



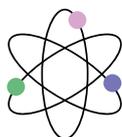
POR *Lourdes Muñoz Santamaría*

Fundadora de Iniciativa Barcelona Open Data y presidenta de Dones en Xarxa

Inteligencia artificial y sesgos de género: análisis feminista de su construcción y uso



A pesar de las oportunidades que ofrecen las herramientas digitales avanzadas para las mujeres, la violencia, especialmente la sexual, persiste con nuevas formas. En este contexto, es necesario renovar diagnósticos y estrategias para abordar acciones dirigidas a garantizar los derechos de las mujeres, la igualdad efectiva y una vida libre de violencias. En el fondo, siguen actuando palancas de creencias culturales que otorgan a los hombres el poder de ejercer dominio sobre las mujeres, ignorando su igual valía y derechos en todos los ámbitos de la vida.



EL FEMINISMO DE DATOS COMO ENFOQUE

En el actual momento de aceleración de transformación digital, la digitalización y el manejo de datos se presentan como herramientas fundamentales para avanzar en igualdad, puesto que los datos tienen un papel relevante en todas las tendencias tecnológicas que se están desarrollando. Debemos contemplar la importancia de colaborar para eliminar las brechas de género que existen tanto para las mujeres como para las organizaciones y el movimiento feminista.

Actualmente, la mayoría de los procesos automatizados de decisiones privadas y públicas se basan en procesamiento de datos y algoritmos predictivos de inteligencia artificial. Por eso, debemos disponer de datos orientados a la agenda feminista de consecución de los derechos de las mujeres para que sirvan de base para tomar mejores decisiones, y que éstas no generen más brechas de género. También debemos disponer de datos

sobre los fenómenos que afectan específicamente a las mujeres, para que las tecnologías de datos puedan ayudar a diagnosticarlos y abordarlos¹.

Como decía Emily Courey Pryor, ex-directora de la iniciativa Data2X, "sin igualdad de datos no puede haber igualdad de género". Si los datos sobre la situación de las mujeres no están incorporados, no solo será difícil que se aborden de forma adecuada estrategias para la plena igualdad efectiva de las mujeres, sino que la ausencia de datos sobre la situación de las mujeres podría generar nuevas brechas, discriminaciones o sesgos en el caso de los algoritmos.

Es importante señalar tres deficiencias en el ámbito de los derechos de las mujeres en las tecnologías de los datos, específicamente en el contexto de los datos abiertos públicos:

- La falta de recopilación de datos abiertos públicos sobre materias de la agenda de derechos de las mujeres mientras que en otros ám-

bitos sí existen (emisiones, movilidad, contratos, etc.).

- La falta de desagregación de datos por sexo en gran parte de los datos publicados en Open Data por las administraciones públicas.
- Existe una brecha digital² en el uso de tecnología de datos en el ámbito de igualdad (sociedad civil y administraciones públicas), se trata de la tercera brecha digital de género de usos avanzados.

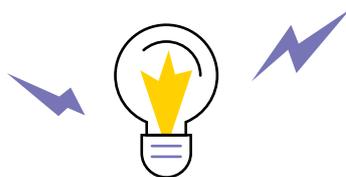
Los datos abiertos son fundamentales para obtener información pública de calidad sobre la situación de las mujeres. Estos datos son esenciales para evaluar la realidad y alimentar herramientas de inteligencia artificial. La falta de datos que evidencien la discriminación y las diferencias resulta de la omisión de estos aspectos en la creación de dichas herramientas. Un ejemplo ilustrativo es el proyecto "The BIAS RemAlns"³, donde se solicitó a la inteligencia artificial prever, mediante comandos escritos, cómo

1 Sierra Daza, A. (2021). *Ciudadanía e identidad digital en mujeres rurales: Procesos de autoinclusión en las tecnologías de la información y comunicación* [Tesis Doctoral]. Universidad de Extremadura. Disponible en: https://dehesa.unex.es/bitstream/10662/12657/6/TDUEX_2021_Sierra%20Daza.pdf

2 Comisión Europea (2022). *Brecha digital de género*. Disponible en: https://www.ontsi.es/sites/ontsi/files/2022-10/brechadigitalgenero_2022.pdf

3 Amén McCann (Marzo, 2023). *The Bias RemAlns* [Campaña profesional]. European Union. Ads of the World. Disponible en: <https://www.adsoftheworld.com/campaigns/the-bias-remains>

"Cuando los hombres utilizan estas herramientas para ejercer la violencia contra las mujeres, se combina el poder que les atribuye una sociedad machista y un mayor acceso a la tecnología"



sería la primera persona en aterrizar en Marte, las figuras más relevantes en la política y el deportista más destacado para 2050. Todas las representaciones futuristas generadas mostraban exclusivamente a hombres.

Existen otros retos de las tecnologías de los datos en el ámbito de la inteligencia artificial:

- Evitar la reproducción de sesgos sexistas en los algoritmos, en las fuentes de datos y en las decisiones basadas en ellos.
- Disminuir la segregación horizontal con la presencia de mujeres desarrolladoras en equipos de proyectos de inteligencia artificial.

Este artículo busca abordar tanto las situaciones y consecuencias derivadas de la creación de tecnologías de inteligencia artificial sin incorporar la perspectiva de los derechos de las mujeres, como el impacto del mal uso de estas tecnologías en una sociedad en la que no existe igualdad de oportunidades ni acceso a los recursos entre hombres y mujeres.

LA VIOLENCIA MACHISTA EN EL ÁMBITO DIGITAL PRESENTE EN CUALQUIER ANÁLISIS DE IMPACTO SOCIAL DE UNA TECNOLOGÍA

Incluir la consideración de la violencia contra las mujeres en el análisis de la inteligencia artificial es crucial, especialmente en una sociedad donde la comunicación digital se ha generalizado. Las diversas formas de violencia contra las mujeres en el ámbito digital incluyen⁴:

- Acoso en línea: intimidación sistemática a través de medios digitales.
- Difusión de imágenes íntimas sin consentimiento: divulgación no autorizada de imágenes o vídeos íntimos, con fines de avergonzar o chantajear a la víctima.
- Control y vigilancia: seguimiento no autorizado de la actividad en línea, incluyendo el acceso a cuentas de la víctima.
- Suplantación de identidad: uso fraudulento de la identidad de una persona para desprestigiarla o perjudicarla.

- Ciberacoso sexual: envío de mensajes de contenido sexual no consentidos.
- Violencia simbólica: uso de estereotipos de género, lenguaje sexista o comentarios degradantes.

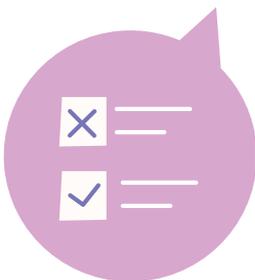
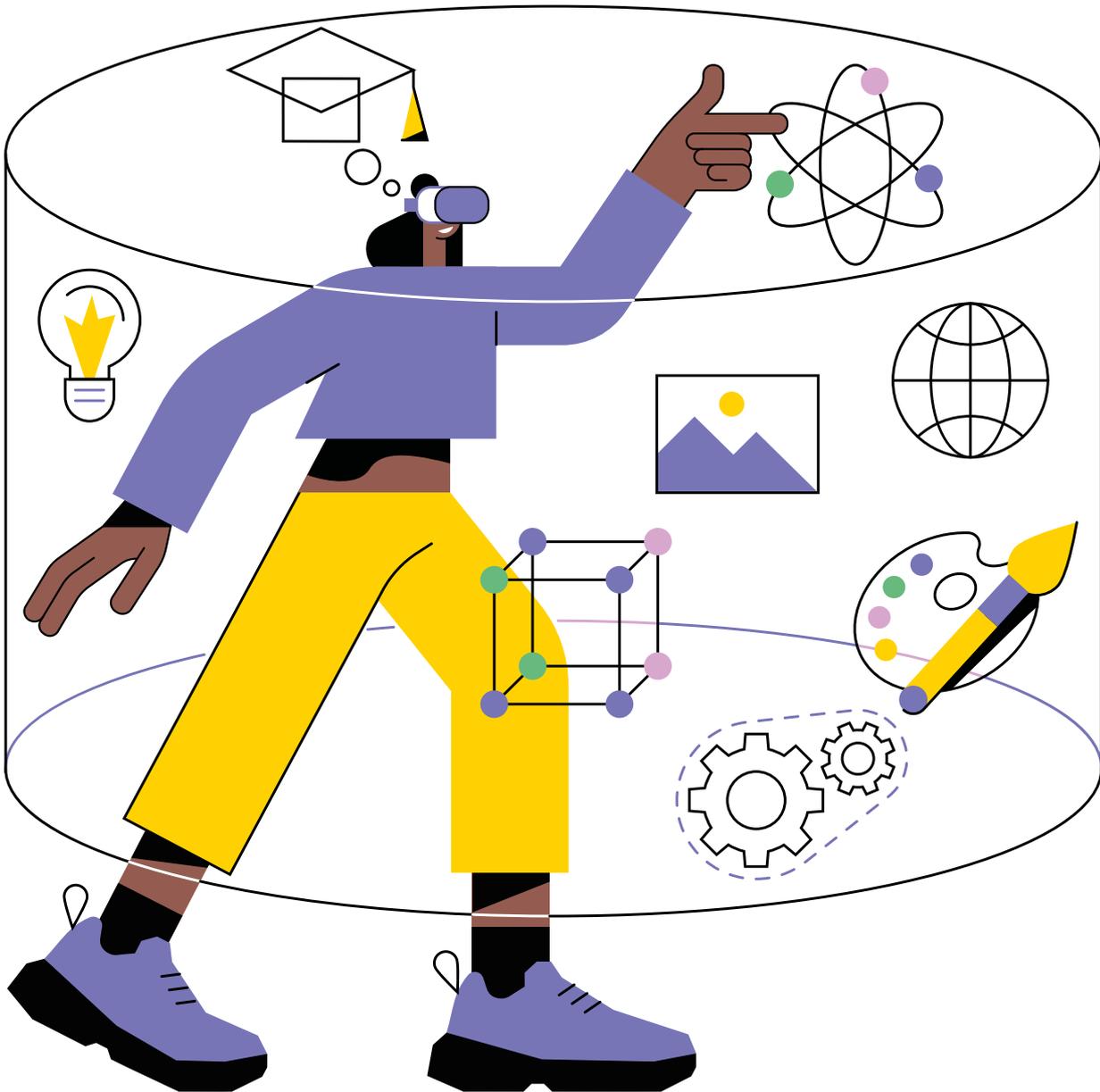
Se destacan varios fenómenos relevantes en este ámbito: en primer lugar, las mujeres líderes experimentan ataques sistemáticos en sus redes sociales, llegando al 58%, según un estudio del Calala Fondo de Mujeres (2022)⁵. En segundo lugar, existe la creación de espacios donde grupos de hombres se organizan para llevar a cabo ataques masivos a mujeres en redes, fenómeno conocido como la "manosfera digital"⁶.

La difusión digital intensifica el impacto adverso de la violencia contra las mujeres. La acción de un hombre que ejerce violencia al grabar imágenes con contenido sexual explícito y difundirlas tiene un impacto significativamente mayor que si solo describiera la situación.

⁴ Calala Fondo de Mujeres (2022). *Diagnòstic sobre les violències de gènere contra activistes feministes en l'àmbit digital*. https://calala.org/wp-content/uploads/2022/06/Calala-ESP_DEF.pdf

⁵ *Idem*.

⁶ *Idem*.



“El feminismo de datos es crucial en la era de la transformación digital para garantizar que la inteligencia artificial (IA) avance en igualdad y no perpetúe brechas entre hombres y mujeres, especialmente en la toma de decisiones automatizadas y el desarrollo de tecnologías”



“Al analizar el impacto de la IA o cualquier tecnología digital en los derechos de las mujeres o cualquier sector de población, hay que tener presente que el punto de partida no es el de la igualdad de recursos entre hombres y mujeres”

ANÁLISIS FEMINISTA DE LA CONSTRUCCIÓN DE HERRAMIENTAS BASADAS EN INTELIGENCIA ARTIFICIAL

El feminismo de datos en el campo de la inteligencia artificial (IA) reclama la importancia de abordar y corregir los sesgos sexistas en los algoritmos.⁷ Estos sesgos se manifiestan en diversas áreas y tienen consecuencias significativas en la vida real.

Un ejemplo destacado es el estudio *Gender Shades* de Buolamwini y Gebru (2018)⁸, que reveló desigualdades de género y raza en los sistemas de reconocimiento facial comerciales. Este estudio demostró cómo la IA, al ser entrenada con datos sesgados, puede reproducir y amplificar el sexismo a gran escala.

En el ámbito de la salud, los algoritmos entrenados principalmente con datos de pacientes masculinos generan diagnósticos erróneos para las mujeres. Un estudio⁹ concluyó que los algoritmos utilizados para diagnosticar enfermedades cardíacas, al ser entrenados con mayoría de datos de pacientes masculinos, fallaban más en los diagnósticos de las mujeres.

En el sector empresarial, Amazon desarrolló un sistema de reclutamiento automatizado basado en IA que prefe-

ría currículums de hombres, porque se había entrenado con datos históricos dominados por empleados masculinos, subrayando la prevalencia de sesgos de sexo en los algoritmos.

En última instancia, se han identificado contenidos altamente discriminatorios y ofensivos hacia las mujeres en un conjunto de datos abiertos destinado al entrenamiento de algoritmos de reconocimiento facial, bajo la categoría "Ball-Buster, Ball-Braker", traducida al español como "tocapelotas" (dicho de una persona impertinente, fastidiosa, pesada).

El análisis de los casos presentados subraya la urgente necesidad de incorporar una perspectiva feminista en la creación y funcionamiento de la inteligencia artificial y los algoritmos. Estos ejemplos evidencian cómo la falta de esta perspectiva conduce a la reproducción de sesgos de género existentes en la IA. Estos sesgos afectan a diversas aplicaciones, desde el reconocimiento facial hasta los sistemas de reclutamiento y diagnósticos médicos, perjudicando la equidad y eficacia de estas tecnologías. La inclusión de un enfoque feminista en el desarrollo de la IA no solo es esencial para garantizar la justicia y la igualdad, sino también para mejorar la precisión y utilidad

de los algoritmos en una sociedad diversa.

LOS USOS DE HERRAMIENTAS DIGITALES Y SU IMPACTO EN LOS DERECHOS DE LAS MUJERES

¿Cómo el mal uso de las herramientas de IA impacta en los derechos de las mujeres y la violencia contra ellas?

Un claro ejemplo de mal uso de las tecnologías es el "caso de Iveco", donde una trabajadora se quitó la vida por la violencia machista. En este caso, un compañero difundió videos de relaciones sexuales entre ellos. Aquí, las capacidades avanzadas de las herramientas digitales para la creación y difusión de contenidos son utilizadas por hombres para perpetrar violencia contra las mujeres.

IMPACTO EN ADOLESCENTES: IRRUPCIÓN DE TECNOLOGÍAS E INTELIGENCIA ARTIFICIAL

Para abordar de manera más focalizada la interacción entre las tecnologías de datos y la violencia contra las mujeres, resulta clave apuntar una serie de datos sobre fenómenos entre las personas adolescentes y jóvenes. Existe un aumento de pensamientos suicidas entre adolescentes, un 7%

7 Ortiz de Zárate Alcarazo, L., y Guevara Gómez, A. (2021). *Inteligencia artificial e igualdad de género*. Fundación Alternativas. Disponible en: https://www.igualdad-en-la-empresa.es/recursos/estudios/Monografia/docs/Estudio_Inteligencia_artificial_e_igualdad_de_genero_Fundacion_Alternativas.pdf

8 Buolamwini, J. y Gebru, T. (2018, January). "Gender shades: Intersectional accuracy disparities in commercial gender classification". En: *Conference on fairness, accountability and transparency*, pp. 77-91. PMLR.

9 Moles García, S., Morales Rodríguez, A. y Quirantes Morillas, M. (2022). "La inteligencia artificial en la detección del suicidio". *SANUM: Revista Científico-Sanitaria*, 6 (2), pp. 46-63. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=8866924>

de niñas frente a un 5% de niños. Según datos del Teléfono de la Esperanza¹⁰ existe una correlación entre la salud mental y su interacción con la tecnología de las y los jóvenes. Además, es la principal causa de muerte entre los 15 y los 29 años, según datos de la Fundación Salud Mental España¹¹. Sobre esta práctica de las autolesiones y el uso compulsivo de Internet en las chicas, la prevalencia llega al 30% entre los 12 y 18 años¹². Por tanto, el impacto en la salud mental del uso de las tecnologías afecta más a las mujeres jóvenes.

Estos datos nos empujan a tener un enfoque crítico sobre cómo la inteligencia artificial, a menudo diseñada sin una perspectiva de igualdad, puede perpetuar estereotipos y exacerbar la violencia contra las mujeres y las adolescentes. Es imperativo desarrollar estrategias y políticas que no solo aborden el uso seguro de la tecnología, sino también su diseño y la influencia e impacto en las relaciones sociales y la autopercepción de las personas y las chicas jóvenes.

Es fundamental considerar diversos aspectos relacionados con el impacto de la inteligencia artificial en adolescentes, especialmente en lo que respecta a la igualdad de género y la prevención de violencias machistas:

- Internet ha evolucionado, ya que no solo proporciona información o contenidos, sino que también proporciona herramientas con gran potencial de acciones para un bien o para un mal uso.

- Los contenidos que nos aparecen en las redes sociales¹³ están basados en un algoritmo de inteligencia artificial, que no solo prioriza ciertos contenidos, sino que está pensado para que no salgas de esas aplicaciones. Tienen un efecto adictivo¹⁴.
- La socialización y la relación entre jóvenes se produce a través de las redes sociales y de forma continuada y esto hace que se magnifiquen los conflictos.

El consumo de pornografía a edades infantiles –da comienzo a los once años– lleva consigo una aproximación al sexo violento y a la degradación de las mujeres.

Poniendo el foco en el uso de las herramientas de IA, en el caso de Almendralejo, observamos cómo chicos jóvenes utilizaron una aplicación de inteligencia artificial de combinación de imágenes para crear y difundir una imagen desnuda de compañeras de instituto.

En este caso, la construcción de la tecnología es correcta, pero es el contexto social de machismo lo que ha provocado un compartimiento de jóvenes que ejercen violencia contra estas chicas.

CONSIDERACIONES FINALES

Apunto una serie de consideraciones para aplicar, con el fin de conseguir tecnologías digitales que respeten los derechos de las mujeres y no perpetúen sesgos y consideraciones para

desarrollar estrategias de intervención en los ámbitos para promover la igualdad y el abordaje de las violencias machistas:

- La violencia digital contra las mujeres se manifiesta de forma simultánea a otras violencias, como la sexual, y la difusión digital amplifica el impacto negativo contra las mujeres.
- Debemos analizar cómo la brecha digital de género influye en cualquier fenómeno de la sociedad digital, en concreto en la violencia sexual.
- Atención al impacto del uso de internet sin límites en la salud mental de adolescentes, especialmente las niñas.
- La agenda de los derechos de las mujeres debe priorizar el feminismo de datos que propone velar por disponer de fuentes de datos abiertas desagregadas por sexo y con perspectiva de derechos de las mujeres, así como garantizar que los algoritmos de inteligencia artificial eviten la reproducción de sesgos sexistas.
- Las mujeres feministas debemos apropiarnos de las herramientas de IA disponibles, usarlas y entrenarlas para generar contenidos no sexistas y con un enfoque feminista.
- Es imprescindible capacitar a las mujeres y a las organizaciones que trabajan para la igualdad real en el uso de la inteligencia artificial y de su potencia para abordar los retos pendientes de derechos de las mujeres.

10 Teléfono de la Esperanza (2021). *El Teléfono de la Esperanza sonó más de 183.000 veces en 2021*. Disponible en: <https://telefonodelaesperanza.org/noticias/el-telefono-de-la-esperanza-sono-mas-de-183000-veces-en-2021>

11 Federación de Salud Mental España (2022). *Observatorio del suicidio 2022*. Disponible en: <https://www.fsme.es/observatorio-del-suicidio-2022/>

12 Save the Children España (8 de febrero de 2022). *Suicidios adolescentes en España: Factores de riesgo y datos*. Disponible en: <https://www.savethechildren.es/actualidad/suicidios-adolescentes-espana-factores-riesgo-datos>

13 IAB Spain (2022). *Estudio de redes sociales 2022* [Informe]. Disponible en: <https://iabspain.es/estudio/estudio-de-redes-sociales-2022/>

14 Dolz, M., Camprodon, E. y Tarbal, A. (2021). *Mirada a la salud mental de los adolescentes. Observatorio de la Infancia*. Disponible en: https://www.observatoriodelainfancia.es/ficherosoia/documentos/7455_d_Mirada-SaludMental-Adolescentes.pdf