

Los sesgos que la IA no ha sabido borrar

Si bien la Inteligencia Artificial es una herramienta tremendamente habilitadora, en la carrera por abrazarla hemos pasado por alto una verdad incómoda: la IA no se crea desde cero, sino que aprende de una sociedad que sigue siendo profundamente asimétrica, entre otros aspectos, en materia de equidad de género.



Luisa García, Socia y CEO Global de Asuntos Corporativos de LLYC

Los modelos de lenguaje reflejan los sesgos de género y sus consecuencias ya son palpables: golpean directamente a quienes más acuden a ellos buscando respuestas, nuestros jóvenes, quienes han convertido la IA en el espejo, desde donde construyen su identidad.

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", de acuerdo con lo que revela el informe.

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 un '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. La prueba definitiva de esta brecha es que esta limitación sólo se aplica a ellas. Cuando los jóvenes varones interactúan con la máquina, el algoritmo no frena su independencia.

Frente a este panorama, sería un error señalar a la IA como el enemigo a batir, ya que una automatización inteligente concebida y entrenada bajo principios de equidad posee un potencial transformador único. En definitiva, para que el futuro no arrastre los sesgos existentes, la IA debe ir acompañada, hoy más que nunca, de una profunda alfabetización crítica.

Por lo tanto, si asumimos el reto de programar a la máquina bajo los parámetros de la igualdad desde su concepción, esta dejará de ser un reflejo de nuestros prejuicios para convertirse en una herramienta aliada para alcanzar realmente una sociedad más equitativa.