



Esta presentación se encuentra disponible
en acceso abierto bajo la licencia Creative
Commons Attribution 4.0 International License

Revista de Derecho Procesal del Trabajo

Publicación Especializada del Equipo Técnico Institucional de Implementación
de la Nueva Ley Procesal del Trabajo del Poder Judicial

Vol. 8, n.º 11, enero-junio, 2025, 72-97

Publicación semestral. Lima, Perú

ISSN: 2708-9274 (En línea)

DOI: <https://doi.org/10.47308/rdpt.v8i11.1133>

El sesgo algorítmico generado por la inteligencia artificial como acto de discriminación en las relaciones de trabajo

**The Algorithmic bias generated by artificial intelligence
as an act of discrimination in employment relations**

**O Viés Algorítmico Gerado pela Inteligência Artificial
como um Ato de Discriminação nas Relações de Trabalho**

KARLA PAOLA CASTILLO-CASTRO

Corte Superior de Justicia de La Libertad
(Trujillo, Perú)

Contacto: kcastillo@pj.gob.pe
<https://orcid.org/0009-0005-3506-1991>

RESUMEN

El presente estudio analiza los nuevos retos planteados para el derecho, en particular, el derecho laboral, ante el avance y la implementación de la inteligencia artificial. El problema de investigación se orienta a examinar el impacto de este sistema inteligente en la toma de

decisiones del empleador durante la relación laboral mediante el uso de algoritmos como sesgos o filtros empleados en la gestión de la fuerza de trabajo. El objetivo principal es determinar si la aplicación de los sesgos algorítmicos generados por la inteligencia artificial constituye algún tipo de acto de discriminación que vulnere el principio de igualdad y de no discriminación en las relaciones de trabajo. La metodología empleada corresponde a un enfoque cualitativo con un diseño de investigación hermenéutico, sin el empleo de instrumentos, dado que se trata de una investigación teórica. Se obtuvo información de internet, de la que se seleccionaron los documentos relevantes para la investigación, consistentes en artículos científicos, libros electrónicos, páginas web y tesis electrónicas a nivel de máster universitario. Todos ellos se encontraban en español y fueron publicados desde 2020 en adelante. Los resultados obtenidos confirmaron la hipótesis de que el uso de los sesgos algorítmicos constituye un tipo de acto discriminatorio en las relaciones laborales, al incidir negativamente en la toma de decisiones empresariales. Se concluye que resulta necesaria la existencia de una protección normativa en nuestro país y en el sistema normativo internacional con el fin de evitar situaciones de indefensión en la protección de los derechos fundamentales de los trabajadores.

Palabras clave: discriminación; discriminación algorítmica; inteligencia artificial; relaciones de trabajo; sesgo algorítmico.

Términos de indización: discriminación; inteligencia artificial; relaciones laborales (Fuente: Tesauro Unesco)

ABSTRACT

This study analyzes the new challenges posed for law and in the field of labor law with the use of artificial intelligence. The research problem is oriented towards examining the impact of this intelligent

system on the employer's decision-making during the employment relationship through the use of algorithms as biases or filters employed in workforce management. The main objective is to determine whether the application of algorithmic biases generated by artificial intelligence constitutes acts of discrimination that violate the Principle of Equality and Non-Discrimination in employment relations. The methodology employed was a qualitative approach with a hermeneutic research design, without using instruments as it is a theoretical investigation. Information was obtained from the internet, selecting relevant documents for the research consisting of scientific articles, e-books, websites, and electronic theses at the master's level, all in Spanish and published from 2020 onwards. The results confirmed the hypothesis that the use of algorithmic biases constitutes discriminatory acts in labor relations by negatively influencing business decision-making. It is concluded that it is necessary for there to be normative protection in our country and in the international normative system to avoid vulnerability in the protection of workers' fundamental rights.

Keywords: discrimination; Algorithmic discrimination; artificial intelligence; labor relations; algorithmic bias.

Indexing terms: discrimination; artificial intelligence; labor relations (Source: Unesco Thesaurus).

RESUMO

Este estudo analisa os novos desafios apresentados para o direito e no campo do direito do trabalho com o uso da inteligência artificial. O problema de pesquisa está orientado a examinar o impacto deste sistema inteligente na tomada de decisões do empregador durante a relação de trabalho, por meio do uso de algoritmos como vieses ou filtros empregados na gestão da força de trabalho. O objetivo

principal é determinar se a aplicação dos vieses algorítmicos gerados pela inteligência artificial constitui atos de discriminação que violam o Princípio da Igualdade e da Não Discriminação nas relações de trabalho. A metodologia empregada foi a abordagem qualitativa com um desenho de pesquisa hermenêutico, sem o uso de instrumentos, uma vez que se trata de uma investigação teórica. Foram obtidas informações da internet, selecionando documentos relevantes para a pesquisa consistentes em artigos científicos, livros eletrônicos, páginas da web e teses eletrônicas em nível de mestrado, todos em espanhol e publicados a partir de 2020. Os resultados obtidos confirmaram a hipótese de que o uso dos vieses algorítmicos constitui atos discriminatórios nas relações de trabalho ao influenciar negativamente a tomada de decisões empresariais. Conclui-se que é necessário que exista proteção normativa em nosso país e no sistema normativo internacional para evitar a vulnerabilidade na proteção dos direitos fundamentais dos trabalhadores.

Palavras-chave: discriminação; discriminação algorítmica; inteligência artificial; relações de trabalho; sesgo algorítmico.

Termos de indexação: discriminação; inteligência artificial; relações laborais (Fonte: Thesaurus da Unesco)

Recibido: 14/01/2025

Revisado: 28/04/2025

Aceptado: 29/04/2025

Publicado en línea: 30/06/2025

1. INTRODUCCIÓN

La inteligencia artificial —en adelante, IA, por sus siglas en inglés— alude a los sistemas o las computadoras que imitan la inteligencia humana para realizar determinadas tareas, así, gracias a la retroalimentación automática que realizan pueden mejorar

sus sistemas a partir de la información que recopilan. Al respecto, Gutiérrez *et al.* refieren que la IA se trata del conjunto de algoritmos, tecnologías y técnicas empleadas para que los *softwares* o las máquinas desarrollen funciones cognitivas orientadas a la toma de decisiones, la solución de problemas, la búsqueda, la comprensión de imágenes, el procesamiento de datos y la interpretación de información (citado por Arguelles, 2023). Por lo tanto, se entenderá por «sistema inteligente» a todo sistema informático que incorpore el uso de la inteligencia artificial. Es comprensible la inquietud generada por el uso de la inteligencia artificial debido a que estos sistemas inteligentes integran modelos y algoritmos con base en datos disponibles, que son capaces de aprender y realizar determinadas tareas cognitivas que originan «resultados como la predicción y la adopción de decisiones en entornos materiales y virtuales», como se expresa con preocupación en el documento de la Unesco titulado *Recomendación sobre la ética en la inteligencia artificial* (2024, p.10).

Actualmente, el rápido avance de la IA ha captado la atención del mundo, causando una gran expectativa en el posible impacto en la economía mundial y, sobre todo, en el mercado laboral. Así, el Fondo Monetario Internacional refiere que este sistema inteligente afectará a casi el 40 % de los empleos en todo el mundo, reemplazando algunos y complementando otros, mientras que el Foro Económico Mundial estima que para el año 2025, 85 millones de puestos de trabajo podrían desaparecer, señalando además que la inteligencia artificial desempeña, a la fecha, un papel crucial en muchas empresas, especialmente en la toma de decisiones, de tal manera que la competitividad en el futuro no consistirá solo en tener la inteligencia artificial más avanzada, sino también con quién utiliza esta tecnología para tomar decisiones estratégicas (Foro Económico Mundial, 2023, párr. 5).

Hoy tenemos conocimiento que la IA, más allá de su uso cotidiano en los medios de comunicación masivos y de las redes sociales digitales, presenta un impacto significativo en muchos ámbitos de nuestra vida:

laboral, educativo, médico, industrial, financiero, deportivo y cultural. En estos campos, al parecer, se están transformando diversas formas de interacción, aprendizaje, procesos de automatización, toma de decisiones, entre otros. Esto implicaría una cierta precisión arbitraria en la utilización de este sistema inteligente, conforme señalan Rezaev y Tregubova, pues algunos aspectos de nuestra vida se modifican, lo cual genera un fuerte impacto en la esfera social de tipo ético y moral (citado por Ramírez, 2023).

Es en este contexto que se plantean nuevos retos para el derecho y, en especial, en el campo del derecho laboral. Esto se debe a que el empleo de la IA en la toma de decisiones del empleador se encuentra presente a lo largo de la relación laboral mediante el uso de algoritmos como sesgos o filtros empleados en la gestión de la fuerza de trabajo como es en la selección de personal, la contratación o en la evaluación de desempeño de los trabajadores, lo cual resulta de gran importancia al identificar si estos sesgos algorítmicos constituyen actos de discriminación en las relaciones de trabajo (por cuestión de sexo, edad, nivel socioeconómico, especialización de labores, tipo de desempeño, entre otros), ya que mermarían la estabilidad de entrada y salida de los trabajadores. Esto causa, de esta manera, desprotección en el ámbito laboral. En este sentido, la presente investigación se encuentra orientada a determinar si los sesgos algorítmicos generados por la inteligencia artificial empleados en la toma de decisiones, resultan ser neutrales para el logro de los fines de la empresa o, por el contrario, responden a conductas discriminatorias que vulneran el principio de igualdad y no discriminación de los trabajadores, el cual debe ser respetado a lo largo de una relación laboral.

Para alcanzar tal objetivo, se ha llevado a cabo un estudio sobre el empleo de la inteligencia artificial, los sesgos algorítmicos y su implicancia en la toma de decisiones del empleador dentro de una relación laboral, realizando una búsqueda de artículos científicos,

libros, revistas y tesis electrónicas a nivel de máster universitario, en repositorios especializados como Scopus, Google Scholar, Scielo y ResearchGate. Se utilizaron las siguientes palabras claves: discriminación; discriminación algorítmica; inteligencia artificial; relaciones de trabajo; sesgo algorítmico —tanto en español, como en inglés y portugués—, las cuales se emplearon para ampliar o restringir los resultados de la búsqueda en los repositorios especializados el operador booleano AND(+). El período de análisis y la elección de artículos fue de septiembre 2023 a noviembre 2024, y se optó por el material con una antigüedad no mayor de cinco años.

Durante la búsqueda bibliográfica se encontró información relacionada con los avances de la IA y su impacto en la gestión de la fuerza de trabajo, en la toma de decisiones del empleador, así como en los problemas generados en la estabilidad laboral de los trabajadores (entrada y salida). En un primer momento, se eligieron 50 textos relacionados con los criterios de búsqueda empleados, los cuales fueron analizados considerando los títulos, las palabras claves, así como las conclusiones, las consideraciones o las reflexiones finales. Finalmente, se seleccionaron 27 documentos: 9 artículos científicos, 7 libros o revistas electrónicas, 3 normas, 7 páginas webs y 1 tesis electrónica, los cuales conforman el presente trabajo.

2. REFERENTES TEÓRICOS: ALGORITMOS, INTELIGENCIA ARTIFICIAL, SEGOS ALGORÍTMICOS Y DISCRIMINACIÓN ALGORÍTMICA

Uno de los motores de cambio social en el mundo son las transformaciones tecnológicas en los modos de producción, en el campo laboral y en la socialización, y traen consigo importantes transformaciones socioeconómicas conocidas como revoluciones industriales. Actualmente, la sociedad se encuentra inmersa en la

Cuarta Revolución Industrial y en el empleo de una de sus tecnologías, como la inteligencia artificial, por lo que resulta innegable en nuestros días sentirse familiarizados en el lenguaje cotidiano con términos como *big data*, también conocidos como *macrodatos* o *datos masivos*, además de *algoritmos* e *inteligencia artificial*, de los cuales no siempre se conoce su significado a plenitud. De cierta forma, estas herramientas facilitan la ejecución de tareas cotidianas, a tal punto que actualmente es imposible no recurrir a la IA para llegar a un lugar determinado, realizar un itinerario de viaje, recomendar un restaurante o un lugar turístico que se desea conocer; inclusive constituye un apoyo en las labores académicas y permite mejorar la productividad en el trabajo.

En ese sentido, resulta necesario precisar algunos conceptos, en primer lugar, se aborda el término *big data*, el cual según el Parlamento Europeo (como se citó en Gómez, 2022) es la «recopilación, análisis y acumulación constante de grandes cantidades de datos, incluidos datos personales (...), objeto de un tratamiento automatizado mediante algoritmos informáticos y avanzadas técnicas de tratamientos de datos, utilizando tanto datos almacenados como datos transmitidos en flujo continuo, con el fin de generar correlaciones, tendencias y patrones», lo que permite colegir que los datos masivos, macrodatos o *big data* se encuentran relacionados con el término *algoritmo*, el cual según la Real Academia Española (RAE), se define como el «conjunto ordenado y finito de operaciones que permite hallar la solución de un problema (Real Academia Española, 2014, definición 1)», y que autores como Osoba y Welser llaman agentes artificiales, que sirven para resolver problemas de la vida cotidiana (citados por Ramírez, 2023). Es importante agregar que el algoritmo es una creación humana, al ser las personas las que elaboran la fórmula que después será ejecutada por una computadora; por lo tanto, debe ser entendida como una secuencia de instrucciones, desarrollada por humanos, que se encargan de realizar tareas y encontrar soluciones de forma automatizada (Eguíluz, 2020, p. 330).

Otro concepto relevante resulta ser la expresión *inteligencia artificial*, que conforme señala Puntriano (2024) «fue primigeniamente empleado(a) en 1950, y se vincula con el cuestionamiento del matemático Alan Turing sobre la posibilidad de que una máquina fuese inteligente, razón por la que redactó el artículo "Computing machinery and intelligence" y el consiguiente "Test de Turing" que sentaron las bases de la inteligencia artificial y sus objetivos (p.13)», y es definida como «máquinas que responden a simulaciones como los humanos, con capacidad de contemplación, juicio e intención». Estos sistemas son capaces de «tomar decisiones que normalmente requieren de un nivel humano de conocimiento que constituyen la esencia de la inteligencia artificial: intencionalidad, inteligencia y adaptabilidad». En este sentido, la IA no tiene una autonomía absoluta, ya que para su funcionamiento requiere de algoritmos, los cuales necesitan de la intervención del ser humano. Esto permite concluir que la autonomía de la IA es relativa.

La autonomía relativa de la IA se debe a que esta aprende con base en los datos que se le proporcionen, de tal manera que, dependiendo de la calidad de los datos que se dispongan para su entrenamiento, será la calidad de los resultados obtenidos lo que implicaría que el resultado sesgado sea producto también del dato sesgado que le da origen (López, 2021, p. 150). De esta manera, se halla un *sesgo algorítmico*, definido como aquellas decisiones erradas propias de un sistema inteligente o, conforme señalan Danesi *et al.*, aquella inclinación o prejuicio de una decisión realizada por un sistema inteligente que afecta a una persona o a un grupo determinado de una forma que se considere injusta y razonable (citado por Sánchez, 2022).

Bajo esta perspectiva, la interacción continua entre los seres humanos y la IA genera importantes implicancias en la toma de decisiones. Así, un estudio reciente reseñado por el Foro Económico Mundial ha descubierto que los responsables de la toma de decisiones en las empresas deciden de manera completamente diferente

utilizando recomendaciones idénticas de la IA en función de sus estilos individuales. Es decir, no bastaría con generar una IA perfecta, ya que, si interviene un ser humano en la generación de algoritmos, se deben considerar los elementos psicológicos que han generado tal programación, y que impulsa, por ende, determinada toma de decisiones. Esta puede verse viciada por cuestiones de género, edad, origen racial, religión, situación socioeconómica, entre otros, lo que significa que existe un «sesgo algorítmico» o a una «discriminación algorítmica», conceptos que algunos autores emplean como sinónimos, posición que en este artículo se considera acertada. Estos términos serán usados indistintamente en el presente trabajo, porque «la selección de características de los datos utilizados para crear y configurar los modelos requeridos en los algoritmos que emplean la IA (robots, redes sociales, plataformas y buscadores digitales) reflejan muchas veces prejuicios de los programadores y un control intencional de estos profesionales que bien pueden desembocar en excluir y perjudicar los grupos más vulnerables de la sociedad, así como en las desigualdades de género y raza (Barrios et al, 2020, párr.21)». Esto implica, según Osornio (2023), que la información que alimenta las nuevas tecnologías, como la IA, contiene los sesgos ideológicos de quienes las crean, replicando la propia naturaleza de la sociedad. En este sentido, si los grupos de programadores encargados del diseño de algoritmos utilizados por la inteligencia artificial no resultan ser lo suficientemente diversos o inclusivos como para reflejar una gran amplitud de visiones, posiblemente ni siquiera podrían percatarse de la existencia de estos sesgos y, por lo tanto, no estarían en condiciones para corregirlos.

Ante ello, resulta necesario abordar este tema mediante la mejora de los conjuntos de datos y algoritmos orientados en el impacto social de la IA, conforme lo señala el profesor del Instituto de Estudios Avanzados en Educación de la Universidad de Chile, Roberto Araya (citado por Aliaga, 2024); debido a las implicancias de su empleo en los ámbitos sociales, económicos, laborales, políticos, entre otros y,

sobre todo, porque, conforme lo señala la Unesco (2024), los humanos somos moralmente responsables de la inteligencia artificial; esto es, cualquiera que se crea, implementa o usa.

3. IMPLICANCIAS JURÍDICAS DE LA DISCRIMINACIÓN ALGORÍTMICA EN LAS RELACIONES DE TRABAJO

Con el fin de intentar explicar esta nueva forma de discriminación originada a partir de la inteligencia artificial, se considerará el estudio de Barocas y Selbst, así como el trabajo de Zuiderveen, citados por Eguíluz, a fin de distinguir hasta cinco formas en las que la toma de decisiones automatizadas podrían conducir, involuntariamente, a la discriminación. Se encuentran problemas en relación con a) la variable objetivo y la definición de etiquetas de clases; b) datos de entrenamiento; c) recopilación de datos de entrenamiento; d) selección de características; e) variables de aproximación (*proxy*) (Eguíluz, 2020, p. 345).

Así, la variable *objetivo* consiste en definir el problema de una forma que el ordenador pueda analizar; mientras que las etiquetas de clase consisten en dividir y etiquetar en diferentes categorías los datos disponibles con la finalidad de organizarlos. En esta etapa se podría distorsionar la programación de algoritmos, lo cual muestra a veces un impacto negativo en las clases sociales vulnerables, al introducir un sesgo dentro del sistema algorítmico, que podría generar discriminación en la población de estudio. Otro problema surge cuando el sistema de IA aprende y es entrenado por un conjunto de datos ya sesgados. En este caso estaríamos frente a datos de entrenamiento, en el cual el algoritmo no introduce un sesgo nuevo, sino que reflejaría los sesgos ya existentes, basados en los datos con los que ha sido entrenado. Es decir que, si el sistema de IA aprende de una muestra anterior que ha sido tomada eligiendo a los sujetos de forma sesgada, se trataría

de una «recopilación de datos de entrenamiento». Otra forma de discriminación algorítmica es la relacionada con la «selección de características», en la que cualquier compañía o entidad, para predecir algo automáticamente, utiliza un sistema de IA que priorice ciertas características sobre otras, lo cual define la programación del algoritmo. Por último, se consideran las variables de aproximación *proxy*, en las cuales la pertenencia a un grupo o colectivo dentro de un conjunto de datos puede correlacionarse con otro grupo o colectivo.

Al respecto, coincidimos con Muñoz al señalar que Barocas y Selbst desarrollan un esquema sobre aquellas consideraciones que condicionan la generación de actos discriminatorios. Es posible resumirlas en la generación de los llamados «sesgos» en datos, algoritmos y modelos. Por su parte, Villegas resume el estudio de Barocas y Selbst, señalado previamente, y logra identificar hasta tres tipos de sesgos algorítmicos más frecuentes y entendibles por el ciudadano común: a) sesgos de selección, los cuales ocurren cuando los datos utilizados para entrenar la IA no resultan ser representativos de la población objetivo; b) sesgos de codificación, que se generan cuando los programadores de la IA incorporan sus propios sesgos en el diseño del sistema; c) sesgos de amplificación, los cuales se originan cuando el sistema inteligente amplifica los sesgos presentes de los datos con los que se entrenó (Villegas, 2024).

El término «sesgo» es empleado en materia legal como sinónimo de prejuicio o trato injusto (Muñoz, 2021, p. 279). Por su parte, autores como Ferrante (citado en Morales, 2023) considera los sesgos algorítmicos como una de las implicancias antropológicas generadas por la inteligencia artificial, entendidas como errores estadísticos, estructurales, cognitivos y sociales que generan desventajas éticamente reprochables al generar resultados discriminatorios o conducir a beneficios a favor de un grupo de individuos frente a otros.

Todo ello pone de manifiesto un nuevo fenómeno generado por la IA, como es la discriminación algorítmica, que podría tener un impacto significativo en los derechos de los trabajadores, al vulnerarse los principios de igualdad y no discriminación por parte de sus empleadores debido al empleo de algoritmos sesgados o sesgos algorítmicos en la toma de decisiones. Estos algoritmos suelen basarse en datos históricos que reflejan patrones de discriminación sistemática hacia ciertos grupos o colectivos, que se clasifican en género, edad, raza, discapacidad, nivel socioeconómico, entre otros. Como consecuencia, se generan decisiones discriminatorias que excluyen a personas pertenecientes a minorías del acceso al empleo, de la obtención de una buena evaluación de desempeño, del derecho de ascenso a un cargo superior o del acceso a cargos directivos dentro de una empresa o institución. Esta situación se ve agravada por la opacidad de los algoritmos, lo cual puede dificultar la responsabilidad y la transparencia en la toma de decisiones. Asimismo, esto obstaculiza la capacidad de los individuos para impugnar decisiones injustas o discriminatorias (Iturmendi, 2023, 262).

En el campo del derecho laboral, el uso de la inteligencia artificial y de los algoritmos se encuentran presentes a lo largo de las relaciones de trabajo, desde el proceso de contratación y selección de personal, pasando por la evaluación de desempeño hasta el cese de un trabajador (estabilidad de entrada y salida). Esto implica un gran desafío para no incurrir en una discriminación algorítmica generada por el sistema inteligente. Así, la IA es empleada en la gestión de personal y los procesos de contratación, ya que permite acceder a un gran volumen de datos, los cuales, con el sistema inteligente, bien pueden ser filtrados de manera mucho más rápida y a menor costo. No obstante, se podría incurrir en la generación de sesgos algorítmicos al momento de definir o evaluar los perfiles que se requieren contratar. Un claro ejemplo son los algoritmos de contratación, en los que se

aprecian preferencias de género. Tal es el caso del proceso de selección de personal convocado para el departamento de paquetería de Amazon el año 2018. Descubrieron que el algoritmo rechazaba a las mujeres, las cuales en muchos casos contaban con una formación superior a la de los hombres (Salvatierra, 2023).

Asimismo, la IA puede ser empleada para evaluar el desempeño laboral de los trabajadores con la finalidad de medir en términos numéricos su productividad, basándose en el desempeño de sus pares; inclusive puede determinarse un índice máximo y mínimo para ser considerado dentro de un «desempeño óptimo» o «deseable». Esto implica que los trabajadores deben agotar su máximo esfuerzo para el logro del piso mínimo de desempeño. Esta forma de evaluación del desempeño ya se viene aplicando en algunas entidades financieras del país, y actualmente su validez está siendo discutida en la sede judicial, debido a la falta de transparencia en su procedimiento. Incluso, en ciertos casos, este tipo de evaluación es utilizado como causa justa de despido, a pesar de basarse en sesgos o parámetros que pueden volverse discriminatorios en su aplicación. Al respecto, no debe perderse de vista que la implementación de la IA en las empresas permite automatizar labores rutinarias y repetitivas, lo que contribuye a liberar tiempo, personal y recursos para que se oriente la fuerza de trabajo en actividades más profesionalizadas. Esto mejora, de esta manera, la productividad de la empresa.

Es importante considerar que un algoritmo por sí mismo constituye una regla, ya que su naturaleza encierra una instrucción que busca mandar, prohibir o permitir, por lo que urge que estos algoritmos generados por la IA permanezcan controlados por los humanos, lo que contribuye en la búsqueda del bien común. Por esta razón, se debe optar por la transparencia de los algoritmos, evitando la generación de sesgos, con el fin de proteger la paz social. Esto no implica que

se restrinja el uso de la tecnología, sino que se establezca su control a través de la promulgación de leyes de la materia (Morales, 2021, p.115).

4. PROTECCIÓN NORMATIVA CONTRA LA DISCRIMINACIÓN ALGORÍTMICA

La evolución tecnológica en la que se encuentra inmersa la sociedad avizora que la IA ha sido creada como una herramienta para el progreso de la sociedad, con el fin de mejorar la calidad de vida de la humanidad. Sin embargo, su empleo ha venido presentando una serie de riesgos, como la discriminación algorítmica. Si bien no se genera de manera intencional, se origina ante la falta de diversidad de datos o algoritmos empleados para entrenar los sistemas de inteligencia artificial. Por lo tanto, aunque resulte utópico mencionarlo, se debe optar por una inteligencia artificial convencional que se base en los instrumentos internacionales de protección de los derechos humanos, en su consolidación en el derecho nacional de los Estados parte de un sistema y en las decisiones judiciales de dichos Estados. Esta IA debe ser capaz de compatibilizar las situaciones de hecho con los derechos humanos, los cuales evolucionan de manera progresiva a partir de la interpretación judicial en casos que implican su vulneración (Figueroa, 2024, p. 152).

En el ámbito internacional, el Convenio 111 de la Organización Internacional del Trabajo (OIT), ratificado por el Perú, trata sobre la discriminación en materia de empleo y ocupación y la regulación de la discriminación algorítmica. Actualmente, es un tema complejo y en constante evolución, por lo que las evaluaciones de impacto algorítmico no resultan uniformes. El Observatorio de Inteligencia Artificial de la Organización para la Cooperación y el Desarrollo Económico (OCDE)

ha constatado variadas políticas públicas y estrategias heterogéneas, instrumentos, normas, guías, pactos, códigos, enfoques, cánones de consideraciones sobre la aplicación y las limitaciones de la inteligencia artificial, en las que destacan, en buenas prácticas en la administración y la gestión pública, los gobiernos de Australia, Canadá, Estados Unidos de América, Japón, Nueva Zelanda, Reino Unido y Singapur (Aguirre, 2022, párr. 45 y 46). Además, la Unión Europea ha publicado una estrategia sobre inteligencia artificial que incluye medidas para garantizar la ética y la transparencia en el uso de algoritmos. Así, en la Carta de Derechos Fundamentales de la Unión Europea se prohíbe la discriminación por motivos de origen racial o étnico, y el Reglamento General de Protección de Datos (RGPD) establece el derecho a la explicación y a la revisión de decisiones basadas en algoritmos.

En Latinoamérica, el artículo 1.1. de la Convención Americana sobre Derechos Humanos establece la obligación de los Estados a respetar y garantizar, sin discriminación alguna, los derechos y las libertades allí reconocidos; específicamente, a nivel nacional. Se protege la igualdad formal a nivel constitucional, específicamente en el inciso 2 del artículo 2 de la Constitución Política del Perú que prescribe: «Toda persona tiene derecho: (...) 2. A la igualdad ante la ley. Nadie puede ser discriminado por motivo de origen, raza, sexo, idioma, religión, opinión, condición económica o de cualquier otra índole»; asimismo, el artículo 26, en su inciso 1, establece: «En la relación laboral se respetan los siguientes principios: 1. Igualdad de oportunidades sin discriminación (...)»; por lo que si nos ceñimos a lo señalado por Nogueira (como se citó en Muñoz, 2021) la clásica definición de *igualdad* desde la perspectiva personal o individual comprende lo siguiente: 1) La igualdad ante la ley, referida a la igualdad de trato en situaciones similares, y 2) la igualdad en la ley, que es la expresión sustantiva del derecho, la igualdad entendida como no discriminación.

Ambos ámbitos del principio de igualdad se encuentran protegidos constitucionalmente.

En el Perú, la implementación de la inteligencia artificial soporta un significativo riesgo de sesgos y discriminación. Si bien a la fecha no se conoce un solo caso de discriminación algorítmica en el país, esto no ha sido óbice para que en el año 2021 se active la Estrategia Nacional de Inteligencia Artificial (ENIA), orientada al impulso de la investigación, al desarrollo y a la adopción de la inteligencia artificial en el país, a partir de las pautas y los lineamientos mundiales que contemplan medidas para reducir los sesgos y la discriminación en su uso. Asimismo, desde el punto de vista legal, el 5 de julio de 2023, el Congreso de la República del Perú aprobó la Ley n.º 31814 que promueve el uso de la inteligencia artificial con el fin de fomentar el desarrollo económico y social del país. Su principal objetivo es promover el uso de la IA en el marco del proceso de la transformación digital, en el que se privilegie a la persona y el respeto de los derechos humanos en un entorno seguro que garantice su uso ético, sostenible, transparente, replicable y responsable. En esta norma se establecen principios para su uso y desarrollo que consisten en lo siguiente: i) Estándares de seguridad basados en riesgos; ii) Enfoque de pluralidad de participantes; iii) Gobernanza de internet, la cual se basa en el desarrollo y la aplicación de principios, normas, reglas, procedimientos de toma de decisión y en programas que determinan la evolución y el uso de internet por parte del Estado, instituciones privadas y la sociedad civil; iv) Sociedad digital; v) Desarrollo ético para una inteligencia artificial responsable, que considera la ética como la base fundamental para identificar de forma precisa el marco de responsabilidades en el uso de este tipo de sistemas inteligentes; vi) Privacidad de la inteligencia artificial.

Asimismo, se reconocen las recomendaciones de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) e identifica

a la Presidencia del Consejo de Ministros mediante la Secretaría de Gobierno y Transformación Digital como autoridad nacional para dirigir, evaluar y supervisar el uso y la promoción de la inteligencia artificial.

Mediante la Resolución Ministerial n.º 132-2024-PCM, del 2 de mayo de 2024, la Presidencia del Consejo de Ministros dispuso la publicación del proyecto de reglamento de la Ley n.º 31814, ley que promueve el uso de la inteligencia artificial en favor del desarrollo económico y social del país, del Decreto Supremo y de su respectiva exposición de motivos, cuyo objetivo es establecer las condiciones que garanticen el desarrollo y el uso responsable de la inteligencia artificial. En esta resolución se determinó que se aceptarían comentarios y sugerencias de manera virtual hasta el 1 de junio pasado, a fin de elaborar el reglamento definitivo, el cual hasta la fecha no ha sido expedido.

El proyecto de reglamento presenta aportes importantes en materia de protección de derechos humanos, al abordar los sesgos algorítmicos —también denominados sesgos de automatización— tanto en su definición (artículo 3, literal n) como en el uso ético de los sistemas basados en inteligencia artificial (artículo 20.2, literal b). De esta manera, el implementador de este sistema debe seguir las mejores prácticas, los estándares o las normas técnicas nacionales o internacionales para identificar y minimizar los sesgos de los algoritmos o las bases de datos usados por el sistema basado en la inteligencia artificial. Se establecen, inclusive, como uno de los principios rectores para el uso y el desarrollo de la inteligencia artificial el principio de no discriminación (artículo 4, literal g), el cual establece: «En todas las etapas del ciclo de vida de un sistema basado en inteligencia artificial, se implementan medidas para evitar la generación, el reforzamiento y la perpetuación de resultados discriminatorios o sesgados a fin de evitar cualquier tipo de discriminación de raza, género, origen étnico,

condición económica o social, religión, discapacidad o por cualquier otro motivo»; asimismo, se orienta el uso de la inteligencia artificial con la participación de grupos diversos con énfasis en las poblaciones vulnerables, con la finalidad de buscar la inclusión y la equidad como otro principio rector de este sistema inteligente (artículo 4, literal b).

Es necesario destacar que la Unesco y la República del Perú, el 3 de mayo de 2024, suscribieron un acuerdo para implementar en el país la Metodología de Evaluación de Preparación (RAM) emanada de la recomendación de la Unesco sobre la ética de la inteligencia artificial, lo cual representaba un gran avance para nuestro país, ya que la metodología RAM proporcionará al país herramientas esenciales para evaluar y fortalecer las iniciativas de gobernanzas sobre marcos regulatorios. De esa forma, se garantiza plenamente que el desarrollo de la inteligencia artificial se alinee con los principios éticos. En este sentido, el Perú se ha convertido en el pionero en la región con la Estrategia Nacional de Inteligencia Artificial (ENIA), durante el período 2021-2026, y con la promulgación de la Ley n.º 31814 cuyo objetivo es establecer las condiciones que garanticen el desarrollo y el uso responsable de la inteligencia artificial.

5. CONCLUSIONES

Uno de los mayores riesgos en el uso de la inteligencia artificial a nivel empresarial es la obtención de resultados sesgados o «sesgos algorítmicos», lo que podría originar un riesgo de discriminación por sexo, raza, edad, procedencia, etc. Esto genera un impacto no deseado sobre ciertos grupos de personas, y afecta, de esta manera, la toma de decisiones de un sistema inteligente (Segovia, 2024). Es decir, compromete no solo a aquellos que serían discriminados por la inteligencia artificial, sino también la reputación y la imagen de las empresas o las instituciones.

En este sentido, si bien resulta innegable reconocer que los sesgos algorítmicos son eliminados totalmente, ya que el factor humano que los genera se torna inevitable. No obstante, se podría intentar minimizarlos. Esto se logra con transparencia y responsabilidad al momento de diseñar y utilizar los algoritmos, ya que al conocer cómo han sido generados se comprenderá cómo se han tomado las decisiones. Así, resulta más fácil detectar y corregir cualquier intento de discriminación de ser el caso.

Según refiere Prieto-Pérez (2023), el empleo de sistemas algorítmicos en la selección de personal o reclutamiento, la contratación, la evaluación de desempeño y la culminación del vínculo laboral de los trabajadores plantea desafíos relacionados con los sesgos algorítmicos; por lo tanto, resulta necesario abordar estos retos de manera responsable, garantizando la protección de los derechos y la equidad en los procesos de toma de decisiones relacionados con el empleo, a fin de no afectar la confianza de los trabajadores en el sistema, para la cual se deberá contar con la apertura suficiente para la corrección de posibles errores o sesgos (p.65).

De este modo, resulta necesario implementar una regulación adecuada para abordar la discriminación algorítmica y proteger los derechos humanos, especialmente, los de los grupos más vulnerables, por lo que se necesitan medidas legales claras que establezcan responsabilidades y obligaciones para los diseñadores y los usuarios de algoritmos, con el fin de que se garantice la transparencia y la rendición de cuentas en su diseño y su uso. Entonces, es necesaria la necesidad de supervisión y las sanciones para asegurar el cumplimiento de las regulaciones y la protección de los derechos fundamentales, conforme lo señala Iturmendi.

En este artículo se coincide con Belloso (2022) al señalar que la «condición algorítmica» está incidiendo en los derechos humanos,

que prescriben la igualdad y prohíben la discriminación; así, en una sociedad en la que todo lo que se hace se transforma en datos procesados por algoritmos que contribuyen a la toma de decisiones, los derechos dependen de cómo se regulen estos avances. En este sentido, se debe evitar generar un círculo vicioso de discriminación en línea y en la vida real. Es en este contexto que el sistema jurídico es el responsable de evitar la desigualdad, la parcialidad y la discriminación. Ello se debe lograr con la redefinición de la protección jurídica y de la teoría de los derechos humanos, ya que en la era de la inteligencia artificial se debe orientar por el diseño de algoritmos equitativos e imparciales que faciliten la explicación y la causalidad de estos, como se viene ejecutando en otros países.

Del mismo modo, resulta fundamental que las empresas y la sociedad en conjunto se comprometan con la ética y la equidad en el uso de la inteligencia artificial, con la finalidad de que al emplear este sistema inteligente se obtengan resultados justos, inclusivos, sostenibles y no discriminatorios.

REFERENCIAS

- Aguirre, J. (2022). Modelos y buenas prácticas evaluativas para detectar impactos, riesgos y daños de la inteligencia artificial. *PAAKAT: Revista de Tecnología y Sociedad*, 12 (23), 1-20. <https://doi.org/10.32870/pk.a12n23.742>
- Aliaga, C. (11 de marzo de 2024). *Inteligencia artificial y sesgos de género: Entrenar a los algoritmos para no perpetuar discriminaciones*. Blog Universidad de Chile. <https://uchile.cl/noticias/213982/avances-y-desafios-sobre-inteligencia-artificial-y-sesgos-de-genero>

- Arguelles, E. y Amaro, M. (2023). Preocupaciones éticas en el uso de la inteligencia artificial, transparencia y derecho de acceso a la información. El caso de los chatbots en el gobierno de México, en el contexto de la COVID-19. *Revista Estudios en Derecho a la Información*, (15), 85-111. <https://doi.org/10.22201/ij.25940082e.2023.15.17472>
- Barrios, H., Díaz H. y Guerra, Y. (2020). Subjetividades e inteligencia artificial: desafíos para 'lo humano'. *Veritas*, (47), 81-107. <http://dx.doi.org/10.4067/S0718-92732020000300081>
- Belloso, N. (2022). La problemática de los sesgos algorítmicos (con especial referencia a los de género) ¿Hacia un derecho a la protección contra los sesgos? En J. Garrido y R. Valdivia (coords.), *Inteligencia Artificial y Filosofía del Derecho* (pp.45-78). Ediciones Laborum. https://www.iltras.com.mx/_files/ugd/e3d2a3_829f4f5ac60346e6811f87566d8a32c6.pdf#page=45
- Eguíluz, J. (2020). Desafíos y retos que plantean las decisiones automatizadas y los perfilados para los derechos fundamentales. *Estudios de Deusto*, 68 (2), 325-367. [https://doi.org/10.18543/ed-68\(2\)-2020pp325-367](https://doi.org/10.18543/ed-68(2)-2020pp325-367).
- Figueroa, E. (2024). Inteligencia artificial (IA) emergente: ¿riesgo potencial para los derechos humanos? *Ius Inkarrí*, 13 (15), 143-167. <https://revistas.urp.edu.pe/index.php/Inkarri/article/view/6464/11719>
- Foro Económico Mundial (4 de octubre de 2023). *Así es como la inteligencia artificial transformará la toma de decisiones*. Foro Económico Mundial. <https://es.weforum.org/agenda/2023/10/la-inteligencia-artificial-transformara-la-toma-de-decisiones-asi-es-como/>

- Gómez, L. (2022): Inteligencia artificial y derechos fundamentales. En J. Garrido y R. Valdivia (coords.) *Inteligencia Artificial y Filosofía del Derecho* (pp. 91-114). Ediciones Laborum https://academikaonline.com/wp-content/uploads/2023/02/Inteligencia-artificial-y-derechos-fundamentales_Gomez-Abeja_compressed.pdf
- Iturmendi, J. (2023). La discriminación algorítmica y su impacto en la dignidad de la persona y los derechos humanos. Especial referencia a los inmigrantes. *Deusto Journal of Human Rights*, (12), 257-284. <https://doi.org/10.18543/djhr.2910>
- López, J. (2021). Desenmascarando datos: Igualdad e Inteligencia Artificial. *Revista IUS*, 15 (48), 137-156. <https://doi.org/10.35487/rius.v15i48.2021.740>
- Morales, D. (2021). Implicaciones jurídicas del algoritmo: derechos intelectuales y privacidad. *Foro: Revista de Derecho*, (36), 111-130. <https://doi.org/10.32719/26312484.2021.36.6>
- Morales, G. (2023). Problemática antropológica detrás de la discriminación generada a partir de los algoritmos de la inteligencia artificial. *Medicina y Ética*, 34 (2), 429-455. <https://doi.org/10.36105/mye.2023v34n2.04>
- Muñoz, C. (2021). La discriminación en una sociedad automatizada: Contribuciones desde América Latina. *Revista Chilena de Derecho y Tecnología*, 10(1), 271-307. <http://dx.doi.org/10.5354/0719-2584.2021.58793>
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [Unesco] (2024). *Recomendación sobre la ética en la inteligencia artificial*. Unesco. https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

- Osornio, A. (23 de septiembre de 2023). *Discriminación digital: algoritmos de IA retratan los prejuicios de sus desarrolladores*. Wired. <https://es.wired.com/articulos/discriminacion-digital-algoritmos-de-ia-retratan-prejuicios-de-desarrolladores>
- Prieto-Pérez, T. (2023). *Implicaciones jurídico laborales de la Inteligencia Artificial*. [Tesis de Máster Universitario, Universidad de León]. <https://buleria.unileon.es/handle/10612/17383>
- Puntriano, C. (2024). *Las tecnologías de la información y de las comunicaciones y las relaciones de trabajo. Relevancia jurídica y operativa de las TIC en el ámbito laboral* (1.^a edición, tomo I). Editorial Gaceta Jurídica S. A.
- Ramírez, R. (2023). Sesgos y discriminaciones sociales de los algoritmos en Inteligencia Artificial: una revisión documental. *Entretextos*, 15 (39), 1-17. <https://doi.org/10.59057/iberoleon.20075316.202339664>
- Real Academia Española (2023). *Diccionario de la lengua española* (23.^a ed.), [versión 23.7 en línea]. Asociación de Academias de la Lengua Española. <https://dle.rae.es/algoritmo>
- Salvatierra, E. (2 de octubre de 2023). *La inteligencia artificial discrimina en los procesos de selección*. Blog Cadena Ser. <https://cadenaser.com/nacional/2023/10/02/la-inteligencia-artificial-discrimina-en-los-procesos-de-seleccion-cadena-ser/>
- Sánchez, M. (2022). Los riesgos de la inteligencia artificial para el principio de igualdad y no discriminación. Planteo de la problemática y algunas aclaraciones conceptuales necesarias bajo el prisma del Sistema Interamericano de Derechos Humanos. *Revista elDial.com, Contenidos Jurídicos*. Editorial Albrematica S. A. <https://acortar.link/gT5nu1>

Segovia, J. (25 de abril de 2024). *Riesgos de sesgo y discriminación en Inteligencia Artificial (IA)*. BDO. [https://www.bdo.es/es-es/blogs-es/coordenadas-bdo/riesgos-de-sesgo-y-discriminacion-en-inteligencia-artificial-\(ia\)](https://www.bdo.es/es-es/blogs-es/coordenadas-bdo/riesgos-de-sesgo-y-discriminacion-en-inteligencia-artificial-(ia))

Villegas, A. (7 de abril de 2024). *Sesgos algorítmicos y discriminación de la inteligencia artificial y formas educativas de abordarlo*. AVD aprendizaje digital. <https://adrianvillegasd.com/sesgos-algoritmicos-y-discriminacion-de-la-inteligencia-artificial-y-formas-educativas-de-abordarlo>.

Fuentes normativas y jurisprudenciales

Ley n.º 31814. Ley que promueve el uso de la inteligencia artificial en favor del desarrollo económico y social del país. *Diario Oficial El Peruano*. (5 de julio de 2023). <https://www.gob.pe/institucion/congreso-de-la-republica/normas-legales/4565760-31814>

Resolución Ministerial n.º 132-2024-PCM. Presidencia del Consejo de Ministros (2 de mayo del 2024). <https://www.gob.pe/institucion/pcm/normas-legales/5516872-132-2024-pcm>

Financiamiento

Autofinanciado

Conflicto de intereses

La autora declara no tener conflicto de intereses.

Contribución de autoría

La contribución se ha dado en todos los aspectos del artículo, por ser la única autora.

Biografía de la autora

Abogada, graduada y titulada en la Universidad Nacional de Trujillo (UNT). Maestra en Derecho con mención en Derecho Civil y Comercial en la misma casa de estudios, magíster en Gestión Pública y doctora en Derecho por la Universidad Privada César Vallejo. Labora en la Corte Superior de Justicia de La Libertad desde octubre de 2006 hasta la fecha. Desde 2014 viene desempeñándose como jueza supernumeraria en los Juzgados de Paz Letrado y Juzgados Especializados Laborales de Trujillo. Se encuentra a cargo de procesos tramitados con la Nueva Ley Procesal del Trabajo.

Correspondencia

kcastillo@pj.gob.pe