

Inteligencia Artificial y Estereotipos de Género: Desafíos y Oportunidades

**ENG Artificial Intelligence and Gender Stereotypes:
Challenges and Opportunities**

Dr. Víctor Cerdán

Universidad Complutense de Madrid, España 

Dr. Daniel Villa Gracia

Universidad Complutense de Madrid, España 

Dra. Sonia Carcelén

Universidad Complutense de Madrid, España 

Dra. Fernanda Nascimento

Pontifícia Universidade Católica do Rio Grande do Sul, Brasil 

<https://dx.doi.org/10.5209/infe.106408>

La acelerada expansión de la inteligencia artificial (IA) en la última década ha introducido transformaciones profundas en la producción de conocimiento, la creación cultural y la comunicación mediática. Herramientas generativas como ChatGPT, Midjourney, DALL-E, Stable Diffusion o Synthesia –cuyo uso cotidiano se ha normalizado en ámbitos educativos, creativos, institucionales y empresariales– han inaugurado una etapa en la que los límites entre lo humano y lo automatizado se desdibujan con rapidez. El potencial democratizador de estas tecnologías convive con nuevas formas de desigualdad, riesgos sociopolíticos y escenarios inéditos de desinformación que afectan de manera especialmente intensa a las mujeres.

La expansión de la inteligencia artificial no puede entenderse únicamente como un avance técnico, sino como un proceso cultural profundamente atravesado por relaciones de poder. Cada modelo entrenado, cada base de datos utilizada y cada decisión algorítmica implican una forma de representar el mundo, de seleccionar qué cuenta como conocimiento válido y qué experiencias quedan fuera del marco computacional. Desde una lectura feminista, esta selección no es neutral: responde a lógicas de visibilidad, exclusión y jerarquización que reproducen las desigualdades de género y raza presentes en la sociedad.

La IA no sólo redefine la producción cultural y los procesos laborales: también reconfigura discursos sociales, percepciones colectivas, subjetividades políticas y narrativas sobre el género. Como han mostrado investigaciones recientes, la IA amplifica sesgos históricos incorporados en los datos con los que se entrena, reproduciendo estereotipos que se infiltran en las imágenes generadas, los textos producidos de manera autónoma, los sistemas de recomendación, los algoritmos de reconocimiento facial y las arquitecturas de desinformación digital que moldean la opinión pública.

El problema no radica únicamente en los sesgos, sino en los marcos epistémicos que definen qué se considera una “solución eficiente” o un “modelo exitoso”. Las métricas de rendimiento y las nociones de objetividad técnica suelen invisibilizar las consecuencias sociales de los algoritmos, mientras priorizan la optimización sobre la ética. Un enfoque feminista propone modificar estas lógicas de evaluación, incorporando la responsabilidad, la diversidad y la rendición de cuentas como dimensiones centrales del desarrollo tecnológico.

El presente prólogo aspira a situar críticamente estos fenómenos y presentar los desafíos centrales a los que se enfrentan los estudios feministas en la sociedad algorítmica contemporánea, siguiendo el enfoque editorial de otras presentaciones recientes en *Investigaciones Feministas*. Esta introducción amplía así las claves conceptuales, éticas y políticas necesarias para comprender el artículo que se abre paso en las páginas siguientes, en diálogo con los debates globales que articulan el campo.

1. La IA como agente cultural y político en un ecosistema saturado de datos

Las tecnologías de IA generativa han dejado de ser herramientas auxiliares para convertirse en agentes culturales con capacidad de producir significado, intervenir en la esfera pública y distorsionar la percepción colectiva de la realidad. Las imágenes que mostraban la supuesta detención del expresidente Donald Trump o la manipulación del Papa Francisco sosteniendo una bandera LGTBI son ejemplos paradigmáticos de una era marcada por la erosión de la factualidad y la crisis de confianza institucional.

Además, la inteligencia artificial interviene en la construcción de imaginarios colectivos sobre el cuerpo, la identidad y la diferencia. Las imágenes que produce —aparentemente neutras o estéticas— reafirman con frecuencia ideales normativos de belleza, juventud y feminidad, reforzando patrones coloniales y androcéntristas. Frente a ello, los feminismos digitales plantean la necesidad de disputar el espacio simbólico, promover nuevas representaciones y crear contraimágenes que desafíen los cánones hegemónicos.

Este ecosistema hiperacelerado obliga a repensar los marcos epistemológicos desde los que analizamos la veracidad y la construcción del conocimiento. Tal y como ocurre con la posverdad y las dinámicas de desinformación descritas en investigaciones recientes, la IA se inscribe en una larga genealogía de discursos que han cuestionado la autoridad epistémica de determinados grupos sociales, entre ellos las mujeres, históricamente situadas en posiciones de deslegitimación testimonial.

La pregunta que enfrentamos hoy no es únicamente cómo verificar lo que vemos, sino quién tiene derecho a ser creído, un problema que el feminismo lleva décadas denunciando a través del concepto de injusticia epistémica.

2. Perspectiva de género y desigualdad estructural en el desarrollo de la IA

La perspectiva de género revela que la inteligencia artificial se desarrolla dentro de un ecosistema profundamente desigual. La presencia femenina en los equipos técnicos y en los espacios de liderazgo sigue siendo minoritaria, lo que condiciona tanto el diseño de los sistemas como los datos empleados para entrenarlos. Esta infrarepresentación no es un detalle anecdótico, sino la continuación de una tradición histórica en STEM que ha excluido voces diversas y ha permitido que los algoritmos hereden sesgos preexistentes.

Cuando los modelos aprenden de bases de datos construidas en sociedades marcadas por estereotipos de género, reproducen esas mismas desigualdades: asignan profesiones de prestigio a hombres, relegan a mujeres a roles de cuidado o sexualizados y presentan mayores tasas de error con mujeres racializadas. Estos resultados no son fallos aislados, sino manifestaciones algorítmicas de estructuras sociales. Por ello, incorporar mujeres y miradas críticas en todas las fases de la IA —del diseño a la auditoría— es imprescindible para evitar que los sistemas profundicen brechas existentes y para avanzar hacia tecnologías más equitativas y representativas. Una reivindicación histórica de los activismos feministas es la de introducir la perspectiva de género y raza en lo digital, proponiendo proyectos de intervención que puedan influir en el desarrollo de iniciativas y transformar sus características ya en las etapas de producción.

La brecha de género en IA no se limita a la representación numérica, sino que se manifiesta en la ausencia de perspectivas críticas durante la definición de problemas técnicos. Los equipos predominantemente masculinos tienden a priorizar aplicaciones de alto rendimiento —como la optimización militar o financiera— sobre soluciones orientadas al cuidado colectivo, como la detección temprana de violencias de género o el apoyo a economías informales lideradas por mujeres. Esta priorización implícita perpetúa un ciclo donde las necesidades específicas de género quedan relegadas a actualizaciones secundarias.

Además, el feminismo tecnológico cuestiona la noción de “escalabilidad universal”, revelando cómo los modelos optimizados para poblaciones privilegiadas fallan sistemáticamente con usuarias de contextos diversos. Desde acentos regionales no reconocidos en asistentes de voz hasta interfaces que ignoran patrones culturales de interacción femenina, estas limitaciones convierten la IA en una tecnología extractiva que consume datos globales, pero devuelve servicios sesgados. Corregir esta dinámica exige codiseño participativo con mujeres de distintos estratos socioeconómicos y geográficos.

Otro eje de análisis ineludible se relaciona con la gobernanza algorítmica y la transparencia de los sistemas de decisión automatizados. El secretismo empresarial en torno a los datasets y los modelos de entrenamiento impide la supervisión ciudadana y dificulta el desarrollo de políticas públicas con enfoque de derechos. La opacidad no solo protege la propiedad intelectual, sino que perpetúa estructuras de desigualdad. La ética feminista demanda abrir estos procesos a la participación social, democratizando el conocimiento técnico y estableciendo mecanismos claros de auditoría y cocreación.

3. Imágenes, textos y estereotipos generados por IA: un nuevo campo de disputa simbólica

La perspectiva de género revela que la inteligencia artificial se desarrolla dentro de un ecosistema profundamente desigual. La presencia femenina en los equipos técnicos y en los espacios de liderazgo sigue siendo minoritaria, lo que condiciona tanto el diseño de los sistemas como los datos empleados para entrenarlos. Esta infrarepresentación no es un detalle anecdótico, sino la continuación de una tradición histórica en STEM que ha excluido voces diversas y ha permitido que los algoritmos hereden sesgos preexistentes.

Cuando los modelos aprenden de bases de datos construidas en sociedades marcadas por estereotipos de género, reproducen esas mismas desigualdades: asignan profesiones de prestigio a hombres, relegan a mujeres a roles de cuidado o sexualizados y presentan mayores tasas de error con mujeres racializadas. Estos resultados no son fallos aislados, sino manifestaciones algorítmicas de estructuras sociales. El panorama incluye además una creciente preocupación por los modelos de hipervigilancia que promueven un control aún mayor sobre los grupos históricamente marginados.

Los textos generados por IA reproducen igualmente patrones misóginos arraigados, asignando agencia narrativa a personajes masculinos mientras confina a las mujeres a roles pasivos o estereotipados. Análisis de outputs masivos revelan que las protagonistas femeninas aparecen desproporcionadamente como objetos románticos o figuras maternales, limitando la diversidad de arquetipos disponibles para creadoras y audiencias jóvenes. Esta simbología automatizada moldea imaginarios generacionales, normalizando desigualdades bajo apariencia de neutralidad creativa.

La hipervigilancia algorítmica adquiere dimensiones particularmente invasivas en el control de cuerpos femeninos, desde apps de fitness que imponen cánones normativos hasta sistemas de reconocimiento que penalizan expresiones no convencionales de género. Mujeres con cabello rizado, tatuajes visibles o vestimentas culturales enfrentan tasas de error que las convierten en “fantasmas digitales”, invisibilizadas por arquitecturas diseñadas desde estándares eurocéntricos. Desafiar esta vigilancia requiere datasets contrahegemónicos y algoritmos que celebren la diversidad. Por ello, incorporar mujeres y miradas críticas en todas las fases de la IA, del diseño a la auditoría, es imprescindible para evitar que los sistemas profundicen brechas existentes y para avanzar hacia tecnologías más equitativas y representativas.

4. IA, desinformación y misoginia digital: riesgos para la democracia

La relación entre inteligencia artificial, desinformación y misoginia digital configura uno de los riesgos más inquietantes del entorno comunicativo actual. Los sistemas automatizados de difusión de contenidos, combinados con algoritmos que priorizan lo emocional y lo viral por encima de la veracidad, facilitan la expansión de narrativas antifeministas y ataques coordinados contra mujeres con presencia pública. En este contexto, la IA no actúa simplemente como un vehículo neutro, sino como un multiplicador de discursos que cuestionan la legitimidad del conocimiento experto, desestabilizan instituciones y refuerzan prejuicios ya arraigados en la cultura digital.

Por último, la integración plena de la inteligencia artificial en la vida cotidiana requiere una pedagogía crítica capaz de conectar ciudadanía, tecnología y justicia social. La alfabetización algorítmica –entendida como la capacidad de leer, interpretar y cuestionar sistemas automáticos– es un componente indispensable de cualquier estrategia democrática. Formar a nuevas generaciones en una cultura digital con perspectiva de género no significa únicamente enseñar a usar herramientas, sino a comprender sus implicaciones políticas, ecológicas y afectivas.

Las campañas de odio, la circulación de rumores fabricados con apariencia de autenticidad y la difusión masiva de mensajes que desacreditan a activistas, periodistas o políticas se ven intensificadas por modelos capaces de generar textos e imágenes indistinguibles de los reales. Esta combinación de automatización y sesgo cultural convierte a la IA en un actor que puede erosionar derechos adquiridos y condicionar la opinión pública. Comprender esta estructura es fundamental para desarrollar estrategias de protección, regulación y alfabetización digital capaces de frenar la propagación de discursos misóginos en el espacio público.

Se añade además la necesidad de investigar el papel de los agentes públicos y de las empresas mediáticas vinculadas a grupos de extrema derecha que no solo promueven campañas fundamentadas en la misoginia, el racismo y la homofobia, sino que también financian el desarrollo de productos cuyo objetivo es difundir la violencia contra poblaciones vulnerables.

5. Oportunidades y líneas de acción: hacia una IA feminista, inclusiva y ética

A pesar de los riesgos que plantea, la inteligencia artificial ofrece también oportunidades para avanzar hacia sistemas más justos si se desarrolla desde una perspectiva ética y de género. La incorporación de mujeres en todas las fases del ciclo de vida de la IA –desde la selección de datos hasta la evaluación y la regulación– es esencial para contrarrestar la tendencia de los algoritmos a reproducir desigualdades estructurales. Esta participación debe ir acompañada de auditorías transparentes, datasets diversos y mecanismos de supervisión que permitan identificar y corregir sesgos con antelación.

El horizonte de una inteligencia artificial feminista no apunta solo a corregir errores del pasado, sino a imaginar futuros tecnoculturales más equitativos, donde la creatividad colectiva y la justicia epistémica se sitúen en el centro. La tecnología, vista desde esta óptica, se convierte en un espacio de corresponsabilidad y de posibilidad, donde lo humano y lo digital coevolucionan bajo principios de cuidado, diversidad y emancipación compartida.

Con la apropiación de herramientas digitales y el hackeo de metodologías empleadas habitualmente con el objetivo de reiterar las desigualdades, los grupos feministas han promovido acciones de resistencia digital. Y disputan los espacios en línea, de manera desigual, en busca de crear significados disruptivos.

El horizonte de una inteligencia artificial feminista no apunta solo a corregir errores del pasado, sino a imaginar futuros tecnoculturales más equitativos, donde la creatividad colectiva y la justicia epistémica se sitúen en el centro. La tecnología, vista desde esta óptica, se convierte en un espacio de corresponsabilidad y de posibilidad, donde lo humano y lo digital coevolucionan bajo principios de cuidado, diversidad y emancipación compartida.

Este planteamiento transforma radicalmente la narrativa dominante sobre la IA, que suele presentarla como un destino inevitable regido por lógicas de eficiencia y escala. En cambio, una visión feminista reivindica la tecnología como terreno de disputa política, donde las decisiones sobre qué priorizar –rentabilidad o equidad, velocidad o sostenibilidad– definen quiénes se benefician y quiénes quedan excluidos de los avances. Así, la IA deja de ser un monolito imparable para convertirse en un proyecto colectivo sujeto a negociación social y ética.

La coevolución humano-digital que proponemos exige, además, romper con la dicotomía entre creadores y usuarios, posicionando a las comunidades marginadas como coautoras activas de los sistemas tecnológicos. Desde hackatones feministas hasta protocolos de datos comunitarios, surgen prácticas que subvierten la centralidad corporativa y democratizan el acceso al conocimiento algorítmico. Este giro no solo corrige desigualdades, sino que enriquece la innovación al incorporar perspectivas múltiples que las métricas tradicionales ignoran.

En última instancia, esta IA feminista se ancla en una ética del cuidado que permea desde el diseño hasta el despliegue, reconociendo que la tecnología no es neutral sino relacional. Cada algoritmo porta huellas de sus creadores y contextos, por lo que orientarlo hacia la emancipación compartida requiere vigilancia constante, alianzas interdisciplinarias y un compromiso inquebrantable con la justicia social. Solo así, la inteligencia artificial cumplirá su promesa de ampliar horizontes humanos en lugar de estrecharlos.

La alfabetización digital crítica ocupa un lugar estratégico en este proceso, ya que capacita a la ciudadanía para reconocer estereotipos, detectar contenidos manipulados y comprender la lógica bajo la cual operan los modelos generativos. Asimismo, la interseccionalidad se convierte en un principio indispensable para garantizar que los sistemas no perpetúen discriminaciones basadas en raza, edad, discapacidad u orientación sexual. Solo mediante esta combinación de vigilancia, diversidad y educación será posible orientar la inteligencia artificial hacia un horizonte en el que actúe como herramienta emancipadora y no como mecanismo de reproducción de inequidades.

La relación entre inteligencia artificial y estereotipos de género se ha convertido en uno de los desafíos más urgentes de la sociedad digital. La IA no es simplemente un conjunto de herramientas innovadoras, sino una infraestructura cultural y política que influye en cómo interpretamos el mundo, cómo nos relacionamos y qué narrativas adquieren legitimidad. De su diseño y de las decisiones que se tomen en torno a su regulación dependerá en gran medida si estas tecnologías contribuyen a reducir desigualdades o, por el contrario, consolidan viejas jerarquías bajo formas nuevas.

Comprender la complejidad de este escenario exige un enfoque interdisciplinario que combine investigación feminista, ética tecnológica, teoría de la comunicación y políticas públicas. Este prólogo sitúa el artículo que sigue en un marco más amplio, destacando la importancia de abordar los sesgos algorítmicos, la violencia digital, los imaginarios producidos por modelos generativos y las tensiones que atraviesan la esfera pública en la era de la desinformación automatizada. La IA puede convertirse en una herramienta de equidad y justicia, pero solamente si se articula desde una vigilancia crítica constante y desde un compromiso real con los derechos humanos y la igualdad de género.

Referencias bibliográficas

- Cerdán Martínez, V. y Padilla Castillo G. (2019). Historia del "fake" audiovisual: "deepfake" y la mujer en un imaginario falsificado y perverso. *Historia y Comunicación Social*, 24(2), 505-520. <https://doi.org/10.5209/hics.66293>
- Cerdán Martínez V., García Guardia M. L. y Padilla Castillo G. (2020). Alfabetización moral digital para la detección de deepfakes y fakes audiovisuales. *CIC. Cuadernos de Información y Comunicación*, 25, 165-181. <https://doi.org/10.5209/ciyc.68762>
- Cerdán Martínez, V., Giménez Sarmiento, Á. y Padilla Castillo, G. (2022). El auge de Vox y el populismo en YouTube antes y durante la pandemia del Covid-19. *Revista de Comunicación de la SEECI*, 55, 17-35. <https://doi.org/10.15198/seeci.2022.55.e751>
- Cerdán-Martínez, V., López-Segura, P., Lucia-Mulas, M. J., Sanz, P. R., & Alonso, T. O. (2024). Male and Female Brain Activity During the Screening of a Violent Movie: An EEG Study. *Journal of Creative Communications*, 19(3), 259-275. <https://doi.org/10.1177/09732586241227592>
- Giménez Sarmiento, Álvaro, Cerdán Martínez, V., & Villa Gracia, A. D. (2025). Artificial Intelligence Applied to the Analysis of Hate Speech: Trump's and Biden During the Capitol Attack in Washington D. C. *Street Art & Urban Creativity*, 11(4), 137-149. <https://doi.org/10.62161/sauc.v11.5813>
- Kolotouchkina, O., Díaz-Bustamante Ventisca, M., Carcelén-García, S., & Díaz-Soloaga, P. (2025). Environmental activism and sustainable communication of B-Corp fashion brands on Instagram: best practices, emerging opportunities, and remaining challenges. *Humanit Soc Sci Commun*, 12, 1674. <https://doi.org/10.1057/s41599-025-05941-2>
- Pedreño-Santos, A., Carcelén García, S., & Viñarás Abad, M. (2025). Study of the boomerang effect in a digital community among young people: Synergies and differences between advertising communication and museum experience. *Street Art & Urban Creativity*, 11(2), 161-174. <https://doi.org/10.62161/sauc.v11.5723>
- Pedreño Santos, A., Viñarás Abad, M., & Carcelén García, S. (2025). The Influence of Opinion Leaders on the Decisions of Online Gamblers . *VISUAL REVIEW. International Visual Culture Review Revista Internacional De Cultura Visual*, 17(4), 253-263. <https://doi.org/10.62161/revvisual.v17.5876>