

*Inteligencia artificial y derechos humanos**

Por Juan G. Corvalán

1. Introducción

La humanidad se encuentra a las puertas de una era en la que los *robots*, *bots*, *androides* y otras formas de inteligencia artificial, cada vez más sofisticadas, parecen ser el motor que desencadenará una nueva revolución industrial que probablemente afecte a todos los estratos de la sociedad. Estas afirmaciones, no provienen de un film de ciencia ficción. Son palabras textuales del Parlamento Europeo a partir de las recomendaciones vinculadas a normas de derecho civil sobre robótica aprobadas a principios de este año¹.

Hacia fines de esta segunda década del siglo XXI, estamos comenzando a formular las siguientes preguntas: ¿quiénes son los responsables por las consecuencias del funcionamiento de las máquinas inteligentes?; ¿cómo garantizar la autodeterminación humana en la era de los algoritmos inteligentes?; ¿cómo es posible “programar” la inteligencia artificial para que pueda incluir un enfoque jurídico, ético y moral?; ¿cómo hacer que la inteligencia artificial no profundice las desigualdades entre las personas?; ¿cómo lograr que las máquinas inteligentes sean compatibles con los derechos humanos? Las respuestas a estas preguntas demandan un esfuerzo trascendente para repensar e innovar acerca de los desafíos de la llamada “cuarta revolución industrial”² que estamos atravesando. Debemos explorar enfoques sustentables, frente a escenarios disruptivos, complejos, inciertos y que, en esencia, presuponen abordajes transversales, interdisciplinarios, multipolares, dinámicos, flexibles e integrales.

Básicamente, nos encontramos frente a un desafío inédito que resumimos bajo los siguientes dos interrogantes: ¿cómo hacer que una inteligencia no humana, beneficie y no perjudique a los seres humanos?; bajo un enfoque estrictamente jurídico, ¿cómo hacer para que el desarrollo de la inteligencia artificial sea compatible con los derechos humanos y con el Estado constitucional? En estas breves líneas, tan sólo intentaremos destacar una posible y larga vía por explorar. El iter secuencial de esta

* El presente es una síntesis de un trabajo de investigación relacionado con el impacto de las nuevas tecnologías en el derecho público. Artículo publicado en el Consejo de Estado Francés. [Versión en francés](#). [Bibliografía recomendada](#).

¹ 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), Introducción, puntos A y B, disponible en www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//ES.

² Tanto el Foro Económico Mundial como la Organización Internacional del Trabajo (OIT), destacan que el mundo está atravesando por una cuarta revolución industrial. Véase, La iniciativa del centenario relativa al futuro del trabajo. Nota Informativa. Oficina Internacional del Trabajo, 2015, p. 2, y El futuro de los empleos. Habilidades y estrategia de la mano de obra para la cuarta revolución industrial. Foro Económico Mundial, 2016, p. 1. Global Challenge Insight Report, disponible en www.weforum.org/reports/the-future-of-jobs; sobre los aspectos conceptuales de esta denominación (cuarta revolución industrial) véase: Schwab, Klaus, *La cuarta revolución industrial*, Barcelona, Debate, 2016.

propuesta, puede ser resumido así: inteligencia humana → inteligencia artificial → problemas/desafíos → enfoque de derechos humanos → principios jurídicos de la IA.

2. Inteligencia humana, inteligencia artificial y algoritmos inteligentes

La inteligencia humana se relaciona con una serie de capacidades³ o cualidades⁴ cognitivas relativamente autónomas que suelen clasificarse en “perfiles de inteligencia” o “inteligencias múltiples”⁵. Ellas son, inteligencia social, inteligencia lingüística (o musical), inteligencia lógico-matemática, inteligencia interpersonal e intrapersonal o emocional⁶, inteligencia fluida⁷, entre otras⁸. Entre los distintos aspectos del concepto “inteligencia”, el elemento común es la *capacidad de procesar información para resolver problemas*⁹. Básicamente, nuestro cerebro controla la capacidad para procesar la información¹⁰ proveniente del entorno y de nuestro propio cuerpo¹¹ que se utiliza para evaluar y elegir futuros cursos de acción. Y aquí entra en escena el proceso de toma de decisiones¹² y la evaluación, que consiste en *seleccionar, recortar y organizar la información disponible*¹³.

A partir de la inteligencia humana, se han desarrollado múltiples y diversas innovaciones tecnológicas. La que aquí nos ocupa tiene que ver con el procesamiento de la información para resolver problemas y tomar decisiones a partir de máquinas o de los llamados algoritmos inteligentes¹⁴. La inteligencia artificial (en adelante, IA) se

³ Ver Gardner, Howard, *La inteligencia reformulada*, Madrid, Paidós, 2010, p. 115.

⁴ Ver Kayser, Daniel, *Diccionario de ciencias cognitivas*, Bs. As., Amorrortu, 2003, p. XXIII.

⁵ Ver Gardner, Howard, *Las cinco mentes del futuro*, Bs. As., Paidós, 2013, p. 17; *Estructuras de la mente: la teoría de las inteligencias múltiples*, México, Fondo de Cultura Económica, 1987.

⁶ Ver Bachrach, Estanislao, *En cambio*, 5ª ed., Bs. As., Sudamericana, 2015, p. 395 y 396.

⁷ Ver Manes, Facundo - Niro, Mateo, *Usar el cerebro*, Bs. As., Planeta, 2017, p. 115.

⁸ Ampliar en Gardner, *La inteligencia reformulada*, p. 62 y siguientes.

⁹ Howard Gardner concibe a la inteligencia como un “potencial biopsicológico para procesar la información que se puede activar en un marco cultural para resolver problemas o crear productos que tienen valor para una cultura” (el destacado me pertenece). Y luego el mismo autor afirma que: “Las inteligencias son potenciales neurales que se activan o no en función de los valores de una cultura determinada, de las oportunidades disponibles en esa cultura y de las decisiones tomadas por cada persona y/o su familia, sus enseñantes y otras personas” (*La inteligencia reformulada*, p. 52 y 53). El director de ingeniería de Google, Ray Kurzweil, define la inteligencia como la “capacidad para resolver problemas con recursos limitados, incluyendo limitaciones de tiempo” (*La singularidad está cerca*, p. 339, Berlín, Lola Books, 2012). Manes y Niro aluden a la inteligencia fluida como el conjunto de recursos con los que cuenta un individuo para adaptarse al medio o la “capacidad para resolver problemas nuevos descubriendo las relaciones que existen entre las cosas e independientemente del conocimiento adquirido a lo largo de la vida” (Manes - Niro, *Usar el cerebro*, p. 114 y 115).

¹⁰ Benítez, Raúl - Escudero, Gerard - Kanaan, Samir - Masip Rodó, David, *Inteligencia artificial avanzada*, Barcelona, UOC, 2013, p. 14.

¹¹ Desde un punto de vista biológico, el ADN es un portador esencial de la información genética. Gerard, Michael - Gerald, Gloria E., *El libro de la biología*, Madrid, Ilus Books, 2015, p. 354.

¹² Sobre las cuestiones acerca del proceso de toma de decisiones en las personas, véase Kahneman, Daniel, *Pensar rápido, pensar despacio*, Barcelona, Debate, 2012, p. 569 y 590.

¹³ Ampliar en Manes, Facundo - Niro, Mateo, *El cerebro argentino*, Bs. As., Planeta, 2016, p. 269, 270, 274, 275 y 301, y de los mismos autores, *Usar el cerebro*, p. 130.

¹⁴ Es importante aclarar dos cuestiones que intuitivamente se suelen pensar cuando se analiza hasta qué punto los robots o las computadoras pueden ser “inteligentes”. En primer lugar, se le quita mérito a la inteligencia artificial porque se sostiene que resulta imposible reproducir el cerebro humano en máquinas, dada su complejidad y la ausencia de datos fiables que permitan determinar a ciencia

sustenta en algoritmos inteligentes¹⁵ o en algoritmos de aprendizaje que, entre muchos otros fines, se utilizan para identificar tendencias económicas o recomendaciones personalizadas¹⁶. Un algoritmo puede ser definido como un conjunto preciso de instrucciones¹⁷ o reglas¹⁸, o como una serie metódica de pasos que puede utilizarse para hacer cálculos, resolver problemas y tomar decisiones¹⁹.

La lógica de estas nuevas tecnologías esencialmente se explica por el proceso de “algoritmizar”²⁰ el procesamiento de datos o de información, a fin de sustituir o mejorar las actividades o tareas humanas que son fruto de su inteligencia. Y desde esta plataforma argumental, podemos plantear tres cuestiones muy importantes a tener en cuenta: 1) hay consenso en afirmar que los organismos, y por tanto los seres humanos, son algoritmos²¹ complejos y refinados que operan mediante sensaciones, emociones/deseos y pensamientos²²; 2) el proceso de innovación en materia tecnológica sostenidamente incrementa su capacidad de volver “algoritmizables” los procesos

cierta su funcionamiento integral. Si bien esto último es correcto, lo cierto es que esta objeción soslaya el hecho de que la inteligencia artificial no tiene que parecerse al cerebro humano para realizar exitosamente ciertas actividades que sólo podían atribuirse a la inteligencia humana.

Además, es importante tener en cuenta que el desarrollo de la IA no consiste –al menos exclusivamente– en imitar o copiar el cerebro humano. La IA utiliza diversos métodos para procesar la información y resolver problemas o tomar decisiones, del mismo modo –si se me concede la analogía– en que los ingenieros en aviación no imitaron o copiaron el método y/o las técnicas de aprendizaje de los pájaros para construir los aviones modernos. En segundo lugar, muchos avances tecnológicos en donde está presente la IA se naturalizan. Si una máquina inteligente comienza a realizar determinadas actividades en las que iguala o mejora la capacidad de procesamiento de información de un cerebro humano, muchas veces se suele afirmar que eso es una cuestión de computación y que no es una verdadera “inteligencia”. Si hoy Siri de Apple nos da respuestas inteligentes, solemos pensar que no es “tan” inteligente porque falla a menudo, o bien porque no es capaz de reconocer lo que expresamos, más allá de que en otros casos brinde respuestas eficientes en menos de un segundo. Lo importante aquí, para poder dimensionar el desafío de regular esta nueva tecnología, es comprender que, así como el cerebro extrae, selecciona, recorta y organiza la información disponible para tomar decisiones, la inteligencia artificial hace lo mismo, con otros métodos y a otra velocidad. Ampliar en Bostrom, Nick, *Superinteligencia*, 2ª ed., España, Tell, 2016, p. 29, y Kurzweil, *La singularidad está cerca*, p. 302.

¹⁵ Ampliar en Domingos, Pedro, *The master algorithm: how the quest for the ultimate learning machine will remake our world*, New York, Basic Books, 2015, p. 1 y ss.; Harari, Yuval N., *Homo Deus*, Bs. As., Debate, 2016, p. 99 a 107; Borruso, Renato, *La ley, el juez, la computadora. Un tema fundamental de la informática jurídica*, “Informática y Derecho”, vol. 5, p. 30 y siguientes.

¹⁶ Ampliar en Palma Méndez, José T. - Marín Morales, Roque, *Inteligencia artificial*, Madrid, Mc Graw-Hill, 2011, p. 683, y Abu-Mostafa, Yaser S., *Técnicas de aprendizaje automático*, “Investigación y Ciencia”, 2013, p. 50 a 53.

¹⁷ Deutsch, David - Ekert, Artur, *Más allá del horizonte cuántico*, “Investigación y Ciencia”, 2012, p. 79.

¹⁸ Benítez - Escudero - Kanaan - Masip, *Inteligencia artificial avanzada*, p. 13.

¹⁹ Ampliar en Domingos, *The master algorithm: how the quest for the ultimate learning machine will remake our world*, p. 1.

²⁰ Borruso acuñó el término “algoritmizable” a fin de trazar una analogía respecto de la computadora como una “máquina algorítmica universal”. Borruso, *La ley, el juez, la computadora. Un tema fundamental de la informática jurídica*, p. 35.

²¹ Harari, *Homo Deus*, p. 350; Kraemer, Felicitas - van Overveld, Kees - Peterson, Martin, *Is there an ethics of algorithms?*, “Ethics and Information Technology”, vol. 13, Netherlands, Springer, 2011, p. 251 a 260.

²² Según Daniel Kahneman, el 99% de nuestras decisiones las toman refinadísimos algoritmos que llamamos sensaciones, emociones y deseos. *Pensar rápido, pensar despacio*, Barcelona, Debate, 2012. En cambio, las máquinas expendedoras funcionan a través de engranajes mecánicos y circuitos eléctricos. Harari, *Homo Deus*, p. 101.

cognitivos²³; 3) en múltiples campos, los algoritmos se vuelven cada vez más precisos y eficientes que los seres humanos²⁴.

Ahora bien, este proceso creciente e incesante de “algoritmizar” la inteligencia, se vincula con la explosión de los datos y de la información²⁵. Para una mejor comprensión de este fenómeno, es preciso destacar dos grandes cuestiones relacionadas con el concepto de IA y con el rol que cumplen. *Primero*. En sentido amplio, cuando hablamos de inteligencia artificial, nos referimos a las “estrategias de programación y máquinas físicas para reproducir de la forma más eficiente y completa posible las tareas cognitivas y científico-técnicas llamadas ‘inteligentes’”²⁶. La IA desarrolla ciertos procesos de forma autónoma, muchas veces sin imitar o simular el proceso que se desarrolla en el cerebro humano, pero obteniendo iguales o mejores resultados en ciertos campos o áreas de conocimiento.

Segundo. Las máquinas inteligentes cumplen un rol cada vez más determinante en la explosión de información y de datos en la que estamos inmersos. La IA resulta esencial porque optimiza el proceso de extracción útil de información o de los datos. La “información” es “conocimiento” en la medida en que las actividades inteligentes puedan encontrar patrones relevantes²⁷. Y aquí llegamos a un punto central que ilustraremos con un ejemplo. Google no necesita saber la estructura gramatical de los diferentes idiomas para realizar traducciones medianamente razonables, que al menos suponen una comprensión básica de información escrita en otro idioma. Aprende de los patrones que extrae de la información y los datos. En conclusión, hay varios caminos no excluyentes entre sí por los cuales la inteligencia artificial iguala, mejora, potencia y/o maximiza los resultados de muchas actividades humanas inteligentes, a través de la optimización continua y exponencial del procesamiento de información.

3. Inteligencia artificial basada en un enfoque de derechos humanos

La inteligencia artificial supone una innovación tecnológica de tal magnitud que, a principios de este año, más de dos mil quinientos expertos (entre ellos, Stephen Hawking²⁸) establecieron el siguiente principio: “Los riesgos que plantean los sistemas de IA, en especial los riesgos catastróficos y existenciales, deberían estar sujetos a esfuerzos de planificación y mitigación, acordes con el impacto esperado. Y en mayor

²³ Borruso, *La ley, el juez, la computadora. Un tema fundamental de la informática jurídica*, p. 35.

²⁴ Kahneman, *Pensar rápido, pensar despacio*, p. 291 y siguientes. En un estudio, un algoritmo informático diagnosticó el 90% de los casos de cáncer de pulmón que se le presentaron, mientras que los médicos sólo acertaron en el 50%. Harari, Yuval N., *Homo Deus*, p. 347, y Tzezana, Roey, *The Guide to the Future*, Haifa, 2013, p. 62 a 64.

²⁵ El siguiente ejemplo permite tomar una acabada dimensión de esta cuestión. El 29 mayo de 2017, en un 1 minuto se enviaron 2.432.440.845 (casi dos mil quinientos millones) de emails, se postearon 458.090 tweets, se publicaron 63.980 fotos en Instagram, se realizaron 3.629.947 de búsquedas en Google, y en la web se procesaron 2.702.994 giga bytes (fuente: Internet Live Stats www.internetlives-tats.com/one-second/#instagram-band).

²⁶ Ampliar en Mira, José, *Aspectos conceptuales de la inteligencia artificial y la ingeniería del conocimiento*, en Méndez Palma, José T. - Morales Marín, Roque, “Inteligencia artificial”, p. 3 a 9; en similar sentido, Ruiz, Francisco E. - Quevedo, Miguel A. C. - Galipienso, María I. A. - Pardo, Otto C. - Lozano Ortega, Miguel A., *Inteligencia artificial*, Madrid, Thomson, 2003, p. 4.

²⁷ Kurzweil, *La singularidad está cerca*, p. 426.

²⁸ Entre muchos otros, suscribieron los principios: Elon Musk, Royal Martin Rees, Smith-Zadeh y Stuart Russell, Sam Harris, Nick Bostrom, Raymond Kurzweil, David Chalmers y Demis Hassabis.

medida, debería estar sujeta a estrictas medidas de seguridad y control”²⁹. Pero además, en el corto plazo existen múltiples desafíos para asegurar la compatibilidad del desarrollo de la IA con el derecho doméstico de los Estados y con el derecho internacional vigente. Veamos algunos problemas que actualmente se generan a partir del uso de algoritmos inteligentes. *Primero*. La IA en múltiples casos no puede ofrecer una explicación detallada acerca de cómo se llega a un determinado resultado. Es decir, existe un déficit de motivación por parte del algoritmo. En ciertos supuestos se desconoce cómo razona la interacción entre la información y los datos para tomar decisiones, establecer resultados o realizar predicciones. O dicho de otro modo, no puede establecerse cómo el algoritmo evalúa y pondera los datos y la información que procesa, lo que genera el fenómeno llamado “caja negra”³⁰. *Segundo*. En el caso de ciertas IA predictivas, la forma o el método por el cual se procesa la información y los datos puede violar el principio de igualdad y de no discriminación, porque las predicciones se basan en un código fuente que toma en consideración distinciones de raza, género, entre otras. Por ejemplo, en el caso “State vs. Loomis”³¹ el apelante sostuvo que el algoritmo inteligente usaba incorrectamente las evaluaciones de género³². *Tercero*. Quienes desarrollan IA suelen ampararse en el secreto comercial y en los derechos de patentamiento.

Podríamos hacer una larga lista de problemas, pero estas breves muestras ponen en evidencia la complejidad y dificultad de abordar el fenómeno de la IA. Como no podremos profundizar en todos estos aspectos en este artículo, nuestra propuesta comienza por los derechos humanos. Desde la teoría del derecho se afirma que los derechos de las personas humanas se definen por cinco notas características: 1) universales, 2) fundamentales, 3) abstractos o indeterminados³³, 4) morales –existencia y validez moral– y, 5) prioritarios³⁴. El “modelo de derechos humanos” se cristaliza a partir de un “paradigma protectorio” que surge de los pactos internacionales, y que en esencia se basa en asignar los siguientes rasgos a los derechos esenciales de la

²⁹ Asilomar AI Principles. Future of Life Institute, disponible en <https://futureoflife.org/ai-principles>.

³⁰ Se llama sistema de “caja negra” a la herramienta computacional en la que uno entiende los datos ingresados y los resultados, pero no entiende el procedimiento subyacente. Aquí el código es inescrutible porque el programa “evoluciona” y los humanos no pueden entender el proceso que siguió la programación para lograr una solución determinada. Véase, Barrat, James, *Nuestra invención final*, México, Paidós, 2014, p. 92.

³¹ Corte Suprema de Wisconsin, 13/7/16, “Estado de Wisconsin v. Eric L. Loomis”, disponible en www.wicourts.gov/sc/opinion/DisplayDocument.pdf?content=pdf&seqNo=171690.

³² Ampliar en Corte Suprema de Wisconsin, 13/7/16, “Estado de Wisconsin v. Eric L. Loomis”, véase especialmente considerandos 17, 28, 34, 51, 93 y 94.

³³ Beitz, Charles R., *La idea de los derechos humanos*, Madrid, Marcial Pons, 2012, p. 244. En similar sentido, en Latinoamérica, véase Rincón Córdoba, Jorge I., *Las generaciones de los derechos fundamentales y la acción de la Administración pública*, 2ª ed., Universidad Externado de Colombia, 2004, p. 335 y 336.

³⁴ Alexy, Robert, *La doble naturaleza del derecho*, Madrid, Trotta, 2016, p. 60 y 61.

persona humana: igualdad³⁵, y su carácter “inalienable”³⁶ o “inderogable”³⁷. Existe una vinculación directa entre estos derechos, la dignidad humana, la paz, la protección de las minorías, de los más vulnerables o de los más débiles³⁸. Es un esquema por el que se obliga a los Estados y a la comunidad internacional³⁹ a garantizar la efectividad de derechos, principios y reglas que se encuentran plasmados en las constituciones, en los pactos internacionales y en las leyes domésticas⁴⁰. Sobre esta plataforma, para que el desarrollo de la IA sea compatible con un “modelo de derechos humanos”, es preciso impulsar una regulación que incorpore una serie de principios que enunciemos a continuación⁴¹.

4. Principios jurídicos para una inteligencia artificial que respete y promueva la vigencia del Estado constitucional y la efectividad de los derechos humanos

Encastrar a la IA en el marco del Estado constitucional presupone enfocar su diseño, desarrollo y uso a partir de respetar la dignidad humana y los derechos humanos. Se trata, básicamente, de hacer más robusto el sistema de protección a partir de incorporar al bloque de juridicidad una serie de principios generales tendientes a regularla. Veamos.

³⁵ Ver, Declaración Americana de los Derechos y Deberes del Hombre, aprobada en la IX Conferencia Internacional Americana, en Bogotá, Colombia, 1948, primer párrafo del Preámbulo; asimismo ver, Preámbulo, considerando primero, Pacto Internacional de Derechos Económicos, Sociales y Culturales, adoptado y abierto a la firma, ratificación y adhesión por la Asamblea General en su resolución 2200 A (XXI), de 16 de diciembre de 1966; Preámbulo, considerando primero del Pacto Internacional de Derechos Civiles y Políticos, aprobado en Nueva York, Estados Unidos de América, el 19 de diciembre de 1966; Preámbulo, considerando primero, párrafo primero del Convenio sobre los Derechos del Niño, adoptada por la Asamblea General de las Naciones Unidas el 20 de noviembre de 1989; considerando primero de la Convención sobre la Eliminación de todas las Formas de Discriminación contra la Mujer (también conocida por sus siglas en inglés CEDAW), aprobada por la Asamblea General de las Naciones Unidas el 18 de diciembre de 1979 y considerando primero de la Convención contra la Tortura y otros Tratos o Penas Cruelles, Inhumanos o Degradantes, aprobada por la Asamblea General de Naciones Unidas el 9 de diciembre de 1975.

³⁶ Ver considerando primero, párrafo primero de la Convención contra la Tortura y otros Tratos o Penas Cruelles Inhumanos o Degradantes y el Preámbulo, considerando primero, párrafo primero del Convenio sobre los Derechos del Niño.

³⁷ Ver segundo considerando, primer párrafo de la Convención Interamericana sobre Desaparición Forzada de Personas.

³⁸ Ferrajoli, Luigi, *Sobre los derechos fundamentales*, en “Teoría del neoconstitucionalismo”, Madrid, Trotta, 2007, p. 73 a 75.

³⁹ La Declaración del Milenio de la ONU afirma que se deben realizar esfuerzos para lograr “la plena protección y promoción de los derechos civiles, políticos, económicos, sociales y culturales de todas las personas en todos nuestros países... Aumentar en todos nuestros países la capacidad de aplicar los principios y las prácticas de la democracia y del respeto de los derechos humanos, incluidos los derechos de las minorías” (punto 25), disponible en www.un.org/es/comun/docs/?symbol=A/RES/55/2.

⁴⁰ Sin embargo, como afirma Charles Beitz, no resulta plausible encontrar un fundamento único o formular una lista de derechos (*La idea de los derechos humanos*, p. 141, 142 y 244).

⁴¹ Es importante considerar el enfoque de la Organización de Naciones Unidas, cuando afirma que todos deben “aprovechar los beneficios de las nuevas tecnologías, en particular de las tecnologías de la información y de las comunicaciones, conforme a las recomendaciones formuladas en la Declaración Ministerial 2000 del Consejo Económico y Social”, resolución aprobada por la Asamblea General 55/2, Declaración del Milenio, apartado III, punto 20, ítem quinto.

Prevención/precaución. Estos principios constituyen dos funciones distintas, con un denominador común: la necesidad de actuar antes de la producción de un daño. En una síntesis radical, se desempeñan sobre distintos tipos de riesgos. Al riesgo potencial, la precaución. Al riesgo verificado corresponde la prevención⁴². El principio precautorio en la IA (de un modo análogo a lo que acontece en el derecho ambiental) se vincula con una falta total o absoluta de certeza científica acerca de la ausencia de riesgos, que en el caso de algoritmos inteligentes debe, al menos, estar asociada a las siguientes circunstancias: 1) un código fuente cerrado o la existencia de un sistema en el que uno entiende los datos ingresados y los resultados, pero no se puede inferir el procedimiento subyacente (“caja negra”); 2) la ausencia de una trazabilidad; 3) la imposibilidad de asegurar un “botón de apagado” o un mecanismo seguro de contención de la IA; 4) cuando la IA afecta de manera actual o potencial los derechos fundamentales de los seres humanos, los derechos de incidencia colectiva, los servicios públicos, los servicios esenciales o los servicios universales; 5) cuando en cualquier fase –diseño, desarrollo o aplicación– se adviertan sesgos relevantes o bien, que no se pueda garantizar la transparencia o el principio de no discriminación.

Autodeterminación algorítmica. La autodeterminación es un derecho fundamental que se deriva de la dignidad de la persona humana. Se trata de asegurar el “libre desarrollo de la personalidad”, a partir de reconocer la autodeterminación informativa que se orienta a garantizar el derecho a elegir –asociado a la libertad de información– el “derecho a saber”, al “conocimiento” y a la “autorregulación de la información”⁴³. Sobre esta base, los Estados y la comunidad internacional, responsablemente deben invertir, y desplegar los máximos esfuerzos de toda índole, para que se pueda garantizar la autodeterminación humana frente al uso de algoritmos inteligentes. Como la IA intermedia cada vez más entre los datos/información y las decisiones de las personas, resulta indispensable proteger sus derechos a partir de promover el respeto de los principios de necesidad, finalidad, proporcionalidad y pertenencia de los datos personales⁴⁴.

Transparencia algorítmica. El diseño, desarrollo y uso de la inteligencia artificial debe ser transparente y abierto. La transparencia algorítmica se vincula con la prohibición de la existencia de “cajas negras” en los algoritmos o la existencia de fallas, frente a daños o lesiones que puedan causar⁴⁵. Es decir, la inteligencia artificial debe ser “transparente” en sus decisiones, lo que significa que se pueda inferir o deducir una “explicación entendible” acerca de los criterios en que se basa para arribar a una determinada conclusión, sugerencia o resultado⁴⁶.

⁴² Allende Rubino, Horacio L., *La acción de prevención en el Código Civil y Comercial. Su relación con el principio de precaución en el derecho ambiental*, Microjuris online, 2016, cita: MJ-DOC-9989-AR | MJD9989.

⁴³ Pitschas, Rainer, *Derecho administrativo de la información*, en “Innovación y reforma en el derecho administrativo”, Sevilla, Global Law Press, 2ª ed., 2012, p. 226, 227 y 236. Este autor habla de un cambio de paradigma del derecho administrativo de la información, en donde cada individuo ha de estar en condiciones de decidir bajo su propia responsabilidad y autonomía entre las posibilidades y los riesgos que genera la libertad de comunicación (p. 236).

⁴⁴ Rodotà, Stefano, *El derecho a tener derechos*, Madrid, Trotta, 2014, p. 182.

⁴⁵ Como se afirma en uno de los 23 principios de Asilomar, si un sistema de IA causa daño, debería ser posible determinar por qué (principio 7, denominado transparencia en la falla).

⁴⁶ La transparencia incluye lo que podemos llamar “motivación algorítmica”. La motivación algorítmica puede entenderse como la descripción, en lenguaje formal e inteligible, de la arquitectura del

Trazabilidad de la inteligencia artificial. La trazabilidad o rastreabilidad es la “apetitud para rastrear la historia, la aplicación o la localización de una entidad mediante indicaciones registradas”⁴⁷. Una IA basada en un enfoque de derechos humanos debe poder explicar, paso a paso, las operaciones técnicas que realiza desde el inicio hasta el fin de un proceso determinado. Como regla, se debe garantizar la inteligibilidad y la trazabilidad del proceso de toma de decisiones de los algoritmos inteligentes. Y como contrapartida, se debe impedir la configuración de “cajas negras” en la IA. La trazabilidad contribuye a garantizar los principios de transparencia y de motivación algorítmica.

Máximo acceso. Derecho de acceso a la información algorítmica. El Estado debe asegurar el acceso al código fuente y a los sistemas que sustentan los algoritmos inteligentes, cuando de ellos pueda resultar una violación a la Constitución Nacional, a los tratados internacionales y a las normas vigentes. Cuando el Estado y las personas públicas no estatales, por sí o a través de terceros, diseñan, desarrollan o utilizan tecnologías de la información o comunicación sustentadas en IA o en algoritmos inteligentes (lo que involucra cualquier tipo de máquina o robot inteligente), deben garantizar el máximo acceso al sistema de procesamiento de información que esas tecnologías realizan⁴⁸.

Principio de no discriminación algorítmica. El diseño y/o implementación de los algoritmos inteligentes deben respetar el principio de no discriminación, que consiste en impedir que las máquinas inteligentes –basadas en sistemas de IA– procesen la información o los datos bajo sesgos o distinciones frente a los seres humanos, por motivos de raza, color, sexo, idioma, religión, opinión política o de otra índole, origen nacional o social, posición económica, nacimiento o cualquier otra condición social (art. 2, inc. 2, Pacto de los Derechos Económicos Sociales y Culturales).

5. Conclusión

Impulsar un marco normativo regulatorio vinculado al desarrollo de la IA, puede asemejarse al fenómeno que se dio a partir del surgimiento de las tecnologías de la información y de la comunicación (TIC) en el ámbito de la Organización de Naciones Unidas (ONU)⁴⁹. En definitiva, esta innovación tecnológica se produce por la optimización artificial del procesamiento de información.

diseño computacional; se trata de que el algoritmo sea capaz de explicitar y hacer inteligible cómo, qué y cuáles son las cuestiones fácticas, jurídicas y/o éticas relevantes que se tienen en cuenta para adoptar tal o cual decisión. La motivación algorítmica permite la auditabilidad humana también se vincula con la regla de la trazabilidad de los sistemas de IA y con el principio de no discriminación algorítmica, que presupone la prohibición de sesgos de máquina.

⁴⁷ Definición según la norma ISO 8402, complemento de la serie de normas ISO 9000.

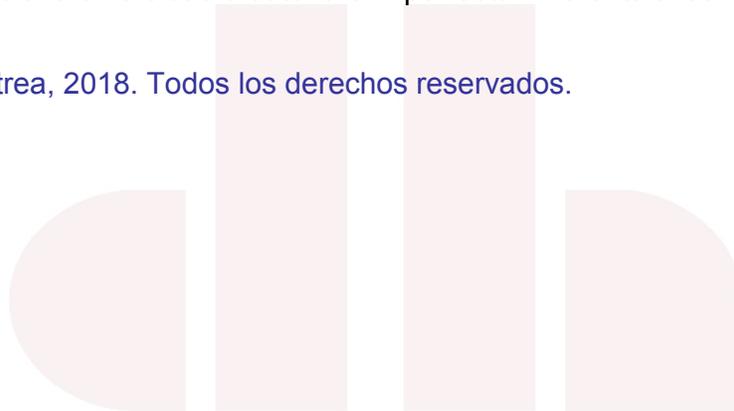
⁴⁸ El mismo acceso debe garantizarse respecto de toda persona humana o jurídica pública o privada, vinculada con fines públicos o fondos públicos recibidos, que haga uso de esas tecnologías, siempre que el diseño o uso de la inteligencia artificial se relacione con los fines públicos o con los fondos públicos recibidos.

⁴⁹ Por ejemplo, desde la UNESCO se incentiva a la creación de un entorno legislativo propicio en el ámbito de las TIC. En la misma línea, la ONU sostiene que hay que apoyar el desarrollo de las tecnologías, la investigación e innovación nacional, garantizando un entorno normativo propicio a la diversificación industrial y la adición de valor a los productos básicos. Además, los Estados se deben abstener de utilizar la tecnología de la información y las comunicaciones en contravención del derecho

A nuestro entender, la tarea de hacer compatible la IA con los derechos humanos, se presenta como el desafío más complejo e indispensable de los próximos años, si tenemos en cuenta la configuración incipiente de dos escenarios: por un lado, cómo garantizar la intervención humana frente a las decisiones o predicciones de algoritmos inteligentes, respetando los principios desarrollados en el punto anterior. Por otro lado, si los expertos en protección de datos están reflexionando acerca de la intervención humana en relación con los algoritmos⁵⁰, entonces, hay que trabajar sobre la siguiente cuestión: *cuánta* intervención del ser humano resulta necesaria para que el resultado del procesamiento de información y de los datos sea legítimo, respetuoso y promotor de la efectividad de los derechos de las personas.

A modo prospectivo, creemos que los desafíos que presenta la IA tienen que ver con nuestra identidad como especie. Si las personas humanas nos caracterizamos por la *diversidad, aleatoriedad e imperfección*⁵¹, estamos ingresando a una era de automatización que podría poner en crisis esos rasgos. Aunque suene improbable, en un futuro no muy lejano, resultará indispensable pensar seriamente en garantizar un derecho fundamental, que podría ser la piedra basal de la era de la inteligencia artificial: el derecho a la diversidad aleatoria e imperfecta inherente al ser humano.

© Editorial Astrea, 2018. Todos los derechos reservados.



internacional. Véase, Asamblea General, resolución n° A/71/307, 5/8/16, considerando 8, p. 4; Asamblea General, resolución n° A/RES/70/1, 21/10/15, considerando 9.b, p. 23 y Asamblea General, resolución n° A/RES/71/101 A-B, 23/12/16, considerando 4, p. 4.

⁵⁰ Rodotà, *El derecho a tener derechos*, p. 302.

⁵¹ Kurzweil, *La singularidad está cerca*, p. 167 y 168.