

Desafíos éticos de la Inteligencia Artificial y de la robótica

José Sols Lucia*

Resumen

Cuando pensamos en inteligencia artificial y en robótica, espontáneamente nos viene al espíritu lo positivo de estas nuevas tecnologías, esto es, las múltiples aplicaciones que ya tienen y que tendrán en el futuro, así como la enorme cantidad de operaciones mentales y físicas que nos ahorraremos gracias a ellas. Sin embargo, hay también otro lado, negativo, inquietante, en el que constatamos que la inteligencia artificial y la robótica cambian la vida humana a peor. En este estudio nos fijaremos de manera especial en algunos desafíos éticos que plantean estas nuevas tecnologías, algo que haremos tras haber explicado la fascinación por ellas, algunas de sus aplicaciones, la ética de la responsabilidad, así como declaraciones y leyes internacionales que llaman a la prudencia y que conviene no desoír, sobre todo teniendo en cuenta que el posible mal causado por ellas, ya sea por mala voluntad de las personas, ya sea por un error técnico, puede tener dimensiones devastadoras tanto para nuestra generación como para las venideras. No podemos reflexionar acerca de la inteligencia artificial o de la robótica de manera binaria, dualista, dicotómica: “o es buena o es mala”. Se trata de una realidad muy compleja, creada por el ser humano, a la que debemos aproximarnos de manera matizada. Las ideas clave son *prudencia* y *discernimiento*.

Palabras clave: Ética de la inteligencia artificial y de la robótica. Ética de la responsabilidad. Prudencia. Discernimiento.

* El Dr. José Sols Lucia es profesor de Ética y de Teología de la Universidad Iberoamericana Ciudad de México, coordinador del Grupo de Inteligencia Artificial del CELAM, del Grupo



Ethical challenges of artificial intelligence and robotics

Summary

When we think of artificial intelligence and robotics, the positive aspects of these new technologies spontaneously come to mind, specifically the multiple applications they already have and will have in the future, as well as the enormous number of mental and physical operations that we will save thanks to them. However, there is also another side, adverse and disturbing, in which we see that artificial intelligence and robotics change human life for the worse. In this study we will pay special attention to some of the ethical challenges posed by these new technologies, after having explained the fascination with them, some of their applications, the ethics of responsibility, as well as international declarations and laws that call for prudence and that should not be ignored, especially considering that the possible harm triggered by them, Whether it is due to people's ill will, or due to a technical error, it can have devastating proportions both for our generation and for those to come. We cannot reflect on artificial intelligence or robotics in a binary, dualistic, dichotomous way: "either good or bad". It is a very complex reality, created by human beings, and which we must approach in a nuanced way. The key ideas are *prudence* and *discernment*.

Keywords: Ethics of artificial intelligence and robotics. Ethics of responsibility. Prudence. Discernment.



CELAM
CONSEJO EPISCOPAL
LATINOAMERICANO

Cuando pensamos en inteligencia artificial, espontáneamente nos viene al espíritu lo positivo de esta nueva tecnología, esto es, las múltiples aplicaciones que ya tiene y que tendrá en el futuro, así como la enorme cantidad de operaciones mentales que nos ahorraremos gracias a ella. Ese lado es indiscutible. Sin embargo, hay también otro lado, negativo, por lo menos inquietante, en el que constatamos que la inteligencia cambia la vida humana a peor. Suele ocurrir que son los ingenieros y científicos los que se entusiasman con lo positivo y lo exponen con todo tipo de detalles (por ejemplo, los *big data*, o macrodatos), mientras que son los teólogos, filósofos y humanistas los que se fijan en lo negativo (por ejemplo, nos dicen que sus alumnos ya no redactan sus trabajos académicos en la universidad, sino que los hace el Chat GPT; incluso que hay profesores que ya usan este programa para corregir trabajos, con lo que ni los alumnos escriben ni los profesores leen, sino que donde antes había una actividad formativa, ahora hay una farsa).

En este estudio nos fijaremos de manera especial en algunos desafíos éticos que plantean la inteligencia artificial y la robótica, algo que haremos tras explicar la fascinación por ellas, algunas de sus aplicaciones, la ética de la responsabilidad, así como declaraciones y leyes que llaman a la prudencia y que conviene no desoír.

Hay autores que prefieren distinguir entre inteligencia artificial y robótica; otros, en cambio, optan por percibir su unidad. Aquí nos ubicaremos en este segundo grupo. Entendemos por *inteligencia artificial* cualquier sistema informático que realice tareas que antes solo lograba llevar a cabo un cerebro humano, de



tal modo que pueda procesar infinidad de datos a gran velocidad y con mucha precisión, algo que no logra hacer la mente humana, todo ello de manera autónoma, sin supervisión humana; y por *robótica*, la ciencia encaminada a diseñar máquinas programadas para realizar tareas de forma automática o para simular el comportamiento humano o animal, tanto en lo mental (inteligencia artificial incorporada al robot) como en lo físico¹. Sintetizando, podríamos afirmar que la inteligencia artificial complementa (o sustituye) el *cerebro* humano (por ejemplo, el cálculo matemático o el catálogo de una biblioteca), mientras que la robótica hace lo propio con el *cuerpo* humano, tanto en lo físico como en lo mental (por ejemplo, la fabricación de coches, una cirugía en el quirófano o cortar el césped del jardín).

1. FASCINACIÓN POR LA INTELIGENCIA ARTIFICIAL

Los antropólogos denominan *fascinación por la tecnología* (“*technological fascination*”)² al fenómeno por el cual nos volvemos acríticos ante una tecnología que no solo es nueva, sino también radicalmente innovadora, de tal modo que solo vemos en ella lo positivo y no tomamos conciencia de sus posibles peligros. A este fenómeno espontáneo contribuyen —de manera ya no tan espontánea, sino abiertamente premeditada— el mercado y el Estado. El mercado desea que consumamos lo que se nos ofrece en él en grandes cantidades y sin espíritu crítico, de tal modo que reduce los ciudadanos a consumidores, mientras que el Estado, obsesionado por el control, reduce los ciudadanos a súbditos. El Estado nos transmite además la sensación de peligro, de amenaza, de inseguridad, y nos asegura que él, y solo él, nos protegerá de esa jungla salvaje que se nos describe en los informativos. Para ello, nos pide que confiemos en él y que le demos información

¹ Sols Lucia, J. y de los Ríos Uriarte, M. E. *Bioética de la inteligencia artificial*. Madrid: San Pablo - Universidad Pontificia Comillas, 2024, pp. 11-12.

² Lambert, D., “The humanization of robots: lethal autonomous weapons systems and ethics”, en *The Humanization of Robots and the Robotization of the Human Person. Ethical Reflections on Lethal Autonomous Weapons Systems and Augmented Soldiers*, Ginebra, The Caritas in Veritate Foundation Working Papers, 2017, editado por A. de La Rochefoucauld y S. Saldi, p. 24.

acerca de nosotros; esa información, procesada con inteligencia artificial, es nuestra jaula. Son incontables las veces que a diario somos grabados por cámaras de seguridad en una ciudad: antes se contaban por decenas de veces, ahora ya por centenares. Las tarjetas bancarias, nuestros accesos a Internet y un sinnúmero de operaciones proporcionan mucha información acerca de nosotros, tal como exponen muy bien algunos de los que antes trabajaron para este sistema en el reportaje *The Social Dilemma (El dilema de las redes sociales)* (2020), de Jeff Orlowski: “Cuando la red te ofrece algo gratis, el producto eres tú”.

Ya sea por la dictadura del mercado que nos impele a consumir de manera acrítica, ya sea por la omnipresencia del Estado que nos invita a arrojarnos en sus *brazos* para luego aplastarnos con el *abrazo* del oso, ya sea por el carácter sorprendentemente innovador de la tecnología, la verdad es que tenemos ante esta una actitud de fascinación casi religiosa, tal como muy bien han descrito los pensadores José Ortega y Gasset³, Martin Heidegger⁴, Jacques Ellul⁵, Jürgen Habermas⁶, Hans Jonas⁷, Gilbert Hottois⁸ o Albert Florensa⁹, entre otros, sin olvidar a los dos últimos papas, Benedicto XVI (*Caritas in veritate*, 2009) y Francisco (*Laudato si'*, 2015, y *Fratelli tutti*, 2020). En el paradigma tecnocrático creemos ingenuamente que la tecnología nos salvará de todos nuestros problemas, sin ser conscientes de que el origen de estos es precisamente la tecnología.

³ Ortega y Gasset, J. *Meditación de la Técnica y otros ensayos sobre ciencia y filosofía*. Madrid: Revista de Occidente en Alianza Editorial, 1998.

⁴ Heidegger, M. *La pregunta por la técnica*. Barcelona: Herder: 2021.

⁵ Ellul, J. *La Technique ou l'Enjeu du siècle*, París: Armand Colin, 1954 (en español: *La edad de la técnica*. Barcelona: Octaedro, 1977); íd. *Le système technicien*. París: Calmann-Lévy, 1977; íd. *Le Bluff technologique*. París: Hachette, 1988 (en español: *Anarquía y cristianismo*, México D.F.: Jus, 2005).

⁶ Habermas, J. *Ciencia y técnica como ideología*. Madrid: Tecnos, 1984.

⁷ Jonas, H. *El principio de responsabilidad*. Barcelona: Herder, 1995.

⁸ Hottois, G. *El paradigma bioético: una ética para la tecnología*. Barcelona: Anthropos, 1991.

⁹ Florensa, A. *La vida humana en el medi tècnic. El pensament de Jacques Ellul*, Barcelona: Claret, 2010; Florensa, A. y Sols, J. (eds.) *Ética de la investigación científica*. Bilbao: Desclée de Brouwer, 2017.



2. ENORME VARIEDAD DE APLICACIONES

No cabe duda de que la inteligencia artificial está constituyendo una revolución en la vida humana a muchos niveles, dado que sus aplicaciones son múltiples y profundas. Algunas de ellas ya forman parte de nuestra existencia cotidiana desde hace años, como las computadoras con las que escribimos a diario o los teléfonos móviles con cámara de video, de fotos, con conexión a Internet y con capacidad para almacenar todo tipos de archivos. Esto no ha hecho más que empezar. Previsiblemente, en el futuro los informativos serán redactados y retransmitidos por personas virtuales, seres inexistentes que podrán “trabajar” 24 horas al día todos los días del año, hablar en el idioma o dialecto que el consumidor prefiera; serán siempre jóvenes y apuestos; nunca envejecerán ni pedirán una baja por maternidad o un aumento de sueldo. No lo harán porque no son personas.

Ya hemos dicho que son programas informáticos los que redactan trabajos académicos en pocos segundos, con lo que los alumnos dejan de aprender, y los que los corrigen, en lugar de hacerlo sus profesores. En el futuro, los cursos serán impartidos por profesores virtuales, tan poco humanos y tan eficaces como los mencionados comunicadores de noticias, y darán clases a miles, millones, de alumnos, en cualquier idioma y a cualquier hora del día, sin nunca presentar una queja.

Los cirujanos dejarán de operar en el quirófano y se limitarán a supervisar lo que hace el robot en su lugar con una precisión miles de veces mayor que la suya (nanotecnologías) y aquí también sin problema de horario, ni de calendario, ni de cansancio, ni de sueldo.

Las sociedades desarrolladas tienen cada vez menos niños (en Barcelona ya hay más perros que menores de catorce años) y más ancianos, lo que comporta un serio problema (en Japón se han disparado todas las alarmas), el de quién cuida de ellos, algo que resolverán los robots, tal como exponemos con detalle en nuestra obra *Bioética de la inteligencia artificial*¹⁰, donde mostramos cómo

¹⁰ Sols y de los Ríos, *Bioética*, 67-72.

los robots ya existentes hoy, todavía muy caros, Spot, CyberDog 2, Zora o Da Vinci, entre otros, cuidan de los ancianos, les dan conversación, les suministran la medicación y avisan al equipo médico en caso de detectar una caída o algo anómalo. Se supone que los ancianos están acompañados por alguien, el robot, cuando en realidad están solos. “Es el progreso”¹¹.

Hay granjas en Australia en las que no trabaja ni una sola persona; el granjero vive en Estados Unidos. Hay fábricas en España en las que no trabaja ni un solo obrero; el ingeniero y el empresario viven en México. Hay series de televisión cuyo guión no ha sido redactado por ninguna persona humana. Hay máquinas expendedoras de medicamentos que sustituyen a las farmacias; uno expone su dolor y la máquina le receta el tratamiento y le suministra la medicación. Hay sinfonías que no han sido compuestas por ningún músico. En definitiva, la industria, los transportes, la comunicación, la medicina, la educación o el arte no son sino algunas de las muchas áreas en las que la inteligencia artificial y la robótica desplazarán a los seres humanos.

3. ÉTICA DE LA RESPONSABILIDAD

Los autores ya mencionados más arriba (José Ortega y Gasset, Martin Heidegger, Jacques Ellul, Jürgen Habermas, Hans Jonas, Gilbert Hottois, Albert Florensa) así como los papas Benedicto XVI y Francisco abogan por una *ética de la responsabilidad*, en la que dejemos de guiarnos por el principio falaz según el cual “hay que hacer todo lo que sea técnicamente posible”. Tenemos una insana tendencia a confundir los adjetivos *nuevo* y *bueno*. Es cierto que lo nuevo no tiene por qué ser necesariamente malo (ideología conservadora de espíritu cerrado), pero es igualmente cierto que lo nuevo tampoco tiene por qué ser necesariamente bueno (ideología ingenuamente pseudoprogresista). Esta última lección nos la enseñaron los nazis de la peor de las maneras, ellos que se autopresentaron como lo novedoso, lo mejor para Alemania y para la humanidad.

¹¹ Cfr. El artículo de María Elizabeth de los Ríos en este número de *Medellín*.



Frente a la cerrazón conservadora y la ingenuidad pseudo-progresista, lo que debemos desarrollar es una ética de la responsabilidad, en la cual hay que alimentar una permanente actitud de *discernimiento*. En esta, las preguntas son varias: ¿Qué modelo de hombre perseguimos? ¿Qué modelo de sociedad pretendemos alcanzar? ¿Qué lugar queremos dejar a la tecnología? ¿Qué espacio humano estamos dispuestos a ceder ante ella? ¿Qué significa *mejorar* cuando decimos que una nueva técnica *mejorará* la vida humana? En definitiva, ¿qué es lo humano? Si logramos contestar esta última pregunta, tendremos relativamente fácil responder las anteriores, pero si no tenemos claro qué es lo humano, entonces aquellas preguntas serán de mal responder. Ahora bien, debemos tener en cuenta que la definición de *lo humano* varía según épocas históricas y culturas.

4. DOCUMENTOS QUE LLAMAN A LA PRUDENCIA

En los últimos años son más y más los actores que van tomando conciencia del peligro que puede comportar la inteligencia artificial para la vida humana y que promueven por ello declaraciones de principios e incluso legislaciones. Hablamos de científicos, empresarios, intelectuales, políticos y líderes religiosos. Ya hemos mencionado el documental *The Social Dilemma*, en el que personas que se dedicaron a construir el sistema de inteligencia artificial aplicado a Internet reconocen ahora haber obrado mal porque contribuyeron a dañar la libertad de millones de personas, algo a todas luces muy grave. En nuestro estudio *Bioética de la inteligencia artificial* presentamos algunos de estas declaraciones y legislaciones. Recordemos aquí brevemente algunas de ellas a modo de ejemplo ilustrativo:

Entre 1941 y 1942, el divulgador científico americano de origen ruso, Isaac Asimov redactó tres principios morales universales para cualquier robot, conocidos como *las Tres Leyes de la Robótica* de Isaac Asimov: 1) un robot nunca podrá causar daño a un ser humano; 2) obedecerá siempre a un ser humano, a menos que este principio colisione con el primero, en cuyo caso seguirá el primero y no el segundo; 3) tendrá derecho a autoprotgerse,

a menos que este principio colisione con el segundo, en cuyo caso seguirá el segundo y no el tercero¹².

En 1994, otro divulgador científico, Roger Clarke trató de mejorar los principios deontológicos de la Ética de la Robótica propuestos por Asimov con su Lista de las Leyes de la Robótica:

La Meta-Ley: Un robot solo podrá actuar si sus acciones están sujetas a las Leyes de la Robótica. Ley Cero: Un robot no podrá causar daño a la humanidad ni podrá permitir por omisión que la humanidad reciba daño alguno. Ley 1: Un robot no podrá herir a un ser humano ni podrá permitir por omisión que un ser humano sea herido, a menos que ello suponga violar una ley de orden superior. Ley 2: Un robot deberá obedecer órdenes dadas por seres humanos, excepto cuando esas órdenes contradigan una ley de orden superior; un robot deberá obedecer órdenes dadas por robots de rango superior, excepto cuando esas órdenes contradigan una ley de orden superior. Ley 3: Un robot tendrá que proteger la existencia de un robot de rango más alto siempre y cuando esa protección no entre en conflicto con una ley de orden superior; un robot deberá proteger su propia existencia siempre y cuando esa protección no entre en conflicto con una ley de orden superior. Ley 4: Un robot deberá llevar a cabo los deberes para los cuales fue programado, excepto cuando ello entre en conflicto con una ley de orden superior. La Ley de Procreación: Un robot solo podrá participar en el diseño o manufactura de otro robot cuando las acciones del nuevo estén sujetas a las Leyes de la Robótica¹³.

En 2011, el Consejo para la Investigación en Ingeniería y Ciencias Físicas (EPSRC - *Engineering and Physical Sciences Research Council*) y el Consejo para la Investigación en Artes y Humanidades (AHRC - *Arts and Humanities Research Council*) del Reino Unido publicaron cinco principios éticos para diseñadores, constructores y usuarios de robots junto con siete mensajes considerados de

¹² Asimov, I., "Runaround", en *Astounding Science Fiction*, Street and Smith, 1948.

¹³ Clarke, R. "An Extended Set of the Laws of Robotics", 1994, rogerclarke.com



suma importancia por sus autores. En ellos se afirman ideas tales como que los robots no deberían ser diseñados para matar o dañar a los seres humanos, salvo en casos de interés nacional (por cierto, inquietante excepción que abre la puerta a una larga lista de posibles horrores); los seres humanos, no los robots, son sujetos responsables; los robots no deberían ser diseñados para explotar a usuarios vulnerables; al contrario, su naturaleza mecánica debería ser transparente; una persona debería ser la responsable legal de un robot¹⁴.

En el 2017, más de 5 700 científicos firmaron en “Asilomar”, centro de eventos situado en la península de Monterrey, California, una lista de *23 Principios sobre Inteligencia Artificial*. En ellos se hacen propuestas tales como que el objetivo de la investigación en inteligencia artificial debería ser crear inteligencia beneficiosa, y que la inversión en inteligencia artificial debería ir acompañada de fondos para investigar acerca de su uso adecuado; debería haber un intercambio constructivo y sano entre los investigadores de inteligencia artificial y los legisladores; una cultura de cooperación, confianza y transparencia debería ser fomentada entre los investigadores y diseñadores de inteligencia artificial; cualquier intervención de un sistema autónomo en una decisión tendría que ir acompañada de una explicación satisfactoria y auditable por parte de una autoridad humana competente; los diseñadores de sistemas avanzados de inteligencia artificial son depositarios de las implicaciones morales de su uso; los sistemas de inteligencia artificial serán diseñados para que sus metas y comportamientos puedan alinearse con los valores humanos, compatibles con los ideales de dignidad humana, derechos, libertades y diversidad cultural; entre otros¹⁵.

En 2018, setenta instituciones educativas, científicas y tecnológicas firmaron en la Universidad de Montreal la *Declaración de Montreal para un Desarrollo Responsable de la Inteligencia*

¹⁴ Bodena, M. et al., 2017. “Principles of robotics: regulating robots in the real world”: *Connection Science* 29/2 (2017), pp. 124-129, www.tandfonline.com

¹⁵ *23 Principios sobre Inteligencia Artificial*, 2017, futureoflife.org

Artificial y condensaron en un *decálogo* los principios por los que debería regirse la inteligencia artificial: principio de bienestar, principio de respeto a la autonomía, principio de protección de la privacidad y la intimidad, principio de solidaridad, principio de participación democrática, principio de equidad, principio de inclusión de la diversidad, principio de prudencia, principio de responsabilidad, principio de desarrollo sostenible¹⁶.

La Iglesia Católica promovió en 2020 la *Llamada a la Ética de la Inteligencia Artificial* (“*The Call*”), firmada por instituciones italianas e internacionales, tanto religiosas, como empresariales, civiles o políticas, tales como la Academia Pontificia para la Vida, Microsoft, IBM, FAO y el Ministerio de Innovación del gobierno italiano en Roma, con la intención de promover un enfoque ético de la inteligencia artificial. En el 2023 se adhirieron un representante del judaísmo y otro del islam. La propuesta contempla tres áreas transversales para las que la inteligencia artificial debería servir y ser desarrollada (ética, educación y derechos) y expone seis principios: transparencia, inclusión, responsabilidad, imparcialidad, fiabilidad y, finalmente, seguridad y privacidad¹⁷.

En 2021, los 193 países miembros de la Unesco redactaron la primera norma mundial en este tema, la *Recomendación sobre la ética de la inteligencia artificial*¹⁸, en la que se promueve una visión ética, humanista e integral de la inteligencia artificial, entendiendo por *integral* el hecho de abordar esta temática en todas las dimensiones del ser humano.

En 2023, el Parlamento Europeo aprobó la *Ley de Inteligencia Artificial*, la primera ley sobre este tema en todo el mundo.

También en el 2023 se dio a conocer la *Declaración de Bletchley de la Cumbre de Seguridad de la Inteligencia Artificial*, firmada

¹⁶ *Declaración de Montreal para un Desarrollo Responsable de la Inteligencia Artificial*, Universidad de Montreal, 2018, declarationmontreal-iaresponsable.com.

¹⁷ *Llamada a la Ética de la Inteligencia Artificial*, Roma, 2023, romecall.org/the-call.

¹⁸ Unesco. *Recomendación sobre la ética de la inteligencia artificial*. París, 2021, unesdoc.unesco.org.



por países de tamaños, culturas, niveles económicos y grados de desarrollo muy diversos, procedentes de los cinco continentes, incluidos algunos grandes como Estados Unidos, la Unión Europea (que son 27 países), China o la India.

Todo esto muestra dos cosas: la primera es que cada vez son más los actores de todo tipo que toman conciencia del impacto que la inteligencia artificial va a tener en la vida humana; y la segunda, que los peligros que conlleva no son menores. De no haberlos, ¿para qué se habrían tomado la molestia de redactar tantas declaraciones y tan altisonantes?

5. ALGUNOS DESAFÍOS ÉTICOS DE LA INTELIGENCIA ARTIFICIAL

El hecho es que los desafíos éticos de la inteligencia artificial no son pocos. Veamos algunos de ellos, sin pretensión de ser exhaustivos.

Desconocimiento humano acerca de la inteligencia artificial.
Cuando hablamos de inteligencia artificial siempre pensamos en programas informáticos realizados por seres humanos que estos conocen y controlan; sin embargo, esto no es del todo cierto. En un soldado robot, por ejemplo, hay muchos programas informáticos, cada uno de ellos realizado por un equipo distinto de ingenieros, desconectados unos de otros. Cada equipo puede, sí, responder acerca del funcionamiento de su programa, pero ignora lo que los demás han introducido. No hay ningún ingeniero en el mundo que conozca con detalle todos los programas que lleva un soldado robot, lo que hace que en definitiva la actuación de este pueda resultar impredecible.

Puede que la ética de los robots no haya recibido la atención debida, por lo menos en Estados Unidos, si tenemos en cuenta que se suele creer erróneamente que los robots solo harán aquello para lo cual han sido programados. Desafortunadamente, esta creencia está desfasada, propia de tiempos en los que los robots eran mucho más simples y

sus programas podían ser creados y dominados por un solo individuo. Actualmente, los programas informáticos de los robots con millones de líneas de código han sido realizados por varios equipos de programadores, ninguno de los cuales conoce la programación completa del robot; por ello, ningún individuo puede predecir el efecto de una programación con certeza absoluta, teniendo en cuenta que algunas áreas de vastos programas pueden interactuar entre ellas de manera inesperada y sin investigación previa alguna¹⁹.

Esto nos lleva a hablar de la denominada *caja negra* de la inteligencia artificial. Los programas informáticos “toman decisiones” sin el conocimiento de sus programadores. De alguna manera podríamos decir que “crean” algo nuevo y de manera imprevista.

*La actividad humana perderá terreno sin que sepamos todavía adónde emigrará*²⁰. Con la inteligencia artificial vamos a delegar en la técnica una gran cantidad de actividad considerada hasta ahora humana. Esa actividad dejará de ser *humana* y pasará a ser *técnica*. ¿Qué haremos cuando nos hayamos liberado de todo ello? ¿En qué consistirá la actividad humana en una sociedad inundada por la inteligencia artificial? Es probable que aumente la pasividad humana, aunque también es posible un efecto muy distinto: haremos más cosas en el espacio de tiempo en el que antes hacíamos muchas menos, con lo que harán falta menos trabajadores, y los que trabajen irán más estresados, algo que han estudiado los filósofos Bjung-Chul Han y Hartmut Rosa. ¿No es una irresponsabilidad delegar actividad humana en la técnica sin saber todavía lo que haremos a continuación? De hecho, hace casi un siglo, el filósofo José Ortega y Gasset hizo ya esta advertencia aludiendo a lo que entonces eran

¹⁹ Lin, P., Bekey, G. y Abney, K., *Autonomous Military Robotics: Risk, Ethics, and Design*, San Luis Obispo, CA, Ethics + Emerging Sciences Group at California Polytechnic State University, 2008, p. 8.

²⁰ Para los siguientes puntos, cfr. Sols Lucia, J. *Ética de la inteligencia artificial. El caso de los soldados robot*. Monterrey: Universidad de Monterrey – Instituto Tecnológico de Monterrey – Universidad Autónoma de Nuevo León – Universidad Regiomontana, 2021, pp. 28-32; íd. *Desafíos éticos en el siglo XXI. Ciclo de conferencias en la Universidad Rafael Landívar*. Ciudad de Guatemala: Universidad Rafael Landívar – Editorial Cara Parens, 2024, pp. 56-59; Sols y de los Ríos. *Bioética*, pp. 54-58. No los citaremos cada vez.



nuevas tecnologías: con ellas se iba a liberar mucho tiempo libre, pero no se sabía qué se iba a hacer con él²¹. Lamentaba también que en los centros especializados y en las universidades se enseñara técnica sin una reflexión antropológica y ética que las acompañara, de tal modo que los nuevos profesionales sabían cómo funcionaba una determinada técnica, pero no qué función tenía la tecnología en la vida humana. El paso de los años contestaría la pregunta de Ortega y Gasset acerca de la derivación del tiempo libre generado por la técnica: desempleo. Lo veremos enseguida.

La toma de decisiones se desplazará del hombre a la máquina, con lo que desaparecerá la responsabilidad moral. Ninguna persona se hará responsable de lo que hagan la inteligencia artificial o la robótica, y las máquinas no podrán asumir esa responsabilidad por el hecho de no ser seres humanos. Vamos entonces hacia una sociedad en la que se tomarán muchas “decisiones”, pero en la que nadie se hará responsable de ellas.

Se producirá una pérdida de conocimiento. La mente humana es capaz de realizar cálculos muy complejos, de retener gran cantidad de información y de formular proposiciones agudas. No obstante, cada vez más delegamos en la inteligencia artificial estas funciones, con lo que, al dejar de desarrollarlas, las vamos perdiendo. No memorizamos fechas de la historia porque podemos dar con ellas en pocos segundos en Internet; no hacemos cálculos matemáticos complejos porque la calculadora los hace por nosotros con mayor precisión y velocidad; no calculamos la estructura arquitectónica de un nuevo edificio porque un programa informático lo hace por nosotros en mucho menos tiempo. Poco a poco olvidaremos el conocimiento que deleguemos. Esto ocurrirá en muchos dominios de lo humano. ¿Es bueno que perdamos conocimiento? Vamos hacia un *homo ignorans*, sustituto del *homo sapiens sapiens*.

Aumentará el desempleo. Ya lo hemos apuntado más arriba. Retomémoslo ahora. Empezó tímidamente en los años ochenta del

²¹ Ortega y Gasset, *Meditación de la Técnica*.

siglo pasado y se extendió como un fenómeno nuevo en la historia de la humanidad en los años noventa hasta el punto de que Jeremy Rifkin afirmó en su libro *El fin del trabajo* (1995) que habíamos entrado en una nueva era caracterizada por la carencia de trabajo. ¿Qué pasará con esa multitud de desempleados? ¿De qué renta vivirán? La propuesta de Rifkin de reducir la jornada laboral para distribuir el trabajo no hizo sino posponer el problema unas décadas. Y el problema es que el trabajo para todos es una realidad que pertenece al pasado.

Se incrementará la desigualdad en la humanidad. Basta con salir a la calle y mirar lo que está pasando para darse cuenta en pocos minutos de que son siempre los más ricos los que adquieren antes las nuevas tecnologías. Más arriba hemos hablado de los robots que cuidan a ancianos (Spot, CyberDog 2, Zora, Da Vinci), algunos de los cuales cuestan veinte millones de dólares, una cantidad imposible para los pobres y para la clase media y media alta, pero asequible para la clase alta. Con ello, las nuevas tecnologías aumentan la desigualdad porque las clases más ricas y los países más desarrollados las implementan enseguida, mientras que las clases más pobres y los países menos desarrollados tardan mucho más en hacerlo, o incluso directamente no llegan a implementarlas nunca. Teniendo en cuenta que la desigualdad está en el origen de la mayor parte de los problemas de violencia que hay en el mundo, es fácil concluir que a mayor desigualdad, habrá aún mayor violencia de la que ya tenemos ahora, algo que no constituye una buena noticia.

*Lo insustituible del ser humano*²². La máquina no puede sustituir la relación médico-paciente, farmacéutico-cliente o psicólogo-paciente. La salud del paciente depende del médico, del farmacéutico, del psicólogo, y estos ejercen su profesión, acompañan a aquel en función de su conocimiento y se ganan la vida con ello. Sin embargo, diversos programas de inteligencia artificial están sustituyendo estas profesiones, de tal modo que el paciente ya no habla con un profesional, sino con una máquina, que es la que le receta un

²² Para los siguientes puntos, cfr. Sols y de los Ríos, *Bioética*, pp. 59-129.



tratamiento e incluso le expende el medicamento adecuado con la dosis necesaria. Sin embargo, las cosas no son tan simples como aparentan. En la relación médico-paciente, farmacéutico-cliente o psicólogo-paciente se establece un vínculo interpersonal de *consentimiento informado*. El paciente da mucha información a su médico, confiando en que este la utilizará adecuadamente. Estamos en el orden de la *sociabilidad*, que desaparece al sustituir al médico por un programa informático. Los malestares que tenemos no son siempre unidimensionales: “¿Le duele la cabeza? Tómese un Ibuprofeno”. El paciente puede estar viviendo emociones muy diversas. Es importante que el médico ayude al paciente a expresarse libremente en un entorno de confianza. En ese marco de confianza y de transparencia, se pueden tomar buenas decisiones entre médico y paciente. Las máquinas están lejos de lograr esto.

En la relación médico-paciente o en otras similares a ella, la *confidencialidad* es importante, pero esta podría estar en peligro por el uso de los macrodatos (*big data*). El médico no puede recordar toda la información acerca de la salud de cada uno de sus pacientes, por lo que introduce todo lo que este le dice en un programa informático que, lo desee o no, está conectado a una red global de información que contribuye al avance de la ciencia: por ejemplo, cuánta gente tiene el Covid o la Influenza (gripe), de qué franja de edad son, etc. En esos macrodatos se acumula información acerca de las prácticas sexuales de las personas, su ritmo de vida, su alimentación, etc., que puede ser usada con fines alejados de la salud. Esta manipulación masiva es mucho más efectiva con la inteligencia artificial.

En el uso de esa información obtenida de los pacientes se pueden dar *hallazgos no deseados*. Tratando de averiguar si un paciente tiene un tumor maligno, el sistema puede dar a conocer otra información que no se buscaba y que el paciente no quería compartir o que desconocía. El paciente se siente agredido en su privacidad, pero nadie sabe quién es el responsable de ese hallazgo no deseado. Ninguna persona en concreto lo es, y el sistema informático no puede ser responsable porque no es una persona.

La inteligencia artificial contribuye a una tendencia que ya se está dando en las sociedades desarrolladas: el *declive de las relaciones interpersonales*. Hemos hablado de la relación médico-paciente y de la “no-relación” interpersonal robot-anciano, pero hay muchas otras. Hay una notable deriva hacia el individualismo, hacia la falta de compromiso comunitario o social. Muchas parejas prefieren no casarse. En algunas, cada uno vive en su casa para no perder su respectiva autonomía. Otras prefieren los perros a los hijos, a los que denominan “perrhijos”, y a sus gatos, “gathijos”. El paso siguiente serán los robots de apariencia progresivamente humana en la forma, en la voz, en el comportamiento, máquinas que nunca nos causarán problemas, que siempre nos darán la razón y conversarán amigablemente con nosotros. Nos harán creer que estamos acompañados, cuando en realidad estaremos completamente solos.

Cada vez se está dando más el fenómeno de la *autonomía de la máquina*. Durante siglos, la máquina era un artilugio mecánico controlado por el ser humano. Ese control está desapareciendo, dado que la inteligencia artificial es cada vez más compleja. Hemos hecho alusión ya a este fenómeno al hablar de los soldados robot. Hasta ahora hemos conocido robots muy simples: uno para la limpieza del hogar, otro para lavar el coche, otro para responder a tus preguntas en el aeropuerto. Sin embargo, los desarrollos tecnológicos siempre siguen la misma lógica: cada nueva máquina tiene una sola función y con el paso del tiempo se van integrando diferentes máquinas en una sola. Los sistemas informáticos se están integrando, al igual que los robots. Eso hace que para diseñar un robot del futuro no baste con un solo equipo de ingenieros, sino que harán falta varios, desconectados entre sí. Cada equipo solo conocerá el sistema que haya introducido, pero no sabrá nada de los demás. Por ello, no habrá ningún ser humano en la Tierra que conozca y domine todo lo que un robot lleva en su interior. Con ello habremos perdido el control de la máquina, que se manejará con progresiva autonomía. Esa autonomía es irreal, pues la máquina no tiene vida propia, sino que se limita a funcionar con los programas que la mueven; no obstante, hablamos de autonomía porque la máquina parece “tomar decisiones” que ningún ser humano controla.



Al no ser el hombre quien toma las decisiones, entonces surge la pregunta acerca de *quién es el responsable de lo que ocurre*. Nadie se hace responsable de los posibles daños que pueda causar un robot, por ejemplo, un soldado robot que dispare a un inocente al interpretar que es un enemigo, o el robot cirujano que mate a un paciente en el quirófano, o el robot industrial que dañe gravemente a un empleado. Entraremos en un limbo moral.

Este tipo de errores de la máquina pueden llegar a tener graves consecuencias para la vida humana. Por ejemplo, en la *modificación genética* en seres humanos puede darse algún error, que no afectará solo al individuo en cuestión, sino también a toda su descendencia biológica, lo cual es muy grave.

Puede parecer de ciencia ficción, pero no lo es. Los *atentados terroristas* serán mucho más fáciles con el uso de la inteligencia artificial. En el futuro los terroristas podrán hacer uso de la inteligencia artificial para causar un daño incalculable en población inocente.

Supongamos que dentro de unas décadas, en un país como España, de unos 50 millones de habitantes, la mitad de la población se trata contra una enfermedad estacional mediante sucesivas dosis que proporciona una tecnología ya introducida en el cuerpo. Durante quince días, a un ritmo de tres veces al día, la inteligencia artificial administra con exactitud matemática esas dosis a unos 25 millones de personas. De pronto, un terrorista hackea el sistema y da orden de que todas las dosis sean suministradas de una sola vez. Con la simple alteración de un algoritmo, el terrorista podrá acabar con la vida de 25 millones de personas, muchas más si hablamos de Estados Unidos, de la India o de la Unión Europea. ¿Nos podemos permitir el lujo de correr ese riesgo? Obviamente, el responsable de las muertes no sería el diseñador del programa bioinformático, sino el terrorista, pero sabiendo que algo así puede ocurrir, tenemos una grave responsabilidad moral si llegamos a implementar este sistema médico²³.

²³ Sols y de los Ríos, Bioética, p. 129.

6. CONCLUSIÓN

No podemos reflexionar acerca de la inteligencia artificial o de la robótica de manera binaria, dualista, dicotómica: “o es buena o es mala”. Se trata de una realidad muy compleja, creada por el ser humano, a la que debemos aproximarnos de manera matizada. Hay dos ingenuidades a evitar: la conservadora (“todo lo nuevo es peligroso”) y la pseudo-progresista (“todo lo nuevo es bueno”). La inteligencia artificial puede realizar aportaciones altamente positivas a la vida humana, pero también presentar riesgos de un volumen incalculable, precisamente debido a sus enormes dimensiones. Se trata de una operación matemática sencilla: si un hombre puede moverse en unidades, el mal que haga también se moverá en esa dimensión; si tiene un gran poder político en un país de millones de personas, su mal afectará a mucha más gente (es el caso de las dictaduras); si un sistema informático se mueve en miles de millones de unidades, el mal que pueda causar también será descomunal, como nunca habíamos visto hasta ahora. Ese mal se dará unas veces por mala voluntad humana (el mercado, el Estado, los terroristas, los *hackers*) y otras por un error, pero en ambos casos será devastador, no solo para nuestra generación, sino también para las venideras. ¿Será la inteligencia artificial la antítesis hegeliana que haga perecer nuestra civilización?

Las palabras clave aquí son *prudencia* y *discernimiento*. A eso apuntan todos los documentos y legislaciones sobre el tema que hemos expuesto. ¿Lograremos regirnos todos por la prudencia y el discernimiento? ¿Todos, sin excepción?

REFERENCIAS BIBLIOGRÁFICAS

23 *Principios sobre Inteligencia Artificial*, 2017, futureoflife.org

Asimov, I., “Runaround”, en *Astounding Science Fiction*, Street and Smith, 1948.

Bodena, M. *et al.* “Principles of robotics: regulating robots in the real world”: *Connection Science* 29/2 (2017), www.tandfonline.com



Clarke, R. "An Extended Set of the Laws of Robotics", 1994, rogerclarke.com

Declaración de Montreal para un Desarrollo Responsable de la Inteligencia Artificial, Universidad de Montreal, 2018, declarationmontreal-iaresponsable.com.

Ellul, J. *La Technique ou l'Enjeu du siècle*, París: Armand Colin, 1954 (en español: *La edad de la técnica*. Barcelona: Octaedro, 1977).

Ellul, J. *Le système technicien*. París: Calmann-Lévy, 1977.

Ellul, J. *Le Bluff technologique*. París: Hachette, 1988 (en español: *Anarquía y cristianismo*, México D.F.: Jus, 2005).

Florensa, A. *La vida humana en el medi tècnic. El pensament de Jacques Ellul*, Barcelona: Claret, 2010.

Florensa, A. y Sols, J. (eds.) *Ética de la investigación científica*. Bilbao: Desclée de Brouwer, 2017.

Habermas, J. *Ciencia y técnica como ideología*. Madrid: Tecnos, 1984.

Heidegger, M. *La pregunta por la técnica*. Barcelona: Herder: 2021.

Hottois, G. *El paradigma bioético: una ética para la tecnología*. Barcelona: Anthropos, 1991.

Jonas, H. *El principio de responsabilidad*. Barcelona: Herder, 1995.

Lambert, D., "The humanization of robots: lethal autonomous weapons systems and ethics", en *The Humanization of Robots and the Robotization of the Human Person. Ethical Reflections on Lethal Autonomous Weapons Systems and Augmented Soldiers*, Ginebra, The Caritas in Veritate Foundation Working Papers, 2017, editado por A. de La Rochefoucauld y S. Saldi.

Lin, P, Bekey, G. y Abney, K., *Autonomous Military Robotics: Risk, Ethics, and Design*, San Luis Obispo, CA, Ethics + Emerging Sciences Group at California Polytechnic State University, 2008.

□ *Llamada a la Ética de la Inteligencia Artificial*, Roma, 2023, romecall.org/the-call.

Ortega y Gasset, J. *Meditación de la Técnica y otros ensayos sobre ciencia y filosofía*. Madrid: Revista de Occidente en Alianza Editorial, 1998.

Sols Lucia, J. *Ética de la inteligencia artificial. El caso de los soldados robot*. Monterrey: Universidad de Monterrey - Instituto Tecnológico de Monterrey – Universidad Autónoma de Nuevo León – Universidad Regiomontana, 2021.

Sols Lucia, J. *Desafíos éticos en el siglo XXI. Ciclo de conferencias en la Universidad Rafael Landívar*. Ciudad de Guatemala: Universidad Rafael Landívar – Editorial Cara Parens, 2024.

Sols Lucia, J. y de los Ríos Uriarte, M. E. *Bioética de la inteligencia artificial*. Madrid: San Pablo – Universidad Pontificia Comillas, 2024.

Unesco. *Recomendación sobre la ética de la inteligencia artificial*. París, 2021, unesdoc.unesco.org.