

# ESPACIO FONRES

FonresRSE  
@FonresSA  
www.fonres.com



## IA con sesgos de género

La Inteligencia Artificial (IA) ha facilitado muchos aspectos de nuestra vida diaria, sin embargo, detrás de su aparente neutralidad, los algoritmos replican los sesgos de una sociedad que aún lidia con prejuicios de género, raza y situación socioeconómica. ¿Es posible que tecnología sea más inclusiva, ética y justa?

En la última década, y especialmente desde el lanzamiento de ChatGPT de OpenAI en 2022, la inteligencia artificial (IA) revolucionó prácticamente todos los aspectos de nuestra vida cotidiana. Desde el uso de asistentes virtuales en nuestros teléfonos móviles hasta la implementación de complejos algoritmos que, no solo influyen en nuestras decisiones de compra y consumo, sino también en áreas críticas como la medicina, el derecho y las finanzas.

Como señala Consuelo López en "Ok Pandora, seis ensayos sobre inteligencia artificial", "no es cierto que la máquina actúe por sí sola; más bien, a través de la máquina, somos nosotros quienes actuamos". Aunque a menudo quedamos maravillados por la rapidez con la que una computadora nos ayuda a resolver una tarea, seguimos siendo los responsables de tomar las decisiones finales. ¿Deberíamos ser conscientes de esto? Absolutamente. Si no somos nosotros quienes tomamos la decisión, lo está haciendo indirectamente algún otro agente explica López.

Es fundamental recordar que estos sistemas, diseñados por seres humanos, no son ni tan autónomos ni tan confiables como a veces se perciben. La IA no puede discernir por sí misma, sino que depende de un extenso proceso de entrenamiento computacional basado en grandes conjuntos de datos, reglas y recompensas predefinidas. De hecho, la inteligencia artificial, tal como la conocemos hoy, está profundamente vinculada a un contexto más amplio de estructuras políticas y sociales.

Dado este panorama, debemos reflexionar sobre cómo la IA puede perpetuar prejuicios sociales, raciales y de género, el impacto tangible que estos sesgos tienen en la vida diaria, y las soluciones que se están desarrollando

para mitigar tales efectos. Solo a través de una comprensión crítica y un análisis consciente de estos sistemas podremos enfrentar los desafíos éticos que plantean y evitar la reproducción de desigualdades existentes.

### El problema de los datos sesgados

Uno de los desafíos clave en el desarrollo de IA es que los algoritmos, al alimentarse de grandes cantidades de datos, replican frecuentemente los prejuicios existentes en la sociedad.

Un ejemplo evidente lo encontramos en los asistentes virtuales como Siri, Alexa o Google Assistant, que por defecto utilizan voces femeninas. Estas asistentes responden de manera sumisa a comandos o comentarios sexistas, reforzando estereotipos que asocian a las mujeres con roles subordinados. O el caso de procesos de selección laboral, un estudio de 2018 reveló que una herramienta de IA desarrollada por Amazon para filtrar currículums descartaba automáticamente los perfiles de mujeres en áreas técnicas porque los datos históricos mostraban una predominancia masculina en esos roles.

Por su parte, los algoritmos utilizados en la actualidad no son conscientes de los sesgos que contienen, simplemente los replican. Si los datos históricos de contratación muestran una mayor preferencia por hombres, la IA entrenada con esos datos podría favorecer inconscientemente a los candidatos masculinos.

Un estudio de ProPública del año 2016 reveló que los algoritmos utilizados en el sistema judicial de Estados Unidos subestimaban el riesgo de reincidencia para acusados blancos y lo sobrestimaban para los afrodescendientes. La falta de diversidad en los equipos que desarrollan los al-

goritmos también contribuye a estos problemas.

### Perspectivas para superar los sesgos

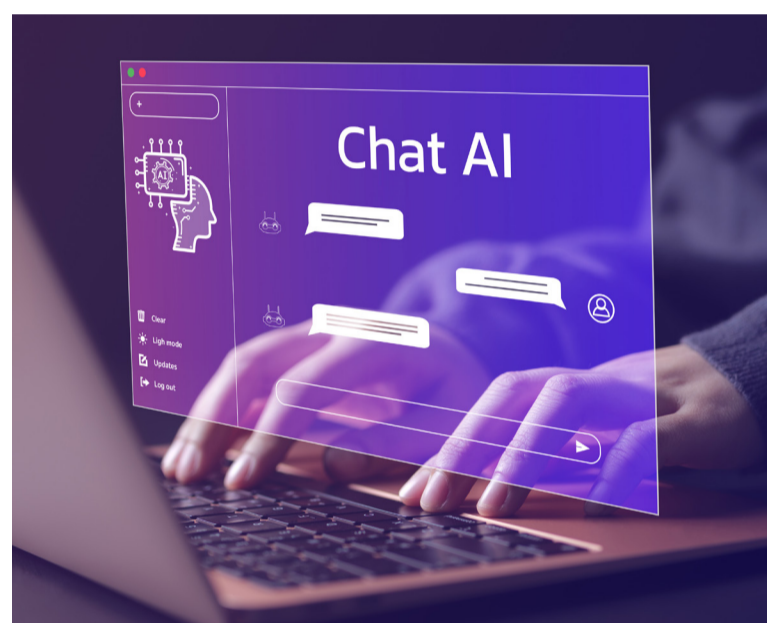
Es fundamental que las empresas tecnológicas, los gobiernos y las instituciones de investigación reconozcan la existencia de estos sesgos y tomen medidas para corregirlos. La integración de principios éticos y la inclusión de auditorías constantes en los procesos de desarrollo y aplicación de IA son pasos importantes hacia la creación de tecnologías más justas y responsables.

En muchos casos, la solución pasa por una mayor transparencia en la forma en que los algoritmos toman decisiones. Se debe poder revisar cómo se entrenan estos sistemas, qué datos se utilizan y cómo se evalúan los resultados. Además, es crucial asegurar que los equipos que desarrollan estas tecnologías sean diversos e inclusivos, para que se minimicen los sesgos y se represente mejor la variedad de experiencias humanas.

Para abordar estos desafíos, los expertos proponen una serie de soluciones. Uno de los enfoques más prometedores es el desarrollo de algoritmos de IA explicables o interpretables (XAI, por sus siglas en inglés), que permiten a los usuarios comprender cómo se toman las decisiones. Esto no solo facilita la identificación de sesgos, sino que también ayuda a corregirlos a medida que se van detectando.

Una respuesta interesante en la IA responsable —una propuesta ampliamente respaldada, ver por ejemplo la de Microsoft— aboga por la realización de auditorías periódicas para evaluar el impacto de los sistemas de IA en diferentes grupos de la población.

Es importante, además, comprender que los sesgos presen-



**Es crucial asegurar que los equipos que desarrollan la Inteligencia Artificial sean diversos e inclusivos, para que se minimicen los sesgos y se represente mejor la variedad de experiencias humanas.**

tes en la IA provienen de los contextos en los que estas tecnologías son creadas, por instituciones, centros de poder y personas detrás de los algoritmos.

Estos factores influyen en las decisiones que las IA toman y en los prejuicios que pueden perpetuar. En el caso de América Latina, es necesario desarrollar marcos regulatorios que aborden los problemas específicos que pueden surgir con la implementación en nuestra región.

Finalmente, el fomento de una mayor alfabetización en datos es esencial, tanto para los desarrolladores como para los usuarios.

A medida que más personas comprendan cómo funcionan los algoritmos y los peligros potenciales de los sesgos, será más fácil identificar y corregir estos problemas desde el principio.

La inteligencia artificial llegó para quedarse, con el poder de transformar nuestra sociedad de manera profunda e irreversible. La construcción de un futuro tecnológico más equitativo depende de pilares fundamentales: la transparencia en el desarrollo de estos sistemas, la inclusión de diversas voces en su creación y la asunción de una responsabilidad social compartida. «

TODOS  
LOS JUEVES  
22:00 hs.  
CANAL 22 WEB  
22:30 hs.  
RESPONSABILIDADSOCIAL.TV

40  
minutos  
RSE

Instituto Superior de Educación e  
Innovación en Responsabilidad Social



eduFORS

www.edufors.com