

INTELIGENCIA ARTIFICIAL

Una aproximación a su entendimiento

ARTURO PRETEL PRETEL

Médico

El otro día recibí un *meme*, como se dice ahora, con un dibujo en el que Mafalda, la genial niña filósofa creación del argentino Quino, reflexiona con una amiga diciéndole: «¿te preocupas por el avance de la inteligencia artificial?». La amiga, después de ardua reflexión, contesta toda tranquila: «no, me preocupa más el retroceso de la inteligencia natural».

Pues eso: ocupémonos de la Inteligencia Artificial (IA), pero no perdamos la perspectiva de lo que supone realmente, no perdamos nuestro interés por el sentido común y el cultivo de nuestros conocimientos y habilidades y hagamos de los avances un instrumento de progreso para la Humanidad. Nuestros esfuerzos deben ir encaminados a procurar una mejora de la calidad de vida de la gente, siempre sobre unas bases y en un contexto ético y moral. Tratemos a la IA como a cualquier avance, aunque no sea uno de tantos, y más si contribuye al desarrollo de materias que mejoren a las personas y a su calidad de vida.

Los adelantos más significativos que se han producido a través de los tiempos han generado en la sociedad de cada época recelos sobre la repercusión que podrían tener sobre el entorno, las formas de vida, el empleo, la salud y muchas otras variables. Sin ánimo de ser exhaustivo, cabe recordar la invención de la imprenta, el ferrocarril, el barco de vapor, la electricidad, el telégrafo, la aviación, la radio, la televisión y un largo etcétera. Todos ellos causaron polémica y recelos en su momento y han tenido unas consecuencias no siempre negativas y casi siempre positivas.

Casos de mala utilización de avances importantes para la Humanidad los podemos encontrar también en muy distintas facetas de la vida. Por ejemplo un fármaco, el Fentanilo, un derivado sintético de la morfina con una enorme utilidad médica que alivia el dolor crónico a millones de seres humanos y que es muy apreciado tanto por anestesistas como por especialistas en cuidados paliativos. Y sin embargo, estamos asistiendo a una auténtica pandemia de marginalidad, muertes y delincuencia alrededor del abuso de esta sustancia fuera de la esfera médica. No se puede demonizar al Fentanilo: quienes lo han hecho malo son las mafias que trafican con él, es el hombre que abusa y hace un uso inadecuado de la sustancia.

Otro claro ejemplo de uso criminal de un específico de gran utilidad clínica es el Misoprostol: de ser un potente antiulceroso y protector gástrico a usarse,

espuriamente, para provocar abortos.

Qué decir de la energía nuclear, del mundo de las radiaciones ionizantes, como paradigma de los distintos usos que se pueden dar a herramientas de progreso puestas en nuestras manos: desde baratos productores de energía a la destrucción masiva en su uso militar; de terribles efectos letales cuando está sin control a patrón en el tratamiento de tumores y en el diagnóstico médico, sin cuya utilización el avance de las ciencias biomédicas que tenemos en la actualidad nunca se hubiese producido; del apocalipsis nuclear a la mejora de las condiciones de vida en nuestro planeta.

Y no digamos del reto actual que supone la clonación de embriones, las técnicas de edición genética, las terapias CART y CRISP. No pararíamos de enumerar avances que, según se utilicen, muestran una cara u otra de la realidad de la vida.

Resulta francamente difícil definir con precisión qué es la inteligencia artificial, qué tipos de inteligencias artificiales existen, en qué punto de su desarrollo estamos y hacia dónde evolucionará en un futuro. Podríamos considerar a la IA como una serie de técnicas en el mundo de la cibernética que combinan información, datos y algoritmos que, mediante las llamadas redes neuronales artificiales, crean programas, sistemas o máquinas que intentan simular la forma de aprender, hacer y pensar de un humano para realizar las tareas que se le hayan asignado. El rendimiento, eficacia y veracidad de uso dependerán en último extremo de los datos de que se haya dotado al sistema (minería de datos) y de la programación de éste para evitar fallos, manipulaciones o sesgos en su utilización práctica (discriminación, perpetuación de prejuicios, señalamiento), sobre todo en lo referido a su uso en aplicaciones que puedan interactuar con la dignidad humana y el derecho natural.

Existen varios tipos y clasificaciones de la Inteligencia Artificial, pero los que más interés práctico nos puedan suscitar están referidos a la IA generativa, la que, en esencia, facilita la producción, mediante programas específicos, de textos, contenido gráfico (producción, alteración y/o manipulación de imágenes), sonoro (imitación de voces, creación de música) y audiovisual (*deepfake*) con una extraordinaria calidad, claridad y «veracidad». Estas técnicas de IA generativa tienen unas implicaciones relevantes en el mundo intelectual, el del entretenimiento, el del periodismo, etc. Lógicamente plantean grandes retos desde el punto de vista ético, legal y social. Los desarrollos que se generan están siendo de una enorme utilidad para la sociedad, pero, en contrapartida, estas aplicaciones pueden mentir muy bien y ser fuente de manipulación social.

Dentro de las clasificaciones, una muy utilizada para intentar abarcar el presente y futuro de esta técnica es la que la cataloga en cuatro grandes grupos:

1.- *IA reactiva*: estos programas proporcionan respuestas predecibles a entradas específicas, sin capacidad para aprender o contemplar acciones pasadas o futuras. Es la primera y más básica en su desarrollo. Operan dentro de las limitaciones de su diseño inicial y no pueden funcionar más allá de las tareas para las que fueron programadas. Por ejemplo, una máquina de jugar al ajedrez.

2.- *IA de memoria limitada*: este tipo puede aprender de experiencias pasadas y utilizar una combinación de datos observacionales e información preestablecida para realizar tareas. Esta forma de IA es la más prevalente en las aplicaciones contemporáneas. Aquí entraría la IA generativa, la de ayuda al diagnóstico médico, la que auxilia los peritajes judiciales etc.

3.- *IA de teoría de la mente*: es la próxima frontera en el desarrollo de estas técnicas y su objetivo sería crear máquinas capaces de tomar decisiones verdaderas y entender las emociones humanas. Se pretende que tengan la capacidad de responder a los estímulos que podrían generar emociones e interacciones con las personas. Tendrían la capacidad de ajustar su comportamiento en base a las señales emocionales, asemejándose mucho a las interacciones humanas.

4.- *IA autoconsciente*: Sería el último eslabón en el desarrollo de la IA. Es un concepto futurista donde los programas y máquinas tendrían un nivel de conciencia e inteligencia similar al de los humanos. Poseerían la capacidad de deducir y reaccionar a sus propios estados mentales y emociones. Este cuarto nivel entra dentro de la distopia y se engarza en las redes del transhumanismo, sin atenerse a consideraciones éticas. No parece previsible que sea posible con los conocimientos actuales, pero seguro que existirán fuerzas que nos quieran llevar a ello.

Existe la idea equivocada de que las máquinas aprenden mejor que nosotros y esto no se corresponde con la realidad. La mente humana no es parecida a las inteligencias artificiales, aprende de una manera muy distinta. Nuestro cerebro está en un proceso constante de aprendizaje a través de los sentidos, de los estímulos que nos rodean, de la más mínima experiencia vivida, de una luz, de un olor, de una sensación. El cerebro humano tiene una enorme plasticidad, tiene «cintura», es orgánico, es rebelde e interactúa constantemente con el entorno con sus propias habilidades y herramientas y aprende por sí solo. Y sobre todo, junto a los demás, aprende de los otros y de lo que nos rodea, tiene una enorme capacidad de adaptación y, mediante la inteligencia emocional, en unas personas se desarrolla más que en otras, es resiliente. Nuestro cerebro aprende de forma integral y está integrado en nuestras necesidades y nuestra personalidad, se adapta a nuestras peculiaridades y personalidad, no está parcelado. Estas características no acompañan a la manera de aprender de las máquinas, ni mucho menos. Estas nunca tendrán la capacidad de juicio y de adaptación de la mente humana por mucho que se avance, y se avanza mucho, en su programación.

La inteligencia artificial será buena o mala según el cristal con que se mire y según el uso o abuso que se haga de ella. Podemos situarnos ante la botella medio llena o medio vacía, no es intrínsecamente buena ni intrínsecamente mala; depende de su uso y de cómo se use. Lo que no parece justificado es meternos en una espiral de negativismo por sus posibles efectos perniciosos, lo que nos llevaría a la melancolía, a la depresión y la desesperanza; tampoco, caer en el pueril infantilismo futurista de que todo va a ser maravilloso gracias a la Inteligencia Artificial. Lo que hay que plantearse es para qué se desarrolla, para qué se quiere usar y quién la usa. Y todo ello, con una perspectiva moral y deontológica que lo sustente y con una sociedad armada de

convicciones y de controles que impidan una deriva repulsiva en su uso. Pero tampoco nos engañemos: la IA no es una innovación más, como han sido dentro del mundo de la cibernética, internet, el blockchain, el metaverso o la realidad virtual. Es algo que va a cambiar no sólo la forma de trabajar en disciplinas heterogéneas (medicina, ingeniería, ciencias sociales etc.), sino que está influyendo en otras áreas hasta ahora no accesibles a este tipo de avances y que van a influir en la forma de ver el mundo y en el desarrollo de las relaciones humanas.

No debemos de perder la perspectiva de que la IA no es más que una herramienta que tenemos a nuestra disposición para ayudarnos a ser más completos, nunca puede sustituir el criterio o la destreza humana, tiene que ser un asistente que nos ayude en la toma de decisiones y en tareas asociadas para hacernos más eficaces y facilitarnos la vida. Su capacidad de decisión y de creación debe estar a nuestro servicio, nunca sustituir nuestro criterio. Ha venido para quedarse y este cambio de paradigma hay que modularlo de tal manera que sepamos aprovecharlo con las precauciones y prevenciones necesarias.

Uno de los mundos en los que la IA puede tener una rentabilidad mayor es el de las Ciencias Biológicas y en concreto la Medicina. Con sus luces y sus sombras, la IA está suponiendo una transformación impresionante en el área sanitaria. En los próximos cinco años esta transformación se va a hacer más evidente y va a llegar en mayor medida al paciente. La IA generativa, las redes neuronales –en definitiva, el uso de algoritmos muy precisos, no tan sólo en el diagnóstico clínico–, va a permitir al profesional médico prestar una atención más eficiente en la pura labor asistencial y en el auxilio para un diagnóstico mucho más preciso. Los modelos generativos de IA están dejando su huella también en los tiempos de desarrollo de nuevos fármacos y la selectividad de éstos ante la patología que deben combatir. Los ensayos clínicos para la validación de nuevas terapias se están acortando en el tiempo y las muestras de pacientes que se precisan para habilitarlas van a ser mejor seleccionadas y con series más cortas, lo que hará que se introduzcan antes y sean más seguras y precisas.

La formación médica también se beneficiará en buena medida con el desarrollo de algoritmos inteligentes que facilitarán el aprendizaje de los futuros médicos con la implementación de modelos virtuales para el ensayo de nuevos procedimientos y técnicas de los ya expertos. Y no hablemos del enorme apoyo que se está produciendo y que se producirá en el futuro en el diagnóstico por imagen de patologías dudosas, de despistajes masivos de enfermedades, de ayuda a la predicción de un pronóstico evolutivo, etc. Por poner un ejemplo, con el soporte de programas de IA, imágenes dudosas o extraordinariamente tempranas en mamografías que se practican para el diagnóstico precoz del cáncer de mama hacen que el radiólogo sea capaz de diagnosticarlas en estadios incipientes, con lo que esto supone en la prontitud de su abordaje y el aumento de la esperanza de vida de la paciente afectada. Ejemplos como los anteriores están implantándose día a día con rotundo éxito. En cualquier caso, en esta área como en tantas otras, estos programas nunca podrán suplantar la labor del profesional, que no solamente evalúa los datos y las predicciones, sino que teniendo en

cuenta la mirada, la cara y la actitud del enfermo, será capaz de dimensionar, con su empatía y calor, la globalidad del padecimiento del enfermo.

No se puede pretender que la IA tenga una ética y que responda a consideraciones morales. La IA son solamente una serie de aplicaciones y técnicas creadas por el hombre y gobernadas por éste y debe de seguir siendo así. El que tiene que tener consideraciones éticas es el diseñador de estos instrumentos, el que les da utilidad y, fundamentalmente, el que los usa. Debemos exigir un comportamiento ético en el uso de la IA, lo contrario sería darle una carta de naturaleza que no tiene ni le corresponde. Ya les gustaría a los transhumanistas una línea de pensamiento en este sentido.

No necesariamente –y este es un error de planteamiento– se ha de concebir a la IA como una copia de la inteligencia humana. La humana es una inteligencia global, genérica, amplia, de aplicación a los diversos aspectos de la actividad humana que se dan en nuestra vida. La IA es y debe seguir siendo una inteligencia especializada y de uso específico, un programa que haya sido alimentado con todos los datos que se requieran para hacer un trabajo concreto y con los condicionantes precisos para evitar los sesgos que lo puedan hacer no ético y no moral. Y siempre bajo la supervisión del usuario que lo utiliza y para los fines, se suponen legítimos, legales y controlados, que se buscan. La última decisión, si el programa en cuestión lleva a una toma de postura, tendrá que estar controlada, supervisada y conocida por una persona.

Lo demás sería pretender desviarse de su motivo fundamental: el ser un instrumento que facilite la vida de las personas y caer en la dualidad hombre-máquina, hombre-tecnología, en el abuso e interacción de estas dos realidades antropológicas distintas a la que nos quieren abocar las tendencias transhumanistas, que con su apariencia de progreso y bienestar en el futuro de la Humanidad encierran tanta mercancía averiada y nos alejan de la realidad del ser humano como portador de valores y constituidos por un cuerpo y un alma trascendente.

La inteligencia artificial no tiene que ser ética. Es una herramienta y, como cualquier herramienta, se puede usar para bien y para mal. No existe una IA buena o mala, solo buenos y malos humanos que la diseñan y la usan para fines buenos o vergonzantes. La IA no es lo suficientemente inteligente para hacer que ella misma sea ética, tenga moral. No le corresponde ese papel. Tenemos que tener claro el presente y el futuro de su uso para mejorar la calidad de vida de los seres humanos. Y nada más.

Los trabajadores de una de las grandes empresas que se dedican a este campo mostraron su preocupación hace unos meses, en pleno proceso de renovación accionarial de la compañía y el cese que luego se revocó de su creador y consejero delegado, que la IA ponía en peligro el futuro de la Humanidad. Esta predicción apocalíptica no se nos puede olvidar, pero no parece posible, sobre todo, si se hace una regulación precisa de su uso legítimo. La IA está más controlada de lo que parece. Otra cosa es que algunos de sus desarrollos, incluso los legales, no nos parezcan lo suficientemente controlados y éticos, pero esto ocurre en tantos aspectos del mundo actual que tendemos a echarnos en brazos del catastrofismo. Sin embargo, los

beneficios superarán con creces a los posibles usos fraudulentos.

Y hay que volver a insistir: dado que las máquinas «inteligentes» no pueden tomar decisiones basándose en consideraciones éticas, la responsabilidad de establecer principios éticos, normas morales y legales y definir las condiciones de uso, los objetivos y los límites de las redes neuronales, de los algoritmos y del uso de la Inteligencia Artificial reside en las personas.

Frente a nuestra arrogancia, hemos de reconocer que la capacidad de una máquina de derrotar al hombre jugando al ajedrez o eligiendo el camino óptimo para llegar a casa, dice más de las limitaciones de la inteligencia humana que de la capacidad de las máquinas. Un poco de realismo nos debería llevar a reconocer que el ser humano no es asombroso por lo que hace, sino por lo que es. En definitiva, no podemos anticipar el futuro, pero sí podemos trabajar por escenarios mejores, con la certeza de que el ser humano se comportará con madurez colectiva cuando disponga de la información necesaria. Y a eso ha de contribuir, como una herramienta más, y parece que bastante decisiva, el desarrollo de programas y técnicas que beban de la Inteligencia Artificial.

El objetivo fundamental de cualquier regulación de la Inteligencia Artificial y de todo lo relacionado con ella es la protección de la dignidad y derechos de las personas, pues su uso fraudulento podría ponerlos claramente en peligro, iguale que puede ocurrir con otras tecnologías. Es necesario delimitar con la mayor claridad posible, tarea no siempre fácil con las nuevas tecnologías en constante evolución, hasta dónde se puede llegar en el tratamiento de los datos en lo relacionado con la intimidad, el derecho a la propia imagen, a la libertad de movimiento, a la toma de decisiones sobre la salud, sobre aspectos laborales, judiciales, etc., que impidan la discriminación por su uso y los resultados obtenidos en tantos aspectos que, aun hoy, no llegamos a imaginar.

Se está abordando, con acierto en mi opinión, la regulación por parte de la Comisión Europea del marco normativo de la IA, simultáneamente con otras instancias supranacionales y estatales. Sería muy prolijo abordar el tema de estas legislaciones en el momento actual. Baste decir que los límites establecidos y que se tendrán que desarrollar en el futuro comprenden la prohibición de la manipulación subliminal (influir en niños o discapacitados, por ejemplo), el establecimiento y limitación de grupos sociales, el control biométrico en espacios públicos, el uso en temas fronterizos, en cualquier proceso de puntuación social y, en general, en todo lo que atente a derechos fundamentales. Un aspecto significativo de esta regulación es la prohibición del control de reacciones humanas con fines de selección laboral, control del desempeño y la atención en el trabajo y del trabajo en línea más específicamente, así como el control de las reacciones y atención en los exámenes que se desarrollen a distancia.

Hay que impedir a toda costa, por referirnos a los aspectos negativos, que esta tecnología y sus derivados pueda dominar al ser humano y que se convierta en una herramienta de control de unos sobre otros, poniendo un especial cuidado en su limitación en los ámbitos gubernamentales y de las grandes corporaciones empresariales, impidiendo el control ideológico y social y todo aquello que atente a la

vida y dignidad de las personas. Y hay que volver a repetir que el problema no es esta tecnología, sino el uso que se le dé para contravenir la ley natural, los derechos de las personas y los colectivos y el bien común.

Declaro y confieso que este artículo ha sido escrito a la antigua usanza, sin usar ninguna de las facilidades que me podría haber aportado la Inteligencia Artificial. No por ello es mejor ni más legítimo, es simplemente una forma de hacerlo. Pero sí que me he aprovechado del uso de tecnologías que, apenas hace 30 o 35 años no habrían estado a mi servicio: tratamiento de texto, consultas a la bibliografía a través de Internet, correo electrónico... Y por ello ni es más ni menos humano, está tan sujeto a errores y a aciertos como si hubiera sido escrito a mano o a máquina de escribir mecánica, o como si hubiera acudido a la sala de lectura de una biblioteca varias mañanas a consultar libros en papel y lo hubiera mandado a la dirección de la revista por correo postal. En próximos artículos, si los hay, me plantearé usar instrumentos de IA para un mejor resultado del objetivo y sin menosprecio de su sinceridad y autenticidad. Es el signo de los tiempos, es la ayuda que nos brinda el progreso y una prueba de que la inteligencia humana puede dar un uso eficaz y ético a la Inteligencia Artificial.