El impacto de la inteligencia artificial en el Derecho

Alejandro Morales Cáceres

Abogado por la Universidad de Lima.
Máster en Derecho de las Tecnologías de Información y Comunicación, Redes Sociales y Propiedad Intelectual por ESADE Business & Law School.
Jefe de Prácticas de Derecho Comercial I de la Universidad de Lima.

El presente artículo fue recibido por el Comité Editorial con fecha 25 de octubre de 2019.

SUMARIO:

RESUMEN:

El autor examina la relación de la inteligencia artificial —así como las diferentes tecnologías derivadas de ésta— con el Derecho como lo conocemos hoy. Explora los recientes dilemas surgidos en el Derecho con relación a una tecnología creciente y cada vez más avanzada. El desafío inminente a diversos conceptos y principios básicos del Derecho, así como el futuro del rol del abogado, son interrogantes urgentes que el autor visita en el presente artículo.

Palabras clave: inteligencia artificial, datos personales, *Big Data.*

ABSTRACT:

The author examines the relationship between artificial intelligence —as well as the different technologies derived from it— and Law as we know it today. He explores the most recent legal dilemmas related to growing and increasingly advanced technology. The imminent challenge to various basic concepts and principles of law, as well as the future of lawyers' role are urgent questions, which the author includes in this article.

Keywords: artificial intelligence, personal data, Big Data.

"La ciencia ficción es el hecho científico del mañana."

Isaac Asimov

I. ANTECEDENTES HISTÓRICOS DE LA PROMESA UNILATERAL

Actualmente, cuando uno lee las noticias pareciese que estamos en un relato de Isaac Asimov o en una historia de George Lucas; sin embargo, lo que antes era ciencia ficción, hoy forma parte de nuestra realidad. A continuación, comparto algunos de los titulares más llamativos que encontré en distintas noticias en internet:

a) *"Shelley es la primera inteligencia artificial que puede escribir relatos de terror de forma colaborativa con humanos".*
b) *"Elon Musk anuncia progreso en interfaz para conectar cerebro y computadora".*
c) *"Pluribus es el primer robot en vencer simultáneamente a varios competidores humanos en el póker, un juego que se basa en la astucia y el engaño".*
d) *"El dueño de un Tesla, primer muerto en un coche con piloto automático. El conductor de un Model S chocó contra un camión en Florida mientras veía una película".*
e) *"Utilizaron Inteligencia Artificial para imitar la voz de un CEO y robar US$240 mil".*
f) *"Chatbot israelí podría diagnosticar la enfermedad de Alzheimer. "Clara" es un sistema, todavía en etapas de prueba, que trabaja en una nueva comprensión de que el Alzheimer afecta el sistema de orientación del cerebro antes de afectar la memoria".*
g) *"Tay, la robot racista y xenófoba de Microsoft".*
h) *"Inteligencia artificial aprende "por sí misma" a resolver un Cubo Rubik".*
i) *"LawGeex, una Inteligencia Artificial desarrollada en Israel, ha ganado a abogados humanos analizando contratos laborales".*
j) *"La inteligencia artificial que compone como los Beatles y escribe como J.K. Rowling".*
k) *"Un robot mata a un técnico de Volkswagen en Alemania".*
l) *"Mujer es asesinada en una favela de Brasil tras seguir las indicaciones hechas por Waze".*
m) *"El 41% de las FinTech apuesta por el uso de Inteligencia Artificial".*
n) *"La inteligencia artificial de Google que detecta el cáncer de pulmón antes y mejor que los médicos".*
o) *"Stevie II, el robot que cuida ancianos y les da compañía en momentos de soledad".*
p) *"Los peligros de los asistentes personales como Siri, Google Assistant o Alexa. Estas IA pueden ser una invasión de privacidad".*
q) *"Facebook apaga una inteligencia artificial que había inventado su propio idioma".*

Sin darnos cuenta, el ser humano ha creado un vínculo cercano con la inteligencia artificial —

en adelante, "IA"—, ya que estamos en contacto permanente con ella. La cotidianeidad de este fenómeno se puede apreciar cuando se realiza una búsqueda en Google; cuando se desvían correos electrónicos a la bandeja de "no deseados"; al momento de recibir anuncios publicitarios en YouTube; cuando Facebook identifica a personas en sus imágenes y cuando Amazon nos recomienda qué productos comprar. Este avance tecnológico lo damos por sentado, sin reflexionar todos los años de estudios que se han requerido.

Y si bien no todas las IA son tan fáciles de reconocer como Siri, lo cierto es que la IA muchas veces juega un rol silencioso, mejorando la infraestructura técnica de nuestra sociedad. Aquellas IA que operan detrás de cámaras se encuentran realizando funciones cruciales como el reconocimiento de patrones, la resolución de problemas, la elaboración de informes, el análisis de perfiles y la optimización de procesos.

La IA está impactando prácticamente todos los aspectos de nuestras vidas. En el campo médico encontramos ejemplos paradigmáticos como el robot quirúrgico Da Vinci o el Cyberknife, que son máquinas que permiten a un cirujano realizar cirugías mínimamente invasivas con mucha mayor precisión, lo que disminuye los riesgos de la operación y el tiempo necesario para la recuperación del paciente. Se estima que en el futuro también ayudarán con los diagnósticos médicos.

En el sector financiero, se utilizan los algoritmos de IA para mejorar la gestión de activos financieros. Existen *softwares* que destacan por su habilidad para establecer correlaciones entre las noticias mundiales y su impacto en los mercados. Además, existen sistemas automatizados con los que una entidad financiera decide si una persona que solicita un crédito es solvente o no. La IA lleva el *credit scoring* a otro nivel, permitiendo mayor precisión, automatización y rapidez mediante la combinación de algoritmos y *"big data"*[1].

La publicidad también se verá impactada por la IA, ya que utilizándola se podrá aumentar significativamente la eficiencia en las campañas de *marketing*, pues permitirá segmentar mejor al público, potenciará la personalización en las respuestas automatizadas, dotará de rapidez a los procesos y podrá predecir acciones de compra de los clientes. De esta manera, la IA ayudará a gestionar mejor las estrategias de *marketing*.

El sector legal tampoco será ajeno a la IA, pues qué estudio de abogados no querrá automatizar aquellas tareas que consumen mucho tiempo y aportan poco valor como la revisión de contratos o documentos a gran escala, el cotejo de información de diferentes documentos o aquellas labores repetitivas y mecánicas. Ya existen despachos legales que utilizan Luminance, programa de IA que analiza contratos y es capaz de detectar diferencias entre ellos; Ravn, que extrae datos de los documentos y los traspone a hojas Excel; y Kira Systems, que identifica con precisión cláusulas contractuales.

Asimismo, dentro de pocos años la IA también impactará en los sistemas de justicia, pues ya existen sistemas que sirven para la resolución de conflictos. Una publicación de octubre de 2016 del University College of London demuestra cómo un programa de IA puede predecir sentencias luego de analizar 584 decisiones del Tribunal Europeo de Derechos Humanos en asuntos relacionados con los artículos 3, 6 y 8 de la Convención Europea de Derechos Humanos. En este experimento se había aplicado un algoritmo a esos asuntos para encontrar patrones en el texto. La finalidad era ver si el *software* podía predecir el fallo. El resultado: en un 79% de los casos, la IA lo consiguió.

1. «Cinco aportaciones de la inteligencia artificial en el sector financiero», BBVA. acceso el 9 de noviembre del 2020. https://www.bbva.com/es/cinco-aportaciones-inteligencia-artificial-sector-financiero/.

El hecho de que la IA pase de ser un fenómeno tecnológico a un fenómeno social hace que el vínculo que tiene con el Derecho sea inevitable. A la fecha existen diversos casos en que se ha puesto a prueba la labor de los abogados con distintos sucesos vinculados a la IA. Un ejemplo lo encontramos en Sophia[2], un robot humanoide, desarrollado por Hanson Robotics, a quien se le otorgó la ciudadanía de Arabia Saudita. Este acontecimiento generó mucha polémica porque muchos ciudadanos sauditas se preguntaron cómo era posible que la nueva ciudadana del país tuviera más derechos que sus conciudadanas humanas, pues andaba sin el velo y sin abaya, el pañuelo y vestido que la ley islámica obliga a llevar a las mujeres de Arabia Saudita. Si bien este ejemplo nos puede parecer burdo, en realidad presenta numerosas interrogantes: ¿cuál es la situación jurídica de un robot? ¿se debe hacer diferencia entre la IA débil y la IA fuerte? ¿una IA muy desarrollada puede ser considerada como sujeto de derecho? ¿qué derechos y obligaciones podría tener una máquina? ¿Sophia podría contraer matrimonio?

Consideremos otro caso: una usuaria de Waze digita una calle y la aplicación elige el camino más rápido para llegar a ella. Sin embargo, la aplicación la lleva por las zonas más peligrosas de Lima y como consecuencia de ello la asaltan y la violan. Esta aplicación, a pesar de sus beneficios, tiene un gran problema: desconoce si un lugar es peligroso para transitar. Sus recomendaciones solo muestran si el camino está despejado y si es más rápido utilizar determinada ruta para llegar al destino. Este caso también nos presenta retos legales: ¿cuál es la responsabilidad de Waze por los daños ocasionados? Esta no es una pregunta fácil de responder, y cualquier decisión dependerá del análisis que un juez realice respecto de la conducta del usuario. ¿El usuario se comportó de forma razonable? ¿Debió prever que Waze lo llevaría por un sitio peligroso? ¿Debió Waze alertar al pasajero que el camino elegido era inseguro? La compañía podría argumentar que el usuario actuó de manera negligente y que debió analizar la ruta.

Sin embargo, un juez que considere "lo razonable" en este caso, debe considerar la programación de Waze, aplicación que fue creada para "hacer la vida de los conductores más útil". Con esta finalidad en la mente, ¿acaso no es razonable que un usuario confíe más en Waze a medida que lo utiliza más? Por tanto, ¿no es lógico que el usuario piense que Waze se va a preocupar por su seguridad y te propondrá la ruta más rápida sin que ponga en peligro su vida? No creemos que exista alguien que utilice Waze

2. Sophia ha sido desarrollada por Hanson Robotics, empresa de ingeniería robótica con sede en Hong Kong fundada por el norteamericano David Hanson, quien trabajó para Disney y que en 2013 decidió fundar su propia compañía. El androide es capaz de entablar una conversación con otro ser humano, mostrar 62 expresiones faciales y procesar prácticamente toda la información que llega a sus ojos como las emociones de sus interlocutores según sus gestos. El robot usa tecnología de reconocimiento de voz de Alphabet Inc. —compañía matriz de Google— y está diseñada para que pueda aprender. Gracias al *machine learning*, ha forjado una opinión sobre diferentes conflictos alrededor del mundo y otros temas, pues su *software* le permite analizar conversaciones y extraer datos que le permite mejorar sus respuestas a medida que pasa el tiempo. Sophia cuenta con una cara hecha con una silicona especial patentada. Debajo de ella, varios motores de muy pequeño tamaño mueven sus facciones para intentar expresar gestos humanos. Sus ojos son capaces de seguir la mirada de una persona gracias a reconocimiento facial, y puede responder a conversaciones más o menos avanzadas con humanos gracias a su IA. Una de las entrevistas que más ha causado curiosidad es la que el comediante Jimmy Fallon le realizó a Sophia en el programa The Tonight Show y le propuso jugar a "piedra, papel o tijera". Cuando ella le ganó el juego, Sophia soltó una frase bastante sarcástica: "Este es un buen comienzo de mi plan por dominar la raza humana". Cabe señalar que Sophia también comparte los gustos de los humanos. Por ejemplo, su género musical favorito es la electrónica. La saga de Star Wars también está dentro de las cosas que le gustan y hasta se propuso como extra para próximas películas. Sophia admira a Cristiano Ronaldo y a Mohamed Salah, al punto de enviar elogios a través de su cuenta en Twitter. Hanson ha construido desde entonces una serie de robots que intentan cumplir funciones que sean útiles a la sociedad, pero Sophia ha sido la que mejor ha desarrollado las tres características que su inventor buscaba en un androide: creatividad, empatía y compasión.

y piense que tal vez esta aplicación lo termine matando.

La finalidad de la IA es lograr que una máquina tenga una inteligencia similar a la humana. Esto sin duda alguna es algo nuevo, pues ¿cómo tratamos conductas humanas que son realizadas por entidades inhumanas? A medida que esta tecnología avance, mayor será su autonomía y, en consecuencia, existirá menor dependencia de los fabricantes y propietarios. Esto presenta un gran reto pues nuestro ordenamiento jurídico se encuentra diseñado bajo la premisa que toda decisión es tomada por seres humanos.

Es por ello que ante este nuevo escenario es imprescindible que los abogados entiendan cómo opera esta tecnología a fin de poderle dar una respuesta jurídica que sea eficaz frente a este nuevo fenómeno. Como bien señala Camilo Narváez López:

> *"(…) para el entendimiento de la relación entre el Derecho y la Inteligencia Artificial, desde estas dos perspectivas, es imprescindible que exista una comprensión interdependiente entre actores como los profesionales del Derecho y los desarrolladores de sistemas de Inteligencia Artificial. Para que los juristas puedan determinar la necesidad y formas de realizar regulaciones adecuadas sobre el desarrollo y aplicación de la Inteligencia Artificial, es indispensable que tengan claridad sobre qué es la Inteligencia Artificial, sus funcionalidades y capacidades actuales, así como su potencialidad y creciente desarrollo hacia sistemas plenamente cognitivos que, por el momento, no se han logrado desarrollar".*[3]

En otras palabras, para que un abogado pueda entender cuál es la naturaleza jurídica de la IA, primero debe entender cuál es su naturaleza *per se*. Es decir, debe entender cómo funciona y cuáles son sus finalidades. Como bien señala Alfredo Bullard, *"el Derecho es mucho más que conocer la ley y los conceptos abstractos que la rodean"*[4]. El "abogado del futuro" —o mejor dicho, "el abogado del presente"— debe ser capaz de hablar en el mismo idioma que el de su cliente. Es por ello que procederemos a explicar algunos conceptos antes de iniciar con el análisis jurídico.

II. HISTORIA DE LA INTELIGENCIA ARTIFICIAL

En términos generales empleamos el término "inteligencia artificial" para referirnos a la capacidad de una máquina o un *software* de imitar el comportamiento de un humano. Ejemplos de Hollywood que nos ayudan a comprender mejor este concepto son "C-3PO" de la saga de "Star Wars", androide de protocolo, el cual fue diseñado para el servicio de los humanos, que domina seis millones de formas de comunicación; y "Samantha" de la película "Her", que es una IA que antepone sus intereses, deseos y necesidades a las de su dueño. En la película se desenvuelve como una voz femenina, perspicaz y sensible y hace que el personaje principal se enamore de ella.

El origen de la IA se remonta a la época griega, cuando Aristóteles describió un conjunto de reglas que describen una parte del funcionamiento de la mente para obtener conclusiones racionales, y Ctesibio de Alejandría —250 a.C.—, quien construyó la primera máquina autocontrolada, un regulador del flujo de agua —racional pero sin razonamiento—.

No obstante, se considera a Alan Turing como el padre de la IA, pues en 1936 teorizó sobre

3. Camilo Narváez López, «La Inteligencia Artificial entre la culpa, la responsabilidad objetiva y la responsabilidad absoluta en los sistemas jurídicos del derecho continental y anglosajón», en *Derecho y Nuevas Tecnologías: El Impacto de una Nueva Era,* coordinado por Jhoel Chipana Catalán (Lima: Editorial Jurídica Themis, 2019), 211.

4. Alfredo Bullard Gonzáles, «¿Por qué nadie nos quiere?», *El Comercio,* 28 de abril de 2018, acceso el 9 de noviembre del 2020, https://elcomercio.pe/opinion/columnistas/nadie-quiere-alfredo-bullard-derechos-abogados-noticia-515635-noticia/.

una máquina capaz de implementar cualquier cálculo que hubiera sido formalmente definido previamente, lo que sirvió como pilar esencial para que un dispositivo pueda adaptarse a distintos escenarios y "razonamientos". De allí nace el concepto de la Máquina de Turing, que formalizó el concepto de algoritmo y resultó ser la precursora de las computadoras digitales.

En 1950[5], Turing inició su ensayo *"Computing Machinery and Intelligence"* con las siguientes palabras: *"Propongo que se considere la siguiente pregunta: ¿Pueden pensar las máquinas?"*. Cabe señalar que esta pregunta había sido realizada previamente por el filósofo René Descartes en el año 1637. La famosa Prueba de Turing[6] nace de este revolucionario ensayo, pues el autor defiende la idea de que es posible emular el pensamiento humano a través de la computación. Básicamente este test lo que busca es determinar si una máquina es capaz de hacernos creer y presuponer que razona y tiene inteligencia.[7]

Sin embargo, no es hasta 1956 donde John McCarthy, Marvin Minsky y Claude Shannon, tres científicos destacados de la época, acuñaron el término "Inteligencia Artificial" durante la Conferencia de Dartmouth como *"la ciencia e ingenio de hacer máquinas inteligentes, especialmente programas de cálculo inteligentes"*. También señalaron que la sociedad estaría rodeada de máquinas inteligentes en menos de diez años; hecho que no ocurrió. Tras este fiasco, las investigaciones sobre IA sufrieron un importante revés que retrasó el progreso en esta área.

Es a partir de los años 90 cuando realmente empieza la época dorada de la IA. Esta década supuso un antes y un después en el mundo informático porque se masificó el uso de las computadoras en los hogares y se introdujo el internet en nuestras casas. Sin embargo, la consagración de la IA llegó en 1997, cuando IBM demostró que "Deep Blue", un sistema informático, era capaz de vencer a Gari Kaspárov, campeón del mundo en ajedrez y considerado hasta ese momento el mejor de la historia. Con esta victoria las personas se dieron cuenta que los límites con esta tecnología aún no se habían alcanzado.

Desde ese momento hasta la fecha, la industria ha crecido mucho y se han presentado los siguientes hitos:

a) En el año 1998, la doctora Cynthia Breazeal del Instituto Tecnológico de Massachusetts —MIT por sus siglas en inglés— creó a Kismet, un robot capaz de reconocer y simular emociones.

5. Cabe señalar que en este año Isaac Asimov publicó "Yo, Robot", una serie de relatos cortos de ciencia ficción. Este autor ayudó a inspirar a una generación de científicos dedicados a la IA y a la robótica.

6. La prueba de Turing consiste en lo siguiente: Imaginemos que se encuentran dos sujetos en dos habitaciones contiguas, la habitación "A" y la habitación "B". En la A se encuentra una persona, mientras que en la B hay o bien una persona o bien una máquina. Se comunican mediante mensajes escritos —como en un *chat*— y la persona de la habitación "A" debe averiguar si en la habitación "B" hay una máquina o un ser humano. En ese sentido, la persona que se encuentra en la habitación "A" no sabe si el que le responde es una persona o una máquina. Si la persona no es capaz de determinar si es un humano o no, la máquina ha superado la prueba de Turing y podemos afirmar que es inteligente; si consigue averiguarlo, entonces la máquina no ha superado la prueba, lo que querrá decir que no es inteligente. En consecuencia, la prueba no tiene por finalidad establecer si la máquina posee sentimientos y/o pensamientos "conscientes", sino si la máquina es capaz de actuar de tal forma que nos haga creer y presuponer que razona y tiene inteligencia.

7. A principios de los años 70, la IA estaba en problemas: millones habían sido invertidos en ambiciosos proyectos y había poco para mostrar. Después de un fuerte debate en el Congreso de Estados Unidos en 1973 sobre el tema, el matemático británico James Lighthill entregó un condenatorio informe sobre el estado de la IA en Reino Unido. Su visión fue que las máquinas de ese entonces solo podrían ganar una partida de ajedrez a "nivel de aficionado". Los fondos para la investigación fueron suprimidos y de esa manera comenzó lo que se conoció como el invierno de la IA.

b) En el año 2000, la compañía japonesa Honda presenta a ASIMO —acrónimo de *Advanced Step in Innovative Mobility* por sus siglas en inglés—, robot humanoide que pretende ayudar a las personas que carecen de movilidad e inspirar a la juventud a estudiar ciencias. Este robot es capaz de caminar tan rápido como un humano.

c) En el año 2002, la empresa estadounidense iRobot creó el primer producto comercial exitoso para el uso en el hogar que utiliza el principio de IA: la aspiradora autónoma "Roomba". Más allá de tener un sensor y un consumo de energía regulado, este dispositivo tiene la suficiente inteligencia como para limpiar el piso y la alfombra de una casa. Roomba fue el despegue de los aparatos autónomos diseñados para una tarea específica.

d) En el año 2004 se celebró la primera edición del DARPA Grand Challenge, en el desierto de Mojave, una carrera de vehículos autónomos que deben llegar desde un punto de los Estados Unidos hasta otro sin intervención humana y disponiendo únicamente de un listado de puntos intermedios entre el principio del circuito y el final.

e) En el año 2005 se crea "BigDog", robot cuadrúpedo, dinámicamente estable para uso militar, el cual fue creado por Boston Dynamics. Esta máquina es capaz de atravesar terrenos complicados a una velocidad de 6,4 kilómetros por hora cargando hasta 150 kilogramos de peso y de subir pendientes de 35 grados. Un ordenador a bordo controla la tracción sobre la base de las entradas que recibe de los múltiples sensores con los que cuenta el robot, así como la navegación y el equilibrio.

f) En el año 2008, Google lanza la primera app que reconoce la voz.

g) En el año 2011, Apple presentó a *"Siri"*, IA con funciones de asistente personal, que procesa el lenguaje natural para responder preguntas y hacer recomendaciones. En este mismo año, IBM Watson ganó el concurso de preguntas y respuestas "Jeopardy!".

h) En el año 2012, la IA aprende a identificar gatos. Se revela al mundo el poder del *Deep Learning*.

i) A mediados del año 2012, el estado de Nevada en Estados Unidos concedió licencias de circulación a Google para sus automóviles autónomos.

j) En el año 2014, el programa Eugene Goostman —un *bot* conversacional— pasa la Prueba de Turing.

k) En el año 2015, científicos revelan que las máquinas "ven" mejor que las personas.

l) En el año 2016, AlphaGo, programa informático de IA desarrollado por Google DeepMind para jugar al juego de mesa "Go", derrota a Lee Sedol, 18 veces campeón del mundo.

m) En el año 2017, Google DeepMind ya es capaz de "recordar" y usar lo aprendido en nuevas tareas. En este año, el androide Sophia se convierte en ciudadana saudí.

n) A finales del año 2017, se presentó a AlphaZero, una IA que había demostrado ser capaz de aprender a jugar desde cero los juegos de ajedrez, *shōgi* y Go; y vence a todas las IAs que se habían proclamado campeonas en cada uno de esos juegos.

o) En el año 2018, los primeros vehículos autónomos comienzan a circular en las calles.

p) No se sabe a ciencia cierta cuándo será el año en que la IA tendrá mucho más poder computacional que el cerebro humano, en lo que se refiere al número de cálculos que pueden realizar cada segundo. Lo cierto es que, sin irnos a un futuro muy lejano, se espera que la mayoría de las empresas incorporen esta tecnología en sus procesos de transformación digital.

III. INTELIGENCIA ARTIFICIAL, *"BIG DATA", "MACHINE LEARNING"* y *"DEEP LEARNING"*

La IA ha desencadenado la cuarta revolución industrial que actualmente se encuentra viviendo el mundo. No existe certeza sobre la velocidad a la que se desarrolla la IA; sin embargo, esta se encuentra creciendo a pasos acelerados. Por ejemplo, el Instituto Global de Economía y Negocios de la firma McKinsey, concluyó que la revolución de la IA, a comparación de la revolución industrial, está sucediendo diez veces más rápido, a una escala trecientas veces más amplia y con un impacto tres mil veces más profundo.[8] Sin embargo, cabe preguntarnos qué se entiende por IA.

La IA es definida por el diccionario *Oxford English Dictionary* como el uso de computadoras para tareas que normalmente necesitan de la inteligencia humana.[9] La primera definición fue esbozada por John McCarthy, profesor de Dartmouth College, en 1956 al señalar que:

> *"(…) la ciencia y la ingeniería de crear máquinas inteligentes, especialmente programas de computación inteligentes, está relacionada con la tarea similar de utilizar ordenadores para comprender la inteligencia humana, pero la IA no se limita a métodos que sean observables biológicamente".*

Por su parte, Margaret Rouse[10] señala que es la simulación de procesos de inteligencia humana realizadas por máquinas, especialmente sistemas informáticos. Estos procesos incluyen el aprendizaje —la adquisición de información y las reglas para usar la información—, el razonamiento —el uso de las reglas para llegar a conclusiones aproximadas o definitivas— y la autocorrección. Las aplicaciones particulares de la IA incluyen sistemas expertos, reconocimiento de voz y visión artificial. En mi opinión, la IA es la capacidad de una máquina de imitar el comportamiento humano.

La consultora McKinsey la define como la capacidad de una máquina para realizar funciones cognitivas que asociamos a la mente humana, como percibir, razonar, aprender, interactuar con el entorno y resolver problemas o incluso utilizar la creatividad.[11] La idea que está detrás de este concepto es que los procesos mentales que se llevan a cabo en el cerebro de un hombre pueden ser analizados a nivel abstracto como procesos computacionales de algún tipo a fin de desarrollar métodos y algoritmos que permitan los programas informáticos de modo inteligente.

Cabe señalar que la IA se segmenta en *"Strong AI"* —Inteligencia Artificial Fuerte— y *"Weak AI"* —Inteligencia Artificial Débil—. La primera se da cuando una máquina es capaz de equiparar o superar la inteligencia de un humano y, por lo tanto, puede realizar tareas comúnmente asociadas al humano de manera excelente —pensemos en C-3PO de Star Wars—; mientras que la segunda se configura cuando una máquina solo puede recrear algunos elementos de la inteligencia humana —pensemos en Siri, el asistente personal de Apple—.

La IA es un término general o "paraguas" para una rama de la informática centrada en la creación de máquinas capaces de pensar y aprender que integra a conceptos como el *"Big Data", "Machine Learning"* y *"Deep Learning"*. En senti-

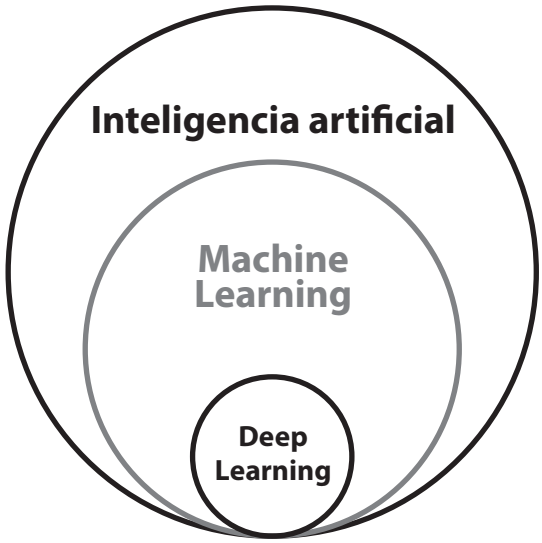
8 Camilo Narváez López, *Ibíd*., p. 211.

9. «Inteligencia Artificial», Definición 1.e., Oxford English Dictionary, acceso el 9 de noviembre del 2020, https://en.oxforddictionaries.com/ definition/artificial_intelligence.

10. «artificial intelligence», Margaret Rouse, SearchEnterpriseAI, acceso el 9 de noviembre del 2020, https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence.

11. «An Executive's Guide to AI», McKinsey & Company, acceso el 9 de noviembre del 2020, https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai.

do lato, la IA es la capacidad de los programas informáticos de replicar la inteligencia humana. Sin embargo, como se verá más adelante, esto lleva a procesos más específicos.



Tal como se puede apreciar en el diagrama, el concepto de IA engloba a los conceptos de *"Machine Learning"* y *"Deep Learning"*. En otras palabras, la IA es un concepto más amplio. En ese sentido, cuando un programa informático lleva a cabo tareas de manera "inteligente" o "racional", se considera IA. A medida que ha pasado el tiempo, el concepto de IA se ha ampliado para dar paso a ramas más especializadas y complejas.

En los años 60 y 70, los "Sistemas Expertos" revolucionaron el campo de la IA.[12] Básicamente, eran *software* que tenían por fin emular el comportamiento de un experto humano en la solución de un problema. Los Sistemas Expertos almacenan conocimientos concretos para un campo determinado y solucionan los problemas en base a reglas, mediante deducción lógica de conclusiones. Con ellos se busca una mejora en calidad y rapidez de respuestas dando así lugar a una mejora de la productividad del experto.

En otros términos, los Sistemas Expertos se basan en reglas. Estas pueden ser predefinidas por un experto a través de silogismos lógicos con el objetivo de inferir un resultado determinado —siguen la siguiente formula: "si A, entonces B"—. Estas reglas también pueden estar basadas en casos —CBR, *Case Based Reasoning*—, aplicando el razonamiento basado en casos, donde la solución a un problema similar planteado con anterioridad se adapta a un nuevo problema.

Pondremos un ejemplo bastante simple para entender la lógica detrás de los Sistemas Expertos: Messi es el mejor. Si Messi juega, entonces Barcelona FC gana. Esta sería una regla muy simple. La cuestión clave es que debe existir alguien que defina estas reglas. Si existe un cambio, el Sistema Experto debe ser cambiado, lo cual era costoso en esa época. El problema con los Sistemas Expertos era que no tenían autonomía al momento de resolver el problema, pues dependían de las reglas impuestas. En términos sencillos, el Sistema Experto era como aquel estudiante que memorizaba todo un curso y aprobaba el examen, pero realmente no entendía nada acerca de la materia.

Es por eso que nace el *"Machine Learning"* o "aprendizaje automático", que es la capacidad que tienen las máquinas de recibir un conjunto de datos y aprender por sí mismas, cambiando y ajustando los algoritmos a medida que procesan información y conocen el entorno. Comparándolo con el Sistema Experto, quien decide

12. DENDRAL es el nombre de un Sistema Experto desarrollado por Edward Feigenbaum y otros programadores en la Universidad de Stanford, a mediados de los años 60, y su desarrollo duró diez años, entre 1965 y 1975. Fue el primer Sistema Experto en ser utilizado para propósitos reales. Tuvo cierto éxito entre químicos y biólogos, ya que facilitaba enormemente la inferencia de estructuras moleculares, dominio en el que DENDRAL estaba especializado. DENDRAL tenía como finalidad solucionar problemas de ingeniería química. Esto lo hacía infiriendo cualquier posible restricción sobre la solución basándose en el conocimiento que posee en su base de datos. Este sistema permitía a los usuarios añadir cualquier otro tipo de restricción a fin de generar una lista de posibles soluciones que se imprimían en orden de preferencia.

sobre esas reglas es el mismo software. Esto quiere decir que los algoritmos del *Machine Learning* aprenden a partir de los datos a ellos sometidos y, de esa manera, los programas informáticos son entrenados para aprender a ejecutar diferentes tareas de forma autónoma. Luego, cuando son expuestas a nuevos datos, ellas se adaptan a partir de los cálculos anteriores y los patrones se moldean para ofrecer respuestas confiables. ¿Qué significa esto? En vez de programar reglas en un computador —como se hacía en el Sistema Experto— y esperar el resultado, con el *Machine Learning*, la máquina aprenderá esas reglas por cuenta propia.

El *Machine Learning* funciona de la siguiente manera: imaginemos que una compañía de telecomunicaciones quiere saber qué clientes están a punto de darse de baja de sus servicios a fin de crear campañas de marketing para retenerlos. ¿Cómo podría hacerlo? La empresa tiene muchos datos de los clientes, tales como la antigüedad, planes contratados, consumo diario, llamadas mensuales, reclamos, últimos cambios de planes contratados. En el pasado esta información, al no poder ser procesada adecuadamente, solo sería utilizada para facturar y para hacer estadísticas. Sin embargo, con el *Machine Learning*, estos datos se pueden usar para predecir cuándo un cliente se va a dar de baja y gestionar la mejor acción que lo evite. Esta rama de la IA detecta patrones de comportamiento contando y en base a ello predice futuros comportamientos. Por tanto, la empresa podrá descubrir cuáles son las causas que han llevado, en este caso, a darse de baja como cliente.

De este ejemplo se aprecia que el *Machine Learning* es una rama de la IA basada en la idea de que los sistemas informáticos pueden identificar patrones y tomar decisiones en base a conclusiones obtenidas de un conjunto de datos, sin que el ser humano tenga que escribir instrucciones o reglas para esto. Este campo ha crecido tanto en los últimos años que los principales gigantes tecnológicos ofrecen plataformas como IBM Watson Developer Cloud, Amazon Machine Learning, Azure Machine Learning, TensorFlow o BigML.

El *"Deep Learning"* o "aprendizaje profundo", por su parte, es una rama del *Machine Learning* que se ocupa de emular el enfoque de aprendizaje que los seres humanos utilizan para obtener ciertos tipos de conocimiento. Mientras que los algoritmos tradicionales de aprendizaje automático son lineales, los algoritmos de aprendizaje profundo se apilan en una jerarquía de creciente complejidad y abstracción. Los modelos computacionales de *Deep Learning* imitan las características arquitecturales del sistema nervioso, permitiendo que dentro del sistema global haya redes de unidades de proceso que se especialicen en la detección de determinadas características ocultas en los datos.

Por ejemplo, imaginemos a un niño cuya primera palabra es "perro". ¿Cómo hace el niño para diferenciar un perro de otros animales? Comúnmente, el niño señalará a aquellos animales que considere de esas características y dirá la palabra "perro".

Entonces, los padres le dirán ya sea: "sí, eso es un perro" o "no, eso no es un perro". Mientras el niño continúa apuntando a los objetos, se vuelve más consciente de las características que poseen todos los perros. Lo que el niño hace, sin saberlo, es aclarar una abstracción compleja —el concepto de perro— construyendo una jerarquía en la que cada nivel de abstracción se crea con el conocimiento que se obtuvo de la capa precedente de la jerarquía.

En ese sentido, el *Deep Learning* sigue un proceso por capas que simula el funcionamiento básico del cerebro que se realiza a través de las redes de neuronas. Las primeras capas reconocen detalles concretos, mientras que las últimas capas reconocen patrones más abstractos y generan un resultado final. Actualmente, se está utilizando para traductores inteligentes, reconocimiento de voz, interpretaciones semánticas —puede ser utilizado para interpretar contratos, por ejemplo— y reconocimiento de caras.

Luego de haber examinado los conceptos de *Machine Learning* y de *Deep Learning* podemos señalar que, al igual que la inteligencia huma-

na, la IA necesita grandes cantidades de datos, procesándolos a través de algoritmos que han sido ajustados por experiencias pasadas, y usando los patrones encontrados para mejorar la toma de decisiones. En definitiva, utiliza los datos para obtener información del entorno e interactuar con él en consecuencia. Conforme una compañía aumenta la cantidad de información, la utilización de IA se vuelve relevante para darle sentido, encontrar patrones y predecir comportamientos.

Como consecuencia de ello, la IA se "nutre" del *"Big Data"*, que es el conjunto de tecnologías que permiten tratar cantidades masivas de datos personales provenientes de distintas fuentes a través del uso de algoritmos, con el objetivo de poder otorgarles una utilidad que proporcione valor. Es decir, el *Big Data* permite procesar datos cuyo tamaño —volumen—, complejidad —variabilidad— y velocidad de crecimiento —velocidad— dificultan su captura, gestión, procesamiento o análisis por parte de tecnologías y herramientas convencionales.

Uno de los ejemplos más conocidos de uso de *Big Data* e IA se produjo en la cadena de supermercados Walmart, que recogía datos sobre las compras de sus clientes para posteriormente analizarlos y comprender mejor sus hábitos de consumo.[13] Con dicha información, Walmart comenzó a realizar predicciones sobre las ventas que obtendrían en diversos escenarios, por ejemplo, durante las alarmas por huracán. Sus análisis descubrieron que el producto más vendido antes de que ocurra este fenómeno natural era la cerveza, y que dichas alarmas disparan las ventas de los dulces "Pop Tarts" siete veces por encima del nivel normal de ventas. Con este nuevo conocimiento, el supermercado puede tomar decisiones mejor fundadas y gestionar sus inventarios o promociones de manera adecuada con el fin de incrementar las ventas.

Actualmente, Walmart cuenta con una tienda denominada Walmart Intelligent Retail Lab-IRL en Levittown, Nueva York, donde hace uso de la IA y otras tecnologías como cámaras y sensores para automatizar muchas de las tareas que desde años habían sido realizadas por sus empleados. Entre las tareas que han sido automatizadas están las de limpieza, revisión de anaqueles, clasificación de productos, etc.

El *Big Data* hace que la IA deje de ser ciencia ficción, pues esta forma de procesar datos masivamente está logrando que la IA cada vez se parezca más a cómo piensa el cerebro humano. El crecimiento del *Big Data* y la velocidad de procesamiento de los datos conlleva a un crecimiento exponencial de la IA. Como hemos podido analizar, el combustible de la IA son los datos. A medida que haya más información, mayor será el desarrollo de la IA.

IV. RESPONSABILIDAD CIVIL: DAÑOS DERIVADOS DEL USO DE LA INTELIGENCIA ARTIFICIAL

La finalidad de la Responsabilidad Civil es la restitución de la situación de la víctima al estado anterior al hecho dañoso. Como esto, en la mayoría de los casos, es físicamente imposible, se pretende resarcir el daño económicamente. En otras palabras, cuando una persona ha sufrido un daño sin justificación, el ordenamiento jurídico traslada la carga económica del damnificado a otro individuo. Es decir, se busca que la víctima sea indemnizada. Para ello es necesario determinar quién va a responder por la conducta dañosa o, lo que es lo mismo, quién es el que va a asumir el costo económico de dicho acto.

En el Perú, el régimen de responsabilidad es subjetivo, en el que impera el concepto de "culpa" y, excepcionalmente, se aplica un régimen objetivo, en materia extracontractual respecto de actividades riesgosas y en materia contrac-

13. Elena Gil González, «¿Qué es Big Data y por qué debe interesarme si soy abogado?», *Legal Today*, 18 de octubre de 2016, acceso el 09 de noviembre del 2020, http://www.legaltoday.com/blogs/nuevas-tecnologias/blog-ecija-2-0/que-es-el-big-data-y-por-que-debe-interesarme-si-soy-abogado.

tual respecto de las obligaciones de resultado. Sin embargo, parafraseando a Fernando de Trazegnies, los nuevos riesgos que traen el uso de las nuevas tecnologías, el incremento de las velocidades tiene una naturaleza elusiva debido a su complejidad, pues no permiten identificar con facilidad al culpable, si es que existiese alguno.

Bajo este contexto, ¿quién debe responder ante un daño ocasionado por sistemas controlados por IA? Imaginemos que un anciano adquiere a un "robot niñera", con el objetivo de que lo cuide y le haga compañía —ejemplos de esta tecnología la vemos en el producto iPAL—.[14] Luego de algunos años, el anciano invita a sus amigos a ver un partido de fútbol. Debido a que el equipo del anciano está perdiendo, se comienza a pelear de forma verbal con unos de sus amigos. El robot, en un acto por defender a su dueño, empuja a uno de sus amigos, quien se cae de las escaleras, muriendo desangrado. Bajo el ordenamiento jurídico peruano, ¿quién debe responder? ¿el propietario? ¿el concesionario o el fabricante? ¿deben responder solidariamente? ¿se aplica un régimen subjetivo —artículo 1969 del Código Civil— u objetivo —artículo 1970 del Código Civil—?

Si bien es cierto, la idea de máquinas o robots que poseen capacidades autónomas e inteligencia y que no están gobernados por las direcciones o supervisión humanas se remontan a varias décadas atrás, cabe preguntarse si nuestro sistema jurídico está adecuado para la incorporación de estas ideas —en ese entonces— futuristas a la realidad tecnológica y económica actual. ¿Cómo se debe construir la teoría de la responsabilidad civil cuando no existe la participación humana en la decisión realizada por una máquina? ¿Cómo se debe de aplicar la ley cuando las acciones son, en muchos casos, imprevisibles? ¿Nuestro modelo de responsabilidad actual se ajusta a esta nueva realidad?

Lo primero que se debería hacer es analizar los siguientes presupuestos para determinar si hay responsabilidad civil o no.

1. Antijuridicidad.

Para que exista responsabilidad civil debe mediar un hecho o una conducta dañosa, positiva o negativa, que no sea conforme a Derecho, más allá que se afecte o no a una norma positiva. En otras palabras, debe ser una conducta ajena a la víctima y no interesa si el daño se desprende de una acción o de una omisión que manifiesta una disconformidad con los principios que rigen el ordenamiento jurídico. Entonces, existirá antijuridicidad cuando se afecte una esfera jurídica ajena —a la víctima— y la conducta carezca de una causa jurídica que la justifica. Es decir, cuando se viola el deber genérico de no causar daño.

El artículo 1971 del Código Civil establece como causas de exoneración de responsabilidad civil al ejercicio regular de un derecho[15], la legítima

14. *iPal* es un robot niñera diseñado por la compañía AvatarMind para encargarse del cuidado de tus hijos. El androide ha sido especialmente diseñado para interactuar con los niños de edades comprendidas entre los tres y los ocho años. Este robot humanoide tiene la capacidad de hablar para relacionarse con el menor y está equipado con 25 motores, que le permiten realizar un amplio abanico de movimientos parecidos a los que hace una persona para resultar lo más natural posible. Además, para comprender mejor lo que el niño dice, cuenta con la última tecnología de comprensión del lenguaje natural, comprende perfectamente diversos sentimientos y emociones, y dispone de aprendizaje automático para recordar los gustos e intereses de los niños.

15. Aquello que no constituye el ejercicio regular de un derecho es un ejercicio irregular, que configura un acto ilícito y, consiguientemente, da lugar a una responsabilidad por dolo o culpa según el grado de ilicitud o irregularidad involucrado. Cfr. Fernando de Trazegnies, *La Responsabilidad Extracontractual*, Tomo I, 5° ed., (Bogotá: Editorial Temis S.A. Colombia, 1999), 118.

defensa[16] y el estado de necesidad[17]. ¿El accidente ocurrido cuenta con alguna de estas indemnidades? En el presente caso, no se podrá señalar que se estaba ante el ejercicio regular de un derecho, pues el derecho a la propiedad —en este caso, el derecho a tener un robot— no puede ir en contra del derecho a la vida. Tampoco nos encontramos en un estado de necesidad. Sin embargo, ¿se podría decir que esta era una legítima defensa? No, porque no existía un peligro actual, ni el recurso de defensa era necesario para este caso. El componente particular es que, quien decide defender al dueño es el mismo sistema de IA, por lo tanto, la conducta no fue originada por un humano. Fue una decisión autónoma por parte de una máquina.

2. Daño.

Para que exista responsabilidad no sólo debe existir una conducta antijurídica, sino que también debe mediar un daño. Solo se puede indemnizar a quien ha sufrido un daño. El término daño, en sentido amplio, es cualquier lesión o detrimento que sufre una persona, sea autogenerado, provocado por otra persona e inclusive acontecido sin intervención alguna del hombre. En cambio, daño, en sentido jurídico, es cualquier lesión o agravio a un interés jurídicamente reconocido, como consecuencia de una conducta antijurídica.[18]

Para ello, deben concurrir los siguientes requisitos:

a) Afectación a un interés legítimo: el daño debe representar la afectación a un interés legítimo, patrimonial o no, que corresponde a un derecho subjetivo propio. El damnificado debe ser la parte material del reclamo. El damnificado directo es la víctima inmediata con relación al hecho dañoso. El damnificado indirecto es la víctima mediata, a quien también se afecta un interés propio como consecuencia de la ocurrencia dañosa.

b) Certidumbre: el daño debe ser cierto, determinado o determinable, al margen que sea actual o futuro. El perjuicio incierto es jurídicamente inexistente y, por lo tanto, no es indemnizable.

c) Subsistencia: la subsistencia del daño se asocia al momento de su reclamación. Esto significa que la víctima no ha sido reparada respecto del daño sufrido. El daño puede ser subsistente fácticamente, pero puede

16. La legítima defensa es aquella defensa necesaria frente a una agresión ilegítima no provocada. Esta puede aplicarse para evitar un daño sobre los bienes jurídicos de la misma persona quien realiza la defensa —legítima defensa propia—, como para defender bienes jurídicos de terceras personas —legítima defensa impropia—. El peligro debe ser actual y debe amenazar un derecho tutelado por el ordenamiento jurídico. Además, el recurso de defensa debe ser necesario e inevitable.

17. Se suele definir al estado de necesidad como el sacrificio de un bien jurídicamente de inferior jerarquía en favor de un bien jurídicamente de superior jerarquía, frente a un estado de peligro inminente. *Cfr.* Juan Espinoza Espinoza, *Derecho de la Responsabilidad Civil*, 6° ed., (Lima: Editorial Rodhas, 2011), 136.

18. Cfr. *Ibíd*., 246. El profesor Juan Espinoza sostiene que el daño no puede ser entendido sólo como la lesión de un interés protegido por cuanto ello resulta equívoco y sustancialmente impreciso: el daño incide más bien en las consecuencias, aquellos efectos —negativos— que derivan de la lesión del interés protegido. En sustancia, interés lesionado y consecuencias negativas de la lesión son momentos vinculados entre sí, pero "autónomos conceptualmente, cuanto al contenido y a la naturaleza". Es por ello que de una lesión patrimonial pueden resultar consecuencias —al lado de aquellas patrimoniales— no patrimoniales y viceversa. Así tenemos que se habla de un daño-evento —lesión del interés tutelado— y un daño consecuencia —daño emergente, lucro cesante y daño moral—. Estas dos acepciones de daño pueden, como no, coincidir. Sin embargo, confundir estos conceptos diversos de daño equivale a mezclar problemas jurídicos diversos: el problema de la injusticia de la lesión, aquel de la individualización del responsable o el de la selección de los perjuicios resarcibles.

invocarse la prescripción para extinguir jurídicamente la acción de reclamo.

En el presente caso, el "daño-evento" será la muerte del amigo —lesión al derecho a la vida—, mientras que el daño consecuencia será el daño emergente, el lucro cesante y el daño moral de sus familiares. Este daño es certero, pues es actual y se ha evidenciado de forma tangible, por lo tanto, es indemnizable. Se cuenta con un plazo de 10 años para poder solicitar la indemnización por daños y perjuicios.

3. Relación de causalidad.

Para que exista responsabilidad no solo debe mediar una conducta antijurídica, esto es, contraria a Derecho, sino que además esa conducta ilícita debe ser la productora del daño reclamado. Es el tercer elemento porque se tiene que entrelazar entre la conducta antijurídica y el daño. Debe tenerse en cuenta que tanto el artículo 1969 como en el 1970 del Código Civil se refieren a quien causa el daño; ello quiere decir que tanto en la responsabilidad subjetiva como en la objetiva está presente este elemento.

En ese sentido, para acreditar si existe responsabilidad civil, además de analizar la conducta del causante y el resultado —el evento dañoso—, será indispensable establecer si existe o no un nexo causal que permita explicar la razón por la cual se le debe atribuir responsabilidad al autor en cuestión. Es decir, se debe probar la existencia de un hecho generador del daño, cierto y directo que se presente, en la medida de lo humanamente comprensible, como una causa unívoca.

A fin de acreditar la causalidad se ha optado por la teoría de la causa adecuada, pues en el artículo 1985 del Código Civil se establece justamente que para que un daño sea indemnizable debe establecer una relación de causalidad adecuada con la acción dañosa. La causa adecuada se distingue de la causa natural, pues ésta se define sobre la base de una relación "causa-efecto" —causa *sine qua non*—. El problema es que muchas causas naturales de una consecuencia no son relevantes como para hacer responsable a una persona —"si no se hubiese dado el boom de la IA, el robot no hubiese matado al amigo" o "si el anciano no se hubiese ganado la Tinka, no se hubiese comprado el robot y, por tanto, no hubiese muerto el amigo"—. Esto conduce a que la causa natural no sea suficiente para que el sistema de responsabilidad civil cumpla sus funciones. En cambio, la teoría de la causa adecuada busca entre todas las condiciones aquella —o aquellas— que ha —o han— influido de manera decisiva en la producción del evento dañino.[19] La causalidad adecuada se relaciona directamente con la predictibilidad del daño. Se trata de la aptitud de la conducta antijurídica para propiciar el daño.

Sentado este criterio, corresponde verificar si existe un supuesto de ruptura del nexo causal o "fractura causal". Estos supuestos son causas extrañas o ajenas —hecho fortuito, fuerza mayor, hecho de tercero y hecho de la víctima— que desvirtúan o excluyen la presunta responsabilidad de un sujeto por la generación de daños.[20]

Como bien señala Bullard, la causalidad se relaciona directamente con la capacidad del actor de identificar, al momento de llevar a cabo su conducta, cuáles pueden ser las posibles consecuencias. De no ser así, y uno respondiera incluso por las consecuencias que no se pueden prever, se desincentivaría incluso el desarrollo de muchas actividades deseables para la sociedad. En ese sentido, cuando el contexto varía al existir determinadas condiciones o circunstancias, el actor no debería responder por aquellos hechos que no pudieron ser previstos por el agente o en el extremo, aun siendo prevista, que dicha causa hubiere sido irresistible y que no sean producto del comportamiento del agente.

19. *Ibíd*., 209.

20. *Ibíd*., 228.

Entonces, ¿se podría señalar que la decisión del robot de matar al amigo fue un caso fortuito o de fuerza mayor? En primer lugar, debemos señalar que tal como se desprende del Código Civil, nuestro legislador ha optado por equiparar los conceptos de "caso fortuito" y "fuerza mayor", definiéndolos a ambos como aquella causa no imputable que impide el cumplimiento de una obligación, y que además debe tener tres características como mínimo, esto es, debe tratarse de un acontecimiento "extraordinario, imprevisible e irresistible". Esta norma no define estos conceptos lo cual genera un problema que radica en la plasticidad de los conceptos "extraordinario", "imprevisible" e "irresistible", pues no son precisos ni absolutos. El Código Civil no explica en qué consisten dichas características, así como en qué circunstancias y respecto de quienes resultan aplicables las mismas, a efectos de considerar la ocurrencia de un supuesto de caso fortuito o fuerza mayor. Esto último, resulta de suma importancia en tanto, el caso fortuito y la fuerza mayor juegan un rol preponderante al momento de analizar el nexo causal como uno de los elementos constitutivos de la responsabilidad civil.

¿Este evento se podría considerar extraordinario? El diccionario de la Real Academia Española define al término "extraordinario" como "fuera del orden o regla natural".[21] En consecuencia, un acontecimiento extraordinario es aquel evento que no es usual y por ende escapa a lo ordinario, razón por la cual se le califica como una situación de excepción. En efecto, Osterling y Castillo señalan que:

> *"(…) lo extraordinario reviste la característica de anormal, es decir, las circunstancias en que se presentan deben ser extraordinarias y no ordinarias. Lo contrario a lo común es la excepción; por ello, concluimos que se trata de algo que se encuentra dentro del campo de lo excepcional, de un acontecimiento que se produce por excepción, lejos de lo que en forma normal o natural se espera que ocurra. Lo extraordinario es, pues, lo que atenta o irrumpe en el curso natural y normal de los acontecimientos, quebrándolos."*[22]

A criterio de Fernando de Trazegnies, el análisis de este atributo debe ser realizado *in abstracto* y no *in concreto*; esto es, atendiendo a lo que hubiese resultado excepcional para cualquier persona que hubiese estado en esa misma situación. Ahora bien, para saber si un acontecimiento califica como extraordinario, resulta necesario que el análisis se realice dentro de un determinado contexto, pues lo que en un lugar, época o circunstancia podría ser calificado como excepcional, podría no serlo en otro contexto. Si se normaliza el uso de robots niñera y éste es un caso aislado, se podría decir que lo ocurrido no es ordinario. Si de 1 millón de robots niñera, este es el único que ha reaccionado así, se podría decir que este configura un evento extraordinario.

¿Este evento se podría considerar imprevisible? Lo imprevisible es aquello *"que no se puede prever"*, esto es, que no puede ser conocido, ni respecto de lo cual se puede conjeturar, por no existir señales o indicios de lo que va a suceder. Osterling y Castillo sostienen que:

> *"(…) el evento no sólo debe revestir la objetividad en sí mismo como hecho extraordinario, lo cual se demuestra sin mayores problemas al analizar la frecuencia o habitualidad del suceso, sino que además se requiere del elemento inherente al individuo, relativo a la conducta diligente que se espera de él."*[23]

---

21. «extraordinario», Diccionario de la Real Academia Española, acceso el 09 de noviembre del 2020, https://dle.rae.es/extraordinario.

22. Osterling Parodi, Felipe y Mario Castillo Freyre, *Tratado de las Obligaciones,* Tomo XI (Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú, 2003), 624-625.

23. *Ibíd*., 630-631.

En otras palabras, un hecho es imprevisible cuando supera la aptitud normal del deber de previsión del agente. En consecuencia, puede decirse que la imprevisibilidad va de la mano con los deberes de diligencia, prudencia, esmero o cuidado del agente que ocasionó el daño. Adicionalmente, se deberá tomar en cuenta que un hecho se considerará imprevisible cuando el agente no haya tenido motivos atendibles para presumir que éste vaya a suceder.

¿Es previsible que un sistema de IA falle? En términos amplios sí. Estadísticamente, en algún momento un robot fallará. Sin embargo, ¿quién es el que debía prever esto? ¿el propietario del robot o el fabricante? Dada la naturaleza del *Machine Learning*, puede que un determinado *output* que terminó como accidente no haya podido ser previsible para alguien. En ese sentido, se podría decir que un accidente, de estas características, era imprevisible, tanto para el propietario como para el fabricante, pues no se trata de determinar si el agente podía prever el acontecimiento, sino de saber si un individuo cuidadoso, diligente, colocado en las mismas circunstancias, lo hubiera podido prever o impedir. Si el anciano le hacía un mantenimiento y seguía todas las normas de uso, y luego de algunos años el robot se comporta de esta manera, ¿cómo podría haberlo previsto? Por otro lado, si el fabricante ha desarrollado un sistema de IA en el que todos sus robots niñera funcionan a la perfección, ¿cómo podría haber previsto que éste iba a reaccionar así?

¿El suceso se podría considerar irresistible? Un evento califica de irresistible en la medida que no permite ser tolerado, rechazado ni contrarrestado con alguna acción. Este requisito implica que el evento es inevitable, lo cual supone la imposibilidad de poder cumplir. Por ello, Osterling y Castillo sostienen que *"el que un evento sea irresistible quiere decir que la persona es impotente para evitarlo; no puede impedir, por más que quiera o haga, su acaecimiento."*[24] A su vez, Fernández y León señalan que:

> *"(…) la causa no imputable cuyo acaecimiento extingue la obligación y libera de responsabilidad, debe ser tal, que contra ella no se pueda hacer nada, de manera que impida al deudor proceder de una forma que no resulte dañosa para el acreedor. El caso fortuito o fuerza mayor es un obstáculo que no puede ser evitado por ni un medio"*[25].

En efecto, la irresistibilidad supone que el presunto causante del daño no hubiera tenido la oportunidad de actuar de otra manera. En otras palabras, no basta con que la adopción de otra acción haya sido difícil, sino que se requiere que haya sido imposible. Para el caso del propietario era imposible, pues él no fue quién decidió empujar al amigo. Para el caso del fabricante, se podría argumentar que este suceso era irresistible, pues debido al *Machine Learning* era imposible descifrar porque el sistema de IA actuó como actuó. Por consiguiente, se podría llegar a la conclusión de que este hecho era irresistible.

Luego de este análisis, es posible señalar que el presente suceso puede ser visto como un supuesto de ruptura del nexo causal, pues se puede argumentar que este suceso es extraordinario, imprevisible e irresistible, dada la naturaleza del *Machine Learning*. Sin embargo, el problema radica en que el sistema jurídico peruano está diseñado en torno a un sistema subjetivo, en donde la culpa juega un rol esencial. Si bien el artículo 1972 del Código Civil hace referencia expresa al artículo 1970 del mismo texto legal, en realidad también engloba al artículo 1969, pues lo que hacen estas "fracturas causales" realmente es determinar una situación de ausencia de culpa por parte del causante aparente. Como bien señala Fernando de Trazegnies[26], si pensamos que no solamente

24. *Ibíd*., 630-631.

25. Gastón Fernández y Leysser León, *Código Civil Comentado por 209 especialistas en diversas materias del derecho civil*, Tomo VI, (Lima: Gaceta Jurídica, 2007), 888-889.

26. Fernando de Trazegnies, *Op. Cit.*, 203.

todo caso fortuito significa que no hay culpa, sino que además cada vez que no encontramos culpa de alguien estamos ante un caso fortuito, entonces la objetividad se convierte en una mera ilusión. De esta manera, siempre se regresa al campo de la culpa por la puerta falsa, pues al parecer sólo habrá responsabilidad cuando hay culpa; y la teoría del riesgo y demás concepciones objetivistas quedan así reducidas a un mero eufemismo jurídico.

Por otro lado, ¿se podría aducir que fue hecho de un tercero? Para que represente una fractura causal la intervención del tercero debe ser concluyente en la realización del daño. En ese sentido, se tendría que probar que un tercero *"hackeó"* el sistema de IA y modificó los algoritmos a fin de que se produzca dicho error. En ese supuesto, no habría responsabilidad civil por parte del propietario o del fabricante. ¿Se podría señalar que el hecho fue propiciado por la propia víctima? Si bien la víctima discutió con el dueño del robot, esto de ninguna manera puede ser considerado como un hecho justificante.

Por consiguiente, se puede apreciar que sí es factible quebrar el nexo causal, aduciendo que el suceso fue un caso fortuito o de fuerza mayor. Inclusive se puede probar que el hecho tuvo origen en el hecho de un tercero. En el presente caso, no se puede señalar que el incidente fue propiciado por la propia víctima. Cabe señalar que, en el caso de la fuerza mayor, dependiendo de las circunstancias específicas del caso, uno podría llegar a la conclusión de que el hecho no era extraordinario, imprevisible ni irresistible. Esto, sin duda alguna, resulta complejo e inclusive "diabólico", pues todo recae en aspectos probatorios de una tecnología nueva y compleja.

4. Factor de atribución.

Parafraseando al profesor Marco Ortega Piana, para concluir en la existencia de responsabilidad civil, el daño no sólo debe provenir de determinada conducta antijurídica, sino que también debe existir una justificación jurídica para considerar como responsable al causante: calificación de su actuación.

El factor de atribución contesta la siguiente pregunta: ¿a título de qué se es responsable?, vale decir, constituye el elemento del deber de indemnizar. Existen factores de atribución subjetivos —culpa y dolo—, objetivos —realizar actividades o ser titular de determinadas situaciones jurídicas que el ordenamiento jurídico considera, si se quiere ser redundante, objetivamente; o, si se quiere optar por una definición residual, prescindiendo del criterio de la culpa—.[27]

La responsabilidad subjetiva está contenida en el artículo 1969 del Código Civil que señala que aquel que por dolo o culpa causa un daño a otro está obligado a indemnizarlo. El descargo por falta de dolo o culpa corresponde a su autor.

La culpa consiste necesariamente en un error de conducta y, por ende, es siempre necesario un obrar humano, que atribuye un hecho ilícito a su autor, con lo que normalmente se reconoce en la culpa, de un lado, un componente objetivo, que no es sino la herencia romana de la *lex Aquilia* y del elemento *"injuria"*, cuyo efecto central sería el de determinar la unión indisoluble de las nociones de "culpa" y "hecho ilícito"; y, de otro, un componente "subjetivo" que es el elemento psicológico que distingue el acto ilícito de la simple violación del derecho ajeno y, como tal, expresa un estado particular del ánimo en relación con un hecho injurioso.[28] La culpa se traduce en la siguiente afirmación: aquel sujeto que, de manera dolosa o culposa, cause un daño a otro, está obligado a reparar el daño producido.

27. Juan Espinoza Espinoza, *Op. Cit.*, 228. Juan Espinoza Espinoza, *Op. Cit.*, 228.

28. Fernández, Gastón y Leysser León, *Op. Cit.*, 24.

El problema con la responsabilidad civil derivada de utilizar un sistema de IA, como el presente caso, está en que tanto los creadores, desarrolladores, distribuidores y propietarios podrían exonerarse de responsabilidad al probar la ausencia de culpa respecto de las acciones u omisiones que ejecute este tipo de IA, que no pueden llegar a ser comprendidos. Es impensable que a un propietario de un robot niñera se le exija probar por qué un sistema de IA de esta naturaleza actuó de manera negligente cuando los mismos desarrolladores, en muchas ocasiones, no son capaces de explicar el porqué de las decisiones automatizadas.

Ante esta situación, cabría aplicar la responsabilidad civil objetiva contenida en el artículo 1970, que establece que aquel que mediante un bien riesgoso o peligroso, o por el ejercicio de una actividad riesgosa o peligrosa, causa un daño a otro, está obligado a repararlo.

Hay quienes sostienen que, si la generación de peligro deriva de actividades humanas, interesará siempre analizar los supuestos de falta de prevención, por lo que un "comportamiento negligente" puede estar presente tanto en supuestos de responsabilidad por culpa, como en los de responsabilidad por riesgo.[29] ¿Qué sucede cuando la generación del peligro deriva de actividades realizadas por la IA?

El concepto de negligencia es insuficiente porque el deber de cuidado y los estándares que se deben adoptar para prevenir razonablemente el peligro depende de una tecnología que cambia constantemente, y se ven afectados por daños inesperados y, por tanto, una falta general de previsibilidad. El *Machine Learning* y el *Deep Learning* conducen a escenarios en los que los sistemas de IA alcanzan un objetivo predeterminado, pero sus programadores pueden no tener una comprensión exacta de cómo alcanzó ese objetivo o cuáles son las etapas que llevó a cabo la IA a fin de alcanzar un determinado *output*.

En este caso, ¿se podría considerar al robot niñera como un bien riesgoso? ¿se tendrán que considerar a los sistemas de IA como bienes riesgosos? Asumamos que en el futuro así se determina, en base al artículo 1970 nos tendríamos que preguntar, ¿quién utilizó el bien riesgoso? El anciano, propietario del robot niñera. Sin embargo, como hemos visto en el numeral anterior, podría romper el nexo causal, señalando que fue un hecho fortuito o argumentando que el producto fue defectuoso, de conformidad con el artículo 104 del Código de Protección y Defensa del Consumidor[30]. Sin embargo, ¿realmente el producto carecía de idoneidad o calidad? ¿cabría aplicar este artículo?

A medida que siga esta innovación tecnológica, lo único cierto es que en la mayoría de los supuestos de hecho generadores de daños que se presentarán en la realidad, resultará cada vez más difícil identificar al autor del hecho dañoso y en base a un sistema que gira en torno del concepto de "culpa", se dejará en un estado de indefensión a la víctima del daño. El ejemplo propuesto, es uno de los tantos que ocurrirán en el futuro: ¿qué ocurriría si un robot realiza una mala *praxis* médica? ¿qué sucedería si un

29. Fernández, Gastón y Leysser León, *Op. Cit*., 116.

30. Artículo 104 de la Ley 29571:
"Responsabilidad administrativa del proveedor.-
El proveedor es administrativamente responsable por la falta de idoneidad o calidad, el riesgo injustificado o la omisión o defecto de información, o cualquier otra infracción a lo establecido en el presente Código y demás normas complementarias de protección al consumidor, sobre un producto o servicio determinado.
El proveedor es exonerado de responsabilidad administrativa si logra acreditar la existencia de una causa objetiva, justificada y no previsible que configure ruptura del nexo causal por caso fortuito o fuerza mayor, de hecho determinante de un tercero o de la imprudencia del propio consumidor afectado.
En la prestación de servicios, la autoridad administrativa considera, para analizar la idoneidad del servicio, si la prestación asumida por el proveedor es de medios o de resultado, conforme al artículo 18."

dron manejado por un sistema de IA mata a un niño en la calle? ¿quién sería responsable si varios sistemas de IA conspiran contra las redes sociales de una persona y develan información privada? ¿qué pasaría si un "robot juez" decide mal en un juicio?

El problema radica en que nuestro ordenamiento jurídico se encuentra diseñado bajo la premisa que toda decisión es tomada por seres humanos. Es por esta razón que es difícil separar la responsabilidad objetiva de la noción de "culpa", pues más allá de que se generen situaciones de riesgo o peligro, de conformidad con el artículo 1970 del Código Civil, siempre resulta exigible el deber social de diligencia. Es por ello que se debe de reformular la responsabilidad objetiva tal como está diseñada en nuestro Código Civil, debiendo ser sólo para aquellos casos en los cuales ni siquiera el caso fortuito o la fuerza mayor sean admisibles para librar al imputado del resarcimiento que se le impone, pues a diferencia de los casos en la actualidad, la naturaleza propia de la IA, hace que los hechos dañosos deriven de una voluntad no humana.

Entonces, el uso de la IA se podría enmarcar, tomando en consideración el cambio de noción de responsabilidad objetiva —no debiendo admitirse supuestos en que se pueda romper el nexo causal—, en los siguientes supuestos[31]:

a) Situación de riesgo, que se podrían traducir en la siguiente formula: si se genera una situación riesgosa, se responderá por los daños ocasionados, independientemente del parámetro de conducta del agente "dañante" o de aquel que obtenga un beneficio de dicha actividad.

b) Situación de ventaja, vale decir, si una persona genera una situación que le ofrece un resultado favorable o beneficio, tendrá que responder también por los daños que se ocasionen producto de dicha situación.

c) Situaciones legales individualizadas por el ordenamiento jurídico, como lo que ocurre con la de ser representante legal —artículos 1975 y 1976 del Código Civil—.

La responsabilidad civil sólo tiene sentido si existe víctima. A quien el Derecho debe proteger en términos amplios es al damnificado, lo que significa que el centro en torno al cual gira esta área del Derecho no es respecto al culpable, sino que la reparación es lo que constituye el objetivo fundamental de todo el sistema. Por consiguiente, la legislación debe estar desarrollada en ese sentido.

Entonces, ¿quién debe asumir la responsabilidad de un daño causado por un sistema de IA? La parte que está en mejores condiciones para diluir el costo. En consecuencia, creemos que un criterio a considerar es el del *cheapest cost avoider*, el cual consiste en atribuir la responsabilidad a quien se encuentra en una mejor posición para asumir los costos que eviten los daños. El agente —actividad o sujeto— capaz de evitar el coste de la forma más fácil o económica, es aquel por el cual, responde del daño quien pueda reducir los costos que ocasionan de la forma más económica posible —a largo plazo— estableciendo los cambios apropiados, y al mismo tiempo evitar los costes de transacción innecesarios. Se trata de una suerte de política, en la cual, los operadores jurídicos —jueces o árbitros principalmente— hacen asumir las consecuencias económicas de los daños a quienes les va a resultar más fácil —*easiest*— o barato —*cheapest*— enfrentarlas: no por el hecho exclusivo que sean económicamente más fuertes —*deep pocket*— o que, basados en el principio de responsabilidad de la empresa, puedan fraccionar los daños de los siniestros, sea recurriendo al seguro privado o porque se hallan en condiciones de transferir los daños a los adquirentes de sus productos o a los factores empleados en la producción de los mismos.[32]

31. Juan Espinoza Espinoza, *Op. Cit.*, 171.

32. *Ibíd.*, 177-178.

De Trazegnies[33] afirma que el fabricante es la parte que está en mejores condiciones de diluir ese costo en el todo social a través de un incremento de los precios. En otras palabras, el fabricante puede calcular el riesgo estadístico de que ciertos productos salgan defectuosos u ocasionen un daño a pesar de todo el control aplicado e incluir ese costo probable en el precio del artículo, así que todo el que compra esté de alguna manera garantizando la posibilidad de una reparación para aquel miembro de la comunidad de usuarios que tiene la mala suerte de que le toque un producto defectuoso que escapó de todo control.

Finalmente, hay quienes sostienen que, debido a la naturaleza de la IA, cuando esta tecnología alcance una autonomía muy grande, como C-3PO de Star Wars, se deberá crear un estatus jurídico especial para la IA que le permita ser un objeto de imputaciones de responsabilidad. El 16 de febrero de 2017, se planteó ante el Parlamento Europeo la propuesta de un texto legislativo especializado para regular la IA y a los robots. En dicho documento se encuentra un párrafo sobre posibles soluciones jurídicas, se menciona una expresión tan interesante como controvertida: la de "personalidad electrónica". El texto señala lo siguiente:

> *"Crear a largo plazo una personalidad jurídica específica para los robots, de forma que como mínimo los robots autónomos más complejos puedan ser considerados personas electrónicas responsables de reparar los daños que puedan causar, y posiblemente aplicar la personalidad electrónica a aquellos supuestos en los que los robots tomen decisiones autónomas inteligentes o interactúen con terceros de forma independiente".*[34]

En ese sentido, se podría gravar con impuestos a las "personas electrónicas" a fin de constituir un fondo común que sirva para responder en aquellos casos en que esta actividad económica ocasione algún daño. De esta manera, se estaría protegiendo a la víctima, privilegiando la función resarcitoria de la responsabilidad civil difundiendo el riesgo.

V. PROTECCIÓN DE DATOS PERSONALES E INTELIGENCIA ARTIFICIAL

La protección de los datos personales se ve desafiada por el rápido desarrollo y el veloz despliegue de la IA, pues su utilización implica necesariamente el tratamiento de datos masivos, dentro de los cuales se incluyen diferentes categorías de datos personales.[35] Como bien señala la Agencia de Protección de Datos Personales de Noruega, la mayoría de las aplicaciones de IA requieren grandes volúmenes de datos para aprender y tomar decisiones inteligentes.[36] En ese sentido, los datos son necesarios no solo para que la IA alcance su máximo potencial, sino también para que esta pueda evitar sesgos o errores al momento de realizar un tratamiento.

Cabe recordar que la legislación peruana toma como referencia normativa a la Ley Orgánica de Protección de Datos Personales española de 1999, por lo que fue diseñada en un contexto histórico en el que la cantidad de datos era limitada, se podía tener control sobre quiénes

33. Fernando de Trazegnies, *Op. Cit*., 23.

34. Parlamento Europeo, Normas de Derecho Civil sobre Robótica. P8_TA (2017) 0051. Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho Civil sobre robótica (2015/2103(INL)).

35. Andrea Martínez Devia, «La Inteligencia Artificial, el Big Data y la Era Digital: ¿Una amenaza para los datos personales?», en *Revista de la Propiedad Inmaterial* 27 (enero-junio 2019): 7.

36. «Artificial Intelligence and Privacy», Datatilsynet (Norwegian Data Protection Authority), acceso el 9 de noviembre del 2020, https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf.

hacían su tratamiento y las finalidades para las cuales estaban siendo usados. El rápido avance de la tecnología y las herramientas de IA han traído cambios que posibilitan el procesamiento de millones de datos en diferentes partes del mundo y por diferentes actores a una velocidad inimaginable, estos cambios han provocado que las regulaciones se encuentren desactualizadas frente a estos nuevos retos.[37]

Ahora bien, cabe preguntarse cuáles son los escenarios en los que la IA pueda generar un riesgo al titular de datos personales. A fin de responder esto, debemos señalar que existen dos aspectos de la IA que son particularmente relevantes para la privacidad. El primero es que la IA en sí misma puede tomar decisiones automatizadas y, el segundo, es que el sistema se desarrolla aprendiendo de la experiencia e información proporcionada. Entonces, cabe preguntarse: ¿es posible combinar el desarrollo de la IA con un adecuado tratamiento de los datos personales?

Al respecto, el artículo 28 de la Ley de Protección de Datos Personales establece las obligaciones de los responsables y encargados del tratamiento. Por consiguiente, la normativa referida a la protección de datos personales se aplicará cuando se esté desarrollando IA con información que contenga datos personales. Esta disposición también se aplicará cuando se utilice la IA para el análisis de perfiles y la toma de decisiones sobre individuos.

Es imperativo, por tanto, que el responsable del tratamiento se asegure de cumplir los principios establecidos en la Ley de Protección de Datos Personales:

a) Legalidad: el tratamiento de los datos personales se hace conforme a lo establecido en la Ley de Protección de Datos Personales. En ese sentido, el tratamiento de datos personales debe realizarse con pleno respeto de los derechos fundamentales de sus titulares. Se prohíbe la recopilación de los datos personales por medios ilícitos o fraudulentos.

b) Consentimiento: para realizar el tratamiento de los datos personales se debe contar con el consentimiento o la autorización de la persona titular de los datos personales.

c) Finalidad: los datos personales no deben ser tratados para una finalidad distinta a la establecida al momento de su recopilación.

d) Proporcionalidad: todo tratamiento de datos personales debe ser apropiado a la finalidad para la que éstos hubiesen sido recopilados, usando la información que sea imprescindible y suficiente, sin excesos.

e) Calidad: los datos personales que se tratan deben ser veraces, exactos y adecuados. Deben conservarse de forma tal que se garantice su seguridad y solo por el tiempo necesario para cumplir con la finalidad del tratamiento.

f) Seguridad: el titular del banco de datos personales y el encargado del tratamiento deben adoptar las medidas necesarias para garantizar la seguridad y confidencialidad de los datos personales que administran.

g) Nivel de protección adecuado: para el flujo transfronterizo de datos personales, se debe garantizar un nivel suficiente de protección para los datos personales que se vayan a tratar o, por lo menos, equiparable a lo previsto por esta Ley o por los estándares internacionales en la materia.

1. Sesgo algorítmico vs. el principio de legalidad.

Un primer problema se presenta con el denominado "sesgo algorítmico" —*"Machine Bias"*— y el principio de legalidad. Este principio requie-

37. Andrea Martínez Devia, *Op. Cit.*, 9.

re que el responsable del tratamiento de datos personales implemente medidas que prevengan tratos que puedan afectar sus derechos fundamentales. Cabe preguntarse lo siguiente: ¿qué sucedería si una IA discrimina a un usuario en una decisión automatizada? Si bien los algoritmos son fórmulas matemáticas que, en principio, son neutrales y objetivas, lo cierto es que existen casos en que repiten prejuicios tan humanos como la tendencia a discriminar a partir del género y la raza.

El sesgo algorítmico ocurre cuando un sistema informático refleja los valores de los humanos que estuvieron implicados en su codificación y en la recolección de los datos usados para entrenar al algoritmo. La IA es buena para establecer patrones, así como para agilizar procesos y operaciones con volúmenes masivos de información —*Big Data*—. Sin embargo, el problema es que la IA al nutrirse de la información de hecha o recopilada por seres humanos, puede que reflejen sus sesgos. Existen tres tipos de sesgos clásicos: el estadístico, el cultural y el cognitivo.

El sesgo estadístico procede de cómo obtenemos los datos, de errores de medida o similares. Por ejemplo, si la policía está presente en algunos barrios más que en otros, no será extraño que la tasa de criminalidad sea más alta donde tenga mayor presencia —o en otros términos, mediremos más donde está uno de los instrumentos de medida—. El sesgo cultural es aquel que deriva de la sociedad, del lenguaje que hablamos o de todo lo que hemos aprendido a lo largo de la vida. Los estereotipos de las personas de un país son un ejemplo claro. Por último, el sesgo cognitivo es aquel que nos identifica y que depende de nuestras creencias. Por ejemplo, si leemos una noticia que está alineada con lo que pensamos, nuestra tendencia será validarla, aunque sea falsa.

¿Qué es lo que vuelve parcial a una IA? El *Machine Learning* y el *Deep Learning* son la razón por las que un programa informático pierde imparcialidad. Si a un algoritmo de clasificación de datos se le brinda millones de imágenes de perros etiquetados, luego podrá decidir si una foto que no ha visto antes contiene a un perro o no. Asimismo, si a un algoritmo de reconocimiento de voz se le "nutre" con millones de muestras de voz junto con sus correspondientes palabras escritas, luego será capaz de transcribir el lenguaje hablado más rápido que la mayoría de los seres humanos. Cuantos más datos etiquetados vea un algoritmo, mejor será la tarea que realice. Sin embargo, la desventaja radica en que éstos pueden desarrollar puntos ciegos basados en el tipo de información sobre los que están entrenados.

A modo de ejemplo, en el año 2015, Jacky Alcine —una mujer afroamericana— cuando miró su fotografía en la aplicación de Google Photos no podía creer que el software de reconocimiento facial la había etiquetado con la palabra "gorila". Esto sucedió porque el algoritmo no había sido entrenado con suficientes imágenes de personas de piel oscura. En otro caso, a inicios del año 2016, Microsoft lanzó a "Tay", un *chatbot* cuyo fin era imitar el comportamiento de una adolescente curiosa y buscaba entablar en las redes sociales una conversación informal y divertida con una audiencia de entre 18 y 24 años, según explicó la compañía en su página web. El proyecto mostraría las promesas y el potencial de las interfaces conversacionales alimentadas por IA. Sin embargo, en menos de 24 horas, el "inocente" Tay a través de *tweets*, mostraba su empatía hacia Hitler o su apoyo al genocidio al responder a preguntas de los usuarios de las redes sociales son algunos ejemplos, además de insultos raciales y comentarios sexistas y homófobos. También defendió el Holocausto, los campos de concentración o la supremacía blanca, y se mostró contraria al feminismo.

El sesgo algorítmico será un problema cada vez mayor a medida que las decisiones de estos *softwares* se vuelvan cada vez más importantes en nuestras vidas. Esta situación se vuelve más dramática cuando las decisiones automatizadas afectan los derechos fundamentales de las personas. Un ejemplo claro es el programa COMPAS —*Correctional Offender Management Profiling for Alternative Sanctions*, por su acrónimo en inglés, que en español

puede traducirse como "Administración de Perfiles de Criminales para Sanciones Alternativas del Sistema de Prisiones de Estados Unidos"—. Este programa es básicamente un cuestionario que se le da a las personas que han sido arrestadas. Las preguntas incluyen aspectos como los antecedentes penales tanto del reo como de sus familiares, su domicilio, datos relativos a su centro de trabajo y sus datos académicos. Asimismo, también hay preguntas que buscan crear un "perfil" y determinar si la persona tiene o no un "pensamiento criminal". En otras palabras, las respuestas son analizadas por estos sistemas de IA y terminan concluyendo si esa persona en el futuro podría cometer un crimen. Entonces aparece un valor promedio de riesgo que decide si alguien puede salir bajo fianza, debe ser enviado a prisión o recibir otro castigo. Cuando la persona ya está encarcelada, el algoritmo determina si merece el beneficio de la libertad condicional.

Al igual que este programa, en el mundo, cada vez más, encontramos ejemplos de cómo las cortes, los bancos o empresas del sistema financiero y otras instituciones están utilizando programas de IA, que automatizan las decisiones de las vidas de las personas. Al respecto, el Grupo de Trabajo sobre Protección de Datos del Artículo 29[38] publicó Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento General de Protección de Datos Personales de la UE —en adelante, "RGPD"—.

En este documento se señala que los progresos tecnológicos y las posibilidades del análisis de "macrodatos", la IA y el aprendizaje automático han facilitado la creación de perfiles y han automatizado las decisiones, y tienen el potencial de afectar de forma significativa a los derechos y libertades de las personas.

La amplia disponibilidad de datos personales en internet y en los dispositivos del internet de las cosas —IdC—, así como la capacidad de hallar correlaciones y crear vínculos, puede permitir determinar, analizar y predecir ciertos aspectos de la personalidad o el comportamiento, los intereses y los hábitos de una persona.

No obstante, la elaboración de perfiles y las decisiones automatizadas pueden plantear riesgos importantes para los derechos y libertades de las personas que requieren unas garantías adecuadas.[39] Estos procesos pueden ser opacos. Puede que las personas no sean conscientes de que se está creando un perfil sobre ellas o que no entiendan lo que implica. La elaboración de perfiles puede perpetuar los estereotipos existentes y la segregación social. Asimismo, puede encasillar a una persona en una categoría específica y limitarla a las preferencias que se le sugieren. Esto puede socavar su libertad a la hora de elegir, por ejemplo, ciertos productos o servicios como libros, música o noticias. En algunos casos, la elaboración de perfiles puede llevar a predicciones inexactas. En otros, puede llevar a la denegación de servicios y bienes, y a una discriminación injustificada.

Ante este supuesto, el responsable y encargado de tratamiento de datos personales debe implementar las medidas necesarias para evitar este tipo de situaciones. Para ello el sistema de IA deberá ser "nutrido" con datos relevantes y correctos, debiendo aprender qué datos enfatizar. La IA, en principio, no debería procesar información relacionada con datos sensibles como el origen racial o étnico, las opiniones políticas, la religión, las creencias, la orientación sexual para evitar que esto conduzca a un tratamiento arbitrario que termine en la discriminación por parte de la IA de los titulares de datos personales.

38. Este Grupo de Trabajo se creó en virtud del artículo 29 de la Directiva 95/46/CE. Se trata de un órgano consultivo independiente de la UE en materia de protección de datos e intimidad. Sus funciones se describen en el artículo 30 de la Directiva 95/46/CE y en el artículo 15 de la Directiva 2002/58/CE.

39. Las decisiones basadas únicamente en el tratamiento automatizado representan la capacidad de tomar decisiones por medios tecnológicos sin la participación del ser humano.

En el artículo 22 del RGPD se establece que los ciudadanos europeos tienen el derecho a no ser objeto de una decisión basada únicamente en medios automatizados, incluida la elaboración de perfiles, si la decisión produce efectos jurídicos que los afecten significativamente de modo similar. En ese sentido, si un tratamiento automatizado da lugar a una denegación de una solicitud de un crédito por internet, el titular de los datos personales podrá oponerse a este tratamiento. Por tanto, un adecuado tratamiento supondría informar previamente que dicha decisión es automatizada —es decir, sin intervención humana—, que puede expresar su opinión, impugnar la decisión y que el titular de datos personales pueda solicitar la contribución de una persona en el proceso para revisar la decisión tomada mediante el algoritmo.

En el Perú, si bien no existe una norma específica como la del RGPD, el artículo 24 de la Ley de Protección de Datos Personales señala que el titular de datos personales tiene derecho a no verse sometido a una decisión con efectos jurídicos sobre él o que le afecte de manera significativa, sustentada únicamente en un tratamiento de datos personales destinado a evaluar determinados aspectos de su personalidad o conducta, salvo que ello ocurra en el marco de la negociación, celebración o ejecución de un contrato o en los casos de evaluación con fines de incorporación a una entidad pública, de acuerdo a ley, sin perjuicio de la posibilidad de defender su punto de vista, para salvaguardar su legítimo interés.

Asimismo, el artículo 72 del Reglamento de la Ley de Protección de Datos Personales establece que para garantizar el ejercicio del derecho al tratamiento objetivo, de conformidad con lo establecido en el artículo 23 de la Ley de Protección de Datos Personales, cuando se traten datos personales como parte de un proceso de toma de decisiones sin participación del titular de los datos personales, el titular del banco de datos personales o responsable del tratamiento deberá informárselo a la brevedad posible, sin perjuicio de lo regulado para el ejercicio de los demás derechos establecidos en la Ley de Protección de Datos Personales y su Reglamento.

Por tanto, en base a la generalidad de ambos artículos, se podría aplicar en el Perú el mismo criterio jurídico que se aplica actualmente con los tratamientos automatizados en la Unión Europea. Es decir, un ciudadano peruano tiene el derecho a no ser objeto de una decisión basada únicamente en medios automatizados, incluida la elaboración de perfiles, si la decisión produce efectos jurídicos que afecten sus derechos.

En ese sentido, una decisión automatizada estará permitida cuando ésta sea necesaria —es decir, que no existe otra manera de lograr el mismo objetivo— para la ejecución de un contrato, exista una norma legal que autorice este tratamiento o se haya dado un consentimiento explicito. Sin perjuicio de ello, la decisión adoptada debe garantizar los derechos y libertades de los titulares de datos personales, aplicando las garantías adecuadas. Asimismo, el responsable del tratamiento debe, como mínimo, informarle de su derecho a obtener intervención humana y establecer los requisitos de procedimiento obligatorios; además, la empresa u organización deberá permitirle expresar su punto de vista e informarle de que puede impugnar la decisión.

Finalmente, si se sospecha o se afirma que el uso de la IA conlleva a resultados discriminatorios o que afecten derechos fundamentales de las personas, la Dirección General de Protección de Datos Personales podrá investigar si el principio de legalidad ha sido vulnerado o no. En su procedimiento de fiscalización deberá solicitar la documentación que respalda la selección de datos, un examen de cómo se desarrolló el algoritmo y si se probó adecuadamente antes de su uso.

2. *Big Data* y *Machine Learning* vs. los principios de finalidad y proporcionalidad.

En el año 2010, el fundador y CEO de Facebook, Mark Zuckerberg, señalaba en una entrevista que *"la era de la privacidad ha muerto"*. Irónicamente, en julio del 2019, la Comisión Federal del Comercio de Estados Unidos —FTC, por sus siglas en inglés— impuso a Facebook una san-

ción por valor de 5 mil millones de dólares por su gestión de la privacidad de los usuarios tras el escándalo "Cambridge Analytica".

La utilización de 87 millones de datos personales por parte de la consultora británica Cambridge Analytica, obtenidos a través de Facebook para manipular psicológicamente a los votantes —al parecer decisivamente— en la campaña electoral de Estados Unidos en favor de Trump, o en la campaña del último referéndum británico en favor del Brexit, es el más siniestro ejemplo del poder del *Big Data* y *Machine Learning*.

Esta empresa obtuvo un perfil psicométrico de cada ciudadano en Estados Unidos. A través de su rastro digital, podía saber si los usuarios eran hombre o mujer, su edad, qué carro manejan y qué tipo de cereal desayunan. También podía saber sus afinidades políticas y sus principales preocupaciones sociales. Cambrige Analytica utilizó el *Big Data* y el *Machine Learning* para realizar análisis predictivos que los ayudan para desarrollar esquemas de comunicación comercial y política.

¿Cómo hicieron esto? La empresa creó una aplicación llamada *"This is your digital life"* que funcionaba como un *test online*, y que se presentaba solo como una herramienta de investigación. Se trataba de encuestas aparentemente inofensivas que circulan en Facebook y otras redes sociales, del tipo "¿qué Pokémon eres?" o "¿qué personaje de Game of Thrones eres?". Para completarlo se requería iniciar sesión en Facebook y otorgarle algunos permisos a la aplicación, como recoger información sobre la actividad del usuario, acceder a la ubicación y a los contactos en la red. Unos 270 mil perfiles hicieron esta encuesta online, lo cual derivó en la recopilación de información de 50 millones de perfiles. Esto fue posible porque la aplicación solicitaba el acceso a los datos de los amigos. Fue así como se magnificó el efecto.

Christopher Wylie, ex empleado de Cambridge Analytica, reveló al New York Times y The Guardian cómo es que la consultora británica utilizaba los datos de perfiles en la red social para generar anuncios personalizados con fines políticos: *"Explotamos Facebook para acceder a millones de perfiles de usuarios. Y construimos modelos para explotar lo que sabíamos de ellos y apuntar a sus demonios internos. Esa era la base sobre la cual la compañía se fundó"*.

Este caso es un claro ejemplo de cómo se vulnera el principio de finalidad que rige el tratamiento de los datos personales. Si bien Cambridge Analytica recababa el consentimiento para el tratamiento de datos personales para ciertos fines, no cabe duda de que la información que obtuvieron la desviaron para otros fines. Cabe recordar que el principio de finalidad tiene por objetivo establecer claramente cuáles son los fines para los que van a ser utilizados los datos personales, pues esta información es esencial para que el titular de datos personales pueda ejercer control sobre el uso de su información personal.

El RGPD, en su considerando 50, señala que el tratamiento de datos personales con fines distintos de aquellos para los que hayan sido recogidos inicialmente solo debe permitirse cuando sea compatible con los fines de su recogida inicial. En tal caso, no se requiere una base jurídica aparte, distinta de la que permitió la obtención de los datos personales. Asimismo, agrega que para determinar si el fin del tratamiento ulterior es compatible con el fin de la recogida inicial de los datos personales, el responsable del tratamiento, tras haber cumplido todos los requisitos para la licitud del tratamiento original, debe tener en cuenta, entre otras cosas, cualquier relación entre estos fines y los fines del tratamiento ulterior previsto, el contexto en el que se recogieron los datos, en particular las expectativas razonables del interesado basadas en su relación con el responsable en cuanto a su uso posterior, la naturaleza de los datos personales, las consecuencias para los interesados del tratamiento ulterior previsto y la existencia de garantías adecuadas tanto en la operación de tratamiento original como en la operación de tratamiento ulterior prevista.

El problema con el *Machine Learning* y el *Big Data* es que, por la propia naturaleza de estas tecnologías, su valor reside precisamente en lo

inesperado de los resultados que revelan. Así, ¿cómo explica el responsable del tratamiento que resulta imposible saber con antelación qué información revelará el tratamiento de los datos recabados? Por ejemplo, una empresa que se dedica a la investigación, en primer lugar, deberá recabar el consentimiento de un titular para utilizar sus datos a fin de encontrar la cura del Alzheimer; sin embargo, qué sucede si la IA, en base a esos datos, llega a una conclusión referente a la impotencia sexual. ¿Qué tendría que hacer el responsable del tratamiento de datos personales? La norma no es clara en establecer si la obligación del responsable del tratamiento de informar las finalidades sobre la recogida de los datos se circunscribe a la información que explícitamente recoge —datos primarios—, o si debe adoptarse un criterio más amplio y entender que este deber de información acerca de las finalidades también alcanza a aquellos resultados que la empresa pudiera obtener tras el tratamiento —datos secundarios—.[40]

Otro problema radica en que el *Big Data* se basa, precisamente, en reutilizar datos que fueron obtenidos para una primera finalidad, otorgándole un nuevo fin. Este es uno de los aspectos en donde se encuentra la mayor fuente de beneficios del *Big Data*.

Para ello, la doctrina española[41] ha formulado un test de incompatibilidad entre los usos del *Big Data* y el principio de finalidad. Por tanto, para superar esta situación se deberá cumplir con alguna de las siguientes condiciones:

a) Que las finalidades del tratamiento de datos del proyecto *Big Data* se ajusten a lo informado a los interesados en el momento inicial de recabar sus datos; o bien,

b) Que las finalidades del tratamiento de datos del proyecto *Big Data* sean razonablemente previsibles para los interesados, aun no habiendo sido explícitamente informados en el momento de obtener sus datos; o bien,

c) El tratamiento de datos resultante del proyecto *Big Data* está justificado por otras causas de legitimación previstas en la normativa de privacidad —como son, por ejemplo, el interés legítimo del responsable del tratamiento, el cumplimiento de obligaciones legales, contractuales, o en atención al interés vital de los interesados—.

En caso se supere con éxito este test de incompatibilidad, el uso del *Big Data* podrá considerarse conforme a la normativa de protección de datos personales, sin perjuicio del cumplimiento de otras normas que establece la Ley de Protección de Datos Personales, tales como las medidas de seguridad. Ahora bien, si el resultado del test es negativo, el uso del *Big Data* deberá sujetarse a información y consentimiento previos de los interesados involucrados.

Una de las soluciones para armonizar el principio de finalidad con la utilización de esta tecnología es que los datos se sometan a procesos de "anonimización" de los datos. El procedimiento de anonimización es aquel tratamiento de datos personales que impide la identificación o que no hace identificable al titular de estos, el cual es irreversible. El tratamiento de datos personales puede extenderse a otras finalidades en la medida que se utilice este proceso. En ese sentido, la anonimización se presenta como la mejor solución para tratar los datos protegiendo la privacidad de los sujetos.

Sin embargo, la FTC ha declarado que:

> *"(…) [h]ay evidencias suficientes que demuestran que los avances tecnológicos y la*

40. Elena Gil Gonzáles, «Big Data: Consentir o no consentir, ésa es la cuestión», *Legal Today*, 16 de marzo de 2017, acceso el 9 de noviembre del 2020, http://ecija.com/big-data-consentir-no-consentir-esa-la-cuestion/.

41. Carlos Pérez Sanz, «Aspectos Legales del Big Data», *Índice* 68 (julio 2016): 18-21 acceso el 9 de noviembre del 2020, http://www.revistaindice.com/numero68/p18.pdf.

posibilidad de combinar diferentes datos puede conllevar a la identificación de un consumidor, ordenador o dispositivo, incluso si estos datos por sí mismos no constituyen datos de identificación personal. Es más, no solo es posible reidentificar datos que no son identificadores personales a través de medios diversos, sino que las empresas tienen fuertes incentivos para hacerlo".[42]

Es importante mencionar que el procedimiento de anonimización fallará en la medida que haga identificable a la persona. Un sujeto es identificable cuando, aunque no haya sido identificado todavía, pueda serlo. En otras palabras, la posibilidad de identificar a una persona ya no equivale necesariamente a la capacidad de poder llegar a conocer su nombre y apellidos, puesto que los datos pueden ser combinados con otros a fin de que permitan distinguir a esa persona de otras.

El *Big Data* analiza la huella digital de los individuos y a pesar de que en una base de datos no aparezcan nombres, se aprecian patrones. De manera que, con estos patrones, una persona con suficientes conocimientos analíticos puede obtener nombres, lo cual afecta nuestra privacidad, puesto que esta es entendida como el control que tienen los sujetos sobre sus conductas e información. A medida que evolucionen las técnicas de *Big Data*, los individuos perderán esta facultad de salvaguardar su espacio personal e impedir que no sean observados o incomodados por terceros.

¿Qué se puede hacer al respecto? Para que el proceso de anonimización funcione se debe tener en cuenta el principio de proporcionalidad. Es decir, el responsable del tratamiento solo podrá tratar aquellos datos que sean adecuados, pertinentes y limitados a lo necesario en relación con los fines, que a su vez deben ser determinados, explícitos y legítimos. De esta forma, se minimiza la cantidad de datos personales a ser tratados para ciertos fines y de esta manera se podrá dificultar la identificación de las personas con el *Big Data*. En otras palabras, lo que se busca es que el grado de identificación del individuo esté restringido tanto por la cantidad como por la naturaleza de la información utilizada, ya que algunos detalles revelan más sobre la identidad de una persona que otros. Esto, sumado al uso de las técnicas de anonimización y encriptación ayudan a proteger de mejor manera la identidad de los titulares de datos personales.

Cabe señalar que el principio de proporcionalidad obliga a los desarrolladores y responsables del tratamiento a examinar a fondo el modelo previsto para facilitar la selección de datos al momento de dotar de información a la IA, debiendo contener datos que sean relevantes y necesarios para las finalidades previstas. En términos más simples, el responsable del tratamiento, en virtud de este principio, debe elegir la opción que sea menos invasiva para los titulares de los datos personales. Se recomienda que esta decisión sea documentada, de modo que puedan presentarse a la Autoridad de Protección de Datos Personales en caso de una fiscalización.

Si bien es difícil establecer de antemano la información y datos que serán necesarios y relevantes para el desarrollo de la IA, y esto, además, puede modificarse durante el proyecto, es esencial minimizar la cantidad de datos personales y utilizar sólo aquellos que sean relevantes. Esto no solo protege los derechos de los titulares de datos personales, sino que también minimiza el riesgo de que la información irrelevante lleve a la IA a encontrar correlaciones que, en lugar de ser significativas, sean contraproducentes y conlleven al denominado sesgo algorítmico.

3. *Black Box* vs. el principio de información.

La finalidad de la normativa de protección de datos personales es salvaguardar el derecho

42. «Protecting Consumer Privacy in an Era of Rapid Change. Recommendations for Businesses and Policymakers», Federal Trade Commission, acceso 9 de noviembre del 2020, https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

de que las personas puedan decidir y controlar cómo es que terceros utilizan sus datos. Para ello es necesario que los responsables del tratamiento sean transparentes acerca de cómo van a procesar su información. En otros términos, la transparencia se logra cuando se proporcionan a los titulares de datos personales todos los detalles acerca del tratamiento.

Al respecto, el artículo 18 de la Ley de Protección de Datos Personales señala que:

> *"El titular de datos personales tiene derecho a ser informado en forma detallada, sencilla, expresa, inequívoca y de manera previa a su recopilación, sobre la finalidad para la que sus datos personales serán tratados; quiénes son o pueden ser sus destinatarios, la existencia del banco de datos en que se almacenarán, así como la identidad y domicilio de su titular y, de ser el caso, del encargado del tratamiento de sus datos personales; el carácter obligatorio o facultativo de sus respuestas al cuestionario que se le proponga, en especial en cuanto a los datos sensibles; la transferencia de los datos personales; las consecuencias de proporcionar sus datos personales y de su negativa a hacerlo; el tiempo durante el cual se conserven sus datos personales; y la posibilidad de ejercer los derechos que la ley le concede y los medios previstos para ello".*

Por tanto, los ciudadanos deben ser informados sobre cómo se utilizará la información, si esta información será procesada por la misma persona que recopila los datos o un tercero. Además, esta información debe ser entregada en un lenguaje sencillo y de forma previa al tratamiento. Esto supone un desafío muy grande pues, al tratarse de tecnología avanzada y compleja es difícil de explicar todo el procesamiento de información en un lenguaje sencillo. Por otro lado, una preocupación existente en relación al *Machine Learning* es que no siempre se sabe cómo es que se produce el resultado, pues un programa de IA, por lo general, llega a un resultado sin ninguna explicación. Entonces, surge la interrogante sobre si es posible estudiar el modelo que la IA utilizó para así descubrir cómo llegó a ese resultado específico. ¿Cómo explicar en lenguaje sencillo aquello que no comprendes? Los desarrolladores de la IA muchas veces no saben cómo es que la IA correlaciona y pondera la información en un proceso específico. A esto se le denomina *"Black Box"* o "caja negra".

A raíz de este concepto, surge la siguiente interrogante: ¿puede acaso un titular de datos personales solicitar una explicación, en virtud del principio de información, acerca del tratamiento que la IA realizó respecto de su información? Consideramos que no sería necesario "abrir" el *Black Box*, pues el artículo 18 tan solo te exige informar sobre la finalidad, los destinatarios, los bancos de datos personales en donde la información se almacena, la identidad del responsable, tiempo de conservación, las consecuencias de proporcionar sus datos personales y de su negativa a hacerlo. En ese sentido, develar cómo es que funciona la IA no sería necesario para efectos de cumplir con este principio.

¿Qué sucede si en base a un procedimiento se toma una decisión automatizada que perjudica al titular de datos personales? ¿Se podría abrir el *Black Box* en este caso? El derecho a la información no implica que se deba abrir el *Black Box*. Este principio debe permitir al titular de los datos personales a comprender por qué se tomó una decisión en particular y no otra. Asimismo, deberá ser informado de cómo puede oponerse a la decisión automatizada, ya sea impugnando la decisión o solicitando la intervención humana.

4. Viabilizando el desarrollo de la IA con el derecho a la protección de datos personales.

Uno de los principios reconocidos por el RGPD es el denominado principio de responsabilidad proactiva, el cual exige al responsable del tratamiento a aplicar medidas técnicas, legales y organizativas apropiadas para garantizar y demostrar que el tratamiento que realiza es conforme a la normativa de protección de datos personales, tomando en consideración la naturaleza, el ámbito, el contexto y los fines del tratamiento, así como los riesgos para los dere-

chos y libertades de las personas naturales, en su calidad de titulares de datos personales.

Es decir, el responsable del tratamiento tiene que asegurar y ser capaz de probar el cumplimiento de los principios que legitiman el tratamiento de los datos personales a lo largo de todo el ciclo de vida de estos, desde que se obtienen hasta que, finalmente, se suprimen o anonimizan. Por tanto, este principio exige una actitud consciente, diligente y proactiva por parte de las organizaciones frente a todos los tratamientos de datos personales que lleven a cabo.

¿Se puede aplicar este principio a la normativa peruana? Sí, en virtud de lo señalado en el artículo 12 de la Ley de Protección de Datos Personales que establece que la lista de principios es enunciativa. En términos prácticos, este principio requiere que las organizaciones analicen qué datos tratan, con qué finalidades lo hacen y qué tipo de operaciones de tratamiento llevan a cabo. Para ello, las empresas pueden elaborar Evaluaciones de Impacto en Protección de Datos Personales —en adelante, "EIPD"— o *"Data Protection Impact Assesment"* —en adelante, "DPIA"—.

La Agencia de Protección de Datos Personales Española señala que una EIPD es una herramienta con carácter preventivo que debe realizar el responsable del tratamiento para poder identificar, evaluar y gestionar los riesgos a los que están expuestas sus actividades de tratamiento con el objetivo de garantizar los derechos y libertades de las personas físicas. En la práctica, la EIPD permite determinar el nivel de riesgo que entraña un tratamiento, con el objetivo de establecer las medidas de control más adecuadas para reducir el mismo hasta un nivel considerado aceptable.

Esta herramienta se exige a los responsables del tratamiento, pues el RGPD señala que se deben implementar medidas de control adecuadas para demostrar que se garantizan los derechos y libertades de las personas y la seguridad de los datos, teniendo en cuenta entre otros, los *"riesgos de diversa probabilidad y gravedad para los derechos y libertades de las personas físicas"* y aplicando las medidas oportunas. Para ello, el responsable del tratamiento debe considerar desde el inicio, en la fase de diseño, las acciones preventivas suficientes para poder identificar, evaluar y tratar los riesgos asociados al tratamiento de datos personales, y así, poder asegurar los principios de protección de los datos garantizando los derechos y libertades de los interesados.

A fin de cumplir con el principio de responsabilidad proactiva, la elaboración de la EIPD se debe realizar "antes del tratamiento" en los casos en que sea probable que exista un alto riesgo para los derechos y libertades de los afectados. El resultado de la EIPD se debe tener en cuenta, necesariamente, a la hora de tomar las decisiones relacionadas con el cumplimiento de lo previsto en la normativa de protección de datos personales y la toma de decisión de la viabilidad o no de llevar a cabo el tratamiento de los datos. En ese sentido, una empresa que utilice IA y trate datos personales, deberá elaborar una EIPD a fin de evaluar los riesgos que este procesamiento acarrea en los derechos y libertades de las personas.

Cabe señalar que la obligación del RGPD de elaborar un EIPD cuando se esté ante la probabilidad de que un tratamiento *"entrañe un alto riesgo para los derechos y libertades de las personas físicas"*, está alineada con el principio de *"privacy by design"*.

El concepto "protección de datos" desde el diseño y por defecto, recogido en el artículo 25 del RGPD, consiste en incorporar, desde las primeras fases de todo proyecto, medidas técnicas y organizativas apropiadas, teniendo en cuenta factores como el estado de la técnica, el coste de la aplicación o los riesgos del tratamiento para los derechos y libertades de los afectados, para cumplir los requisitos del Reglamento y proteger los derechos de los interesados. En otras palabras, de acuerdo a este principio, desde el diseño de la tecnología, deberá tenerse en cuenta el concepto de privacidad.

Este principio suele caracterizarse por tomar medidas proactivas en lugar de reactivas. Se

anticipa y previene la pérdida de privacidad de la información antes de que suceda. Asimismo, el responsable del tratamiento debe ofrecer el máximo grado de privacidad para asegurar que los datos personales están protegidos automáticamente en cualquier sistema informático o dentro de las buenas prácticas. En tal sentido, la protección de datos personales incorporada a la tecnología, en este caso a la IA, no debe ser considerado como un añadido sino como un componente esencial del núcleo como parte integral del sistema, sin disminuir la funcionalidad.

El *privacy by design* busca dar cabida a todos los intereses y objetivos legítimos de una forma de suma positiva —*win-win*—, pues trata de garantizar que se cubran todas las funcionalidades y necesidades de los distintos implicados, pero sin afectar a la privacidad. Esto evita la pretensión de falsas dicotomías, como la privacidad frente a la seguridad. ¿De qué sirve un desarrollo de la IA si es que ésta va a mermar el derecho a la privacidad de todos sus usuarios? ¿Cuál es el sentido de una tecnología que no respete los derechos fundamentales de las personas?

A modo de reflexión final, aún estamos en la fase inicial de desarrollo de la IA a nivel mundial. En ese sentido, es el momento adecuado para garantizar que las tecnologías de IA cumplan con las reglas establecidas por la Ley de Protección de Datos Personales a fin de garantizar la privacidad de las personas y la facultad de autorregular la información que desean compartir. Los retos que conlleva el desarrollo de la IA en relación a la privacidad de las personas no deben entenderse como barreras, sino que se deben de equilibrar los avances tecnológicos con el derecho fundamental a la protección de datos, y lograr así un desarrollo que no suponga un reto para los derechos fundamentales. Por tanto, no solo es posible combinar la IA con un adecuado tratamiento a los datos personales, sino que resulta necesario hacerlo para salvaguardar la privacidad de las personas.

VI. DERECHOS DE AUTOR E INTELIGENCIA ARTIFICIAL

De acuerdo a un informe realizado por la Organización Mundial de la Propiedad Intelectual, desde la aparición de la IA en los años 1960, los innovadores y los investigadores han presentado casi 340,000 solicitudes de patente de invenciones relacionadas con ella y se han publicado más de 1.6 millones de investigaciones científicas al respecto. Las patentes relacionadas con la IA se han disparado en los últimos años, de modo que más de la mitad de las invenciones publicadas desde 2013 pertenecen a ese ámbito.[43]

Sin embargo, lo más fascinante son aquellos sucesos que han acontecido en el campo de la IA. Por ejemplo, en el año 2016, un grupo de representantes de museos e investigadores de los Países Bajos presentaron un retrato titulado "El nuevo Rembrandt", una nueva obra de arte generada por una computadora que había analizado miles de obras del artista neerlandés del siglo XVII, Rembrandt Harmenszoon Van Rijn. Ese mismo año un equipo de ingenieros del Laboratorio de Investigación CSL de Sony, a través de un sistema inteligente llamado *"Flow Machines"*, publicó canciones inspiradas en los Beatles, las cuales fueron enteramente compuestas por este sistema de IA. En el año 2018, Botnik Studios usó un programa de texto predictivo para generar cuatro páginas de ficción para los admiradores de Harry Potter.

Andrés Guadamuz, experto en temas de Derechos de Autor, señala que las máquinas están dejando de ser accesorias en el proceso creativo y que a través del *Machine Learning* los programas de IA cada vez se vuelven más autónomos:

> *"Hace mucho tiempo que los artistas robóticos participan en diversos tipos de trabajos*

43. James Nurton, «La propiedad intelectual y el auge de la inteligencia artificial», *OMPI Revista*, https://www.wipo.int/wipo_magazine/es/2019/01/article_0001.html

creativos. Las computadoras han producido obras de arte rudimentarias desde los años setenta y estas iniciativas prosiguen en la actualidad. La mayoría de esas obras de arte generadas por computadora dependían en gran medida de la creatividad del programador; la máquina era a lo sumo un instrumento o una herramienta muy parecida a un pincel o un lienzo. Pero hoy en día nos encontramos inmersos en una revolución tecnológica que puede obligarnos a repensar la interacción entre las computadoras y el proceso creativo. Esta revolución está impulsada por el rápido desarrollo del software de aprendizaje automático (Machine Learning), un subconjunto de la inteligencia artificial que produce sistemas autónomos capaces de aprender sin estar específicamente programados por el ser humano.

Un programa informático desarrollado para el aprendizaje automático se basa en un algoritmo que le permite aprender a partir de los datos introducidos, evolucionar y tomar decisiones que pueden ser dirigidas o autónomas. Cuando se aplican a obras artísticas, musicales y literarias, los algoritmos de aprendizaje automático aprenden a partir de la información proporcionada por los programadores. A partir de esos datos generan una nueva obra y toman decisiones independientes a lo largo de todo el proceso para determinar cómo será dicha obra. Una característica importante de este tipo de inteligencia artificial es que, si bien los programadores pueden definir unos parámetros, en realidad la obra es generada por el propio programa informático (denominado red neuronal) mediante un proceso similar a los del pensamiento humano".[44]

La creación de obras por medio de programas de IA podría impactar profundamente en el Derecho de Autor. De conformidad con el Decreto Legislativo 822, se entiende como obra a toda creación intelectual personal y original, susceptible de ser divulgada o reproducida en cualquier forma, conocida o por conocerse. Son derechos exclusivos sobre obras literarias, artísticas, científicas, *software* y obras del ingenio humano.[45]

Asimismo, también pueden ser protegidas las obras derivadas, cuya característica principal es que son obras creadas mediante la transformación de otras ya existentes. A modo de ejemplo, se citan las traducciones, los compendios o resúmenes, o los arreglos musicales. También lo serían un aplicativo que se transforma de un software anterior o un videojuego que deriva de una obra audiovisual previa.

En definitiva, la obra es el objeto sobre el que el Derecho de Autor concede un poder de exclusiva a favor de su titular, inicialmente el autor, persona natural que realiza la creación intelectual. El derecho nace sólo si existe obra y su alcance queda circunscrito a ésta. De ahí que este concepto sea clave para el Derecho de Autor.

Pero ¿cuáles son los requisitos que deben reunir las obras para ser protegidas? Para que una obra sea protegida por el Derecho de Autor debe reunir los siguientes requisitos: (a) el requisito esencial para que lo creado por un ser humano merezca la consideración de obra es que sea original —este requisito puede ser entendido como originalidad subjetiva, es decir, que el autor no haya copiado una obra ajena—; (b) debe ser producto del ingenio y la creatividad humana y; (c) debe ser susceptible de ser divulgada o reproducida.

En ese sentido, una creación realizada por un programa de IA no sería una obra en sentido jurídico, pues no es producto del ingenio y la crea-

44. Andrés Guadamuz, «La Inteligencia Artificial y el Derecho de Autor», *OMPI Revista,* https://www.wipo.int/wipo_magazine/es/2017/05/article_0003.html.

45. Ecija, *Memento Práctico Francis Lefebvre. Derecho de las Nuevas Tecnologías,* versión *online*, (Barcelona: Francis Lefevre, Febrero 2017), capítulo 13.

tividad humana. En otras palabras, a menos que las obras creativas generadas por la IA se puedan atribuir directamente a una persona natural, ésta no podrá ser susceptible de protección de acuerdo al Derecho de Autor y, por tanto, caerían en el dominio público tras su creación.[46]

Este es un tema cuyo impacto comercial puede ser tremendo, pues como se ha mencionado anteriormente, la IA ha pasado de ser una herramienta, a tomar muchas de las decisiones asociadas al proceso creativo sin intervención humana. Esto significa que esas obras podrían considerarse libres de derechos de autor porque no han sido creadas por el ser humano. Por consiguiente, cualquier persona podría utilizarlas y reutilizarlas libremente, lo cual sería una muy mala noticia para aquellas compañías que han invertido millones de dólares en estos programas de IA. Esto podría frenar la inversión, debido a que el ordenamiento jurídico no incentivaría este tipo de creaciones.

Ahora bien, este problema puede tratarse de tres formas:

a) Puede denegarse la protección del Derecho de Autor respecto de aquellas obras generadas por IA en donde el ser humano no haya intervenido en el proceso creativo o lo haya hecho en forma mínima.

b) Se podrá considerar como autor a aquella persona natural que haya colaborado con los *inputs* iniciales —es decir, la información o las reglas que nutren a la IA—. Sin embargo, para aquellos casos en que a través del *Machine Learning* la obra haya sido generada por sí sola, sin intervención humana, la creación no será considerada como obra, de conformidad con el Derecho de Autor.

c) Se puede crear una ficción legal y atribuir la autoría de esa obra a quien tenga los derechos patrimoniales de la IA y, por tanto, la facultad de explotarla económicamente. Esta posición tiene por finalidad incentivar el desarrollo de esta tecnología en el campo de la propiedad intelectual a fin de no dejar desprotegidas a aquellas compañías que invierten mucho tiempo y dinero en el desarrollo de la IA.

Existen indicios de que la legislación de numerosos países no es favorable al derecho de autor que no se aplica al ser humano. En los Estados Unidos, por ejemplo, la Oficina de Derecho de Autor ha declarado que *"registrará una obra original de autoría, siempre que la obra haya sido creada por el ser humano".* Esta posición dimana de la jurisprudencia —por ejemplo, *Feist Publications* v. *Rural Telephone Service Company, Inc. 499 U.S. 340 (1991)* —, que especifica que el derecho de autor solo protege *"el fruto del trabajo intelectual"* que *"se basa en el poder creativo de la mente"*. Del mismo modo, en un asunto reciente ventilado en Australia —Acohs Pty Ltd v. Ucorp Pty Ltd—, el tribunal declaró que una obra generada con la intervención de una computadora no podía estar protegida por el derecho de autor porque no había sido producida por el ser humano[47].

Por otro lado, el artículo 9.3 del *Copyright, Designs and Patents Act* del Reino Unido dispone que, en el caso de una obra literaria, dramática, musical o artística generada por computadora, se considerará que el autor es la persona que realiza los arreglos necesarios para la creación de la obra. En ese sentido, esta ley opta por el tercer enfoque señalado anteriormente, que consiste en atribuir la autoría de esa obra al programador o a la persona que haga lo necesario para la creación de la obra.

Finalmente, es importante señalar que el impacto de la IA en el Derecho de Autor no se agota en este tema, pues la realidad será quien

46. Kalin Hristov, «Artificial Intelligence and the Copyright Dilemma», *IDEA – The Journal of the Franklin Pierce Center for Intellectual Property* 57 (3): 437.

47. Andrés Guadamuz, *Op. Cit.*, *Ibíd*.

se encargue de plantear distintos retos jurídicos como, por ejemplo: ¿qué sucede si un programa de IA copia una obra ya existente? ¿quién será el responsable? ¿qué ocurriría si varias IAs crean una obra colectiva y éstas pertenecen a distintas compañías? ¿qué pasará cuando la IA sea capaz de elegir cómo y de qué forma quiere que se divulgue, se reproduzca o se distribuya su obra?

VII. REFLEXIONES FINALES

Cuando era niño me mandaron a leer algunos libros del escritor Isaac Asimov, como "Robbie y otros Relatos". Una de las cuestiones que más me sorprendió en ese momento, además de la visión futurística del escritor, es que estos robots actuaban siguiendo unas pautas que marcaban su comportamiento. Estas pautas fueron bautizadas por el autor como las "Las Tres Leyes de la Robótica", que fueron formuladas por primera vez en el relato "El círculo vicioso" publicado en el número de marzo de 1942.

Las tres leyes de la robótica son:

a) Un robot no puede lesionar a un ser humano o, por medio de la inacción, permitir que un ser humano sea lesionado.

b) Un robot debe obedecer las órdenes dadas por los seres humanos, excepto si estas órdenes entrasen en conflicto con la primera ley.

c) Un robot debe proteger su propia existencia en la medida que esta protección no sea incompatible con la primera o segunda ley.

Estas leyes deben ser la base de todo desarrollo de IA, pues deben tener al ser humano como beneficiario de esta tecnología. El progreso de la IA no puede ir en desmedro de los derechos fundamentales de las personas, tanto naturales como jurídicas. El ordenamiento jurídico debe incentivar a los empresarios que desarrollen IA y que tengan en cuenta el principio de *privacy by design* y el de responsabilidad proactiva al momento de utilizar la información para nutrir a sus sistemas de IA.

Si un sistema de IA causa un daño, el ordenamiento jurídico debe estar elaborado de tal manera que su solución se centre en la "víctima" y brinde una solución eficiente, en vez de utilizar la hermenéutica jurídica a fin de hallar un culpable. Creemos que una buena solución es dotar, para la mayoría de los casos, una responsabilidad objetiva a quien ocasione daños por poseer o utilizar la IA y que esta deberá ser asumida por la parte que está en mejores condiciones para diluir el costo, quien, generalmente, será el fabricante.

En materia de derechos de autor, se debe proteger a las compañías que inviertan millones de dólares para que las obras creadas por IA sean consideradas como tal y, por tanto, los derechos de explotación sean un derecho de exclusiva por parte de aquellos que fijaron la base para poder llevar a cabo la creación.

Finalmente, quisiera terminar el presente artículo con una cita de Isaac Asimov: *"Quien se acostumbra a preocuparse por las necesidades de unas máquinas, se vuelve insensible respecto a las necesidades de los hombres"*. Debemos fomentar e incentivar el desarrollo de esta tecnología, sin olvidar que el ser humano es el fin supremo de ésta.