, ni ninguna ética artificial, sino la ética humana de los diseñadores, fabricantes y usuarios de robots”.

El objetivo de la IA es crear sistemas que emulen al cerebro humano y tengan conciencia de sí mismos y del mundo que les rodea. Muchas de estas tecnologías basadas en el *machine learning* (aprendizaje automático) ya se han introducido silenciosamente en nuestras vidas: desde asistentes personales como Siri (Apple), Cortana (Microsoft) o Alexa (Amazon), hasta las sugerencias de plataformas como Netflix o Spotify, pasando por empresas que atienden a sus clientes a través del teléfono


El impacto de la inteligencia artificial podría ser muy positivo y en muchos ámbitos si se logra que la sociedad confíe en ella, dice el directivo de Microsoft David Carmona


sin necesidad de tener a una persona al otro lado de la línea.

Por este motivo, David Carmona, director general de desarrollo de negocio en inteligencia artificial de Microsoft, explica en conversación telefónica desde Redmond (EE.UU.) –a unos 20 kilómetros de distancia de Seattle–, donde está la sede central de esta empresa, que en los próximos años la IA “podría convertirse en uno de los acontecimientos más grandes de la historia de la humanidad”.

El objetivo, explica Carmona, es que cada vez más las máquinas hagan cosas que antes sólo podían hacer los humanos. Por ejemplo, automatizar la producción mundial de alimentos o detectar un posible cáncer de

piel mucho antes que el dermatólogo más experimentado. La IA también podría ayudar a millones de personas con algún tipo de discapacidad a tener una mejor interacción con su entorno, de la misma manera que se convertirá en una herramienta vital para prevenir el cambio climático. Es decir, su impacto podría ser muy positivo, “siempre y cuando consigamos que la sociedad confíe en ella”, concluye Carmona.

No obstante, la invasión de los robots buenos suscita interrogantes. Según indica Carme Torras desde el Instituto de Robótica e Informática Industrial del CSIC y la Universitat Politècnica de Catalunya, en Barcelona, cada vez más abundan las


Los robots no dejan de ser máquinas; quienes debemos ser éticos somos los humanos que los utilizamos, advierte la experta en robótica Carme Torras

máquinas que pueden tomar decisiones por sí mismas, pese a que hasta ahora los únicos robots que tenían algo parecido a la ética humana eran Astroboy y otros conocidos héroes de ficción. “Al final, se trata de adaptar la ética humana a la robótica”, señala Torras. “Muchas de las nuevas máquinas autónomas –subraya esta experta en roboética– no dejan de ser electrodomésticos más versátiles que nos liberan tiempo de trabajo para dedicarlo a otras cosas. No dejan de ser máquinas. Quienes debemos ser éticos somos los humanos”.

No obstante, la proliferación de sistemas basados en la IA ha hecho saltar algunas alarmas. En marzo del 2018, por ejemplo,

se produjo en Tempe (EE.UU.), el primer atropello mortal protagonizado por un coche sin conductor. Por otra parte, desde que los drones transformaron la manera de hacer la guerra, cada vez más países disponen de robots bélicos teledirigidos. Pero... ¿quién controlará que no cambien de bando o que cumplan con las convenciones internacionales? ¿En qué casos podría estar justificado que una máquina tuviera derecho a matar? “En ningún caso”, responde Torras, que respalda la campaña Stop Killer Robots, para prohibir las armas autónomas.

Otro tanto opina Ferro. “Una de las líneas rojas de Pal Robotics es utilizar la IA para fines militares. Por este motivo, me →

QUIÉN PAGA SI EL ROBOT FALLA

→ gusta la denominación ‘robots colaborativos’ porque da a entender que los robots no son Terminators, como nos ha hecho creer Hollywood”, aclara.

El hecho de que la IA se encuentre todavía en una fase incipiente ha propiciado algunos experimentos fallidos. En el 2015 Google tuvo que pedir disculpas después de que el algoritmo de Google Photos etiquetara como “gorilas” la foto de dos afroamericanos. En cuanto a Amazon, canceló un robot que evaluaba currículums por un sesgo machista que le llevaba a decidir que los mejores candidatos eran casi siempre masculinos.

Más relevante todavía fue lo que le ocurrió a Microsoft con Tay, un programa de IA diseñado para entablar conversaciones como si fuera una encantadora joven de entre 18 y 24 años. El objetivo era que Tay aprendiera de la interacción con los usuarios. Y eso hizo, sólo que, en cuestión de minutos, comenzó a publicar en Twitter frases como “odio a las feministas, deberían morir todas y pudrirse en el infierno” o “Hitler habría hecho un trabajo mejor que el mono (Barack Obama) que tenemos ahora”, lo que provocó que Microsoft cerrara la cuenta de la tuitera Tay a las 16 horas de abrirla.

“Lo que ocurrió con Tay –explica David Carmona a toro pasado– es que recibió un ataque coordinado de muchísimos usuarios que empezaron a envenenarla con expresiones que eran ofensivas e intolerantes y que luego Tay repetía en otras conversaciones. Lo que hicimos

En el supuesto de que un robot inteligente dañe a una persona... ¿a quién habrá que pedir responsabilidades? ¿al fabricante?, ¿al creador del software?, ¿al usuario? “Para mí, el responsable desde el punto de vista mecánico es el fabricante”, responde Francesco Ferro, director general de Pal Robotics. “Ahora bien, si este bolígrafo que sirve para escribir se lo meto a alguien en el ojo, ¿el fabricante debería responder por ello?”, cuestiona para dar a entender la complejidad del asunto. Más preguntas: ¿cómo se defenderá un coche autónomo ante un tribunal para probar su inocencia? Para un grupo de expertos británicos en robótica, una posible solución sería que los robots dispusieran de una caja negra, similar a la de los aviones. Sin embargo, hasta la fecha, indica Silvia Leal, asesora de la Comisión Europea, “sigue estando en el aire quién será el responsable, algo que preocupa a las compañías de seguros”. Si las citadas compañías pusieran el listón muy alto y exigieran que la inteligencia artificial no fallase nunca, la innovación podría verse lastrada. En previsión de ello, el Parlamento Europeo estudia cómo evitar que los presuntos culpables se pasen la pelota ante posibles perjudicados que puedan recibir una compensación.

fue retirarla del mercado, aprender de los errores y publicar los resultados para que otras empresas puedan prepararse para estas nuevas amenazas que no estaban presentes en los sistemas informáticos del pasado”.

En previsión de que estos excesos se generalicen en el futuro, la empresa española Acuilae comercializa desde esta primavera un módulo informático denominado Etyha que, según explica Cristina Sánchez, su fundadora, tiene por objetivo evitar que un sistema de IA se corrompa durante el contacto

con seres humanos. “¿De qué manera? Pues enseñándole de lo que tiene que aprender y de lo que no. Pero deberá ser el cliente quien decida, en última instancia, de qué quiere protegerse”, precisa esta informática.

Por lo que se observa entre bambalinas, aunque todos los problemas importantes del ser humano se pueden abordar cada vez más usando la IA, la privacidad de los datos y la ética tras las decisiones y acciones siguen siendo los principales escollos.

Tal vez por ello, David Cuartielles, investigador asociado al

grupo Internet of Things and People en la Universidad de Malmö (Suecia) y cofundador de la plataforma de software libre Arduino, explica en conversación telefónica que la pregunta es “si estamos preparados como sociedad para admitir el hecho de que algunas máquinas decidan por nosotros”. Según este ingeniero, la solución pasa por que los humanos sigan teniendo algún tipo de soberanía algorítmica.

“Ya no hay humanos por un lado y máquinas por el otro, sino una mezcla, aunque, obviamente,

tenemos que prevalecer como quienes controlan la situación”, apunta. “Hemos de estar preparados para una fase de transición, en la que convendría no echarse atrás a las primeras de cambio, mientras establecemos los parámetros que hagan que las máquinas funcionen como deseamos”, concluye.

Otro tanto opina Carmona al señalar que el gran objetivo de la IA “es poner a las personas en el centro, en lugar de reemplazarlas”. A título informativo, el directivo de Microsoft explica que su empresa ha lanzado en



El investigador David Cuartielles se pregunta “si estamos preparados como sociedad para admitir el hecho de que algunas máquinas decidan por nosotros”

China un agente conversacional llamado Xiaoice que, gracias a sus capacidades emocionales, puede mantener conversaciones telefónicas sin que su interlocutor llegue a darse cuenta de que es un robot. Tanto es así que los ciudadanos de China (donde este agente conversacional cuenta ya con 660 millones de usuarios) buscan su consejo en todo tipo de temas personales y le hablan de su familia, de la elevada contaminación, de sus desengaños amorosos...

Aunque en Japón ya existe un robot creado a imagen y se-

mejanza de la actriz Scarlett Johansson y otro llamado Michihito Matsuda, que se presentó en el 2018 como candidato a la alcaldía de un distrito de Tokio y quedó en tercera posición con 4.013 votos, se trata de episodios pintorescos. La batalla para que la IA despliegue su potencial camina por otros derroteros y se centra, según los expertos, en garantizar la privacidad de los datos y en reforzar el mensaje de que detrás de cada robot hay mujeres y hombres interesados en que las cosas vayan mejor que ahora.○



El aprendizaje automático de las máquinas sofisticará cada vez más su inteligencia, pero los expertos mantienen que los humanos no deben renunciar a tener el control