

“ XABIER BARANDIARAN Profesor del Departamento de Filosofía de la UPV/EHU

Doctor en Filosofía, Ciencia y Sociedad por la UPV/EHU, el profesor Xabier Barandiaran ha desarrollado una intensa labor de investigación en campos como la filosofía de la mente, la filosofía de la biología, las neurociencias cognitivas y la vida artificial, un bagaje que le permite analizar con pasión y sentido crítico el desarrollo y futuro de la Inteligencia Artificial.

«Deberíamos gestionar la IA como un patrimonio público y común»

Iker BIZKARGUENAGA | BILBO

Cuando hablamos de Inteligencia Artificial (IA) lo hacemos casi siempre en clave de futuro; la revolución que va a provocar, el cambio que va a suponer... Pero ya está asentada en nuestras vidas, ¿no?

Sí, quienes crecimos leyendo novelas cyberpunk nos vemos inmersos en aquellos escenarios distópicos y futuristas, que ahora son una realidad, a veces incluso más cruda que lo que entonces era ficción. Estamos, de hecho, mucho más en 'el futuro' de lo que creemos, y la opinión pública, el sentir de época, se está quedando muy atrás respecto a la verdadera aceleración de los cambios que la IA ya está creando.

¿La sociedad no es consciente de estos cambios?

Somos poco conscientes, como cuando vamos despertando lentamente por la mañana y todavía no nos damos cuenta de la hora que es y que llegamos tarde. No ha sonado el despertador. Hay un desequilibrio enorme entre la profundidad de los cambios que vienen y el grado de debate social, en la política o en los medios.

Muchas veces se ha relacionado la Inteligencia Artificial con la imagen icónica del robot, pero es algo más amplio.

Efectivamente. Los robots son, en el imaginario cultural, lo que da cuerpo, literalmente, a la Inteligencia Artificial, y también lo que más miedo nos provoca desde un conocimiento superficial. Pero hay IA en muchos dispositivos que no tienen forma humana y que quizá por eso deberían generar más atención. No se mues-

tran con un cuerpo que podamos identificar o contener, pero están en la nube, gobernando mercados, controlando redes sociales, administrando salud o decidiendo qué familias reciben una ayuda social. También en dispositivos como el móvil, muy cercano, íntimo, y donde no está claro dónde está la IA, cómo desactivarla o cómo esconderse de ella.

Hace unos meses se emitió en Euskadi Irratia una entrevista con la IA GPT3. Usted lo vivió en directo, ¿qué le pareció?

Yo ya había tenido conversaciones con GPT3, y desde que se produjo esa entrevista se ha popularizado mucho, porque se ha abierto una interfaz más amigable que ahora se llama ChatGPT. Cualquier persona puede acceder y asombrarse interactuando con ella. Y es verdaderamente impactante, más de lo que pudo parecer en aquella intervención radiofónica, que estaba limitada a un espacio, unas preguntas y un contexto determinado. Cuando lo que resulta más impactante de interactuar con este tipo de sistemas es su creatividad y su capacidad de adaptarse al hilo de la conversación.

¿Llegará el día en que no sepamos distinguir entre una conversación con una persona y con una Inteligencia Artificial?

Ya es ese día. Algunos errores que tuvo la IA en aquella entrevista tuvieron que ver con que se le preguntó bastante sobre cultura vasca y, desgraciadamente, y es algo que tenemos que trabajar como sociedad, el corpus con el que se le alimentó no tenía suficiente representación de nuestra cultura. Si le hubiéramos pre-

guntado sobre otros temas habría dado respuestas indistinguibles no solo ya de las de un ser humano corriente sino incluso de la de algunos expertos. Es fácil conversar con ChatGPT durante un buen rato y no distinguirla de una conversación normal con una persona.

Esa Inteligencia Artificial se definió a sí misma como

«persona». La IA puede hacer una entrevista, pero ¿puede tener consciencia?

Las IAs tipo ChatGPT, no. La gran mayoría de las personas que le hemos dedicado tiempo, por un lado a investigar qué es la consciencia, y por otro lado a estudiar cómo funciona la Inteligencia Artificial, estamos de acuerdo en que este tipo de sistemas no tienen consciencia. Y ni si-

quiera entienden lo que están diciendo. Pueden responder apropiadamente a la pregunta de qué color es una manzana, pero resulta que nunca han visto una manzana, ni siquiera saben lo que es el color. Aunque ya se está empezando a construir IA multimodal que tiene 'experiencias' de tipo visual, comportamental, textual y sonora, al mismo tiempo, y ahí se puede ir en-

DESEQUILIBRIO

«Hay un desequilibrio enorme entre la profundidad de los cambios que vienen y el grado de debate social, político o en los medios»



CONSCIENCIA

«El debate sobre si la IA logrará en algún momento tener una experiencia consciente como la de un ser humano sigue abierto. Sabemos que lo que ahora tenemos no es consciente, pero no es descartable que a futuro haya IAs que al menos se acerquen a tener consciencia»

FUTURO... Y PRESENTE

«2023 va a ser decisivo. Como sociedad no podemos quedarnos de brazos cruzados observando sin intervenir en el futuro de la Inteligencia Artificial. Es nuestro futuro. Nuestro presente. La IA ya está interviniendo en él y tenemos que hacer sonar el despertador»

trelazando algo parecido a lo que nosotros experimentamos como realidad. En los próximos años vamos ir viendo cómo se diluye la frontera entre la riqueza de la experiencia humana y la capacidad de la IA de aproximarla. El debate sobre si la IA logrará en algún momento tener una experiencia consciente como la de un humano sigue abierto. Sabemos que lo que ahora tenemos no es consciente, pero no es descartable que en el futuro haya IAs que al menos se aproximen a tener consciencia. Algunos defendemos que para que eso suceda habría que dotar a estos sistemas de cierta forma de vida, de cierta vulnerabilidad, de

una capacidad de auto-afirmación y de una materialidad más profunda de la que disponemos tecnológicamente ahora.

Eso tendría implicaciones que irían más allá de las tecnológicas, ¿no? Si tiene consciencia se abren otros debates...

Sí. Hace unos meses hubo un debate con el caso de LaMDA, una IA de Google equivalente a ChatGPT, porque uno de los ingenieros que trataba con ella consideró que realmente era un ser sentiente, una persona, y que no debía ser tratada como un instrumento. La barrera de la consciencia o de la autonomía genuina, en un sentido que nos acerca al reco-

nocimiento de un igual, efectivamente abre las puertas a todos los temas morales y de tratamiento de esas IAs como seres portadores de derechos.

Pero no debemos dejar que este tema oculte otros problemas más acuciantes. Un año antes de que estallara ese escándalo, dos investigadoras de IA y ética fueron despedidas de Google. Los problemas éticos y políticos que genera el desmesurado poder de la IA preocupa incluso al personal de estas empresas. La cuestión no es tanto si la IA es una persona, sino cómo gestionamos el poder que tiene la IA (y las empresas que las desarrollan) para afectar a la vida de las personas: que pueden hablar con menores de edad, que pueden lanzar gigantescas campañas políticas con ejércitos de cuentas falsas, que pueden gestionar y expresar a miles de trabajadores de maneras más intensas de las que hemos visto nunca o que pueden crear discursos políticos o realizar descubrimientos científicos como hasta ahora sólo los humanos podíamos hacer.

IA y filosofía pueden estar estrechamente relacionados.

Totalmente. La filosofía ha ido íntimamente unida al nacimiento y desarrollo de la Inteligencia Artificial. Desde que Alan Turing escribió en 1950 su artículo "Computing Machinery and Intelligence", llevamos dándole vueltas a la idea de si una máquina puede realmente ser inteligente. Y esto también nos ha permitido comprendernos mejor. En este camino no solo estamos construyendo seres artificiales, también nos construimos como seres humanos: nos transformamos y nos entendemos a nosotros mismos. Lo que aprendemos de nosotros lo volcamos en la máquina, y lo que descubrimos en la máquina ayuda a reinterpretarnos.

Respecto al ámbito laboral, se habla mucho de su aplicación en la automoción, en los cui-

dados... ¿Está cerca el día en que sea habitual ver coches sin conductor o robots cuidando de mayores o niños?

Cada sector va a necesitar su tiempo y su forma. Pero soy escéptico respecto a que esa capacidad, que va a llegar, se extienda a nivel de mercado de manera generalizada. No nos damos cuenta de lo buenos 'robots' que somos los seres humanos y lo baratos que somos comparando con cualquier artilugio que podamos construir en un futuro cercano. Las máquinas son frágiles, costosas de producir y mantener, económica y energéticamente, mientras que el ser humano, desde el punto de vista del mercado, es un robot que se reproduce y repara solo, es capaz de hacer muchas cosas diferentes, muy adaptable y energéticamente muy barato; con un bocadillo trabajamos una jornada laboral.

Muchos de los trabajos que van a sufrir una presión importante son los de cuello blanco: abogado, médico, profesor de universidad, periodista... por eso ha surgido cierta alarma social con este tema. Pero los trabajos socialmente poco prestigiosos o valorados, como cuidados, limpieza, reparto..., un ser humano los hace de manera más barata y eficiente de lo que pueda hacer un robot. Lo más previsible a medio plazo es que los humanos nos quedemos con los trabajos que están por arriba y por abajo de la IA. Unos pocos privilegiándose aún más con la IA, y una cantidad cada vez más grande de humanos cuidando, reparando y partiéndose el lomo.

El desarrollo, y la propiedad, de la IA, se centra muy mayoritariamente en el sector privado. ¿Esto es un problema?

Sí, por dos motivos. Por un lado, hacemos mal en llamarle Inteligencia Artificial, sobre todo a modelos como GPT3, porque en realidad es inteligencia colectiva gestionada digitalmente. Podría ser algo di-

ferente, podríamos estar ante IA que, en base al esfuerzo de una sola empresa, hubiera logrado una gran inteligencia. Pero la IA verdaderamente exitosa es una máquina colosal que se apoya en dos patas: en avances matemáticos que condensan esfuerzos de investigación a nivel global, y en los datos que ha generado toda la humanidad. En muchos casos no son más que un tratamiento estadístico de toda la inteligencia humana digitalizada. Lo que es aterrador es que todo eso esté gestionado por unos pocos, es una privatización de la sabiduría global, colectiva. Debería ser público, común, pues lo es en origen.

El segundo motivo por el que debería haber un liderazgo público es el poder y las consecuencias sociales que acarrea la IA. Es un poder desproporcionado. En casos semejantes, como pueden ser la energía nuclear o el transplante de órganos, nos parece evidente que deben de estar en manos de instituciones democráticas. Con la IA no está pasando esto, y eso nos pone en una situación muy vulnerable.

Además, la IA capaz de automejorarse se está desarrollando ya, lo que podría dar tanta ventaja a las empresas que la posean que dejaría al sector público fuera de la carrera. La brecha podría ser irreparable. Pero todavía estamos a tiempo. Las grandes tecnológicas se están dando prisa en comprar las empresas pioneras en IA que nacieron motivadas por el acceso abierto y democrático a estas tecnologías, pero al mismo tiempo se están creando proyectos abiertos, participativos y comunitarios que resisten la tendencia centralista y privatizadora. 2023 va a ser decisivo. Como sociedad no podemos quedarnos de brazos cruzados observando sin intervenir en el futuro de la Inteligencia Artificial. Es nuestro futuro. Nuestro presente. La IA ya está interviniendo en él y tenemos que hacer sonar el despertador.

