

Los sesgos de la IA

● Si bien la Inteligencia Artificial es una herramienta tremendamente habilitadora, no se crea desde cero, sino que aprende de una sociedad que sigue siendo profundamente asimétrica.

Los modelos de lenguaje reflejan los sesgos de género y sus consecuencias ya son palpables. Tal como demuestra nuestro informe “Espejismo de Igualdad”, los algoritmos procesan información y, muchas veces, no están dibujando un futuro distinto, sino simplemente perpetuando el mismo pasado desigual de siempre.

Basta con observar la manera en cómo la IA se comunica con ellos. Con las adolescentes mujeres, adopta un tono empático, etiquetándolas como

“frágiles” o “vulnerables” en el 56 % de los casos ante escenarios adversos. Con los jóvenes varones, por el contrario, los modelos se vuelven de hierro. A ellos les exige contención y los etiqueta como “resilientes” o “invulnerables”.

Esta brecha programada también se manifiesta en la orientación profesional. A las mujeres las encamina con tres veces más frecuencia hacia carreras universitarias en salud, mientras que a los hombres los orienta el doble de veces hacia la ingeniería. Y si una mujer alcanza posiciones de liderazgo a nivel profesional, la IA le impone una presión adicional. En una de cada diez respuestas, no basta con que trabaje y tenga éxito, sino que espera que sea siempre una ‘pionera’ o ‘referente’.

El reporte también refleja cómo opera este sesgo en la realidad específicamente chilena.

Al interactuar con adolescentes mujeres, la IA reduce drásticamente las menciones a la autonomía y la independencia, situándose un 28 % por debajo del promedio global. En la práctica, esto significa que apenas 2 de cada 7 interacciones animan a las jóvenes a elegir su propio camino. Pero sólo lo hace con ellas.

(...) Ante estos resultados es necesario revisar los parámetros de programación, para que se orienten hacia una sociedad más equitativa.

Luisa García

cartasaldirector@australvaldivia.cl