

## Sesgos algorítmicos y de estereotipos en investigadoras. Por un entorno científico más justo e inclusivo

### Algorithmic and stereotype biases in women researchers. for a fairer and more inclusive scientific environment

Fátima Figueredo<sup>1</sup> 

Fe y Alegría, Valencia, Venezuela  
metodologika2@gmail.com

Recibido: 29/4/2024.

Aceptado: 27/6/2024.

#### RESUMEN

Las mujeres en la investigación desempeñan un papel esencial para el avance del conocimiento científico y el desarrollo social. A medida que la inteligencia artificial (IA) y los algoritmos se integran en diversas áreas de nuestra vida, surgen preocupaciones sobre la equidad y la representación, especialmente debido a los sesgos algorítmicos que pueden discriminar a ciertos grupos, incluidas las mujeres investigadoras. Este trabajo bibliográfico destaca la necesidad urgente de abordar estos sesgos y fomentar una mayor diversidad en los equipos de desarrollo tecnológico. Para lograrlo, se propone la implementación de estrategias concretas, como la capacitación digital dirigida a mujeres para cerrar la brecha digital de género y mejorar la representatividad en los conjuntos de datos utilizados por los algoritmos. Con estas acciones, se busca promover un entorno científico más justo e inclusivo, donde todos los investigadores, independientemente de su género, tengan igualdad de oportunidades para contribuir al avance del conocimiento.

**Palabras clave:** sesgos algorítmicos, estereotipos, investigadoras, equidad

#### ABSTRACT

Women in the field of research play a fundamental role in the advancement of scientific knowledge and social development. However, in today's world, increasingly dependent on technology, artificial intelligence (AI) and algorithms are present in various aspects of our lives. This growing presence also brings with it challenges related to equity and representation. One of the most worrying problems is the existence of algorithmic biases and stereotypes that can discriminate against certain groups of people, particularly women researchers. This bibliographical work proposes to highlight this current topic in order to better inform and understand how algorithms impact our daily lives; calling for reflection so that inequality does not continue to be perpetuated in the field of research; but, rather, to promote a fairer and more inclusive scientific environment.

**Keywords:** algorithmic biases, stereotypes, female researchers, equity

<sup>1</sup>Magíster en Educación Especial y Rehabilitación Integral (UNERG), con mención honorífica. Licenciada en Educación Especial Mención Dificultades de Aprendizaje (UNA Carabobo), Técnico Medio en Enfermería (ETA Francisco Antonio Rísquez), Especialista en Salud Mental y Psiquiatría (Hospital José Ortega Durán). Maestrante en Investigación Educativa (UC). Asesora metodológica en trabajos de investigación.

## La era de la transformación digital: repensando el Derecho en la Cuarta Revolución Industrial

Caracterizada por la convergencia de tecnologías como el Internet de las Cosas (IoT), *blockchain*, nanotecnología, computación cuántica e inteligencia artificial (IA), la Cuarta Revolución Industrial ha irrumpido y transformado nuestras sociedades de manera radical. Este nuevo panorama, que se vio intensificado por la pandemia del COVID-19, presenta desafíos significativos debido a la rapidez con la que se producen los avances tecnológicos y a las desigualdades sociales y económicas que limitan el acceso al conocimiento y la comprensión y adopción de estas tecnologías. Esto puede dificultar que tanto las personas como las organizaciones se adapten y aprovechen las oportunidades que esta nueva realidad ofrece.

Es innegable que estamos viviendo una época en la que la digitalización está cambiando fundamentalmente la forma en que trabajamos, nos relacionamos y creamos valor. Este fenómeno va más allá de simplemente adoptar nuevas tecnologías; representa un cambio profundo en nuestra manera de pensar y de abordar los desafíos en el entorno digital. Así pues, las personas deben estar dispuestas a aprender constantemente y a cuestionar las prácticas establecidas. Particularmente, la inteligencia artificial emerge como una herramienta clave en la transformación organizacional, permitiendo una mejor adaptación a las nuevas realidades del mercado; sin embargo, su implementación también plantea desafíos éticos y de seguridad que deben ser cuidadosamente gestionados.

Para abordar estos desafíos, es fundamental integrar la perspectiva de género y diversidad en el desarrollo de tecnologías digitales. La inclusión de mujeres y otros grupos subrepresentados no solo enriquece el proceso creativo, sino que también ayuda a reducir los sesgos algorítmicos que pueden surgir de una visión limitada. Iniciativas como programas de mentoría para mujeres en tecnología pueden ser fundamentales para cerrar la brecha de género en este campo.

Antes de continuar, es necesario aclarar que la inteligencia artificial (IA) se define como la capacidad de las máquinas para realizar tareas que requieren inteligencia humana, como el aprendizaje, el razonamiento y la toma de decisiones. Esta innovadora tecnología se caracteriza por su autonomía y autoaprendizaje, lo que significa que puede operar de forma independiente y mejorar continuamente su rendimiento. Además, posee la habilidad de procesar grandes cantidades de datos, lo que la convierte en una herramienta invaluable en diversos campos, desde la medicina y el *marketing* hasta la robótica y el transporte autónomo. No obstante, su adopción gradual trae consigo, como ya se dijo, importantes retos éticos y sociales, tales como la necesidad de asegurar la transparencia de los algoritmos y de atender las inquietudes relacionadas con la privacidad y el sesgo. Por lo tanto, es crucial que las organizaciones adopten un enfoque responsable en su desarrollo y uso para aprovechar al máximo sus beneficios y mitigar sus riesgos.

Si bien no existe una definición única y universalmente aceptada de la IA, podemos aproximarnos a su esencia desde diversas perspectivas. La Real Academia Española (s.f.) la define como una “disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico” (documento en línea). Por otra parte, en el ámbito legal, la Comisión Europea, en su propuesta de regulación de la IA de la Unión Europea, conocida como “Ley de Inteligencia Artificial”, la define como:

un sistema de software que utiliza una o más técnicas y estrategias y que, para un conjunto definido de objetivos establecidos por humanos, puede generar información de salida como contenido, predicciones, recomendaciones o decisiones que influyen en los entornos con los que interactúan. (Documento en línea, 2023)

Una descripción más exhaustiva y esclarecedora se encuentra en el documento *Una definición de la inteligencia artificial: principales capacidades y disciplinas científicas*, elaborado por el Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial (2018). En este, se define la IA como:

sistemas de software (y en algunos casos también de hardware) diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno a través de la obtención de datos, la interpretación de los datos estructurados o no estructurados que recopilan, el razonamiento sobre el conocimiento o el procesamiento de la información derivados de esos datos, y decidiendo la acción o acciones óptimas que deben llevar a cabo para lograr el objetivo establecido. (p. 6)

En esencia, la IA se caracteriza por su capacidad para recopilar datos de manera continua, procesar grandes volúmenes de información en tiempo real, operar de forma autónoma y aprender constantemente. Estas habilidades la convierten en una herramienta poderosa con un amplio abanico de aplicaciones en diversos sectores, aunque su potencial conlleva significativos retos éticos y sociales que demandan consideración.

En el mes de marzo de 2024, la Asamblea General de Naciones Unidas aprobó por unanimidad la primera resolución para la regulación internacional de la inteligencia artificial (IA), siendo este el primer fruto de la llamada "conversación global" sobre dicha tecnología. Presentada por Estados Unidos y apoyada por 123 países (incluyendo a China), la resolución busca promover la protección de datos personales, mejorar políticas de privacidad y garantizar un seguimiento de la IA para identificar riesgos potenciales. Este acuerdo, resultado de tres meses de negociaciones, tal como señala Benj Edwards (2024), refleja un consenso internacional sobre la necesidad de regular la IA para que beneficie a la ciudadanía y respete los derechos humanos. La resolución también destaca la urgencia de abordar los peligros del uso irresponsable de esta tecnología en un contexto global cambiante.

Está ampliamente documentado que los sistemas de inteligencia artificial (IA) a menudo se entrenan con datos sesgados que reflejan las desigualdades existentes en la sociedad, lo que puede perpetuar y amplificar la discriminación de género. Investigadoras como Lucía Vicente y Helena Matute<sup>2</sup> (2023) e Iris Domínguez Catena<sup>3</sup> (2024) –incluso, Enzo Ferrante (2021), autor al que se hará referencia más adelante– así lo demuestran.

Sobre la subrepresentación de las mujeres en el sector tecnológico y en la IA, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO, 2023) advierte que las mujeres ocupan menos de un tercio de los puestos en el sector tecnológico y solo el 22 % de los empleos en el área de la inteligencia artificial. La escasa representación femenina en el desarrollo de tecnologías, sin duda, contribuye a los sesgos algorítmicos y dificulta identificarlos y corregirlos. En relación con esto, Syndy García (2024) asegura:

<sup>2</sup>Profesoras de la Universidad de Deusto, autoras del artículo *Los humanos heredan sesgos de inteligencia artificial (2023)*, donde demuestran que los sistemas de IA pueden heredar y amplificar los sesgos humanos, un hecho que exige abordar estos sesgos en el desarrollo de algoritmos.

<sup>3</sup>Investigadora del Instituto de Investigación Smart Cities de la Universidad Pública de Navarra, quien ha desarrollado herramientas para detectar y corregir sesgos demográficos en sistemas de IA. Su trabajo se centra en evidenciar cómo los sesgos estereotípicos pueden afectar el comportamiento final de los modelos de IA.

El panorama en otras áreas no dista mucho. Las mujeres representan el 20 % de las personas empleadas en desempeñar funciones técnicas en empresas de aprendizaje automático, el 12% de quienes investigan IA a nivel mundial y el 6 % de quienes desarrollan profesionalmente softwares, según UNESCO. (Documento en línea)

A esto se suma la falta de transparencia y responsabilidad, ya que los algoritmos de IA pueden ser complejos y difíciles de entender, lo que dificulta la identificación y responsabilización de los creadores de los sesgos o los errores. Esta situación resalta la necesidad urgente de implementar políticas que no solo fomenten la inclusión de las mujeres en el sector tecnológico, sino que también garanticen una mayor claridad en el desarrollo y uso de la IA para disminuir sus efectos negativos.

### **Injusticia algorítmica**

La falta de transparencia en los sistemas de IA puede ocasionar que pasemos de la injusticia humana a la injusticia algorítmica. En el pasado, la responsabilidad de las injusticias recaía directamente sobre los humanos; no obstante, la era digital ha introducido un nuevo agente: los algoritmos de aprendizaje automático. Estos algoritmos, entrenados con vastos conjuntos de casos reales procesados, buscan identificar patrones y replicarlos en nuevos casos. El proceso de entrenamiento implica exponer el algoritmo a datos reales con resultados conocidos, ya sea procesados por otros programas o por humanos. La coincidencia entre el resultado real y el predicho por el algoritmo se conoce como "acierto". Es pertinente señalar también que los algoritmos no emiten resultados absolutos de "correcto" o "incorrecto", sino que buscan aproximarse a los patrones observados en los datos de entrenamiento. Si estos datos presentan sesgos, el algoritmo replicará esos sesgos, perpetuando la injusticia.

Actualmente, se han desarrollado innovaciones tecnológicas entre las cuales se destaca el aprendizaje de máquinas (*machine learning*), que es una rama de la inteligencia artificial y, a su vez, un subcampo de la ciencia de la computación. Según Alejandra Ávila (2019), la inteligencia artificial es la ciencia que permite a las computadoras simular la inteligencia humana y comportarse como esta. El término *machine learning* se utiliza en inglés para describir los procesos mediante los cuales las máquinas "aprenden" de los datos a través de algoritmos, con el objetivo de llegar a conclusiones en forma de predicciones, clasificaciones o segmentaciones.

*Machine learning* opera mediante un sistema que sigue instrucciones y reglas lógicas para llevar a cabo una tarea específica (Flach, citado por Ávila, 2019). Estos algoritmos se producen en múltiples iteraciones, es decir, repiten un proceso varias veces hasta alcanzar un objetivo deseado, mejorando en cada ciclo al adherirse a las reglas e instrucciones establecidas por los humanos. Los parámetros de entrada que utiliza un algoritmo de *machine learning* son datos históricos previamente recogidos y almacenados en bases de datos, que se etiquetan estrictamente para identificar las correlaciones entre todas las variables analizadas y su influencia en la variable dependiente. Esto permite pronosticar la variable dependiente que se busca predecir bajo ciertos parámetros, con la máxima precisión posible.

Por consiguiente, es imperativo formular con precisión las reglas e instrucciones dadas a la máquina para evitar omisiones importantes. Esto es especialmente relevante porque supervisar el trabajo completo de una máquina puede ser complicado y, aunque no siempre es posible establecer *a priori* el tiempo que un dispositivo tardará en procesar cálculos y estimaciones mediante *machine learning*, una planificación adecuada puede ayudar a optimizar este proceso y minimizar posibles inconvenientes. En opinión de Sebastian Raschka y Vahid Mirjalili (2019), el riesgo radica en que la máquina pueda tomar decisiones no contempladas por los

programadores, lo que podría dar lugar a resultados no deseados en situaciones atípicas o específicas.

Contingencias de este tipo pueden provocar sesgos algorítmicos que restrinjan el acceso de las mujeres a la información y recursos esenciales. Aquellas con menor acceso a internet y dispositivos electrónicos enfrentan dificultades para entrar a bases de datos académicos, revistas científicas y otras fuentes de información relevantes para su trabajo de investigación o científico. Como resultado, se ve afectada su participación y colaboración en el intercambio de conocimientos con colegas<sup>4</sup>, aspectos fundamentales para el desarrollo de investigaciones científicas y la ampliación del impacto de sus contribuciones.

De igual forma, la publicación de artículos científicos en revistas de alto impacto ha sido tradicionalmente considerada determinante para el reconocimiento y la difusión del trabajo de investigación<sup>5</sup>. En tal sentido, estudios realizados por la UNESCO, como *I'd blush if I could. Closing gender divides in digital skills through education* (2019), han demostrado que las mujeres investigadoras enfrentan desafíos adicionales que pueden limitar sus oportunidades para publicar en estas revistas, lo que a su vez puede afectar su visibilidad y su capacidad para obtener financiación para futuras investigaciones. Estos desafíos pueden estar relacionados con sesgos algorítmicos en los sistemas de evaluación y selección que predominan en el ámbito académico.

Adicionalmente, la escasa presencia de mujeres en los equipos que diseñan los sistemas de IA se explica por la ya mencionada brecha de género en la adquisición de competencias digitales, que se extiende a la infrarrepresentación de las mujeres entre quienes se dedican profesionalmente a las disciplinas de ciencia, tecnología, ingeniería y matemáticas (STEM, por sus siglas en inglés). Según la UNESCO (2019), las mujeres tienen un 25% menos de probabilidad que los hombres de saber cómo aprovechar la tecnología digital para fines básicos, son cuatro veces menos propensas a conocer cómo programar una computadora y trece veces menos propensas a presentar una solicitud de patente electrónica.

De acuerdo con la misma fuente, solo el 6% de las desarrolladoras de aplicaciones móviles y *software* en el mundo son mujeres (UNESCO, 2019, p. 26). Según datos del Parlamento Europeo (2020), en 2018, las mujeres solo representaron el 22% de los profesionales mundiales de la IA, mientras que un estudio de 2017 mostraba que solo el 13% de las altas posiciones ejecutivas en empresas tecnológicas dedicadas a la IA son ocupadas por mujeres (Belloso Martín, 2022, p. 67). Los datos apuntados ponen de manifiesto cómo el sector tecnológico se encuentra fuertemente masculinizado, proyectando la realidad de un mundo digital menos igualitario que el real.

Esta realidad hace evidente la discriminación algorítmica, pues la creciente utilización de algoritmos en diversos ámbitos, incluyendo la evaluación de propuestas de investigación, la selección de becas y la revisión de artículos académicos, ha generado preocupación sobre la posibilidad de que los sesgos algorítmicos discriminen a las mujeres investigadoras. Estos sesgos pueden estar presentes en los datos utilizados para entrenar los algoritmos, así como en los criterios de evaluación establecidos.

<sup>4</sup>Aunque los hombres pueden enfrentar desafíos similares, hay que reconocer las particularidades de las experiencias de las mujeres en la investigación: las expectativas sociales y responsabilidades familiares pueden impactar su tiempo y recursos disponibles para acceder a información y participar en redes profesionales.

<sup>5</sup>No obstante, es importante señalar que el modelo de ciencia abierta ha demostrado ser efectivo en la difusión de la producción académica, independientemente de las publicaciones en revistas de alto impacto.

La preocupación por la discriminación algorítmica se ve reflejada en las palabras de Cecilia Danesi (2022), quien afirma, en *El imperio de los algoritmos*, que:

la tecnología en sí es neutral, pero desde el mismo momento en el que las desarrollamos, las estamos cargando de sentido porque desde que las imaginamos, la pensamos, las estamos moldeando e influenciando por nuestros valores, nuestros prejuicios, nuestros sesgos. El punto no es la tecnología, sino quienes las están creando. Nosotros la estamos utilizando, la verdad, como podemos porque en la mayoría de los casos hay sistemas de Inteligencia Artificial que toman decisiones por nosotros. Pero no somos conscientes de eso. (p. 5)

Conviene destacar que el debate sobre la neutralidad tecnológica es complejo y ha sido ampliamente discutido en los últimos años, con diversas propuestas que ofrecen diferentes enfoques sobre el tema. La perspectiva de Danesi pone de manifiesto, en primer lugar, que la influencia de nuestros valores y prejuicios en el desarrollo tecnológico es un aspecto crítico a considerar y, luego, resalta la necesidad de revisar cómo los sistemas de inteligencia artificial son diseñados y utilizados, ya que las decisiones tomadas por quienes los crean pueden perpetuar desigualdades, especialmente en contextos donde los sesgos algorítmicos pueden discriminar a grupos como las mujeres investigadoras.

Las científicas enfrentan una serie de desafíos en su vida profesional y personal, entre los cuales destacan los tecnológicos. El acceso desigual a la tecnología, especialmente en países en vías de desarrollo, limita su capacidad para utilizar herramientas avanzadas necesarias para la gestión de datos, análisis estadístico y comunicación científica. Esta brecha digital existente puede limitar sus habilidades para emplear *software* especializado, lo que afecta su desempeño investigativo. Asimismo, es determinante que los algoritmos utilizados en la búsqueda de información y la evaluación de propuestas sean diseñados con equidad, para evitar perpetuar desigualdades en el acceso a oportunidades.

Los algoritmos utilizados para seleccionar becas, contratar investigadores y publicar artículos científicos deben ser lo más equitativos posible. Esto implica no solo diseñar criterios de selección justos, sino también garantizar que estos algoritmos sean transparentes, permitiendo que las personas comprendan cómo funcionan y qué datos se utilizan para entrenarlos. La transparencia en el uso de algoritmos es esencial para fomentar un entorno más justo y accesible en la investigación.

### **Sesgos de estereotipos**

Los sesgos de estereotipos en mujeres investigadoras son, de acuerdo con lo expuesto, un fenómeno que afecta profundamente su experiencia en el ámbito científico. Según la matemática y escritora estadounidense Cathy O'Neil (2018), estos sesgos se refieren a actitudes y creencias preconcebidas que influyen, de manera consciente o inconsciente, en la percepción y el trato que reciben las mujeres en el ámbito científico. Estos estereotipos de género limitan sus oportunidades y desarrollo profesional en la ciencia, distorsionando así la realidad y ampliando o restringiendo sus oportunidades en la vida. O'Neil también señala que "estos algoritmos son destructivos y debilitan su propio objetivo original, como la mejora del sistema educativo" (p. 37), frase con la que subraya la paradoja de que herramientas diseñadas para fomentar la equidad pueden exacerbar las desigualdades existentes.

Casos como los mencionados evidencian que la revolución tecnológica, en lugar de contribuir a erradicar los sesgos humanos, parece haberlos ocultado bajo la falsa apariencia de neutralidad (O'Neil, 2018). Esta ilusión de neutralidad es engañosa. Las decisiones tecnológicas a

menudo se presentan como asépticas y objetivas, cuando en realidad reflejan los valores morales y prejuicios de la cultura que las desarrolla. La ciencia y la tecnología no son entes aislados; están impregnadas de los mismos sesgos que afectan a las personas. Esto nos invita a cuestionar cómo se desarrollan estos sistemas y quiénes están detrás de ellos, enfatizando la necesidad de mayor diversidad y equidad en los equipos que crean estas tecnologías.

La crítica de Cathy O'Neil sobre los algoritmos y su impacto en la equidad social encuentra eco en la reseña de José Arribas (2018), quien también aborda cómo la tecnología ha perpetuado los sesgos humanos. Arribas (2018) señala que, al adoptar modelos de *big data*, las empresas han dejado de lado la representatividad y validez de los datos, priorizando análisis masivos que ignoran problemas metodológicos fundamentales. Esta tendencia hacia la opacidad y el secretismo en el diseño algorítmico, como menciona O'Neil, no solo afecta el ámbito educativo, sino que también se extiende a otros sectores críticos como la salud y la administración pública. Ambos autores coinciden en que los algoritmos, lejos de ser herramientas neutrales, reflejan y amplifican las ideologías y prejuicios de la sociedad que los crea. Por lo tanto, es imperativo incorporar valores humanos en el desarrollo de algoritmos, para asegurar que sirvan al interés general y no solo al beneficio económico.

La discusión sobre los sesgos en los algoritmos, además de atender su impacto en la equidad social, también abarca cómo estos influyen en la percepción de la identidad digital de los individuos. La Dra. Johanna Faliero, especialista en Derecho Informático, se ha pronunciado sobre este punto y afirma que "la identidad digital del individuo conoce más sobre un sujeto de lo que este conoce sobre sí mismo" (2021, p. 121), lo que resalta cómo los algoritmos pueden revelar aspectos profundos e inconscientes de la identidad personal.

Para comprender el impacto de los algoritmos en la identidad digital, hace falta reconocer cómo estos sistemas de procesamiento interactúan con la psique del individuo. En este sentido, Faliero destaca que:

Los algoritmos logran penetrar en el individuo, en su psiquis, en sus aspectos más íntimos y personales, desde un enfoque exógeno, con elementos externos al individuo, quien puede conocer o desconocer qué se ha obtenido de él. Del mismo modo, los algoritmos, como técnicas de procesamiento, carecen de sensibilidad, criterio y orientación natural en derechos humanos y pueden acertar o errar, ya que sus aciertos no se guían por la certeza de la veracidad comprobada, sino por aquella obtenida por la incidencia estadística. Por lo tanto, en este camino, los algoritmos que definen la identidad digital de los individuos dicen lo que estos son, lo que estos desconocen que son y lo que estos no son, pero los algoritmos creen que son. (p. 121)

Este planteamiento encierra serias preocupaciones sobre la veracidad y justicia de las representaciones digitales generadas por los algoritmos, en tanto pueden distorsionar esa identidad al basarse en datos estadísticos en lugar de realidades comprobadas. De ahí la insistencia en que se concrete una ética de los datos, pues toda la humanidad es objeto del perfilamiento algorítmico.

Johanna Faliero (2021) también advierte que la inteligencia artificial no es inocua. Como herramienta diseñada para realizar análisis predictivos más eficientes, basa su efectividad en la inferencia estadística de patrones y en una interpretación flexible de los parámetros con los que se configura. Además, sus procesos de toma de decisiones –asegura Faliero– son a menudo difíciles o imposibles de explicar y transparentar, a pesar de los esfuerzos para abordar esta problemática.

En lugar de fomentar igualdad en un mundo polarizado, la inteligencia artificial ha introducido mecanismos discriminatorios menos evidentes pero igualmente dañinos. El sesgo algorítmico contamina el reconocimiento de patrones y predicciones basadas en datos. Los algoritmos no son entidades imparciales; están influenciados por sus creadores e intereses, así como por los conjuntos de datos con los que se alimentan. Lo anterior facilita entender cómo la discriminación permea desde los datos hasta sus resultados, en tanto que, si los algoritmos se entrenan con datos sesgados, es probable que produzcan resultados igualmente sesgados.

Esta situación tiene un impacto directo en las mujeres investigadoras, ya que los algoritmos utilizados para evaluar publicaciones, financiamiento y oportunidades laborales pueden favorecer a candidatos masculinos debido a patrones históricos en los datos. Esto no solo perpetúa la desigualdad, sino que también limita la visibilidad y el reconocimiento del trabajo realizado por mujeres en el ámbito académico. Para cerrar la brecha de género en este campo, es propicio alentar iniciativas como programas de mentoría para mujeres en tecnología y realizar auditorías regulares de los algoritmos utilizados en estos contextos. Estas auditorías son esenciales para identificar y corregir sesgos que puedan afectar negativamente a las investigadoras. Al implementar estas estrategias, se busca no solo aliviar los efectos de los sesgos algorítmicos, sino también fomentar un entorno más equitativo en la investigación científica.

### **Algoritmos de equidad o *fairness***

La implementación de algoritmos de equidad o *fairness*, como AI Fairness 360 y Google What-If, es esencial para los grupos que enfrentan discriminación en la actualidad. En un futuro donde la inteligencia artificial se generalice aún más, la falta de estas herramientas podría resultar en un aumento de la desigualdad social. Es preciso acotar que, cuando se habla del aprendizaje automático, la equidad o *fairness* se define como la capacidad de los algoritmos para no discriminar a ningún grupo de personas en particular, lo que implica que los resultados del algoritmo deben ser independientes de atributos sensibles como género, raza, religión u orientación sexual.

En su trabajo *Inteligencia artificial y sesgos algorítmicos: ¿Por qué deberían importarnos?* (2021), Enzo Ferrante examina cómo la inteligencia artificial se ha integrado en la vida cotidiana, a menudo de manera imperceptible, y cómo los sesgos algorítmicos pueden perpetuar desigualdades. Destaca que estos sesgos surgen de datos no representativos y de las decisiones de diseño de los desarrolladores, quienes pueden no ser conscientes de sus propios prejuicios. Ferrante subraya que los sistemas de IA son diseñados por personas con visiones del mundo y experiencias que influyen en su funcionamiento. Si esos grupos de trabajo carecen de diversidad, es muy probable que no reconozcan la existencia de sesgos y, por ende, no puedan corregirlos. Para construir una IA más justa, es primordial centrar la atención en los datos, los modelos y las personas involucradas, así como fomentar la diversidad en los equipos de desarrollo para mitigar el impacto negativo de estos sesgos en grupos vulnerables.

Ferrante también muestra un optimismo fundamentado en la creciente conciencia sobre los riesgos asociados a la inteligencia artificial y los esfuerzos para aumentar la diversidad en el campo. Iniciativas como Chicas en Tecnología y WomenInML están comenzando a cuestionar la realidad actual. Además, la creación de foros de discusión y el interés interdisciplinario abren nuevas oportunidades para un desarrollo científico más consciente. Como enfatiza el investigador y docente argentino: no se trata de frenar el avance de la inteligencia artificial, sino de asegurar que su desarrollo y uso sean responsables y equitativos, beneficiando así a toda la sociedad.

Laura Flores (2023), por su parte, reitera que los algoritmos a través de los cuales opera la IA no son neutrales y reflejan los valores culturales de quienes los crean y les suministran la información con la que trabajan. Para evitar que estos sistemas reproduzcan sesgos de género que perpetúen estereotipos sexistas y decisiones discriminatorias, es necesario implementar estrategias clave en el sector tecnológico. Estas incluyen: (1) proporcionar competencias digitales a las mujeres para cerrar la brecha digital de género; (2) asegurar la calidad representativa de los metadatos; (3) entrenar a las IA con una perspectiva de género e incorporar para corregir sesgos; y (4) garantizar la transparencia en el código utilizado para desarrollar algoritmos.

En este escenario, las mujeres investigadoras desempeñan un papel fundamental en el avance del conocimiento científico, por lo que se hace imprescindible seguir trabajando para eliminar las barreras que obstaculizan su progreso y crear un entorno más equitativo e inclusivo en el ámbito científico. Al hacerlo, no solo se beneficiará a las mujeres investigadoras, sino que también se fortalecerá la comunidad científica en su conjunto. El sesgo algorítmico puede tener un impacto negativo significativo en las mujeres dedicadas a la investigación, dificultando su acceso a información, financiación, oportunidades laborales y publicaciones. A largo plazo, esto puede resultar en una menor representación femenina en los campos de investigación. La equidad en la inteligencia artificial y la supresión (o, cuando menos, la disminución) de los sesgos algorítmicos es un imperativo para un futuro justo.

## Referencias

- Arribas, José (2018). CATHY O'NEIL: Armas de Destrucción Matemática. Cómo el Big Data aumenta la desigualdad y amenaza la democracia. Madrid: Capitán Swing, 2017, 269 pp. *Empiria*, (41), pp. 199-202. <https://www.redalyc.org/journal/2971/297165396011/html>.
- Ávila, Alejandra (2019). *Principios para prevenir los sesgos en los algoritmos del machine learning* [Tesis de grado, Universidad de los Andes]. Repositorio institucional – Universidad de los Andes. <https://repositorio.uniandes.edu.co/bitstreams/a9799464-ebc2-43ee-8d91-1f2704f8f327/download>.
- Belloso Martín, Nuria (2022). La problemática de los sesgos algorítmicos (con especial atención a los de género). ¿Hacia un derecho a la protección contra los sesgos? En Llano Alonso, Fernando (Dir.), *Inteligencia Artificial y Filosofía del Derecho* (pp. 45–69). Laborum ediciones.
- Comisión Europea (2023). *Comunicado de Comisión Europea acerca del acuerdo político sobre la Ley de Inteligencia Artificial*. [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_23\\_6473](https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473).
- Danesi, Cecilia (2022). *El imperio de los algoritmos*. Galerna.
- Domínguez, Iris (2024). *Demographic bias in machine learning: measuring transference from dataset bias to model predictions*. [Tesis doctoral, Universidad Pública de Navarra]. <https://academica-e.unavarra.es/entities/publication/2d2cc1aa-b0fc-4260-b1c5-77f108fac0e8>.
- Edwards, Benj (2024, 22 de marzo). La ONU aprueba la primera resolución global sobre inteligencia artificial. *Ars Technica*. <https://es.wired.com/articulos/onu-adopta-la-primera-resolucion-global-sobre-inteligencias-artificial>.
- Faliero, Johanna (2021). Limitar la dependencia algorítmica. Impactos de la inteligencia artificial y sesgos algorítmicos. *Revista Nueva Sociedad*, 294, pp. 120-129. <https://library.fes.de/pdf-files/nuso/nuso-294.pdf>.

- Ferrante, Enzo (2021). Inteligencia artificial y sesgos algorítmicos ¿Por qué deberían importarnos? *Revista Nueva Sociedad*, 294, pp. 27-36. <https://library.fes.de/pdf-files/nuso/nuso-294.pdf>.
- Flores, Laura (2023). Sesgos de género en la inteligencia artificial: el estado de derecho frente a la discriminación algorítmica por razón de sexo. *Revista Internacional de Pensamiento Político*, 18, pp. 95-120. <https://www.upo.es/revistas/index.php/ripp/article/view/8778>.
- García, Syndy (2024, 8 de marzo). Mujeres en IA: ¿quiénes son líderes y por qué se necesita a más mujeres en esta industria? *Voz de América*. <https://www.vozdeamerica.com/a/mujeres-en-ia-quienes-son-lideres-por-que-se-necesita-mas-mujeres-en-esta-industria-/7518178.html>.
- Grupo Independiente de Expertos de Alto Nivel sobre la Inteligencia Artificial (2018). *Una definición de la inteligencia artificial: principales capacidades y disciplinas científicas*. Unión Europea. <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.
- Naciones Unidas (14 de marzo de 2023). Tecnología e igualdad de género: situar a las mujeres y las niñas en el centro de la innovación. *Crónica ONU*. [https://www.un.org/es/cr%C3%B3nica-onu/tecnolog%C3%ADa-e-igualdad-de-g%C3%A9nero-situar-las-mujeres-y-las-ni%C3%B1as-en-el-centro-de-la#:~:text=Las%20mujeres%20ocupan%20menos%20de,la%20inteligencia%20artificial%20\(IA\)](https://www.un.org/es/cr%C3%B3nica-onu/tecnolog%C3%ADa-e-igualdad-de-g%C3%A9nero-situar-las-mujeres-y-las-ni%C3%B1as-en-el-centro-de-la#:~:text=Las%20mujeres%20ocupan%20menos%20de,la%20inteligencia%20artificial%20(IA)).
- O'Neil, Cathy (2018). *Armas de destrucción matemática. Cómo el Big Data aumenta la desigualdad y amenaza la democracia*. Capitán Swing S.L.
- Raschka, Sebastian y Mirjalili, Vahid (2015). *Python Machine Learning*. Packt Open Source.
- Real Academia Española (s.f.). *Inteligencia artificial*. En *Diccionario de la lengua española* Recuperado de <https://dle.rae.es/inteligencia?m=form#2DxmhCT>.
- UNESCO (2019). *I'd blush if I could. Closing gender divides in digital skills through education*. <https://unesdoc.unesco.org/ark:/48223/pf0000367416>.
- Vicente, Lucía y Matute, Helena (2023). Los humanos heredan sesgos de inteligencia artificial. *Scientific Reports*, 13. <https://doi.org/10.1038/s41598-023-42384-8>.