



De la Inteligencia Artificial a la inmortalidad del alma

N. 269. Septiembre del 2023. Suplemento del Cuaderno CJ n. 234
Cristianisme i Justícia. Roger de Llúria, 13, 08010 Barcelona
93 317 23 38 • info@fespinal.com • www.cristianismeijusticia.net

A finales de 2022 se dio a conocer al gran público el llamado ChatGPT, un *software* programado para reproducir el lenguaje humano y con la capacidad de responder a cualquier pregunta que se le plantee. El hecho de que este *software* haya aprendido a dar respuestas no prefijadas, gracias a un entrenamiento algorítmico mediante redes neuronales, sumado a la cantidad de información que ha llegado a digerir han puesto la expresión *inteligencia artificial* (IA) en boca de todos. Tenemos una herramienta capaz de escribir desde un poema romántico inédito al estilo de Goethe hasta un trabajo académico en el que se comparen dos autores que nunca han sido estudiados juntos y de hacerlo mejor que el 99% de los mortales. Las preguntas que esto nos suscita son muchas: ¿Es realmente inteligente la IA?, ¿dónde radica su novedad?, ¿qué consecuencias puede tener su popularización?, ¿ayudará a mejorar nuestro mundo?,

¿qué podemos esperar? Nos aventuramos, a continuación, a compartir algunas respuestas distintas de las que nos proporcionaría ChatGPT.

El problema no reside en la verdad, sino en la verosimilitud

Uno de los debates más importantes en torno a ChatGPT —y a todos los modelos de lenguaje extenso (LLM, por sus siglas en inglés: *Large Language Models*)— no es si lo que dice es verdad o mentira, si se equivoca mucho o poco, sino que ha logrado una verosimilitud total, logrando que la conversación con el chat sea indistinguible de una conversación real. El teórico Ramón López de Mántaras lo escribía hace poco: el problema con ChatGPT es su antropomorfismo, que nos hace caer en la falsa sensación de realismo. Nos lo creemos no porque diga

la verdad o se acerque, sino porque imita a la perfección la conversación humana, sus respuestas están muy bien escritas y transmite credibilidad. Pero tengamos siempre presente que la IA ni sabe por qué sabe lo que sabe ni entiende lo que dice o responde. Su respuesta busca simular el lenguaje humano y, por tanto, después de una frase como «El mejor jugador de fútbol de la historia es...» nos responderá aquello que haya aprendido de frases similares escritas en internet, añadiendo a la frase una palabra detrás de otra siguiendo criterios probabilísticos y garantizando que lo que dice tiene sentido, pareciendo incluso que se ha forjado una opinión sobre ello. Pero lo que hay detrás es, como ha descrito la profesora de lingüística computacional Emily Bender, «un loro estocástico».

En verdad, ante una herramienta como esta la confusión está garantizada y la posibilidad de utilizarla para aprovecharse de nuestras vulnerabilidades aún lo está más. La IA aumenta de forma notable lo que el mundo digital y las redes sociales iniciaron: la posverdad. Las posibilidades de manipulación, de utilización no contrastada de información, de chantaje emocional, nos obligan a una formación ciudadana que tenga criterios de discernimiento para poner bajo sospecha aquello que lee y haya interiorizado suficientemente el sentido común para distinguir si se está siendo o no manipulado mediante la información que se recibe. En definitiva, deberemos ser más escépticos y menos confiados.

La irrupción de la IA puede dar lugar a un giro epistémico definitivo en nuestra escala de valores, en la cual la verdad deja de tener importancia y lo que cuenta es la verosimilitud, es decir, la apariencia de verdad.

Su insostenibilidad material

Gran parte del debate generado en torno a la IA tiene que ver con el peligro de sus sesgos. Nos preocupa que, siendo la IA una caja negra (no sabemos por qué dice lo que dice o hace lo que hace), acabe reproduciendo aquellas discriminaciones que se dan *de facto* en la sociedad. Pero es evidente que, si la IA es racista o machista, lo será porque nuestra sociedad lo es. Ni la IA es autónoma ni podrá sustituir nuestro juicio moral; menos aún mejorarlo. La preocupación por el sesgo de los algoritmos —pero sobre todo de los datos de los que se nutre— participa de una gran ingenuidad tecno-optimista: la creencia de que la tecnología ha venido para resolver nuestros dilemas morales, para convertirnos en mejores personas. No podemos esperar más de la IA de lo que esperamos de nosotros mismos.

La segunda preocupación más extendida tiene que ver con el hecho de que la IA pueda sustituir definitivamente el trabajo humano. Un artículo reciente publicado por *Fortune* exponía lo que se conoce como la *paradoja de la productividad*: el cambio digital introducido desde mediados de los noventa, contrariamente a lo que se esperaba, no ha provocado grandes aumentos en la productividad. Es decir, disponer de tecnología digital no nos hace más eficientes ni elimina realmente lugares de trabajo; solo los transforma. Más bien, la pregunta que debemos formularnos es quién se beneficia de estas tecnologías y de qué manera permiten acumular cada vez más poder en menos manos.

Aquello que realmente tendría que ser motivo de debate, y es uno de los temas de los que menos se ha hablado, es la insostenibilidad de la IA. Cuando se

comenta la cuestión con informáticos, lo que más les preocupa es la capacidad computacional necesaria para entrenar y mantener la IA en funcionamiento: los cálculos utilizados para entrenar redes neuronales se han multiplicado por cien millones en los últimos diez años. A nosotros nos parece inocuo, casi mágico, pero pedir a ChatGPT que te explique el chiste más divertido que conoce implica una utilización de servidores, un consumo de energía y una cantidad de cálculos que hacen inviable su uso extenso y universal. Un divertimento como este se convierte en un coste computacional imposible de asumir a gran escala.

Una lectura antropológica y teológica

La lectura más interesante del fenómeno siempre cae del lado humano. Y nosotros, ¿cómo nos relacionaremos con la IA? Hace poco, con los alumnos de un seminario sobre pensamiento científico vimos un capítulo de la serie *Black Mirror* titulado «Be right back». Tiene exactamente diez años. Una mujer joven pierde a su pareja. Una empresa ha desarrollado una tecnología que permite, recuperando la información de la persona en videos, fotos y redes sociales, simular una conversación por chat con el muerto, propuesta que más adelante se convierte en una conversación telefónica y, finalmente, en un robot idéntico a la persona muerta. Como podéis imaginar, el supuesto retorno del muerto no llega a satisfacer en ningún momento a la persona viva, más bien al contrario: la aísla, la desestabiliza y la incapacita para reprender su vida. La hunde aún más en un pozo sin luz.

Esto ya es posible hoy. Los LLM como los de ChatGPT permiten, dándole información al respecto, simular una conversación por chat con cualquier persona fallecida, por ejemplo, con John Fitzgerald Kennedy o Freddie Mercury. Otras herramientas ya desarrolladas pueden tomar muestras de sus voces con cortes de tres segundos y simular a partir de ellas una conversación sonora real. ¿Y ahora qué? Si esto ya es posible, ¿estamos obligados a desarrollarlo? Imaginad la cantidad de dinero que podría conseguir un tipo de reclamo así: «¡Vuelve a hablar con tus muertos por teléfono!». ¿Cuánta gente pagaría por este servicio? Estamos a pocos meses de saberlo por macabra que nos suene la idea.

En este sentido, la mejor reflexión al respecto me la compartió un alumno: «Es que a la mujer no se le ha dado ni la oportunidad de superar el duelo». Aquí radica la cuestión más relevante: la tecnología nos permite huir y no tener que enfrentarnos a aquello que nos da miedo, nos amenaza o supone un reto vital. Pero eso parte de una mala comprensión de la antropología humana: recordando a Hölderlin, «donde hay adversidades nace aquello que nos salva». Queremos ahorrar al humano tener que ser humano. No hacemos ningún favor a nadie negándole la posibilidad del duelo en aquello que forma parte central de una vida. La muerte de un ser querido puede ser la peor de las desgracias, pero la vida va exactamente de esto. Como afirma Josep Maria Esquirol, no se trata de cerrar la herida infinita que nos constituye, sino de aprender a «acompañar y responder a su exceso». Con cada muleta tecnológica que añadimos a nuestro día a día, nos empequeñecemos y hacemos más débiles, más incapaces; en definitiva, menos humanos.

No cabe ninguna duda de que la IA participa de un cierto gnosticismo. Se trata de la vieja herejía cristiana según la cual la salvación se puede alcanzar a través del conocimiento y la iluminación. A la vez, se rechaza el mundo material e incluso el cuerpo, por imperfecto. El culto a los datos, la posibilidad de trascender nuestra vida mortal vaciándonos «en la nube» es, de hecho, la versión contemporánea de la herejía gnóstica. Y es sobre todo la ilusión de que una vez el cuerpo acaba nuestra alma se perpetúe, aunque sea a base de bits. ¿Será la IA la puerta de entrada a la inmortalidad del alma? La *tecnolatría* como religión es el gnosticismo del siglo XXI. Sea como sea no es lo mismo la resurrección de la carne que la inmortalidad del alma. La primera no es ni será nunca posible; la segunda, ya la tenemos al alcance gracias a un chat que puede hacer creíble la conversación con personas muertas. Deberíamos protegernos de ello. En tiempos de secularización extendida, donde nada está provisto de lo sagrado, la posibilidad de trascendencia se ha trasladado al campo digital. Buscamos la salvación a través de la tecnología. No la obtendremos. Lo que tendremos, más bien, será una pesadilla de confusión, una desconexión de nuestro ser natural y la incapacidad de entender la vida como seres finitos que somos.

Regular la IA y dotarnos de herramientas para el discernimiento

Nadie niega que la IA puede ser muy útil en aplicaciones de diagnóstico médico, de predicción climática o de prevención de accidentes en la conducción. Ahora bien, no a cualquier precio. Expresiones como «la IA ha venido para quedarse» o

«la IA es neutra; todo depende del uso que le des» forman parte de una mirada miope a lo que es e implica la tecnología para nuestras vidas. La fascinación que genera la tecnología a menudo viene acompañada de un discurso ingenuo, alimentado, a su vez, por el interés de tanto dinero invertido que busca rentabilidad. La buena noticia: aún estamos a tiempo de limitar las peores consecuencias, evitando cualquier negocio con las vulnerabilidades del alma humana. La mala noticia: una vez la tecnología está desarrollada transforma nuestro medio, y su simple existencia ya ha cambiado la manera que tenemos de relacionarnos con el mundo.

El antídoto personal será cultivar nuestra capacidad de discernimiento; esto es, saber lo que hacemos y por qué lo hacemos: «Adónde voy y a qué», que se preguntaba San Ignacio. Solo nuestra intencionalidad, el hecho de que nuestras acciones y palabras estén dotadas de orientación, puede hacernos hoy diferentes a las máquinas. Y, como sociedad, más nos vale que empeceemos a regular con legislaciones y protocolos éticos cada nuevo desarrollo digital, porque hay muchos millones invertidos esperando su retorno económico y a punto para aprovecharse de todas nuestras debilidades. ¿Que quieres volver a hablar por WhatsApp con tu hermana que murió el año pasado? Lo tenemos a un clic. ¿Dejaremos que esto se desarrolle como un negocio? Espero que seamos lo suficientemente inteligentes, perdón, lo suficientemente humanos como para no permitir que esto suceda.

Xavier Casanovas Combalia
Profesor de la Cátedra de Ética de IQS
(URL)