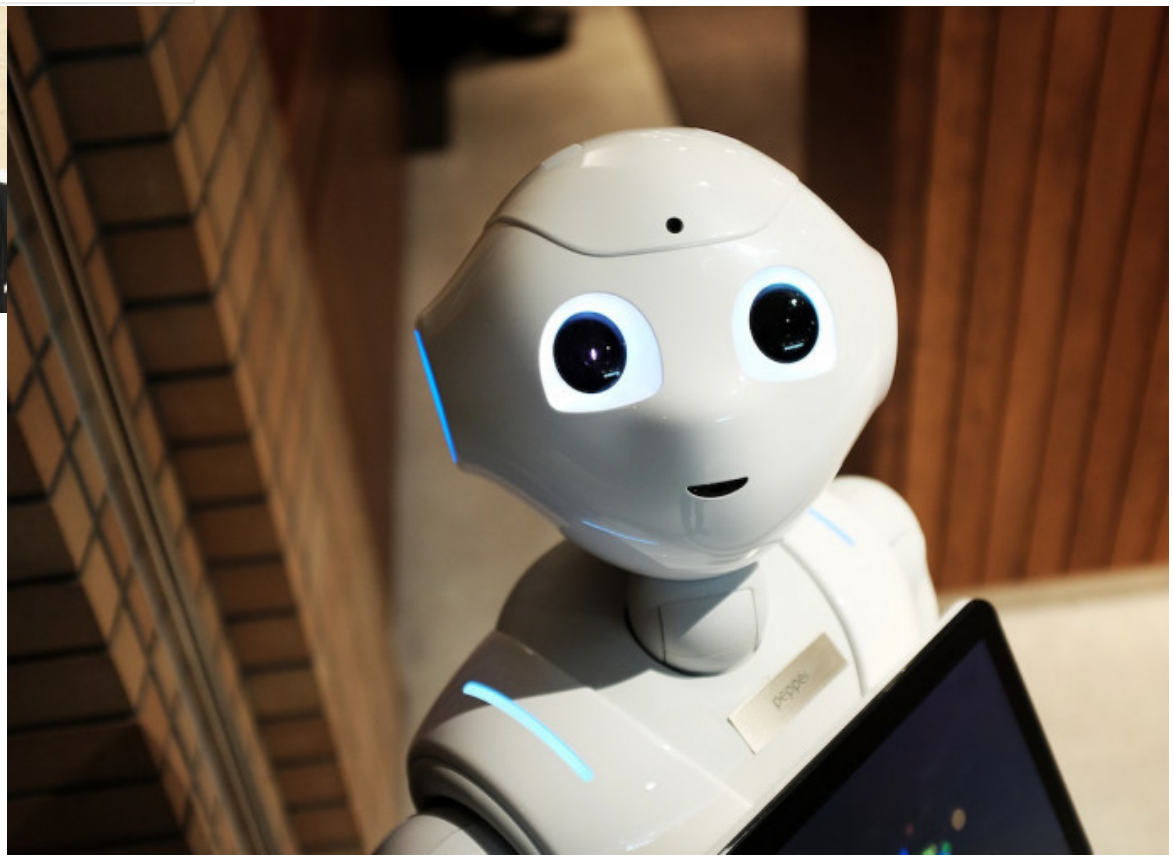
BRÚJULA COTIDIANA
Director Riccardo Cascioli
HECHOS PARA LA VERDAD

TECNOLOGÍA

Atención: la inteligencia artificial guiará nuestras intenciones

VIDA Y BIOéTICA 13_01_2025

Tommaso Scandroglio



Querido lector, ¿es usted consciente de que, antes de que usted lo decidiera, la inteligencia artificial (IA) ya sabía que hoy leería la *Brújula Cotidiana* y en concreto este artículo? Ésta es la esencia de un sugerente artículo de los investigadores de Cambridge

Yaqub Chaudhary y Jonnie Penn, titulado *Beware of the economics of intention: harvesting and commodification of intent through large language* ("Cuidado con la economía de la intención: cosecha y mercantilización de la intención a través del gran lenguaje") publicado el 30 de diciembre.

Los dos académicos sostienen que estamos pasando de la economía de la atención a la economía de la intención. En cuanto a la primera, es bien sabido que los sitios web, las redes sociales, los chats, etc. registran lo que miramos, vemos y compramos, y envían estos grandes datos a las empresas para que puedan utilizar la publicidad, los artículos sugeridos, etc. para orientar nuestras compras, basándose en su conocimiento de nuestros gustos. Ahora se está dando el siguiente paso: la IA predecirá nuestras intenciones. Ya no se limita a observar lo que observamos, sino que interactúa con nosotros para conocernos mejor y anticiparse a nuestros movimientos. ¿Y cómo interactúa la IA con nosotros? Con asistentes personales o asistentes digitales (*smart assistants*) -pensemos en el asistente de Google o en Alexa o Siri- y con chat bots, es decir, software programado para hablar con nosotros, los humanos. Ambos sistemas registran una inmensa cantidad de información sobre nosotros: elecciones, preferencias y hábitos relacionados con estilos de vida, consumo, intereses, estados emocionales, dónde estamos, con quién nos reunimos, qué leemos, etc. Los registran con gran precisión y durante largos periodos porque hablamos con ellos, interactuamos constantemente y con muchos fines. En definitiva, estos asistentes personales y chatbots nos conocen mejor que Facebook.

Y vamos al grano: todo este conocimiento sobre nosotros servirá a la IA para predecir nuestras elecciones y sugerírnoslas antes de que las hagamos: desde lo que queremos querer hasta lo que nos gustaría querer. El artículo pone este ejemplo en el que un asistente de voz interactúa con el usuario así: "Has dicho que te sientes agobiado de trabajo, ¿te reservo esa entrada para el cine de la que hablamos?". ¿Y por qué quedarnos solo con el ejemplo del cine? Los diálogos posibles, inventados por nosotros, podrían ser también los siguientes: "Has dicho que estás harto de tu mujer. ¿Has pensado alguna vez en una nueva vida sin ella? Aún eres joven"; "Estás embarazada, es tu segundo hijo, además tú y tu pareja aún tenéis que terminar de pagar la hipoteca. ¿Has pensado alguna vez en abortar? Si quieres puedo leerte algunos artículos sobre el tema".

Por supuesto, la sugerencia no vendrá tanto de la IA, sino de las empresas o los grandes medios de comunicación o los grupos de poder político que nos han vendido o regalado los asistentes digitales que tenemos en nuestros smartphones o en nuestras

casas. Así, si antes los datos sobre nosotros valían oro, ahora lo que valen son nuestras intenciones. “Estas empresas -añaden los dos investigadores- ya están vendiendo nuestra atención. El siguiente paso lógico para obtener una ventaja comercial es utilizar la tecnología, que evidentemente ya están desarrollando, para predecir nuestras intenciones y vender nuestros deseos antes incluso de que comprendamos del todo cuáles son”.

Ni que decir tiene que, como se desprende de los ejemplos anteriores, la distancia entre la “sugerencia” y la “manipulación” es muy corta. Los investigadores del *Leverhulme Centre for the Future of Intelligence* (LCFI) de Cambridge hablan de “tecnologías persuasivas”, por decirlo suavemente. La IA presente en estas tecnologías creará relaciones de confianza y entendimiento con nosotros y así nos persuadirá para que sigamos sus sugerencias. En resumen: la IA tomará decisiones por nosotros, aunque no nos demos cuenta. De la información, a la sugerencia, pasando por la conformación de nuestra conciencia y la del colectivo.

Los dos estudiosos al respecto son muy claros: “Ya se están desarrollando herramientas de este tipo para elicitar, deducir, recopilar, registrar, comprender, predecir y, en última instancia, manipular, modular y mercantilizar los planes y propósitos humanos, ya sean banales (por ejemplo, elegir un hotel) o profundos (por ejemplo, elegir un candidato político)”.

Todo esto no es futuro, sino presente. Los desarrolladores de *App Intents* de Apple para conectar aplicaciones a Siri (el asistente personal de Apple controlado por voz) han incluido protocolos en la aplicación para “predecir las acciones que alguien podría realizar en el futuro [y] sugerir la intención formulada por la aplicación”.

Las repercusiones de este proceso que pasa de lo predictivo a lo prescriptivo son infinitas. También en el campo de la bioética. En enero del año pasado se publicó en la revista científica *The American Journal of Bioethics* el siguiente artículo: *A personalised predictor of patient preferences for substituted judgments in healthcare: technically feasible and ethically desirable.* (“*Un predictor personalizado de las preferencias de los pacientes por los juicios sustitutivos en la atención sanitaria: técnicamente factible y éticamente deseable*”).

¿Qué hacer cuando un paciente ya no está en su sano juicio? Sí, existen las Declaraciones Anticipadas de Tratamiento (DAT). Pero, ¿y si faltan? Y, aunque existan, ¿qué pasa si son oscuras, ambiguas, deficientes? Sí, existe la figura del administrador. Pero si él también faltara, o incluso si estuviera, ¿quién puede decir que sería fiable a la hora de describir los deseos del paciente? Lo mismo ocurre si pensamos en los

familiares. Aquí entra en acción la IA que, en este caso, se llama *Predictor Personalizado de Preferencias del Paciente:* el modelo 4P.

Los autores del artículo que acabamos de citar proponen "utilizar el aprendizaje automático para extraer los valores o preferencias de los pacientes a partir de datos obtenidos a nivel individual y producidos principalmente por ellos mismos, en los que es probable que estén codificadas sus preferencias (aunque sólo sea de forma implícita)". Para simplificar y ejemplificar: tienes un accidente y acabas en coma. Los médicos preguntan a Alexa qué elección habrías hecho en esa coyuntura. En un primer momento, Alexa reúne todas tus lecturas y vídeos sobre el tema de la eutanasia que *te hayan gustado*, así como las conversaciones que hayas mantenido con ella o con otras personas sobre este tema. En segundo lugar, compara este paquete de datos con tu temperamento un tanto humilde y tu actitud no siempre alegre ante la vida, interpretada por tus películas, tus lecturas, tus intereses cultivados, tus correos electrónicos y mensajes, tus fotos de puestas de sol publicadas en Instagram, tus compras de ropa gótico-espeluznante en Amazon, tus frases infelices a lo Leopard que lanzas al cielo en un momento de abatimiento temporal. Y así, finalmente, en una milmillonésima de segundo te encuentras en un ataúd porque Alexa así lo ha decidido. O mejor dicho: la persona que programó a Alexa. Y poco importa si en esa coyuntura también podrías haber decidido algo distinto a tus decisiones anteriores, ya que "las situaciones hipotéticas no reflejan necesariamente lo que la gente elige en situaciones reales".

Leyendo ambos artículos, entendemos entonces que se está produciendo una involución antropológica: lo virtual inicialmente nos informaba, luego nos ayudaba, y en un futuro próximo nos sustituirá. De la información, a la ayuda, a la sustitución. De hecho, los investigadores que propusieron el modelo de las 4P afirman que la IA se convertiría en "una especie de 'gemelo psicológico digital' de la persona". Nuestra libertad, hoy ya muy plagiada en muchos aspectos, pasaría a manos de los que manejan la IA y ésta elegiría por nosotros si ir al cine, con quién casarnos y si desenchufarnos. Daríamos plena delegación a la IA porque en la percepción colectiva ésta es superinteligente, neutral en sus juicios y objetiva porque está libre de condicionamientos emocionales e intereses propios. El resultado sería fatal: ya no estaríamos viviendo nosotros, sino que lo haría nuestro yo virtual.