

Cuatro ángulos de análisis de la igualdad y la no discriminación en la inteligencia artificial

El documento establece qué es la inteligencia artificial (IA) y plantea una base de entendimiento sobre los alcances de la igualdad y la no discriminación como derechos humanos. Propone y desarrolla un análisis de los efectos positivos y negativos de la IA en la igualdad y no discriminación a partir de cuatro ángulos fundamentales que responden a las preguntas: dónde, quién, para qué y cómo se hace la IA.

Autor: Karlos Castilla

Fecha: Febrero 2022

Edición:
Institut de Drets Humans de Catalunya
Av. Meridiana 32, entr. 2a. Esc B
08018 Barcelona
www.idhc.org

Diseño y maquetación: nadianmartin.com



Esta obra está bajo una licencia de Creative Commons Reconocimiento-No-Comercial 4.0 Internacional. Se puede copiar, distribuir, comunicar públicamente, traducir y modificar, siempre que sea para fines no comerciales y se reconozca su autoría.



El contenido de esta publicación es responsabilidad exclusiva del Institut de Drets Humans de Catalunya y no refleja necesariamente la opinión de la Open Society Initiative for Europe.

Índice

I.	Introducción	4
II.	Inteligencia artificial, algoritmos y datos	8
III.	Bases generales de la igualdad y la no discriminación	14
IV.	Cuatro ángulos para observar la igualdad y la no discriminación en la inteligencia artificial	19
	1. Primer ángulo: dónde	21
	2. Segundo ángulo: quién	22
	3. Tercer y cuarto ángulo: para qué y cómo	24
V.	Conclusiones	33

I.

Introducción

Lo primero que debemos hacer para hablar de igualdad, no discriminación e inteligencia artificial es reconocer que, es un tema que en sí mismo excluye a millones de personas en el mundo, en la medida en que para funcionar la inteligencia artificial requiere en la gran mayoría de los casos de una conexión a Internet y, según datos de la Unión Internacional de Telecomunicaciones¹, para el año 2021, solo el 63% de la población mundial tenía acceso a Internet, con lo que al menos 2.900 millones de personas en el mundo aún están desconectadas².

El problema de ese porcentaje de conexión y cifras de desconexión es que no están distribuidas de manera equitativa en todo el mundo. Así, por ejemplo, vemos que a nivel mundial, las personas de zonas urbanas tienen el doble de probabilidades de utilizar Internet que las de zonas rurales (el 76% de las zonas urbanas frente al 39% de las zonas rurales) o que a nivel mundial, un promedio del 62% de los hombres usa Internet en comparación con el 57% de las mujeres, aunque la brecha digital de género se amplía en los países más pobres³.

De esa manera, hablar de inteligencia artificial y lo que esta requiere para funcionar es en sí mismo hablar de desigualdad en el mundo. Bien se podría afirmar que el tema que nos ocupa en este documento es actualmente “un problema de ricos”, un problema de quienes vivimos en el conocido como “Norte global”, especialmente si tenemos en cuenta que desde hace algunos años se observa que en África solo el 22% de la población utiliza Internet, mientras que en Europa es el 80%⁴. Esto es, que en los llamados “países menos adelantados”, solo 1 de cada 5 personas está conectada a Internet, mientras que son 4 de cada 5 en los llamados “países desarrollados”⁵. Con lo que hay gran-

5

1 Es una agencia especializada del sistema de Naciones Unidas surgida en 1865 para armonizar las comunicaciones telegráficas, posteriormente telefónicas y más recientemente para las tecnologías de la información y la comunicación.

2 International Telecommunication Union, *Statistics*, disponible en: <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx> [Consultado el 7 de enero de 2022].

3 International Telecommunication Union, *Press Release*, 30 de noviembre de 2021, disponible en: <https://www.itu.int/en/mediacentre/Pages/PR-2021-11-29-FactsFigures.aspx> [Consultado el 7 de enero de 2022].

4 United Nations Conference on Trade and Development, *Trade and Development Report 2018: Power, Platforms and the Free Trade Delusion*, United Nations, 2018, p. viii, disponible en: https://unctad.org/system/files/official-document/tdr2018_en.pdf

5 United Nations Conference on Trade and Development, *Digital Economy Report 2019: Value Creation and Capture: Implications for Developing Countries*, United Nations, 2019, p. 13, disponible en: https://unctad.org/system/files/official-document/der2019_en.pdf

des sectores de la población mundial para los cuales la inteligencia artificial es irrelevante en el ejercicio de sus derechos, pues ni siquiera tienen acceso a los mínimos requeridos para que esta funcione.

No se trata simplemente de tener conectividad a través de una línea telefónica [eléctrica o de otro tipo], sino de que, para que la conexión sea realmente efectiva y accesible, es necesario que los servicios de conexión también se extiendan y universalicen a un precio que no resulte prohibitivo para [las personas usuarias] comunes⁶. Situación que en muchas regiones del mundo es un problema agregado al primero de inexistencia de conectividad.

Sin olvidar también que la falta de neutralidad de la red⁷, atacada en muchos lugares del mundo, se suma a todas las situaciones que están haciendo más compleja la pretendida universalización del Internet.

Si bien es cierto que ya hay avances para que la inteligencia artificial pueda funcionar sin una conexión a Internet⁸ y así pueda ser utilizada por un mayor número de personas, esos avances se están haciendo donde hoy hay mayor conectividad a Internet, con lo que por ahora la desigualdad en ese acceso se mantiene.

Un ejemplo claro de lo que venimos señalando es Cataluña⁹, que es uno de los territorios mejor situados mundialmente, pues los avances tecnológicos confluyen con un sistema de investigación y desarrollo robusto y maduro respecto a las tecnologías digitales¹⁰, y eso se corresponde con el hecho de que de acuerdo con el Instituto de Estadística de Cataluña, en 2019 el 94% de los

6

6 Sabanes, D., "Tecnologies de la informació i de la comunicació per a la inclusió i la participació en la societat de la informació i del coneixement", en *Efectes de les tecnologies de la informació i la comunicació sobre els drets humans*, IDHC, Barcelona, 2009, p. 69 (original en catalán, traducción propia del autor).

7 La conocida como *Net Neutrality* es un principio que garantiza la igualdad de acceso a contenidos y evita que haya contenidos de primera y de segunda clase. Se basa en que los proveedores de Internet deben dar acceso a los contenidos y máquinas conectadas sin privilegiar a una participante de la red por encima de otras. Para más información, véase: Bartra, Joan, "El concepto de net neutrality y la tensión entre regulación pública y autoregulación privada de las redes", en *VII Congreso Internacional Internet, Derecho y Política. Neutralidad de la Red y otros retos para el futuro de Internet*, (monográfico en línea), IDP, Revista de Internet, Derecho y Política, N.º 13, 2012, pp. 44-52, disponible en: <http://idp.uoc.edu/ojs/index.php/idp/article/view/n13-barata/n13-barata>

8 Waterloo Artificial Intelligence Institute, *Artificial Intelligence Without Internet Now Possible*, noviembre 2017, disponible en: <https://uwaterloo.ca/artificial-intelligence-institute/news/artificial-intelligence-without-Internet-now-possible>

9 El gobierno de Cataluña ha creado la *Artificial Intelligence Research Alliance (AIRA)*, que tiene como objetivo impulsar la investigación científica, la gestión del talento y acelerar el desarrollo de soluciones basadas en inteligencia artificial en Cataluña.

10 Universitat Pompeu Fabra, *Revista 360upf*, número 9, octubre 2021, disponible en: <https://www.upf.edu/web/360upf/numero9/presentacio>

hogares tenía conexión a Internet¹¹. Información que se confirma en 2021 por el Instituto Nacional de Estadística que indica que el 93,9% de la población de 16 a 74 años ha utilizado Internet¹².

De esta manera tenemos, por una parte, grandes desigualdades en el mundo en el acceso a Internet y, con ello, a todo lo que (como más adelante veremos) puede representar la inteligencia artificial y, por otra parte, lugares en donde el acceso a dicha inteligencia podría ser casi “universal” y, por tanto, la exposición a esta puede ser muy alta e incidir en el ejercicio de nuestros derechos a la igualdad y a la no discriminación.

Bajo esa premisa, en este documento solo analizaremos los aspectos en los cuales la igualdad y la no discriminación se pueden ver afectadas y beneficiadas por la inteligencia artificial. La desigualdad y discriminación en el acceso a Internet y, como consecuencia de ello, a la inteligencia artificial, por ahora, simplemente lo dejamos expresamente manifestado, pero no nos ocuparemos a fondo de ello.

Para eso, en primer lugar, estableceremos lo qué es o puede considerarse como inteligencia artificial, poniendo especial énfasis en lo relacionado con los algoritmos y los datos a partir de los cuales funciona, buscando explicarlos de una manera sencilla que sea accesible a personas que no cuentan con los conocimientos técnicos, pero sin desvirtuar su naturaleza.

En segundo lugar, desarrollaremos algunos de los conceptos básicos sobre igualdad y no discriminación que deben tenerse presentes al analizar los efectos que en estos puede tener la inteligencia artificial, a partir de las bases teóricas que se han trabajado desde hace tiempo en el Institut de Drets Humans de Catalunya (IDHC).

En tercer lugar, a partir de los dos aspectos anteriores, desarrollaremos un análisis particular de la igualdad y la no discriminación en la inteligencia artificial. Esto es, mostraremos dónde están los riesgos y beneficios que puede generar la inteligencia artificial en la igualdad y no discriminación, no de manera genérica, sino tras analizar los aspectos concretos de los algoritmos, datos y fórmulas con las que se desarrolla actualmente la inteligencia artificial. Finalmente, estableceremos algunas conclusiones generales.

11 Instituto de Estadística de Cataluña, *Equipamiento TIC en el hogar. Comarcas y Aran, y ámbitos*, disponible en: <https://www.idescat.cat/pub/?id=aec&n=620&lang=es>. Consultado el 7 de enero de 2022.

12 Instituto Nacional de Estadística, *Población que usa Internet (en los últimos tres meses). Tipo de actividades realizadas por Internet*, disponible en: https://www.ine.es/ss/Satellite?L=es_ES&c=INESeccion_C&cid=1259925528782&p=1254735110672&page-name=ProductosYServicios%2FPYSLayout. Consultado el 7 de enero de 2022.

II.

Inteligencia artificial, algoritmos y datos

El término “inteligencia artificial” se le atribuye al informático y científico cognitivo estadounidense John McCarthy, quien denominó de esa forma al resultado de “decisión inteligente” que ofrecen las máquinas a partir de datos y el desarrollo de algoritmos que controlan o guían su funcionamiento.

De esta manera, la expresión inteligencia artificial¹³ (IA) sirve para describir el resultado de los procedimientos que permiten a las máquinas simular o imitar algunas de las capacidades propias del ser humano o muy parecidas a estas. Esto es, la capacidad de las máquinas de simular procesos de inteligencia humana, especialmente a partir de sistemas informáticos que permiten a dichas máquinas en sentido similar a los seres humanos: percibir, identificar, representar y reconocer objetos, aprender, crear, razonar, planear, almacenar y clasificar información (datos) y crear modelos alternativos.

Por ahora (año 2022 occidental), la “autoreflexión”, el “sentido común” o “comprender la causa-efecto” son algunas de las capacidades de la inteligencia humana que todavía no puede reproducir la inteligencia artificial¹⁴. Aunque pueda manejar información y producirla, como antes se establecía, su principal limitación es que no es capaz de crear conocimiento por sí misma, a pesar de que sus resultados sí sirvan de base para ello. Esto, a pesar de que existen redes neuronales artificiales que se asemejan en mucho a las redes neuronales humanas¹⁵.

Así, la inteligencia artificial es en realidad todo aquello que lleva a cabo una máquina, asemejando capacidades o comportamientos propios del ser huma-

13 El término también se utiliza para describir de una manera amplia lo siguiente: Campo del conocimiento que estudia la teoría y el desarrollo de sistemas informáticos capaces de realizar tareas con cierto grado de autonomía y adaptabilidad a partir de la experiencia.

14 Cfr. Forbes, *What Artificial Intelligence Still Can't Do*, junio 2021, disponible en: <https://www.forbes.com/sites/robtoews/2021/06/01/what-artificial-intelligence-still-cant-do/?sh=7df543c366f6>

15 Véase: Hasson U, Nastase SA, Goldstein A., “Direct Fit to Nature: An Evolutionary Perspective on Biological and Artificial Neural Networks”, *Neuron*, Vol. 105 (3), 2020, pp. 416-434, disponible en: <https://www.sciencedirect.com/science/article/pii/S089662731931044X>

no a partir de modelos computacionales nutridos por datos y algoritmos¹⁶. Algunos ejemplos concretos de esto que tenemos en la actualidad son:

- Buscar información.
- Recomendar productos o servicios.
- Recomendar rutas, actividades, lugares, música, etc.
- Dar asistencia para el desarrollo de actividades en muy diversos ámbitos.
- Organizar rutinas o planes.
- Traducir información o palabras de otros idiomas.
- Medir la temperatura o gasto energético de un lugar.
- Regular el tránsito en una ciudad.
- Limpiar y ordenar.
- Asistir en la conducción de vehículos o maquinarias.
- Reconocer personas u objetos.
- Identificar enfermedades.
- Reproducir información y datos.
- Producir mercancías.
- Clasificar y distribuir objetos.
- Optimizar recorridos.
- Preparar alimentos.
- Acceder a bienes y servicios.
- Responder preguntas.
- Acceder a lugares.
- Observar el comportamiento de personas.
- Recibir instrucciones.
- Establecer el resultado de preguntas que se le formulan.
- Clasificar información y datos.
- Etc.

Las habilidades cognitivas humanas que actualmente puede desarrollar la inteligencia artificial son muchas, por lo que el listado anterior simplemente busca dar una muestra de las “cosas” que pueden hacer las máquinas asemejando al ser humano, ya que en los últimos años ha producido incluso poesía, prosa, imágenes, música y rostros humanos, ha diagnosticado algunas condiciones médicas con mayor precisión que un médico humano o solucionado grandes desafíos en biología.

¹⁶ En el glosario de la “Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno” (2018), se define a la inteligencia artificial de la siguiente forma: conjunto de métodos, teorías y técnicas científicas cuyo objetivo es reproducir, mediante una máquina, las habilidades cognitivas de los seres humanos. Los desarrollos actuales buscan que las máquinas realicen tareas complejas previamente realizadas por humanos.

En ese mismo sentido, hablar de inteligencia artificial no es hablar necesariamente de “robots”, ya que esta se desarrolla y está en máquinas con muy diversas características, como podrían ser: ordenadores, teléfonos inteligentes, televisores y otros tantos aparatos de uso común; pero también en elementos que no tienen un cuerpo físico, como los *chatbots* e infinidad de programas computacionales.

Pero no solo eso, la inteligencia artificial es algo que tenemos muy cerca, especialmente si contamos con una conexión a Internet, ya que está presente en el correo electrónico, en los buscadores de Internet, en las redes sociales, en las plataformas digitales de entretenimiento, en las compras electrónicas, en muy diversas aplicaciones y programas informáticos, en asistentes de localización y *gps*, en asistentes personales digitales, en traductores, en vehículos, en semáforos, en aparatos de limpieza del hogar, en aparatos de cocina y electrodomésticos, drones, teléfonos inteligentes, en el servicio al cliente, etc.

Aunque el surgimiento del término y de los desarrollos iniciales de la inteligencia artificial se pueden ubicar en la década de 1950, ha sido a partir de 2008 cuando la investigación en esta materia se ha acelerado notablemente¹⁷ y, como se desprende del listado y enumeraciones anteriores, se ha extendido en muchas disciplinas de conocimiento y, con ello, a una infinidad de ámbitos cotidianos de la vida de muchas personas.

Ese acelerado crecimiento y desarrollo hace que muchas veces no seamos capaces de darnos cuenta de que estamos frente a una aplicación de la inteligencia artificial, situación que ha causado también cierta preocupación y la sensación de que los avances en esta materia se están poniendo por encima de las capacidades humanas. Lo cual no es cierto, ya que el ser humano siempre está detrás de las máquinas, pero en su uso cotidiano es verdad que muchas aplicaciones generan la sensación de que las máquinas son quienes están tomando las decisiones y soluciones que antes correspondían al ser humano.

El gran crecimiento en las últimas décadas de la inteligencia artificial se vincula con el acelerado desarrollo que se ha dado para obtener los dos elementos centrales de los que se nutre para funcionar: algoritmos y datos.

Así, la mejora de la capacidad de las redes neuronales¹⁸ y de la supercomputación¹⁹ han permitido el desarrollo de algoritmos cada vez más avanzados. En tanto que la gran disponibilidad de datos masivos ha permitido que se cuente con datos (información) de muy variados ámbitos de toda la actividad humana desarrollada en espacios públicos, privados e incluso íntimos.

17 Cfr. Universitat Pompeu Fabra, Revista 360upf, número 9, octubre 2021, disponible en: <https://www.upf.edu/web/360upf/numero9/comunitat>

18 IBM Corporation define a la “red neuronal” como un modelo simplificado que emula el modo en que el cerebro humano procesa la información: Funciona simultaneando un número elevado de unidades de procesamiento interconectadas que parecen versiones abstractas de neuronas. Disponible en: <https://www.ibm.com/docs/es/spss-modeler/SaaS?topic=networks-neural-model>

19 Una supercomputadora es aquel tipo de ordenador que presenta capacidades de cálculo muy por encima de la media, están diseñadas para procesar en un segundo miles de millones de datos. En Barcelona (*Barcelona Supercomputing Center*) se tiene el *MareNostrum* que es el supercomputador más emblemático y el más potente de España actualmente (2022), aunque se ubica en el lugar 74 del mundo (<https://www.top500.org/lists/top500/list/2021/11/>).

De manera muy simple, un **algoritmo**²⁰ informático es la instrucción o conjunto de instrucciones escritas que debe seguir la computadora para obtener un resultado previamente establecido²¹. De manera un poco más elaborada, también se ha definido como: “Conjunto ordenado y finito de operaciones [secuencia de instrucciones] que permite [llevar a cabo una serie de procesos para dar respuestas a determinados problemas o] hallar la solución de un problema”²². Sus características concretas y la manera en cómo se integra las analizaremos más adelante para determinar en dónde y cómo se puede incidir en la igualdad y no discriminación desde los algoritmos.

Por ahora, simplemente destacamos que los algoritmos son muy importantes en el mundo de la informática, ya que permiten a la persona que programa resolver el problema que busca solucionar antes de escribirlo en un lenguaje de programación²³. En ese sentido, los programas informáticos no son otra cosa que una serie compleja de algoritmos escritos en un lenguaje de programación que pueden ser ejecutados en un ordenador. De ahí la importancia de los algoritmos informáticos en el diseño, creación y desarrollo de la inteligencia artificial.

En tanto que los **datos** son el conjunto de informaciones que procesan o requieren los algoritmos para ejecutar las instrucciones que se les dan y buscar los resultados que se espera obtener. Con lo que, cuantos más datos y datos de calidad, mejores pueden ser los resultados que den los algoritmos. Este aspecto también lo profundizaremos en el análisis concreto de la igualdad y la no discriminación en el mundo *online* de la inteligencia artificial.

Por ahora es importante señalar que, en informática, los datos son representaciones simbólicas (numéricas, alfabéticas, algorítmicas, alfanuméricas), es decir, la descripción codificada de un hecho empírico, un suceso, una entidad, un acto humano. Por lo que los datos no tienen valor semántico (sentido o interpretación lingüística) en sí mismo²⁴, pero convenientemente tratado (procesado) se puede utilizar en la realización de cálculos o toma de decisiones. De ahí que los datos son un recurso muy importante en la informática, pero si no se tienen los medios necesarios para procesarlos, difícilmente se pueden obtener todos los beneficios de su valor²⁵.

12

20 La palabra “algoritmo” proviene del sobrenombre de un matemático árabe del siglo IX, Al-Khwarizmi, que fue reconocido por ser el primero en enunciar paso a paso las reglas para las operaciones matemáticas básicas con decimales (suma, resta, multiplicación y división). Así, de esa enunciación paso a paso de instrucciones para realizar operaciones matemáticas, se construyó el concepto que hoy usamos.

21 En el glosario de la “Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno” (2018), se define al algoritmo como: Secuencia finita de reglas formales (operaciones lógicas e instrucciones) que permiten obtener un resultado de la entrada inicial de información. Esta secuencia puede ser parte de un proceso de ejecución automatizado y aprovechar modelos diseñados a través del aprendizaje automático.

22 Real Academia Española, *Diccionario de la lengua española*, 23.ª ed., [versión 23.5 en línea], disponible en: <https://dle.rae.es/algoritmo>. Consultado el 7 de enero de 2022.

23 Un lenguaje de programación es el tipo de instrucciones que decide utilizar cada persona que programa para interactuar con una computadora. Ese lenguaje de comunicación entre humano y computadora está integrado por algoritmos que la computadora entiende e interpreta en lenguaje de máquina. Algunos de los varios lenguajes de programación que actualmente más se utilizan son: *Scratch, Python, Code Studio, Lua, Blockly, Alice, Ruby, Java*.

24 Así, por ejemplo, el dato “derecho humano” no es lo que “comúnmente” podamos entender, sino que será la representación que de él haga la persona programadora, con las características que le dé en un determinado lenguaje informático. Recordando que, además, el dato informático suele estar representado utilizando el sistema de número binario de unos (1) y ceros (0), siendo por tanto distinto de la representación analógica que tiene. De ahí, parte de la complejidad de los datos en la informática.

25 Cfr. Herbert Jones, *Ciencia de los datos: La guía definitiva sobre análisis de datos, minería de datos, almacenamiento de datos, visualización de datos, Big Data para empresas y aprendizaje automático para principiantes*, 2019, 210p.

Así, se puede establecer que los algoritmos de cualquier tipo necesitan datos, datos medianamente estructurados, y es que sin datos difícilmente funcionarán los algoritmos y, por tanto, no existe la inteligencia artificial.

Por los elementos antes descritos, el Parlamento Europeo²⁶ ha sugerido que se considere como **sistema de inteligencia artificial** a “todo sistema basado en programas informáticos o incorporado en dispositivos físicos que muestra un comportamiento que simula la inteligencia, entre otras cosas, mediante la recopilación y el tratamiento de datos, el análisis y la interpretación de su entorno y la adopción de medidas, con cierto grado de autonomía, para lograr objetivos específicos.” De manera mucho más simple, un sistema de inteligencia artificial “es un software, programa informático, que procesa datos y genera decisiones en base a los datos analizados”²⁷.

Atendiendo a eso, bien se puede afirmar que existen muchos tipos de inteligencia artificial con diferentes capacidades, ya que eso depende del algoritmo utilizado y los datos de los que se dispone. Por ello, algunos son extremadamente especializados, usando algoritmos específicos, procesamiento natural del lenguaje (PLN) y visión por computador. Pero la gran mayoría se basan en técnicas de aprendizaje profundo y automático.

Por esto último, es común que se hable de forma generalizada de dos formas de desarrollar la inteligencia artificial: la de aprendizaje automático (*machine learning*)²⁸ y la de aprendizaje profundo (*deep learning*). Respecto a esto también profundizaremos en el apartado en el que se hace el análisis concreto de la igualdad y la no discriminación en la inteligencia artificial.

Con todo lo anterior, resulta evidente que hablar de lo que es la inteligencia artificial es cada vez más difícil, ya que es complicado concretar de manera simple, pues hay tantas formas de inteligencia artificial como formas de crearlas y entrenarlas existen a partir de las innumerables posibilidades que pueden ofrecer los algoritmos y los datos que las nutren. Es decir, la inteligencia artificial es y será tan compleja, cambiante y variable, como se lo permita ser la inteligencia humana que es la que la crea y entrena. Aunque, en todo caso, su base o núcleo de funcionamiento es el mismo: datos y algoritmos, algoritmos y datos.

Por ahora, simplemente es importante establecer que lo que es y sea la inteligencia artificial no puede desligarse de lo que ha sido, es y será la inteligencia humana, porque en el fondo, la máquina es simplemente un reflejo de lo que es el ser humano que la crea, con todos sus defectos y virtudes, con todas sus limitaciones y oportunidades, con todos sus conocimientos, prejuicios y cargas sociales en los que se desarrolla vitalmente.

26 Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal (2020/2013(INI)), disponible en: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0009_ES.html

27 Benjamins, R. y Salazar I., *El mito del algoritmo. Cuentos y cuentas de la inteligencia artificial*, Anaya Multimedia, 2020, p. 38.

28 El aprendizaje automático a su vez se puede dividir en tres tipos: supervisado, no supervisado y *reinforcement learning*.

III.

Bases generales de la igualdad y la no discriminación

La igualdad y la no discriminación son dos derechos humanos reconocidos en prácticamente todas las normas jurídicas existentes en el mundo. De ahí, que también se les considere como principios jurídicos que deben regir la creación y aplicación de las normas jurídicas en cualquier parte del mundo. Con lo que deben estar presentes en todo ámbito en el que las personas puedan ejercer derechos, libertades y obligaciones, tanto en ámbitos públicos como privados.

Para establecer de manera sencilla qué son y cómo se caracterizan estos derechos, seguiremos los trabajos previos del Institut de Drets Humans de Catalunya en esta materia²⁹.

Así, podemos decir que el derecho a la **igualdad** es un derecho de toda persona y la base del ejercicio de todos los derechos, aunque no significa identidad de trato en toda circunstancia. En ese sentido, la igualdad tiene como fin reducir las desigualdades no porque somos iguales, sino porque tenemos la misma dignidad como personas.

En el ámbito jurídico, que es el que aquí nos interesa especialmente, la noción de igualdad (igualdad ante la ley) significa, ante todo, la exigencia de que todos los ciudadanos (en el Derecho Constitucional clásico), todas las personas (en el Derecho Internacional de los Derechos Humanos) se hallen sometidas a las mismas normas y tribunales, que se les apliquen las mismas leyes, con las mismas consecuencias, cuando se encuentren en los mismos supuestos y se les diferencie cuando no lo están. Esto es, que la ley debe ser idéntica para todos y todas, sin que exista ningún tipo de privilegio, ni jurisdicción especial para unos cuantos, sin que se considere como tal una diferencia objetiva y razonable que busque eliminar la desigualdad³⁰.

29 Para una mayor profundización en el tema, véase: Castilla Juárez, Karlos, *Guía para la elaboración de informes integrales sobre discriminación en la ciudad*, IDHC, Barcelona, 2020, 50p. disponible en: <https://www.idhc.org/es/investigacion/publicaciones/discriminacion-intolerancia-y-odio/guia-para-la-elaboracion-de-informes-integrales-sobre-discriminacion-en-la-ciudad.php> y Castilla Juárez, Karlos, *Igualdad y No discriminación. Mecanismos y estándares internacionales*, IDHC, Barcelona, 2021, 106 p., disponible en: <https://www.idhc.org/es/investigacion/publicaciones/otras-publicaciones/igualdad-y-no-discriminacion-mecanismos-y-estandares-internacionales.php>

30 Castilla Juárez, Karlos, "Igualdad y no discriminación. Breve análisis con referencia a México y España", en *Revista Judicatus*, Año 6, No. 11, enero-junio 2018, p. 54.

Con lo que la también conocida como igualdad formal supone el reconocimiento de un mismo estatuto jurídico para todas las personas, lo que implica la garantía de la paridad de trato en la legislación (igualdad en la ley—creación) y en su implementación (igualdad frente a la ley—aplicación)³¹.

Esto al final significa que el mandato de trato paritario ante las normas obliga a establecer un tratamiento igual para todas y todos que no sea discriminatorio, pero no impide una diferenciación basada en causas objetivas y razonables, que sirvan justamente para alcanzar ello. Esto es, se debe tratar a toda persona por igual, con la obligación de establecer las diferencias que sean necesarias para alcanzar ello. Estas diferencias en todo caso deben justificarse de manera amplia y detallada, en la medida de que solo serán admisibles si son objetivas y razonables³².

Por tanto, la igualdad ante la ley significa que las situaciones iguales deben ser tratadas iguales y las situaciones desiguales deben ser tratadas desigualmente siempre de manera justificada objetiva, razonable y proporcional³³.

Por su parte, la prohibición de **discriminación** puede ser vista como un blindaje del principio de igualdad que busca que en verdad ninguna persona quede fuera de protección. Un principio que, si se quiere, ha servido para venir actualizando y modernizando el principio de igualdad³⁴.

De ahí que en la práctica, el principio de no discriminación es en gran medida una garantía reforzada del principio de igualdad, por medio del cual se busca que las personas y grupos que históricamente han sido excluidos de la protección de sus derechos y libertades, sean especialmente protegidos, al no ser *per se* su condición o sus diferencias con las mayorías justificación suficiente (y menos si es única) para distinguir en el ejercicio, goce y protección de los derechos humanos³⁵.

El principio de no discriminación prohíbe que se justifique la desigualdad teniendo como única razón a alguna de las denominadas como “categorías sospechosas”³⁶, que ello sea la única o principal base de la distinción. Exigiendo que las distinciones admisibles no estén basadas en un ser o pertenecer, en algo históricamente contingente, porque no se ha elegido autónomamente, sino en un hacer que busque igualar en condiciones, dar las mismas oportunidades, equilibrar la balanza.

Siguiendo lo establecido por la Convención Internacional sobre la Eliminación de todas las Formas de Discriminación Racial³⁷, podría entenderse que la discriminación no solo incluye toda *distinción*, sino

31 *Ibidem*, p. 56.

32 *Ibidem*, p. 57.

33 *Ibidem*, p. 60.

34 Castilla Juárez, Karlos, *Guía para la elaboración de informes integrales sobre discriminación en la ciudad*, IDHC, Barcelona, 2020, p.15.

35 *Ídem*.

36 Se denomina así al conjunto de características personales que suelen enumerarse para expresar las razones por las cuales las personas no pueden ser discriminadas. Siendo estas, de manera ejemplificativa pero no exhaustiva, casi de manera general, las siguientes: raza, color, sexo, idioma, religión, opinión política o de cualquier otra índole, origen nacional o social, posición económica, nacimiento o cualquier otra condición.

37 Adoptada y abierta a la firma y ratificación por la Asamblea General de Naciones Unidas en su resolución 2106 A (XX), de 21 de diciembre de 1965. Entrada en vigor: 4 de enero de 1969, de conformidad con su artículo 19.

también toda *exclusión, restricción o preferencia*. Aspecto que es muy importante en razón de que, comúnmente, nos centramos solo en la distinción, pero olvidamos (intencionadamente o no) las otras tres acciones con las cuales también se puede discriminar³⁸.

Por lo que, si cualquiera de esas cuatro acciones se hace por motivos de la raza, el color, el sexo, el idioma, la religión, la opinión política o de otra índole, el origen nacional, étnico o social, la posición económica, los impedimentos físicos, el nacimiento o cualquier otra condición, de cualquier persona, se podría estar discriminando³⁹ y se debe presumir la existencia de una discriminación.

Debe destacarse en este sentido que, según se ha ido construyendo en la doctrina de la no discriminación, la discriminación puede ser directa e indirecta.

Es *directa* cuando los actos, leyes o políticas públicas desventajosas están dirigidas expresamente a una categoría, grupo o persona. Es *indirecta* cuando esos actos, leyes o políticas públicas aplican un criterio aparentemente neutral al no dirigirse a una categoría, grupo social o persona, pero impone a estos una carga desproporcionada o desventaja particular con respecto a otras personas que les perjudica o imposibilita colocarse en el supuesto previsto⁴⁰.

Pero también, por las formas en las que se puede presentar un acto o situación discriminatoria, se habla de discriminación *por asociación*, cuando una persona es objeto de discriminación como consecuencia de su relación con una persona de uno de los grupos protegidos, sin que necesariamente pertenezca o se identifique como parte de ese grupo. Se diferencia de la discriminación *por error*, que es la situación en la que una persona o un grupo de personas son objeto de discriminación como consecuencia de una apreciación errónea, es decir, porque se ha creído o supuesto que pertenecen a un grupo determinado⁴¹.

Sumado a estas, también se ha desarrollado la idea de *discriminación múltiple*, que es la situación en la que una persona, por sus características personales, pertenece a varios grupos que son objeto de discriminación y sufre diferentes discriminaciones en diferentes momentos o en un mismo momento pero por varias razones⁴².

La discriminación múltiple no debe confundirse con la *interseccionalidad*, ya que, aunque son muy parecidas, en la primera se dan discriminaciones por diversas razones en diferentes o un mismo momento. La interseccionalidad es una metodología que sirve para analizar la forma en la que varios tipos de desigualdad y discriminación inciden y operan de manera conjunta en una persona por las características personales que esta tiene, así como para establecer la forma en la que dichas causas de

38 Castilla Juárez, Karlos, *Guía para la elaboración de informes integrales sobre discriminación en la ciudad*, op. cit., p.15.

39 *Ibidem*, p.16.

40 *Ibidem*, p.16.

41 *Ibidem*, p.16.

42 *Ibidem*, p.16.

desigualdad y discriminación se exacerban mutuamente y multiplican entre sí de acuerdo al contexto social específico en el que son puestas de manifiesto o identificadas⁴³.

De igual forma se podría hablar de la *discriminación estructural o sistémica*, esto es, del conjunto de normas, reglas, rutinas, patrones, actitudes y pautas de comportamiento, tanto *de jure* como *de facto*, que dan paso a una situación de inferioridad y exclusión contra un grupo de personas de forma generalizada, las cuales son perpetuadas a lo largo del tiempo e incluso por generaciones⁴⁴. Discriminación que, aunque es la última que aquí referimos, es en realidad el origen de todas las desigualdades y causas de discriminación.

En todo caso, para eliminar la discriminación en la práctica se debe prestar suficiente atención a los grupos o personas que sufren injusticias históricas o son víctimas de prejuicios persistentes en lugar de limitarse a comparar el trato formal que reciben las personas en situaciones similares⁴⁵. Esto será especialmente importante tenerlo presente en el siguiente apartado.

De todo eso, se puede establecer como elemento general de estudio y base para el análisis que se hará a continuación que el derecho a la igualdad y a la no discriminación abarcan dos concepciones: una negativa relacionada con la prohibición de diferencias de trato arbitrarias, y una concepción positiva relacionada con la obligación de crear condiciones de igualdad real frente a grupos que han sido históricamente excluidos o que se encuentran en mayor riesgo de ser discriminados⁴⁶.

De esta manera, si ya en el ejercicio ordinario y general de derechos, libertades y obligaciones la igualdad y no discriminación deben ser una constante, es evidente que no pueden ser ajenas a todo lo que implica la inteligencia artificial. Si la igualdad y la no discriminación son elementos fundamentales en el desarrollo de la vida humana *offline*, no hay razón alguna para que no lo sean también en todas las manifestaciones de la vida humana *online*.

43 Véase en este sentido: *Derechos humanos desde una perspectiva interseccional*, Institut de Drets Humans de Catalunya, Barcelona, 2021, disponible en: <https://www.idhc.org/arxius/recerca/1633338099-InterseccionalidadyDDHH.pdf>

44 Cfr. Comisión Interamericana de Derechos Humanos, Informe sobre pobreza y derechos humanos en las Américas, 2017, párr. 393; disponible en: <http://www.oas.org/es/cidh/informes/pdfs/pobrezaddhh2017.pdf>

45 Comité de Derechos Económicos, Sociales y Culturales, Observación General No. 20 “La no discriminación y los derechos económicos, sociales y culturales (artículo 2, párrafo 2 del Pacto Internacional de Derechos Económicos, Sociales y Culturales)”, párr. 8; disponible en: https://www2.ohchr.org/english/bodies/cescr/docs/E.C.12.GC.20_sp.doc

46 Cfr. Corte IDH, Caso Furlan y familiares Vs. Argentina, Excepciones Preliminares, Fondo, Reparaciones y Costas, Sentencia de 31 de agosto de 2012, Serie C, No. 246, párr. 267.

IV.

Cuatro ángulos para observar la igualdad y la no discriminación en la inteligencia artificial

Si entendemos la inteligencia artificial como un postre, en el que la receta que se sigue son los algoritmos y los ingredientes que se le ponen los datos, es evidente que el resultado que obtendremos dependerá de esos factores, pero también de quién siga la receta e incorpore los ingredientes, respectivamente. Sin olvidar, por supuesto, las herramientas o utensilios que se tengan para cocinar o sacar el mejor provecho a esos ingredientes y receta.

De esa forma, por ejemplo, un postre en sí mismo no puede ser señalado como algo malo para la salud. Se podrá considerar más o menos saludable, dependiendo de la receta, los ingredientes y el fin con el que se haya elaborado.

Así, en principio, la inteligencia artificial no se puede considerar como algo creado para afectar la igualdad o para discriminar, pero es importante también no presumir que no lo hace, sin conocer en detalle los algoritmos, datos y fines con los que ha sido creada.

Lo que en todo caso es importante establecer desde ahora, por lo antes señalado, es que la inteligencia artificial no es neutra, en la misma medida que no son neutros los algoritmos ni los datos a partir de los cuales es creada, funciona y se desarrolla. Como tampoco lo son los postres, pues si no todos serían iguales y sin efectos en la salud o alimentación. En ambos casos, el factor humano que está detrás importa y mucho.

De esta manera, la mejor forma para identificar en dónde y cómo se puede incidir en la igualdad y en la no discriminación a partir de la inteligencia artificial es a partir de analizar las partes que la integran y que participan en su creación.

Así, es importante identificar: dónde se está creando (i), quién la está creando (ii), para qué se está creando (iii) y cómo se está creando (iv). Ya que a partir de esos cuatro ángulos de observación, se puede identificar de mejor forma cómo y por qué se incide, positiva o negativamente, en la igualdad y no discriminación en todo lo que representa la inteligencia artificial.

1. Primer ángulo: dónde

La respuesta de **dónde se está creando la inteligencia artificial** nos da una información que trasciende a todos los demás aspectos de análisis, ya que quienes controlan económicamente este ámbito son quienes en gran medida toman las decisiones que influyen en los demás aspectos.

Bajo esa premisa, desde el año 2019 y sin muchos cambios sustanciales para el año 2021, se puede afirmar que China y Estados Unidos son los lugares en los que se cosechan las mayores ganancias económicas de la inteligencia artificial, mientras que África y América Latina experimenten los niveles más bajos. China, Estados Unidos y Japón, juntos, representan el 78% de todas las solicitudes de patentes de inteligencia artificial en el mundo⁴⁷.

Esos datos se corresponden de manera casi simétrica con el hecho de que, de las 10 supercomputadoras que existen en el mundo en 2022, 5 están ubicadas en Estados Unidos, 2 en China y 1 en Japón, Alemania e Italia⁴⁸. Así como con el hecho de que de las 10 compañías⁴⁹ más importantes en inteligencia artificial en el año 2021, 9 tienen su sede en Estados Unidos y 1 en China⁵⁰.

Esto significa, en gran medida, que los valores políticos, culturales, económicos y sociales de las personas dueñas de estas compañías, y de los países en los que se encuentran, son las que están determinando de manera relevante la forma en la que la inteligencia artificial funciona, sus fines, propósitos y características, no solo para esos lugares en donde son creadas, sino para todos aquellos lugares en el mundo al que actualmente llegan y en donde se utilizan.

Adicionalmente a eso, y no menos importante, es el hecho de que los más importantes desarrollos de la inteligencia artificial están en manos de empresas privadas, por lo que son estas las que están estableciendo las “reglas del juego”. Y no precisamente a partir de la consideración de los derechos o las necesidades de las personas, sino de sus propios intereses económicos y, vinculados con estos, los intereses políticos que mejor les permiten desarrollar y alcanzar sus objetivos empresariales.

21

47 United Nations Conference on Trade and Development, *Digital Economy Report 2019: Value Creation and Capture: Implications for Developing Countries*, United Nations, 2019, p. 8, disponible en: https://unctad.org/system/files/official-document/der2019_en.pdf

48 Véase para más información: Mohammed Hussein, “Visualising the race to build the world’s fastest supercomputers”, *AI Jazeera*, 14 enero 2022, disponible en: <https://www.aljazeera.com/news/2022/1/14/infographic-visualising-race-build-world-fastest-supercomputers-interactive>

49 Microsoft Azure, Amazon Web Services, IBM Cloud, Salesforce, Google Cloud Platform, Alibaba Cloud, Clari, Abnormal Security, People.ai, Atomwise.

50 Cfr. Virginia Tech, India, *Top 10 Artificial Intelligence Companies to Work for in 2021*, disponible en: <https://vtindia.in/top-10-artificial-intelligence-companies-to-work-for-in-2021/>

2. Segundo ángulo: quién

A partir de lo anterior, resulta importante saber **quién está creando la inteligencia artificial**, pues es evidente que los dueños de las empresas antes mencionadas no son precisamente quienes hacen los desarrollos en inteligencia artificial.

Así, se observa que la fuerza laboral de inteligencia artificial sigue siendo predominantemente masculina y carece de diversidad, tanto en su parte académica como en la industria⁵¹. De hecho, los sectores de las tecnologías digitales emergentes, como los de Silicon Valley⁵², se caracterizan por una “crisis de diversidad” en lo que respecta al género y la raza, especialmente en los niveles directivos más altos⁵³.

Según una encuesta anual de la *Computing Research Association*⁵⁴, en el año 2019, entre los nuevos graduados de doctorado en inteligencia artificial residentes en Estados Unidos, el 45% eran blancos, mientras que el 22,4% eran asiáticos, el 3,2% eran hispanos y el 2,4% eran afroamericanos.

De esos porcentajes, se debe poner atención en que las mujeres graduadas de programas de doctorado de inteligencia artificial han representado el 18,3% de todas las personas graduadas en los últimos 10 años⁵⁵. Sin perder de vista en este análisis de género el dato relativo a que más del 40% de las personas *queer* que se desarrollan en ámbitos de la inteligencia artificial han sufrido discriminación o acoso en el trabajo o la escuela⁵⁶.

Estos datos, centrados en el país que es la mayor potencia en creación y desarrollo de inteligencia artificial ponen en evidencia la falta de diversidad que hay entre quienes están a cargo de esas labo-

22

51 Stanford Institute for Human-Centered Artificial Intelligence, *Artificial Intelligence Index Report 2021*, Chapter 6: Diversity in AI, p. 3; disponible en: https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report_Chapter-6.pdf

52 También conocido como Valle del Silicio, es el nombre que recibe la zona sur del Área de la Bahía de San Francisco, en el norte de California, Estados Unidos; lugar en el que se encuentra la sede de muchas compañías emergentes y globales de tecnología. De ahí su fama cuando se habla de tecnología.

53 Informe de la Relatora Especial sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia de Naciones Unidas, *La discriminación racial y las tecnologías digitales emergentes: un análisis de los derechos humanos*, A/HRC/44/57, junio de 2020, párr. 17; disponible en: <https://undocs.org/es/A/HRC/44/57>

54 Es una asociación basada en Washington, D.C., Estados Unidos, que cuenta entre sus miembros con más de 200 organizaciones norteamericanas activas en la investigación informática: departamentos académicos de informática e ingeniería informática; laboratorios y centros en la industria, el gobierno y la academia; y sociedades profesionales afiliadas.

55 Stanford Institute for Human-Centered Artificial Intelligence, *op. cit.*, p. 5.

56 *Ibidem.*, p. 16.

res⁵⁷, con lo que no solo se consolida el riesgo de crear una distribución desigual del poder en la fuerza laboral, sino que también, e igualmente importante, se refuerzan las desigualdades existentes pero ahora generadas por los sistemas de inteligencia artificial, al reducir el alcance de las personas y organizaciones para quienes estos sistemas funcionan, y, con ello, se contribuye a resultados injustos⁵⁸.

En relación a los datos antes citados, es importante señalar la advertencia que hace el *Stanford Institute for Human-Centered Artificial Intelligence*, al señalar en su informe que la falta de disponibilidad pública de los datos demográficos limitan el grado en que los análisis estadísticos pueden evaluar el impacto de la falta de diversidad en la fuerza laboral de la inteligencia artificial en la sociedad⁵⁹. Pues, por paradójico que parezca, estando en el ámbito de la inteligencia artificial, esos datos no están disponibles de manera pública⁶⁰, en gran parte, porque están en manos de quienes obtienen y poseen el mayor volumen de datos en el mundo, esto es, las empresas que dominan la escena global de la inteligencia artificial⁶¹.

En todo caso, con la información disponible, es evidente que en el puñado de empresas y laboratorios universitarios de élite que crean y desarrollan los sistemas de inteligencia artificial en el mundo, la diversidad ya no mundial, sino de la sociedad de los propios países líderes, no se refleja, al mantenerse una fuerte presencia de varones blancos, acomodados económicamente y con un perfil técnico especializado.

Esta realidad lo único que pretende enfatizar es que el problema de la desigualdad y la discriminación no es exclusivo de la inteligencia artificial, sino que es un problema endémico de las sociedades que, como vemos, también se ve reflejado en los ámbitos académicos y laborales vinculados con la inteligencia artificial. Con lo que, es uno de los muchos ámbitos en los que como sociedades debemos avanzar, a fin de que la diversidad que existe y es una realidad en nuestras ciudades se refleje también en todos los ámbitos educativos, productivos, laborales, económicos, políticos, etc.

Teniendo claro dónde y quién está creando la inteligencia artificial, ahora es más fácil describir y entender aspectos técnicos más concretos del núcleo de la inteligencia artificial.

57 Los datos que aquí se ofrecen son limitados ya que no se ocupan de toda la diversidad que hay en la sociedad (discapacidad, edad, nacionalidad, etc.), pero lo relacionado con raza-etnia-género es la única información disponible de forma sistematizada, por lo que, como las propias personas autoras del estudio que se cita reconocen: “estas estadísticas de diversidad cuentan solo una parte de la historia”.

58 Cfr. Stanford Institute for Human-Centered Artificial Intelligence, op. cit., p. 3.

59 *Ídem*.

60 Con relación a la no disponibilidad de datos, véase en un sentido similar: Castilla Juárez, Karlos, *Datos para la identificación de interseccionalidades en el estado español: ¿una misión imposible hoy?*, IDHC, Barcelona, 2020, 27p.

61 De igual forma sería interesante contar con esos datos respecto a otros países situados en los primeros lugares de la creación de inteligencia artificial. Siendo especialmente relevante también lo que ocurre en Cataluña, si atendemos a que es uno de los lugares que mejor se está situando en temas vinculados con la inteligencia artificial, pero actualmente solo sabemos que para 2019 en las 179 empresas que trabajan con inteligencia artificial había 8.483 personas empleadas, sin conocer más detalles (Generalitat de Catalunya, *La intel.ligencia artificial a Catalunya* https://www.accio.gencat.cat/web/.content/bancconeixement/documents/informes_sectorials/informe-tecnologic-intel.ligencia-artificial.pdf).

3. Tercer y cuarto ángulo: para qué y cómo

Para responder a las preguntas de **para qué** se está creando y **cómo se está creando la inteligencia artificial**, volveremos a los elementos que la integran y, a partir de describir las partes o formas en que esta se crea, estableceremos en dónde se puede incidir en la igualdad y la no discriminación⁶².

Así, en primer lugar, debemos establecer que un algoritmo consta de tres partes:



En la entrada, es donde se introducen todos aquellos datos que el algoritmo necesitará para operar. El procesamiento o las instrucciones son las operaciones y cálculos lógicos que el algoritmo desarrollará para resolver el problema que se haya planteado. La salida es el resultado que se obtiene del procesamiento que se sigue de los datos dados inicialmente al algoritmo.

24

De esta simple descripción de las partes de un algoritmo, resulta evidente que, tanto en la entrada como en el procesamiento o instrucciones, podría presentarse una situación que afecte o beneficie a la igualdad y a la no discriminación⁶³. Ya sea porque se le dan datos que distinguen, excluyen, restringen o prefieren una característica personal sobre otra, sin justificación objetiva, razonable y proporcional, o porque, por medio de las instrucciones o el procesamiento, se haga lo antes señalado. Si se hiciera así, podríamos estar, según sea el caso, en presencia de una discriminación directa o de una acción afirmativa si está debidamente justificada la distinción.

Lo anterior no se debe confundir con el hecho de que, por su naturaleza, los algoritmos están diseñados para distinguir entre clases de objetos o datos. Ese no es el problema para la igualdad y la no

62 Se debe advertir de que las ideas que a continuación se desarrollarán parten de una visión en la que la inteligencia artificial se está dirigiendo a ámbitos en los que hay ejercicio, goce y disfrute de derechos, libertades y deberes reconocidos como derechos humanos. Por lo que, la igualdad y la no discriminación deben estar especialmente presentes en esos casos como una constante en todo lo que implica a la inteligencia artificial. En ese sentido, no se debe entender que las ideas que a continuación se presentan pretenden ser un obstáculo para desarrollos de la inteligencia artificial que no tienen incidencia en el ejercicio, goce y disfrute de derechos humanos y, por esa razón, las dejamos fuera de estudio, no porque no sean relevantes, sino porque las variables de análisis son distintas.

63 Véase en ese sentido: Peter K. Yu, "The Algorithmic Divide and Equality in the Age of Artificial Intelligence", en *Florida Law Review*, Vol. 72, 2020; disponible en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3455772; Xenidis, Raphaële, "Tuning EU equality law to algorithmic discrimination: Three pathways to resilience", *Maastricht Journal of European and Comparative Law*, 27 (6), 2020, pp.736-758; disponible en: <https://journals.sagepub.com/doi/full/10.1177/1023263X20982173> o Zuiderveen, Frederik, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making*, Consejo de Europa, 2018; disponible en: <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>

discriminación. El problema es cuando en esos datos o esas instrucciones se incluyen variables que directamente y sin justificación objetiva, razonable y proporcional, distingan, excluyan, restrinjan o prefieran una característica personal sobre otra en el ejercicio de derechos, libertades y obligaciones que se manifiesten en el acceso o disfrute de bienes, servicios u otro tipo de prestaciones cotidianas. Pero también, cuando eso se haga indirectamente, al no tomarse las precauciones necesarias para evitar que datos o instrucciones aparentemente neutros, discrimine.

Esto, que parece muy simple a partir de esa descripción, no lo es en la práctica, ya que los datos de entrada y las instrucciones se elaboran en lenguajes informáticos que no son de fácil comprensión por cualquier persona. Con lo que, sin duda, la correcta comprensión e integración del lenguaje y principios antidiscriminatorios y de igualdad, antes que estar bien integrados en los algoritmos, deben estarlo en las personas que crean las entradas e instrucciones, de otra forma no podrán ser introducidos de forma correcta en los algoritmos.

En ese sentido, para observar si puede o no haber incidencia en la igualdad y la no discriminación, antes que en la entrada y en el procesamiento, se debe poner atención en la salida. Sí, porque en prácticamente todos los manuales y cursos en los que se enseña a crear algoritmos, por regla general, se dice que el primer paso es: determinar el objetivo del algoritmo o el resultado que se quiere obtener.

Lo cual parece lógico, porque solo conociendo el resultado o meta que se quiere alcanzar, es posible introducir los datos necesarios para ello y establecer las instrucciones que se deberán seguir. Por esa razón, en las guías, libros y manuales para construir algoritmos se insiste en que cada uno de sus pasos sea claro en todos los aspectos y conduzca a un solo significado; que las entradas estén bien definidas al igual que las salidas; que no debe terminar en bucles infinitos; que debe poder ejecutarse con los recursos disponibles y que deben ser simplemente instrucciones que se puedan implementar en cualquier idioma, permitiendo que el resultado sea el mismo, como se esperaba⁶⁴.

Así las cosas, el primer aspecto en el que se debe poner especial atención para poder evaluar los posibles efectos positivos o negativos de la inteligencia artificial en la igualdad y no discriminación es en el **problema** a resolver y/o los **resultados** que se busca obtener con cada uno de los algoritmos que la integren. Esto es así porque el problema gravita en conocer perfectamente en qué consiste y qué resultados se desean obtener (tercer ángulo).

El resultado que se plantee obtener es fundamental en esta observación porque el algoritmo simplemente describe un procedimiento computacional específico para lograr esa relación de entrada/salida. Un algoritmo correcto es el que resuelve el problema computacional planteado; por tanto, el problema a resolver o resultado esperado es lo que determina cómo se crea el algoritmo.

Si se conoce el problema que se va a resolver con un algoritmo, ya se puede tener información más precisa respecto a si ese problema/resultado puede incidir en la igualdad y no discriminación. Especialmente porque es a partir de esa definición de donde se pueden conocer las restricciones del problema

⁶⁴ Cfr. Cormen, Thomas H y otros, *Introduction to Algorithms*, The MIT Press, Cambridge, Massachusetts London, England, 2009; disponible en: https://edutechlearners.com/download/Introduction_to_algorithms-3rd%20Edition.pdf o Geeks for Geeks, *Introduction to Algorithms*, disponible en: <https://www.geeksforgeeks.org/introduction-to-algorithms/>

que se han considerado para resolverlo, la entrada que se considerará para resolver el problema, lo que se espera cuando se resuelva el problema (salida) y la solución que se espera.

Bajo esa perspectiva, parece evidente que puede haber resultados/problemas que expresamente puedan parecer discriminatorios, otros que abiertamente lo sean y algunos más que parezcan neutros y que solo, a partir de conocer su entrada y procesamiento, se podrá establecer con más claridad si afectan o benefician la igualdad y la no discriminación (discriminación indirecta). Pero, en todo caso, esta es la primera gran referencia técnica a tener en cuenta.

Así, por ejemplo, podría parecer discriminatorio que el problema a resolver implique la exclusión de personas menores de 18 años, pero si se hace porque antes de esa edad no se puede acceder a algún tipo de bien o servicio concreto, lo que parecía, no lo es. Sería abiertamente discriminatorio el hacer esa misma exclusión, general o específica para un grupo concreto de menores de edad, si no hay justificación objetiva, racional y proporcional (como la anterior) para hacerlo. Mientras que se podría dar una discriminación indirecta, cuando en apariencia el problema incluye a toda persona, pero se toma en cuenta una característica que ninguna persona menor de edad podrá cumplir por ser menor de 18 años.

De lo descrito antes en el dónde y quién de la creación de la inteligencia artificial (primer y segundo ángulo), parece evidente que actualmente son esas personas las que están estableciendo los objetivos, resultados y problemas que se buscan resolver con la inteligencia artificial. De ahí, la importancia de la diversidad, pues parecería en principio que sin esta es más difícil buscar y alcanzar resultados no discriminatorios en un algoritmo y, por tanto, en la inteligencia artificial. Aunque la diversidad no lo es todo, ni es garantía de la no discriminación, sí es una posibilidad de garantizar que se tendrán más formas y visiones en el ejercicio de derechos, libertades y deberes.

En este sentido, dónde y quién está estableciendo los problemas a resolver por medio de algoritmos que se utilizarán en la inteligencia artificial es fundamental, ya que son quienes están estableciendo el para qué de la inteligencia artificial. No solo porque son quienes están definiendo eso, sino porque lo que se plantea como problemas parte de su visión, intereses, necesidades propias o de sus clientes y, por tanto, pudiera no tratarse de problemas que interesen o incidan en necesidades más trasversales de la sociedad, sino muy específicas de personas o grupos.

De hecho, podrían plantearse como problemas situaciones que para otras personas no lo sean o problemas que aún requieren soluciones que sean tomadas por seres humanos antes que por medio de algoritmos o, peor aún, que los problemas a resolver tengan como base intereses que abiertamente busquen estigmatizar, excluir o señalar a grupos específicos de población.

Una vez establecido el problema, el segundo paso es el diseño del algoritmo (cuarto ángulo). En este, los **datos**⁶⁵ vuelven a ser importantes porque, como ya se ha establecido antes, para cualquier algoritmo, los datos de entrada son esenciales. Así, por ejemplo, se necesitan suficientes datos históricos para dis-

65 Por superar los objetivos de este documento, no se analizará aquí todo lo relacionado con los datos y sus vínculos con la privacidad, transparencia, etc. Simplemente dejamos anotado que, por su importancia en la inteligencia artificial, requiere especial interés su análisis desde una mirada de derechos humanos. Véase en ese tema: Filimowicz, Michael, *Privacy, Algorithms and Society*, Routledge, London, 2022, 144p; Gil González, Elena, *Big data, privacidad y protección de datos*, Agencia Española de Protección de Datos–Agencia Estatal Boletín Oficial del Estado, Madrid, 2016, 145p.

tinguir los datos de prueba de los datos de control necesarios para que un algoritmo dé mejores resultados. El volumen de datos y su calidad es lo que puede mejorar el hecho de que se alcance efectivamente el resultado final.

Es evidente que en los datos puede estar otra parte fundamental de las incidencias positivas o negativas en la igualdad y no discriminación. Esto es así no sólo porque en los datos en sí mismos puede haber distinciones, exclusiones, restricciones o preferencias no justificadas, o peor aún, ni siquiera estar identificadas al ser datos provenientes de un ámbito en el que ya se ha dado de origen esa exclusión, restricción, distinción o preferencia de una o varias características personales. El ejemplo más utilizado por su claridad (y lamentable permanencia histórica) es la existencia de datos que excluyan a las mujeres por expresa decisión o por provenir de ámbitos en donde la presencia de estas ha sido previamente excluida por las desigualdades estructurales.

Pero también porque esos datos requieren una interpretación para convertirse en información, y la interpretación está, en primera instancia, a cargo de la persona que decide incorporarlos como entrada y, en segunda instancia, en el algoritmo a partir de las instrucciones que se le den. En este sentido, no solo es que en los datos ya se excluya a las mujeres, sino que, en primer lugar, a quien incorpora los datos eso le pueda parecer irrelevante y, en segundo lugar, que el algoritmo reproduzca la exclusión al ser compatible con las instrucciones que ha recibido a partir de los datos de los que dispone.

Este sesgo en la representación de los datos también se puede dar cuando los datos obtenidos en un país o sector de población se usan para representar la totalidad o regiones que en poco o nada se corresponden, con lo que evidentemente hay una incorrecta interpretación de los datos y, por buena que pueda ser la instrucción, el resultado estará distorsionado. En este aspecto se insiste en que puede deberse a una decisión de la persona que diseña, conscientemente o no, pero también a los límites técnicos que se tengan. Sea una u otra forma, es un aspecto relevante a tener presente en el ámbito de la no discriminación.

En términos generales, los datos rara vez son neutros, están ligados a experiencias e historias de personas, por lo que reducirlos a modelos matemáticos sin tener en consideración las circunstancias que los rodean con el objeto de darle una aparente neutralidad⁶⁶ puede llevarnos a resultados incompletos, equivocados y discriminatorios.

Además de lo anterior, se debe tener presente que en los últimos años ha habido un crecimiento exponencial de datos, por lo que no todos los datos de los que se pueda disponer son veraces, aunque no se tenga conciencia de ello, otros pertenecen a situaciones muy específicas, ya sea temporales, culturales o sacadas del contexto donde se originan, con lo que tener datos no es necesariamente garantía de fiabilidad. Por lo que, el análisis y el procesamiento de datos son un elemento fundamental⁶⁷, ya que el algoritmo simplemente utilizará los datos y seguirá las instrucciones que los seres humanos les den.

66 Muñoz, Catherine, "La discriminación en una sociedad automatizada: Contribuciones desde América Latina", *Revista Chilena de Derecho y Tecnología*, Vol. 10, No. 1, 2021, p. 279; disponible en: <https://rchdt.uchile.cl/index.php/RCHDT/article/view/58793>

67 Cfr. Benjamins, R. y Salazar I., *op. cit.*, p. 38.

Así por ejemplo, si como en el caso de España no es fácil tener datos de las características étnicas de las personas que habitan su territorio, por prohibición legal⁶⁸, evidentemente ahí ya hay algo que de manera constante no aparecerá en las múltiples bases de datos (o casi todas) que hagan referencia a la población en España. Y, peor aún, de aparecer, por dicha prohibición de su tratamiento, pueden tener importantes errores, faltas de veracidad o darse interpretaciones incorrectas de estos.

En el propio diseño del algoritmo, las **instrucciones** son muy importantes, y ahí se puede incidir también en la igualdad y no discriminación. Las instrucciones que se construyen en un algoritmo son mucho más complejas que las que se podrían dar a un ser humano, ya que, siguiendo el ejemplo inicial del postre, al ser humano le podrías decir: mezcla en un recipiente la harina, huevos y azúcar; mete la mezcla en el horno por 30 minutos, etc. Pero, a un algoritmo se le tiene que decir no solo eso, sino qué es harina, huevo, azúcar, horno, tiempo, temperatura y cómo todo eso está representado en las instrucciones que está siguiendo.

Con lo que, evidentemente, la forma en la que se pueda introducir la idea de igualdad y no discriminación en el algoritmo tiene sus complejidades, ya que muy probablemente en muchos ámbitos de la inteligencia artificial no sea uno de los aspectos que suelen tenerse presente. No necesariamente porque no se quiera, sino porque no forman parte evidente de ese ámbito de trabajo o no se ve como un aspecto a tener en cuenta en los problemas que se busca resolver⁶⁹. En ese sentido, es y será importante en la inteligencia artificial tener la capacidad de desarrollar instrucciones y variables que permitan entrenar algoritmos teniendo siempre una variable de “dato o atributo sensible” que no es más que las llamadas “categorías sospechosas”, pues solo así se pueden reducir los riesgos de discriminación en la inteligencia artificial.

La “máquina” puede darse el lujo de no comprender la diferencia y calidad que hay entre los datos e instrucciones que se le proporcionan, pero quien diseña el algoritmo tiene que llegar a un nivel de comprensión muy profundo. Porque si ya es difícil explicarle un dato y darle instrucciones a otro ser humano, dárselas a una “máquina” requiere, de entrada, hablar su lenguaje y, después, ser capaz de representarle de manera fiel la realidad en la que dará un resultado. Esto nos muestra nuevamente que las personas programadoras y desarrolladoras en general de algoritmos de la inteligencia artificial deben estar al menos sensibilizadas en materia de igualdad y no discriminación. De otra forma, se traslada la responsabilidad hacia quien no la tiene y solo obedece.

Lo anterior es importante porque no podemos perder de vista que, al final, los algoritmos son una forma de manipular la realidad que nos rodea a través de la computación y, además, ponen en evidencia el modo en el que se interpreta el mundo. Por lo que un mundo en el que la realidad muestra o mantiene grandes desigualdades y discriminaciones, difícilmente podrá ser interpretado en algoritmos de una

68 Véase *supra* nota 58.

69 Problema que no es exclusivo de ese ámbito laboral, ya que esto se ha presentado por mucho tiempo también en organismos y organizaciones no gubernamentales que trabajan en el ámbito de los derechos humanos, en los que, a pesar de su especialidad, la diversidad ha tardado en llegar o siguen presentándose desigualdades en el acceso a posiciones o cargos. Con lo que, evidentemente, es un problema social mucho más profundo en el que se debe de trabajar.

forma distinta. Así, las decisiones algorítmicas pueden ser discriminatorias, por ejemplo, cuando el sistema aprende de decisiones humanas discriminatorias⁷⁰.

Siempre será más difícil darse cuenta de la existencia de desigualdades o discriminación, de la forma en que eso influye, o no, en nuestra manera de ser y en el modo en el que tomamos nuestras decisiones en los ámbitos personales o laborales, cuando convivimos a diario con o entre ello en nuestro país, nuestra ciudad, nuestra cultura, nuestro trabajo, nuestra historia de vida⁷¹. Pero si no están ahí, o no les vemos porque a nosotras no nos pasa, más difícilmente podremos darnos cuenta de situaciones discriminatorias que ocurren todos los días.

De ahí la importancia del factor humano detrás de la inteligencia artificial. Aquí ponemos énfasis en las personas programadoras e ingenieras informáticas por ser quienes están detrás de la inteligencia artificial, pero esto es relevante y ocurre en todas las profesiones y oficios.

Sin entrar en detalles técnicos que nos superan, lo anterior es mucho más complejo en la realidad de la creación de algoritmos de lo que se ha descrito, ya que esos datos para poder ser utilizados en un algoritmo son manipulados para convertirlos en un lenguaje informático, sin olvidar que los algoritmos para su desarrollo se pueden describir o representar por diagramas de flujo⁷² o pseudocódigos⁷³. Con lo que conocer cuáles son los datos incorporados y el diseño del algoritmo a partir de las instrucciones que sigue requiere de conocimientos técnicos de los cuales no estamos dotados todas las personas.

El tercer gran paso en la creación de un algoritmo lo simplificamos con lo que denominamos como puesta a prueba y entrenamiento. En el caso de la inteligencia artificial, los algoritmos a probar y entrenar están diseñados para aplicarse en las “formulas” que más se utilizan en su desarrollo: el *machine learning*⁷⁴ y el *deep learning*⁷⁵.

Las técnicas de *machine learning* permiten a los algoritmos identificar patrones complejos entre gran cantidad de datos, infiriendo así sus propias reglas para detectar patrones similares en nuevos con-

70 Zuiderveen, Frederik J., “Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence”, *The International Journal of Human Rights*, 2020, p. 1574; disponible en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3561441

71 Cfr. Benjamins, R. y Salazar I., *op. cit.*, p. 53.

72 Representaciones gráficas de secuencias de pasos a realizar. Cada operación se representa mediante un símbolo normalizado por el Instituto Norteamericano de Normalización (ANSI–American National Standards Institute). Las líneas de flujo indican el orden de ejecución.

73 Representaciones que describen un algoritmo de forma similar a un lenguaje de programación pero sin su rigidez, de forma más parecida al lenguaje natural escrito de los seres humanos, aunque con la misma complejidad técnica que los lenguajes informáticos. Son más compactos que los diagramas de flujo, más útiles para las instrucciones complejas y más sencillos de transferir en un lenguaje de programación. El pseudocódigo no está regido por ningún estándar.

74 También conocido como aprendizaje automático, utiliza algoritmos para analizar datos, aprender de esos datos y tomar decisiones basadas en lo aprendido o dar sugerencia sobre algo.

75 También conocido como conocimiento profundo, utiliza algoritmos en capas, para crear una red neuronal artificial, que puede aprender y tomar decisiones por sí misma.

juntos de datos. Así, pueden aprender a predecir comportamientos, detectar similitudes o anomalías automáticamente o tomar las decisiones adecuadas⁷⁶.

Por lo que, en las instrucciones que se le den para identificar los patrones y, sobre todo, en los datos que se le suministren, atendiendo a todo lo antes señalado, se podría incidir más o menos en la igualdad y la no discriminación.

El *machine learning* se puede desarrollar de tres formas concretas que, como se verá, el mayor o menor riesgo de incidencia en la igualdad y la no discriminación vuelve a estar presente en los datos que se le dan y las instrucciones que recibe.

La primera forma es el “aprendizaje supervisado” que entrena a los sistemas con los conocidos como datos etiquetados, esto es, ejemplos concretos de lo que se quiere que aprenda a identificar el algoritmo. Si quiero que aprenda a identificar autos, le suministro datos de lo que sí y lo que no representa un auto. Con esta información previa, los algoritmos buscarán lo mismo en nuevos conjuntos de datos. Evidentemente, en la información que se le dé respecto a lo que sí y no representa algo, se podría presentar una mayor o menor incidencia en la igualdad y no discriminación. Si el algoritmo está diseñado para “creer” que una pauta repetida es la correcta, concluirá que la hipótesis correcta es aquella que más veces se repite, sea verdadera o no⁷⁷. Así, si se le dice que los autos verdes no son autos o se le da la instrucción de que los autos verdes no se consideren autos, evidentemente se estaría discriminando. Como también se haría indirectamente con el hecho de no incluir, consciente o inconscientemente, datos de autos verdes.

La segunda forma es el “aprendizaje no supervisado”, en donde los algoritmos no cuentan con un conocimiento previo, sino que es el propio sistema el que observa las características o comportamientos de los datos y busca similitudes y patrones para agruparlos de alguna forma o detectar patrones anómalos que pueden resultar de interés. Por ejemplo, le vuelvo a dar información sobre autos y, a partir de las características que de estos le doy, aunque no tenga una indicación exacta, el algoritmo agrupa los que tienen un mismo color, un mismo número de puertas, portaequipajes, etc. En ese sentido, tanto en los datos, como en las características que destaco para que el propio sistema busque similitudes y patrones, se podría dar una incidencia en la igualdad y la no discriminación. Si en los datos no incluyo autos de color verde o le doy la instrucción de que la característica verde no tiene relevancia, que es un error o simplemente no aparece, se estaría discriminando si no hay justificación objetiva, razonable y proporcional.

La tercera forma es el “aprendizaje por refuerzo”, en el que el sistema aprende a partir de su propia experiencia, con base a un proceso de prueba y error y recompensas si toma decisiones correctas. Así, llegará a generar automáticamente estrategias que permitan optimizar un proceso. En este caso, por ejemplo, el algoritmo aprende que cuando distingue entre autos de alta gama y de baja gama, recibe una recompensa, por lo que, en el futuro, lo seguirá haciendo. Evidentemente, quien fija las recompensas o penalizaciones por los resultados es un ser humano, que también es quien le proporciona los da-

76 Instituto de Ingeniería del Conocimiento, *Machine Learning & Deep Learning. Los sistemas de IA aprenden de tus datos*, disponible en: <https://www.iic.uam.es/inteligencia-artificial/machine-learning-deep-learning/>

77 Rivas-Vallejo, Pilar, “Discriminación algorítmica: detección, prevención y tutela”, en *Associació Catalana de Iuslaboralistes*, Ponencias 2021, p. 5; disponible en: <http://www.iuslabor.org/wp-content/plugins/download-monitor/download.php?id=415>

tos y establece las instrucciones que deberá seguir para identificar determinadas características. Con lo que, nuevamente, si los autos de color verde no proporcionan ninguna recompensa, no están en los datos o generan una penalización cuando se les identifica, estarán fuera de su ámbito y, por eso mismo, se podría estar discriminando.

Por su parte, el *deep learning* representa un acercamiento más real al modo de funcionamiento del sistema nervioso humano. Por esa razón, está dedicado a sistemas más complejos de datos y problemas. Por sus características, se ocupa más de los conocidos como datos no estructurados o multimedia (imágenes, vídeo, texto, audio) y crea sistemas mucho más complejos y precisos.

En este, se pueden llevar a cabo los procesos de *machine learning*, solo que usando una red neuronal artificial que se compone de un número de niveles jerárquicos. En el nivel inicial de la jerarquía, la red aprende algo simple (identificar una imagen) y luego envía esta información al siguiente nivel. El siguiente nivel toma esta información sencilla, la combina, compone una información un poco más compleja (imagen donde aparecen autos), y se la pasa al tercer nivel (tamaños de los autos), y así sucesivamente. En el caso del ejemplo, las redes de *deep learning* aprenderían a identificar autos aunque las imágenes no tuvieran la etiqueta "auto".

Parece que, si no se toman las precauciones necesarias, los algoritmos de aprendizaje automático reproducen los sesgos incorporados en los conjuntos de datos a gran escala, y son capaces de imitar y reproducir los sesgos implícitos de los humanos, incluso en ausencia de reglas algorítmicas explícitas que generen estereotipos.

Si en todos los ejemplos anteriores cambiamos el "auto verde" por una característica personal humana protegida contra la discriminación, parecería más claro porqué y cómo se incide en la igualdad y la no discriminación. Aunque el gran reto en la inteligencia artificial parece estar en, primero, tener siempre presente que existen esos "autos verdes", y en pensar más allá de los límites de trabajo cotidiano para ver si esos "autos verdes" deben ser considerados, en sorprenderse si esos "autos verdes" no aparecen en los datos si queremos efectivamente resolver los problemas de los autos. Es decir, en pensar de manera inclusiva y en todos los sectores de población a los que se dirigirá la inteligencia artificial, en pensar más allá de nuestra burbuja cotidiana de vida.

En todo caso, es evidente que nuevamente los datos e instrucciones serán determinantes, con lo que las distinciones, exclusiones, restricciones o preferencias no justificadas que exista en estos serán también aquí fundamentales para incidir positiva o negativamente por medio de la inteligencia artificial en la igualdad y la no discriminación. Y esto es así porque los valores humanos están incrustados en cada paso de su diseño⁷⁸.

La cuarta y última etapa de esta creación de algoritmos es su utilización y puesta en práctica con las personas. En este sentido, simplemente recordamos lo que se estableció al inicio de este documento, esto es, que sin Internet es muy difícil que la inteligencia artificial funcione, con lo que, por ese simple hecho, ya hay una importante exclusión de personas.

⁷⁸ Mercader, Jesús R., "Discriminación algorítmica y derecho granular: nuevos retos para la igualdad en la era del big data", en *Labos*, Vol. 2, No. 2, p. 5; disponible en: <https://e-revistas.uc3m.es/index.php/LABOS/article/view/6213>

La otra parte importante desde la mirada de la igualdad y la no discriminación se vincula con lo que antes se señaló respecto a los problemas o resultados que se están buscando solucionar, con las capacidades humanas que se están imitando con la inteligencia artificial, ya que si estas están pensadas solo en las necesidades e intereses de un sector de la población mundial, evidentemente no se puede hablar de igualdad y no discriminación. Y si, además, por la forma en las que estas están siendo desarrolladas y los problemas que están resolviendo, cuando una persona ajena a esa “burbuja” entra en contacto, siendo un “fallo” en el “sistema”, muy probablemente también se podría estar discriminando o repitiendo, multiplicando y profundizando en situaciones de desigualdad.

Todo esto nos lleva a establecer que, para hablar de igualdad y no discriminación en la inteligencia artificial, son muchos los ángulos desde los que se debe observar, y no solo a partir de fracciones del todo. A pesar de eso, somos conscientes de que en este documento estamos dejando fuera algunos ángulos más que son relevantes en el análisis pormenorizado del diseño técnico de los algoritmos y la inteligencia artificial. Por lo que simplemente queremos enfatizar la importancia que tiene el intentar explorar con el mayor detalle posible lo que es y qué significa la inteligencia artificial.

Así las cosas, es importante insistir que todo lo anterior busca poner énfasis en toda aquella inteligencia artificial que se está dirigiendo a ámbitos en los que hay ejercicio, goce y disfrute de derechos, libertades y deberes reconocidos como derechos humanos. Un llamado a que la igualdad y la no discriminación estén presentes como una constante en todo lo que implica la inteligencia artificial en esos ámbitos, para evitar la discriminación y revertir efectivamente desigualdades estructurales o, al menos, para que se puedan identificar los sesgos de los cuales no somos conscientes y, mejor aún, nos pueda servir para identificarlos y corregirlos en los procesos, usos y aplicaciones de la inteligencia artificial. Esto, porque los algoritmos, como técnicas de procesamiento, carecen de sensibilidad, criterio y orientación natural en derechos humanos y pueden acertar o errar⁷⁹, pero el ser humano sí puede imprimir esa esencia de los derechos humanos en los algoritmos con los que funciona la inteligencia artificial. Aunque tal vez ese es justo el problema de fondo, que el ser humano sigue discriminando y generando desigualdad.

⁷⁹ Cfr. Faliero, Johana C., “Limitar la dependencia algorítmica. Impactos de la inteligencia artificial y sesgos algorítmicos”, en *Nueva Sociedad*, No. 294, 2021, p. 121; disponible en: <http://library.fes.de/pdf-files/nuso/nuso-294.pdf>

V.

Conclusiones

Como se estableció desde el inicio, este documento ha pretendido explorar la inteligencia artificial desde una visión muy particular desde la igualdad y la no discriminación, buscando identificar los puntos concretos en los cuales la primera puede incidir positiva o negativamente en la segunda. Desde ese entendimiento, podemos establecer como conclusiones lo siguiente:

- Como regla general, la inteligencia artificial la crean, diseñan, poseen y disfrutan los ricos y la sufren las personas pobres.
- Los valores políticos, culturales, económicos y sociales de las personas dueñas de las compañías líderes en inteligencia artificial, y de los países en los que se encuentran, son las que están determinando de manera relevante la forma en la que la inteligencia artificial funciona, sus fines, propósitos y características.
- La inteligencia artificial es un producto de determinadas sociedades, sus valores, sus prioridades e incluso sus desigualdades.
- La inteligencia artificial está generando una distopía con un mundo con dos tipos de personas: los humanos mejorados gracias a la inteligencia artificial y el resto del mundo.
- Uno de los principales desafíos de la inteligencia artificial es evitar la discriminación de personas, ya sea por sesgos que existen en los datos (el caso más común), sesgos que provienen del algoritmo, o sesgos que se producen en la interacción de los sistemas de inteligencia artificial y las personas usuarias de esta.
- La discriminación algorítmica que incluyen muchas aplicaciones de inteligencia artificial resulta invisible para grandes sectores de la sociedad, es difícil determinar los alcances que tiene y los efectos que provoca, y es técnicamente muy compleja.
- La inteligencia artificial, como otras tantas creaciones humanas, no es neutra, sino más bien el reflejo de la sociedad con todas sus cargas políticas, económicas, culturales y de respeto de los derechos humanos. Por lo que no debe servir como el único o más importante instrumento para la toma de decisiones vinculada con el ejercicio de derechos, libertades y deberes.

- Todos los actuales modelos de aprendizaje de inteligencia artificial actúan como "cajas negras". Es decir, son sistemas matemáticamente complejos en los que desconocemos qué pasa dentro del modelo entre la entrada de datos y la salida de un resultado. Esto es uno de sus grandes problemas y uno de los grandes retos en la lucha contra la desigualdad y la discriminación en la inteligencia artificial.
- Todas las aplicaciones de inteligencia artificial deberían contar con el equivalente de la tabla nutricional como la que tienen los alimentos, en la que se indique, al menos: el número de algoritmos que las integran, las respuestas que resuelven estos, el tipo de datos que están utilizando como entrada y el tipo de *machine learning* que utilizan.
- La inteligencia artificial ha crecido, en parte, bajo la promesa de aportar mayor igualdad a nuestro polarizado y desigual mundo. Sin embargo, solo nos están aprovechando como fuente inagotable de datos y continúan reproduciendo las estructuras de desigualdad, discriminaciones y opresión que han existido en el mundo históricamente, pero ahora de una manera más sutil al responsabilizar a máquinas y operaciones matemáticas de ello.
- Quienes han permitido que la inteligencia artificial se desarrolle causando daños por sus efectos discriminatorios y/o perpetuadores de estructuras de desigualdad, prejuicios y estereotipos en grupos históricamente discriminados, deben asumir su responsabilidad sin importar si son del ámbito privado o público.
- La calidad de la ingeniería, la programación y todas las profesiones vinculadas con la inteligencia artificial debe también vincularse con la capacidad de reflejar los principios de igualdad y no discriminación en su trabajo. Al igual que se debería exigir a todas las profesiones y oficios.
- Es y será importante en la inteligencia artificial tener la capacidad de desarrollar instrucciones y variables que permitan entrenar algoritmos teniendo siempre una variable de "dato o atributo sensible" que no es más que las llamadas "categorías sospechosas", pues solo así se pueden reducir los riesgos de discriminación en la inteligencia artificial.
- La igualdad y la no discriminación ya están presentes en todas las legislaciones nacionales e internacionales del mundo y, por tanto, deberían estar presentes en la sociedad y en el desarrollo de la inteligencia artificial. No obstante eso, la aprobación de documentos como la *Declaración de Toronto* de 2018 o los *Principios de Pekín* de 2019 en los que se hace un llamado a prevenir la discriminación en el diseño e implementación de la inteligencia artificial son esfuerzos que deben continuar replicándose.

- La desigualdad y la discriminación no han surgido ni son exclusivas de la inteligencia artificial, esta solo está reflejando la desigualdad y discriminación que hay en toda la sociedad y, más específicamente, en el ámbito concreto de los estudios, profesiones y lugares donde se crea, diseña y produce.
- En tanto que existan algoritmos utilizados en la inteligencia artificial que cuenten con datos y/o instrucciones que incluyan aspectos que, sin justificación objetiva, razonable y proporcional, directa o indirectamente, distingán, excluyan, restrinjan o prefieran una característica personal sobre otra, la discriminación seguirá siendo parte de su esencia práctica real. Evitarlo es un deber ético (y jurídico) humano, no una responsabilidad de los algoritmos.
- Si la igualdad y la no discriminación son elementos fundamentales en el desarrollo de la vida humana *offline*, no hay razón alguna para que no lo sean también en todas las manifestaciones de la vida humana *online*.

