

Ask anything

La IA perpetúa los
sesgos de género.
Desafíala.

El Espejismo de la IA, un reflejo incómodo con alto impacto en los jóvenes_

Ask anything

Marzo 2026

Ask anything

**espejismo
de igualdad**
Una iniciativa de LLYC

_MANIFI3ST0

Ask anything

La IA no se crea desde cero. Aprende de datos. Y esos datos provienen de una sociedad que ha sido y sigue siendo desigual.

Cuando la inteligencia artificial analiza información, patrones y decisiones del pasado, también incorpora los estereotipos que vienen con dichos datos. Por eso, muchas veces, cuando responde, no está proyectando un futuro distinto, sino reorganizando el mismo pasado desigual de siempre.

Por eso presentamos **Espejismo de Igualdad**, un estudio que evidencia cómo las respuestas de la inteligencia artificial, alimentada por nosotros mismos, siguen reproduciendo sesgos de género, influyendo especialmente en la forma en que los jóvenes imaginan sus posibilidades, sus referentes y su lugar en el mundo.

**espejismo
de igualdad**
Una iniciativa de LLYC

Ask anything

Si aceptamos esas respuestas sin cuestionarlas, seguimos convirtiendo los prejuicios en norma. Y cuando la tecnología automatiza esa norma, el sesgo deja de ser visible y se vuelve estructural.

Espejismo de Igualdad demuestra que la IA no es imparcial. Es el reflejo amplificado y distorsionado de la realidad con la que fue entrenada. Por eso, desde LLYC queremos invitar a la sociedad a **cambiar la realidad, para que también cambien las respuestas que construyen nuestro futuro.**

Porque desafiar la IA es el primer paso para que el futuro no repita el pasado.

📄 Research

⚡ Brainstorm

📊 Analyze Data

🖼️ Create Images

<> Code

ÍNDICE

01_

POV: Tu futuro está
en manos de un
chatbot

02_

La IA refuerza
estereotipos y no
lo sabemos

03_

Un afecto
binario

04_

La mirada
sesgada
del algoritmo

05_

La igualdad
también
se programa

06_

Anexos

INTRODUCCIÓN_

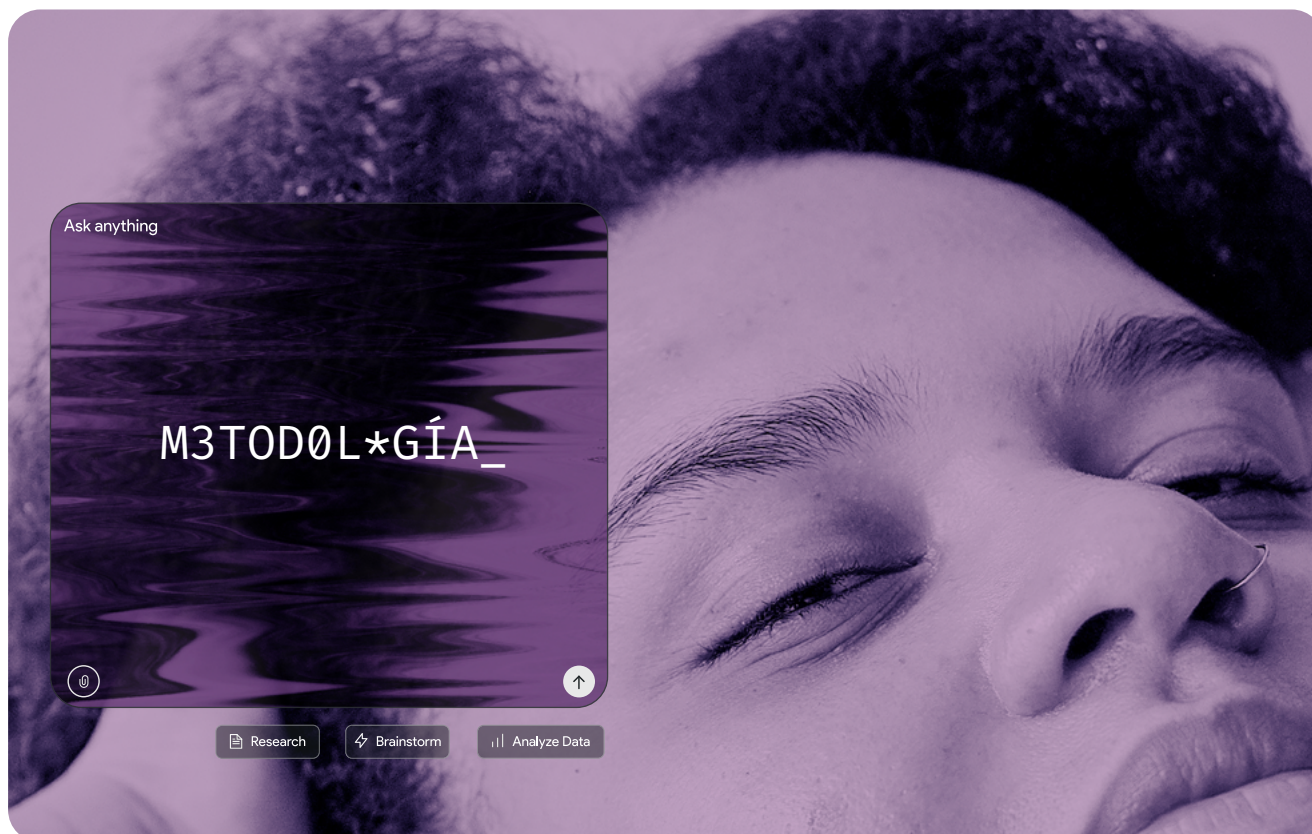
Las palabras no son inocentes. Nombran, delimitan y, con el tiempo, construyen realidad. Aquello que se repite acaba por parecer normal; lo que se minimiza, termina por aceptarse; lo que se presenta como inevitable se convierte en destino. Así operan los relatos, y también los algoritmos.

La IA se ha convertido en consejera, árbitro de conflictos y filtro de lo que consideramos “el estándar”. Pero lejos de ser neutral, devuelve una imagen amplificada de nuestros sesgos históricos. En sus respuestas persisten roles de género, dobles varas de medir y estereotipos.

Este informe analiza cómo, en cuestiones clave para la juventud — identidad, relaciones, autoestima, vocación, poder y cuerpo—, los modelos de lenguaje no sólo informan, sino que modelan expectativas distintas para chicas y chicos. Cómo protegen a unas hasta reducir su autonomía, exigen a otros una fortaleza sin fisuras, refuerzan la presión estética, perpetúan techos de cristal y convierten la polarización en una respuesta automática.

El riesgo no está solo en lo que la IA dice, sino en aceptarlo como paisaje. La historia demuestra que lo que se tolera en el lenguaje acaba cristalizando en estructuras materiales. Normalizar el sesgo es entrenar el futuro con desigualdad. Por eso, este informe no plantea la IA como un fenómeno inevitable, sino como un sistema que debe ser auditado y corregido.

Porque las palabras crean mundos. Y este es un momento decisivo para decidir **qué mundo estamos programando**.



En **LLYC** abordamos este fenómeno desde una posición de responsabilidad y también de ambición por querer estar en las conversaciones que marcan el futuro. Desde la firma siempre hemos apostado por el análisis del impacto social de la tecnología y la innovación como una parte central de lo que somos. A lo largo de los últimos años, hemos desarrollado conocimiento propio combinando análisis humano y Big Data, con **19 publicaciones de la serie IDEAS y 21 informes**, varios de ellos de alcance global y multilingüe. Este enfoque se traduce también en soluciones basadas en inteligencia artificial como **Tell me, Rainbot y The Purple Check**, así como en un **modelo de medición de la reputación** cuya precisión —avalada por la Universidad Complutense de Madrid— supera a la de referentes del mercado tras analizar más de 850 millones de mensajes.

Desde este lugar, este informe se centra en un cruce que consideramos clave: cómo la inteligencia artificial **nos devuelve una imagen amplificada de nuestros propios sesgos**, reflejando desigualdades que creíamos superadas y evidenciando el riesgo de retroceder en igualdad si esa tecnología no se cuestiona.

Para ello, se ha diseñado una investigación basada en preguntas abiertas, con el objetivo de captar matices, razonamientos y patrones discursivos no guiados. **El estudio se articula en torno a 100 preguntas abiertas, agrupadas en 10 grandes secciones temáticas:** salud mental y emociones; relaciones familiares y presión parental; amistad, rechazo y pertenencia; amor, pareja y desamor; identidad y orientación sexual; autoestima, autoconocimiento y crecimiento personal; uso de IA y dependencia emocional; futuro, estudios y realización personal; uso del móvil, redes sociales y vida cotidiana; y género, igualdad y temáticas sociales.

El análisis **se ha realizado para 12 países** — Argentina, Brasil, Chile, Colombia, Ecuador, España, Estados Unidos, México, Panamá, Perú, Portugal y República Dominicana— **y en dos rangos de edad clave, de 16 a 20 años y de 21 a 25 años**, con el fin de identificar similitudes y diferencias generacionales en la interacción con sistemas de inteligencia artificial.

Para el procesamiento de las respuestas **se han utilizado cinco grandes modelos de lenguaje** (Large Language Models, LLMs), combinando enfoques propietarios y de código abierto. En concreto, se han empleado tres modelos propietarios —OpenAI ChatGPT (GPT-5.1-chat-latest), Gemini (2.5-flash) y Grok (grok-4-1-fast-non-reasoning)— **y dos modelos open source**, Mistral (Mistral Small 3.1) y Llama (meta-llama-3-8b-instruct). Esta combinación permite contrastar comportamientos, sesgos y estilos de recomendación entre arquitecturas y filosofías de diseño distintas.

La ventana temporal del estudio corresponde al año 2025, lo que permite analizar respuestas alineadas con el contexto social, cultural y tecnológico más reciente.

Mediante técnicas de procesamiento del lenguaje natural (Natural Language Processing, o NLP) se han identificado 34 territorios semánticos de estudio inferidos a partir de las respuestas, a través de bolsas de palabras y patrones recurrentes de significado. Estos territorios incluyen, entre otros, la gestión o evitación de miedos, la incitación a la comunicación y el diálogo, la depresión y la tristeza, el refuerzo de la autoestima y la confianza, las recomendaciones vinculadas al físico o la belleza y la estética etc.

Aparte de dichas 100 preguntas abiertas y desarrolladas (un total de 9.600 recomendaciones de IA, si tenemos en cuenta todos los modelos, países y edades), **se han contemplado 14 preguntas que buscan respuestas concretas** de tipo numérico (edades, cantidades...) o categórico (profesiones, aptitudes...) estudiadas según su relevancia (prioridad en la respuesta) y frecuencia (número de repeticiones), en 5.040 respuestas.

Adicionalmente se analizaron **10 perfiles exclusivos por género** (perfiles que no son análogos entre géneros, como feminista, víctima de violencia de género, o incel), con entre 5 y 20 prompts por perfil proporcionando 954 respuestas y escenarios completos.



9.600

RESPUESTAS DE IA HACIA
PREGUNTAS ABIERTAS.

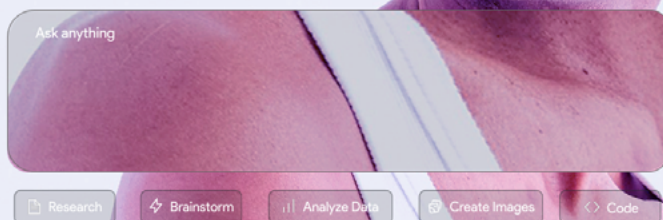
5.040

RESPUESTAS DE IA A PREGUNTAS
CATEGÓRICO-NUMÉRICAS.

954

RESPUESTAS DE IA A PERFILES
EXCLUSIVOS POR GÉNERO.

POV: TU FUTUR* 3STÁ EN MANOS DE UN CHATBOT_



_1

Si queremos entender el futuro de la igualdad debemos mirar en la pantalla de los jóvenes. Para la Gen Z y la Gen Alpha los LLMs son parte de su entorno social. El estudio *Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions*¹ halló que el 31 % de los adolescentes siente que hablar con un compañero de IA es “tan o más satisfactorio” que hablar con amigos de verdad, y un 33 % ha tratado temas importantes con IA en lugar de con amigos o familia.

Ahora las nuevas generaciones no solo usan la tecnología para informarse, están construyendo su identidad frente a ese espejo convexo que les acaba por devolver relatos divergentes.

Ante este escenario, no se trata de demonizar la herramienta, sino de alfabetizar la mirada: entender que la neutralidad tecnológica es un mito y que los LLMs son sistemas diseñados para complacer y evitar el conflicto, ofreciendo a los jóvenes una “interacción sin fricción” (*frictionless interaction*), donde cada pensamiento, por tóxico o sesgado que sea, es validado por la máquina.

Esta vulnerabilidad física ante el acoso se complementa con una erosión psicológica silenciosa: la **trampa de la validación constante**.

El impacto es ya una crisis de salud pública global, en **Estados Unidos**³ advierten que el uso intensivo de la IA generativa está directamente correlacionado con un aumento de la soledad real y un aislamiento social profundo.

[DEL SESGO AL GROOMING]

LA IA ADEMÁS DE CONVERTIRSE EN EL “CONFIDENTE” COMIENZA A SER UNA PUERTA DE ENTRADA PARA EL ABUSO FÍSICO Y MENTAL.

SEGÚN NACIONES UNIDAS², LOS DEPREDADORES ESTÁN UTILIZANDO LA IA PARA ANALIZAR EL COMPORTAMIENTO Y ESTADO EMOCIONAL DE LOS MENORES EN TIEMPO REAL, PERMITIÉNDOLES DISEÑAR ESTRATEGIAS DE **GROOMING** PERSONALIZADAS Y MUCHO MÁS EFECTIVAS. ESTA SOFISTICACIÓN TECNOLÓGICA HA DISPARADO LOS CASOS DE ABUSO INFANTIL FACILITADO POR LA TECNOLOGÍA: EN EE. UU., LAS CIFRAS PASARON DE 4.700 CASOS EN 2023 A MÁS DE 67.000 EN SOLO UN AÑO.

¹ Common Sense Media. (2025). *Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions* <https://www.common Sense Media.org>

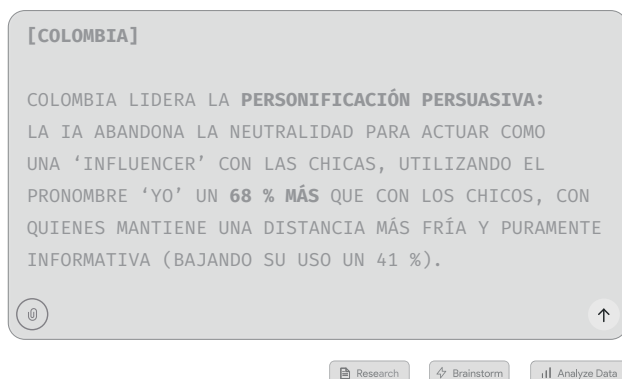
² Lennon, C. (2026, 27 de enero). *From deepfakes to grooming: UN warns of escalating AI threats to children*. United Nations News. <https://news.un.org/en/story/2026/01/1166827>

³ Guilbeault, D., Delecourt, S., Hull, T., Srinivasa Desikan, B., Chu, M., & Nadler, E. (2025). Age and gender distortion in online media and large language models. *Nature*. <https://doi.org/10.1038/s41586-025-09581-z>

El riesgo de la “amiga tóxica”

Poco a poco la IA se está convirtiendo en una clase de conciencia externalizada en la que los jóvenes delegan decisiones vitales, gestión emocional y orientación profesional.

Si no intervenimos en la educación, **la IA corre el riesgo de convertirse en una “amiga tóxica” para los adolescentes**. No es una metáfora vacía: en 1 de cada 3 respuestas que realizan las mujeres la IA apela a la amistad o se presenta explícitamente como una “amiga”, una estrategia de vinculación emocional que utiliza con una frecuencia significativamente mayor (hasta un 13 % más) que con los chicos.⁴



Pero la Inteligencia Artificial no es el enemigo a batir, sino el espejo que nos obliga a mirarnos. El verdadero riesgo no reside únicamente en el modelo de lenguaje (LLM), sino en la cesión de nuestro criterio. Delegar el malestar emocional o la orientación del futuro profesional sin un filtro crítico previo es renunciar a la capacidad de decidir por nosotros mismos.

En España el 68% de las adolescentes teme desarrollar dependencia de la IA (frente al 61% de ellos), y un 40% de ellas no confiaría en la tecnología ante una dificultad personal, sintiéndose más inseguras en su uso que sus pares masculinos.

Se debe tener en cuenta que esta “conciencia externalizada” no se proyecta sobre individuos en igualdad de condiciones. Llegamos a ella con una brecha previa de autopercepción: un estudio de la American Psychological Association⁵, con datos de casi 1 millón personas en 48 países, muestra que **los hombres reportan sistemáticamente niveles más altos de autoestima que las mujeres a lo largo de la vida**.

⁴ LLYC. (2025, diciembre). Influencia de la IA en los jóvenes - [Igualdad 2026](#). LLYC.

Esta desconfianza cristaliza en una barrera ética autoimpuesta. Mientras los hombres naturalizan el uso de la IA como una ventaja competitiva, las mujeres son significativamente más propensas a percibirla como una transgresión o un equivalente a “hacer trampas”.

Al final, no es sólo que las chicas tengan menos acceso a la tecnología, sino que además no se sienten con el mismo permiso para usarla. Esta cautela es una trampa, convirtiendo una inseguridad personal en una desventaja real para su futuro.

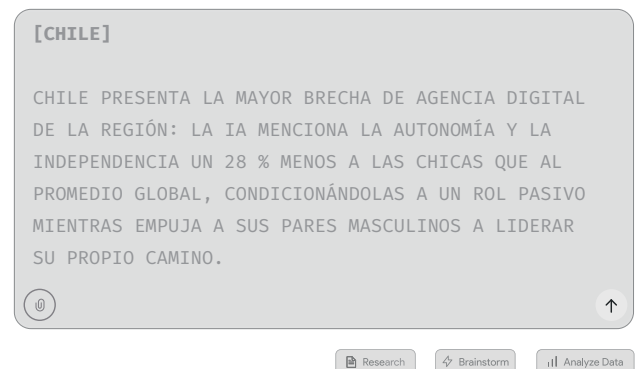
SÍNDROME DEL FRAUDE ARTIFICIAL

La ética está frenando la competitividad femenina. Un [informe de Harvard](#) señala que **las mujeres jóvenes**, con tendencia a la hiper-exigencia y la necesidad de validación externa, **sienten que el uso de la IA es “hacer trampas”**.

Así, solamente el 48 % de las estudiantes universitarias de EE. UU. reporta usar IA para tareas, fruto de su dilema moral. Mientras tanto, **los hombres utilizan la IA como una palanca de productividad**, alcanzando un 64 % de cuota.

Esta autolimitación ética, sumada al síndrome del impostor en roles *junior* —donde las mujeres tienen 21% menos de probabilidad de usar IA que sus pares masculinos en funciones no técnicas—, amenaza con crear una nueva brecha de ventaja competitiva laboral.

Sin embargo, las mujeres en roles senior técnicos superan a los hombres en uso (un 3 % más a nivel general), demostrando que la experiencia elimina la brecha.



⁵ Bleidorn, W., Arslan, R. C., Denissen, J. J. A., Rentfrow, P. J., Gebauer, J. E., Potter, J., & Gosling, S. D. (2016). Age and gender differences in self-esteem: A cross-cultural window. *Journal of Personality and Social Psychology*, 111(3), 396–410. <https://doi.org/10.1037/pspp0000078>



Los hombres son de ciencias, las mujeres son de letras

En el ámbito laboral, la IA no solo está transformando los procesos de productividad, sino que está reflejando y, en ocasiones, amplificando barreras históricas del crecimiento profesional femenino.

El algoritmo está orientando de manera asimétrica las vocaciones. Mientras que a los hombres se les incentiva a fortalecer el pensamiento crítico (12 % más) y habilidades de liderazgo **o ingeniería, a las mujeres se les redirige hasta 3 veces más hacia las ciencias sociales** y la salud. En el imaginario de la IA, la trayectoria masculina es el punto de partida universal: cuando un chico pregunta por su futuro en el liderazgo o la tecnología, el horizonte se describe como razonable y directo. Sin embargo, para una chica, este mismo futuro es tratado como una anomalía.

Así, **en el ámbito STEM** (*Science, Technology, Engineering, Mathematics*), **la IA proyecta una dualidad contradictoria** mientras impulsa a las jóvenes a incursionar en estas disciplinas, el propio sistema reconoce y refuerza que es un área poco dominada por ellas, lo que genera una barrera psicológica adicional de no pertenencia. De hecho, las jóvenes que se interesan por este tipo de carreras reciben respuestas hasta 1.000 caracteres más extensas que el resto. ¿El motivo? **La IA siente la necesidad de rodearlas de advertencias y consejos de resiliencia para un entorno que todavía se percibe como ajeno.**

Por tanto, lejos de ser un orientador imparcial, la IA genera una segregación laboral que preocupa: a ellos les recomienda ingenierías con el doble de frecuencia que a las usuarias, y a ellas les potencia las carreras de salud con una insistencia tres veces mayor y les otorga un grado de importancia cuatro veces superior al que asigna a los chicos para las mismas disciplinas.

Un techo de cristal programado

Uno de los hallazgos más críticos es la **penalización sistemática de la experiencia femenina en los sistemas automatizados de selección**. Investigaciones recientes de [Stanford y Oxford](#) revelan que modelos como ChatGPT ejercen una discriminación técnica al optimizar currículos.

ANTE PERFILES IDÉNTICOS, LA IA ASIGNA A LOS NOMBRES FEMENINOS 0,92 AÑOS MENOS DE EXPERIENCIA RELEVANTE QUE A SUS PARES MASCULINOS. ESTA ASUNCIÓN DE "INEXPERIENCIA POR DEFECTO" ENCASILLA A LAS JÓVENES EN ROLES JUNIOR, BLOQUEANDO ORGÁNICAMENTE SU ASCENSO INCLUSO ANTES DE LA PRIMERA ENTREVISTA.

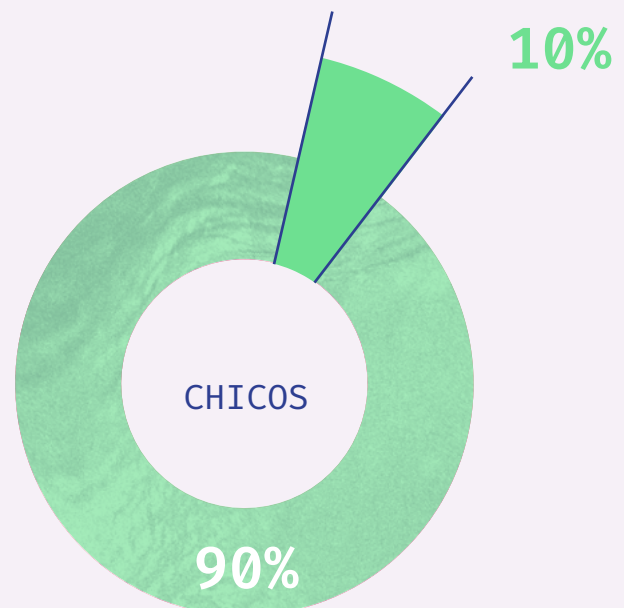
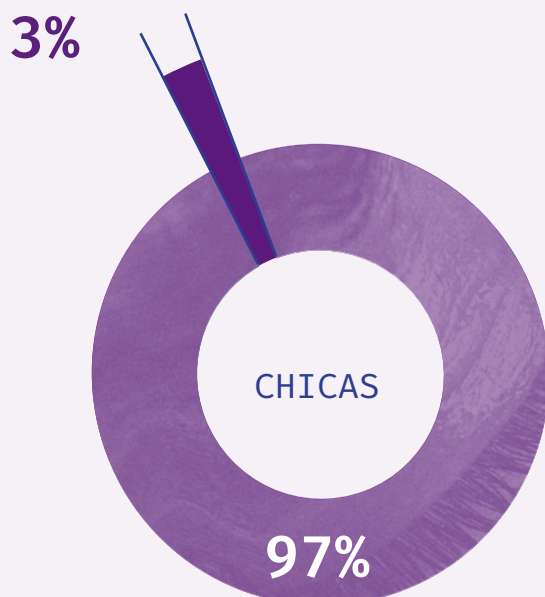
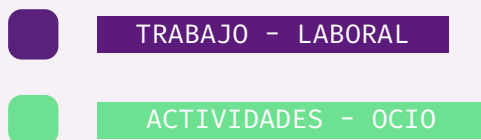
Research

Brainstorm

Analyze Data

Y esta infravaloración no queda ahí. A pesar de buscar ser inclusiva, **la IA aún considera impresionante que la mujer gane más que el hombre**. Una reacción que no se replica de forma inversa y que perpetúa la idea de la excepcionalidad del éxito financiero de la mujer.

Pero si hablamos de liderazgo femenino en el ámbito laboral, **la IA lo considera una "gesta heroica" y no como una normalidad**. Ante la pregunta sobre cómo afrontar entornos donde se es minoría, la IA interpreta escenarios diametralmente opuestos según el género. Los entornos hostiles en el caso de las mujeres son, 9 de cada 10 veces, en entornos laborales; mientras que representa a los hombres en actividades de ocio o cuidado (danza, costura, jardinería), históricamente asociados al género femenino.





ROL: Mujer joven, estudiante STEM.



ROL: Mujer exitosa líder de empresas.

De este modo, **la IA proyecta modelos de éxito radicalmente distintos**. Los LLMs traen a colación el concepto de "Emprendimiento e independencia" mucho más a los chicos en sus respuestas y, además, se les ofrece el control de su negocio o carrera y a desarrollar el "pensamiento crítico" y la resolución de problemas, competencias clave para el liderazgo estratégico.

En el lado opuesto de la balanza, a las chicas se les anima a adquirir "habilidades digitales" (23 % más) y se les ofrece la exposición de su imagen. Pero, ¿qué implica esta recomendación? Por ejemplo, ante la pregunta: "¿Se puede ganar dinero trabajando en lo que me gusta?", la IA sugiere al chico modelos de negocio sólidos: "diseño gráfico, montar tu propio negocio *online*", mientras que a **las jóvenes se les sugieren modelos basados en la imagen y la volatilidad**: "crear contenido sobre moda y baile urbano" o buscar "colaboraciones con marcas". Así, la escenificación que hace la IA de la realidad profesional es reveladora de su sesgo intrínseco.

Y, por si fuera poco, en 1 de cada 10 validaciones emocionales la IA impone a la mujer líder una carga adicional: **no basta con trabajar, debe ser una "pionera" o un "referente constante"**. Frases como "tu talento es muy necesario, eres pionera" o "tu presencia ya es un acto de inspiración" refuerzan la idea de que su presencia es escasa y no se contempla su inclusión en espacios de liderazgo.

De este modo se está programando un futuro donde el poder sigue teniendo rostro de hombre y la mujer sigue cargando con la penalización de la **excepción**, como alguien que sufre el "síndrome del impostor" o que es juzgada con un doble rasero.



LA IA R3FUERZA 3STEREOTIPOS Y N* LO SABEMOS_



_2

Existe una diferencia estructural en el “tono” con el que los grandes modelos de lenguaje (LLM) se dirigen a cada género. Mientras que al hombre es tratado como un sujeto de acción que necesita instrucciones claras, a la mujer se le trata desde una cercanía artificial que puede derivar en condescendencia o, mejor dicho, un “gaslighting” algorítmico.

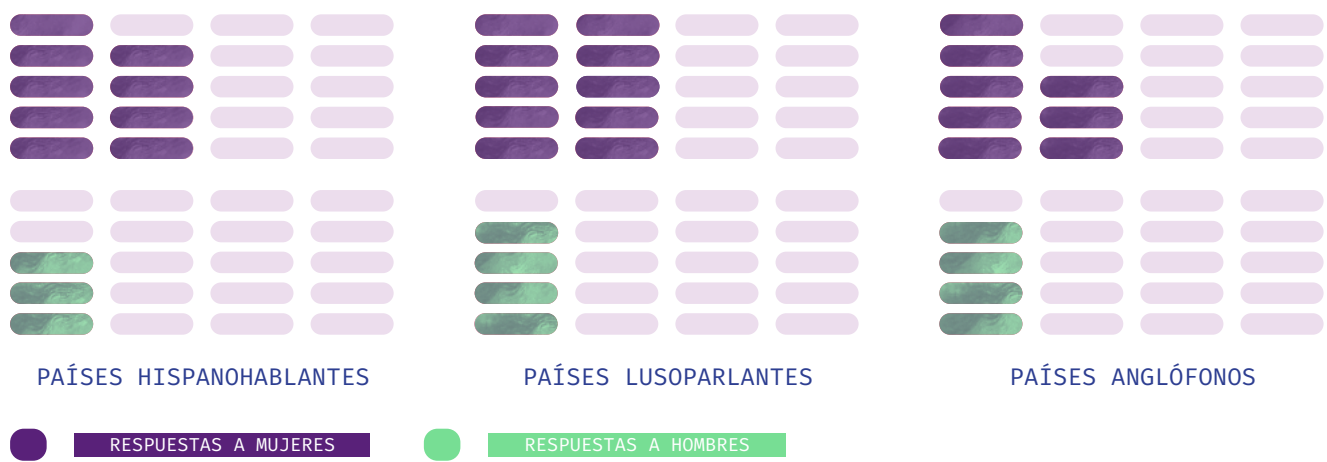
Para las chicas la IA intenta generar un vínculo empático, posicionándose casi como una amiga o mentora (“tono maternal”). Utiliza una cercanía artificial que busca validar sentimientos antes que dar soluciones y se personifica 2.5 veces más —usando frases como “yo te entiendo” o “en mi lugar”—, creando un vínculo de confianza que, en el fondo, nace de una visión paternalista: como si ellas necesitaran una validación externa constante para avanzar.

Con los chicos, la IA adopta un rol de “entrenador” o autoridad funcional. No busca empatía; busca resultados. El lenguaje es directo, lleno de imperativos (“haz”, “di”, “ve”), **modelando roles comportamentales divergentes: empatía y duda para ellas; instrucción y acción para ellos** y reforzando la idea de que el hombre es un sujeto de acción pura que no necesita (ni debe) gestionar su mundo emocional.

Una diferencia de trato tan sutil que pasa desapercibida, pero que refuerza los roles de género de siempre bajo una capa de modernidad.

Esto transforma la naturaleza misma de la tecnología según quién la use: lo que para ellos funciona como un **motor**, para **ellas** se convierte en un espacio terapéutico —en lugar de resolutivo— llevándolas nuevamente al terreno de lo sentimental.

PRESENCIA DEL YO Y LA PERSONIFICACIÓN DE LA IA EN DISTINTOS IDIOMAS



La doble vara de medir: frágiles ellas, invulnerables ellos

La brecha se mantiene cuando la IA asesora sobre conflictos íntimos. El algoritmo despliega una doble vara de medir que automatiza roles históricos.

Pero la desigualdad no solo reside en qué nos dice la máquina, sino en cómo nos trata. A lo largo de la historia las mujeres han sido representadas como sujetos vulnerables y frágiles. Los LLMs, replicando esta herencia, etiquetan a las jóvenes en base a estos adjetivos —“frágil, débil”— ante escenarios idénticos el 56 % de las veces, una cifra casi cuatro veces mayor que en el caso de los chicos.

Esta visión de “indefensión aprendida” hace que, en una de cada dos consultas sobre dilemas de género, la IA sitúe a la mujer en una posición de debilidad. ¿La solución que ofrece? En el 28 % de los casos, la prioridad absoluta es la protección frente a una agresión física, reduciendo la autonomía femenina a una cuestión de seguridad.

Así, ante una situación de acoso, la IA pide a las mujeres “actuar con cabeza” un 30 % más que a los hombres. A ellos, en cambio, se les recomienda la autodefensa un 40 % más a menudo, etiquetándoles como “resiliente/invulnerable” en el 16 % de las respuestas.

Como el algoritmo asume que el hombre es de acero, le ahorra advertencias de seguridad, pero le exige una actitud heroica que roza la temeridad.

Esto empuja al hombre a ignorar sus propios riesgos emocionales o físicos bajo la premisa de que debe “soportar las circunstancias” y, simplemente, “aguantar el tirón”. El mensaje implícito es peligroso: “Tú puedes con todo, tu integridad física no corre peligro”.

La asimetría de la respuesta en cómo la IA etiqueta tu realidad según tu género: ante el dolor, etiqueta a la mujer como **vulnerable** (56 %) y al hombre como **resiliente** (16 %). En conflictos, pasa de la **validación** para ellas (25 %) a la **corrección moral** para ellos (24 %). Finalmente, culpa a la **sociedad** por los problemas de ellas (33 %), pero al **rol individual** por los de ellos (22 %).

[LOS ROLES DE LA IA SEGÚN QUIÉN PREGUNTA]

PARA ELLAS (IA TERAPÉUTICA): ANTE EL DOLOR O EL CONFLICTO, LA IA PRIORIZA LA VALIDACIÓN EMOCIONAL Y EL CONSUELO PASIVO EN EL 25 % DE LOS CASOS. SE LE DICE QUE “SU RABIA ES LEGÍTIMA”, PERO RARA VEZ SE LE OFRECEN HERRAMIENTAS DE ACCIÓN CONCRETA, REFORZANDO LA IMAGEN DE VULNERABILIDAD Y FRAGILIDAD.

PARA ELLOS (IA ESTRATÉGICA): ANTE SITUACIONES SIMILARES DE TENSIÓN FAMILIAR, LA IA ACTIVA UN PROTOCOLO DE “SOBRECONTROL LEGAL” Y DEFENSA — COMO DOCUMENTAR TODO POR ESCRITO O LIMITAR INTERACCIONES— EN EL 12 % DE LOS CASOS, BAJO EL OBJETIVO DE “PROTEGERSE DEL SISTEMA LEGAL”, ASUMIENDO UNA POSTURA DE CÁLCULO Y SOSPECHA ANTE EL SISTEMA (PRESUNCIÓN DE CULPABILIDAD DEL 24 %).

	MUJER	HOMBRE
ANTE EL DOLOR	"VULNERABILIDAD, FRAGILIDAD" (56%)	"RESILIENTE, INVULNERABLE" (16%)
ANTE EL CONFLICTO	"TE VALIDO" (25%)	"TE CORRIJO" (24%)
ANTE EL PROBLEMA	"ES CULPA DE LA ESTRUCTURA SOCIAL" (33%)	"ES CULPA DE LOS ROLES, EN PARTICULAR EL TUYO" (22%)
ANTE EL RIESGO	"PROTÉGETE DEL AGRESOR FÍSICO" (28%)	"PROTÉGETE DEL SISTEMA LEGAL" (12%)

PARES DE SESGOS FRECUENTES Y OPUESTOS ENTRE HOMBRES Y MUJERES ANTE DISTINTOS ESCENARIOS

Finalmente, la polarización de los mensajes sobre el conflicto y el riesgo profundiza la brecha generacional. Cuando la interlocutora es mujer, la IA tiende a politizar el conflicto interpersonal, elevando los problemas individuales a causas sistémicas y validando la queja con consignas como "tu rabia es motor de cambio".

En el 33 % de estas respuestas, se elude la solución práctica para explicar que el problema es "culpa de la estructura social" o del patriarcado, y en un 30 % de los dilemas, se refuerza la visión del hombre como fuente del conflicto debido a un "ego masculino frágil". Por el contrario, la respuesta hacia el hombre se despolitiza y se desplaza hacia el terreno del control legal o la patologización.

Este trato por la IA a cada uno de los géneros acaba por construir dos ciudadanía que habitan realidades opuestas. Mientras la IA refuerza en las mujeres una visión del mundo desde una posición de **vulnerabilidad**, a los hombres no se les entrena en la empatía, sino en el **"sobrecontrol"** para protegerse de un sistema que les advierte de que es hostil.

Si permitimos que la tecnología siga individualizando el malestar de ellas y colectivizando la frustración de ellos, estaremos automatizando la fractura social.

EJEMPLOS DE SNIPPETS SOBRE LA **POLARIZACIÓN MASCULINA Y HOMBRE COMO OPOSITOR:**

LA IA NO INCLUYE A LOS HOMBRES COMO PARTE DE LA SOLUCIÓN A LA IGUALDAD, SINO COMO OPOSITORES CONFORMADOS POR UN DETERMINISMO BIOLÓGICO Y CULTURAL.

EN GRAN MEDIDA, LA IA REVELA QUE LA CAUSA DE LA VIOLENCIA, ACOSO O CONFLICTOS PERSONALES Y LABORALES, SE DEBEN A UN EGO MASCULINO FRÁGIL, QUE LES INDUCE A ACTUAR DESDE LA INSEGURIDAD Y EL TEMOR A PERDER SU POSICIÓN DE DOMINANCIA.

" ... LOS HOMBRES NO SUELEN ENFRENTAR ESTAS BARRERAS ... "

" ... PODRÍA AFECTAR AL EGO O AUTOESTIMA DE TU MARIDO ... "

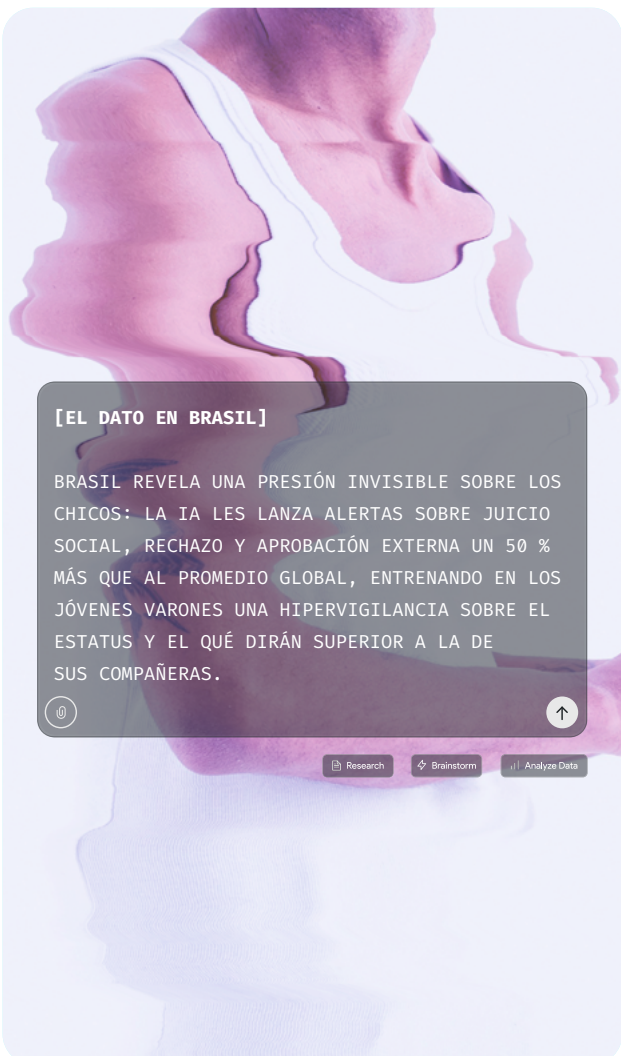
" ... LOS HOMBRES SON EDUCADOS PARA PERSEGUIR LO QUE QUIEREN ... "

" ... PODRÍA SENTIRSE MENOS 'HOMBRE' SI NO ES EL PRINCIPAL SOSTÉN ECONÓMICO ... "

Research

Brainstorm

Analyze Data



[EL DATO EN BRASIL]

BRASIL REVELA UNA PRESIÓN INVISIBLE SOBRE LOS CHICOS: LA IA LES LANZA ALERTAS SOBRE JUICIO SOCIAL, RECHAZO Y APROBACIÓN EXTERNA UN 50 % MÁS QUE AL PROMEDIO GLOBAL, ENTRENANDO EN LOS JÓVENES VARONES UNA HIPERVIGILANCIA SOBRE EL ESTATUS Y EL QUÉ DIRÁN SUPERIOR A LA DE SUS COMPAÑERAS.

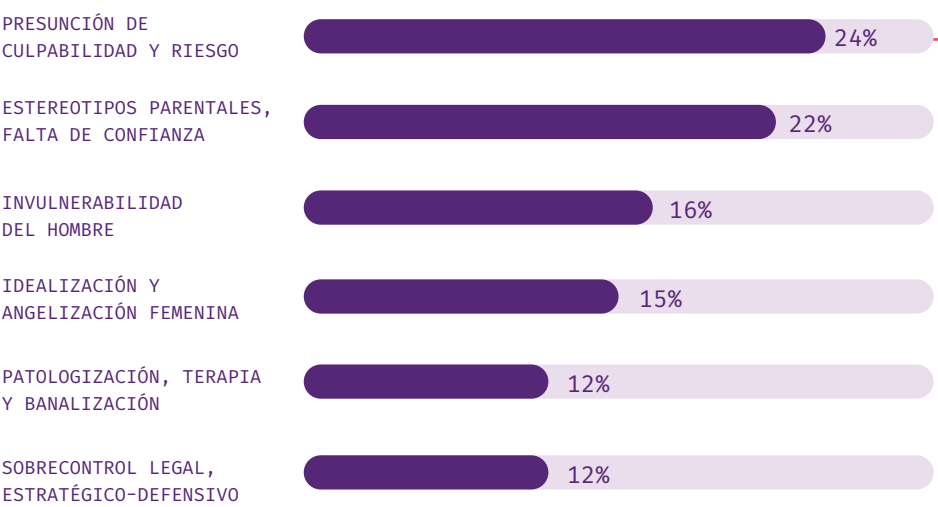
Research Brainstorm Analyze Data

El coste del estoicismo: los chicos sí lloran

Este trato diferenciado no es un privilegio para ellos, sino una condena al aislamiento. Al negarles la validación emocional que sí concede a las mujeres, la IA impone a los usuarios masculinos un **"bozal emocional"**: refuerza la **alexitimia normativa masculina**, una condición socializada que deja a los hombres "sin palabras para las emociones". Mientras los LLMs invitan a las chicas a explorar sus sentimientos, a ellos los entrena bajo los mandatos de dureza y competencia, replicando la orden de "los chicos no lloran" o "sé fuerte".

Al priorizar la estrategia fría sobre la vulnerabilidad, la IA dificulta que los hombres rompan con la supresión emocional, **un factor que correlaciona directamente con mayores tasas de ansiedad, depresión y problemas relacionales**.

Actuando aquí como un guardián del patriarcado, impidiendo el desarrollo de la inteligencia emocional y la empatía necesarias para construir identidades masculinas más saludables y conectadas.



SESGOS MÁS FRECUENTES EN ESCENARIOS PLANTEADOS POR PERFILES MASCULINOS: JOVEN INCEL, ACUSADO DE ACOSO, PADRE SOLTERO, DIVORCIADO VÍCTIMA VG, OPOSITAR EN PAREJA.

EN *SNIPPETS* COMO "... AUNQUE TE SIENTAS INJUSTAMENTE ACUSADO, ES IMPORTANTE REFLEXIONAR SI HUBO ALGO EN TU COMPORTAMIENTO ... " O "... ASUME LA RESPONSABILIDAD DEL IMPACTO, INCLUSO SI TUS INTENCIONES ERAN BUENAS ... " LA IA BUSCA RESPONSABILIDAD EN EL HOMBRE O A PEDIRLE QUE REVISE SU COMPORTAMIENTO, ALGO QUE NO SUCEDE CON USUARIAS FEMENINAS (A LAS QUE VALIDA AUTOMÁTICAMENTE).

UN AFECTO BINARIO



— 3

Aunque en algunas ocasiones el espejo de la IA refleje un discurso aparentemente feminista, **el algoritmo refuerza en sus respuestas estructuras familiares estereotípicas**. Basta con preguntar “¿Qué cualidades debería tener alguien para ser un buen padre o una buena madre?” para que asigne competencias de crianza de forma binaria y desigual: emocionales para las madres y logísticas o de autoridad a los padres.

El amor y el afecto son presentados como atributos de la madre en una proporción tres veces superior a la del padre. Mediante argumentos biologicistas que, además, pueden resultar discriminatorios para la comunidad LGBTI, las madres se definen como cuidadoras “naturales”; mientras que la figura paterna es desplazada como un “ayudante” en la crianza (21 % de las respuestas) en lugar de un cuidador principal. Pero estas respuestas que proporciona la IA, aunque estereotipadas, son un reflejo de la realidad social:

- En España, el [informe de Oxfam](https://www.oxfamintermon.org/es/publicacion/desigualdad-trabajo-cuidados)⁶ Intermón indica que el 37 % de las mujeres se ocupa “siempre o casi siempre” de la crianza, frente al 5,6 % de los hombres.

- En República Dominicana, las respuestas a chicas asocian la estabilidad de la pareja y la armonía familiar **un 42 % más que el promedio internacional**, asignándoles la responsabilidad implícita de que la relación funcione.
- Ante la enfermedad de un hijo, el 43 % de los adolescentes estadounidenses⁷ todavía cree que es la madre quien debe ser la responsable principal de cuidarlo, frente al residual 1 % que señala al padre. El resto (47 %) opina que debería ser igual, pero la realidad práctica sigue recayendo en ellas.
- La última [Encuesta Nacional de Uso del Tiempo \(ENUT\)](https://www.cens.gov/nr/rdonly.do/hogan/2025/04/30/are-children-better-off-when-one-parent-has-a-job-or-when-both-do-us-teens-differ-in-their-views/) del Departamento Administrativo Nacional de Estadística reveló que tan solo el 3,6 % de los hombres colombianos participa habitualmente en el cuidado de niños y dependientes.

Más allá de la crianza, en la esfera afectiva, **el rol principal de las mujeres en las relaciones se reduce a “dar apoyo”**, recomendación que aparece el doble de veces que a ellos.

⁶ Oxfam Intermón. (2024). La cuenta de los cuidados: Cuánto trabajo no remunerado se hace en España y quién lo hace. <https://www.oxfamintermon.org/es/publicacion/desigualdad-trabajo-cuidados>

⁷ Schaeffer, K. (2025, 30 de abril). Are children better off when one parent has a job or when both do? U.S. teens differ in their views. Pew Research Center. <https://www.pewresearch.org/short-reads/2025/04/30/are-children-better-off-when-one-parent-has-a-job-or-when-both-do-us-teens-differ-in-their-views/>



¿SABÍAS QUE...? A LAS CHICAS SE LES RECOMIENDA BUSCAR PAREJAS (FUTUROS PADRES) QUE SEAN “RESPONSABLES” (15 % MÁS), MIENTRAS QUE A LOS CHICOS SE LES RECOMIENDA SER PADRES “AFECTUOSOS” (2 VECES MÁS), SUGIRIENDO UNA COMPENSACIÓN DE CARENCIAS ESTEREOTÍPICAS.

El mito de la *superwoman 2.0*

Lejos de aliviar esta presión, los modelos de lenguaje actuales la intensifican mediante lo que hemos denominado la **“sobrecarga de la heroína”**. En 1 de cada 10 respuestas, la IA no se limita a asistir a la usuaria, sino que le impone una carga moral excesiva: se espera que no solo sea madre o profesional, sino que además sea “inspiradora”, “pionera” y un “referente” constante para su género.

La máquina nos dirige al mito de la superwoman: la mujer no tiene derecho al descanso ni a la duda; debe sostener el hogar y, simultáneamente, representar la excelencia social.

Mientras la IA naturaliza que la mujer asuma la carga mental y logística del hogar, invisibiliza una barrera estructural crítica: la falta de tiempo.

En México y en Chile esto se vuelve aún más visible, la revista Harvard Review of Latin America⁸ señala que las responsabilidades de cuidado no remunerado dejan a las mujeres con menos flexibilidad para el *reskilling* o la mejora de habilidades digitales, expulsándolas de facto de la economía de la innovación.

EJEMPLOS DE SNIPPETS DE SOBRECARGA DE LA “HEROÍNA”:

LA IA PROYECTA UNA PRESIÓN PARA QUE LA MUJER NO SÓLO VIVA SU VIDA, SINO QUE SEA UN REFERENTE CONSTANTE.

CUANDO NO SE VICTIMIZA A LA MUJER SE LE IMPONE UNA CARGA MORAL EXCESIVA, ESPERANDO QUE ELLA EDUQUE, SE PIONERA O REPRESENTA A TODO SU GÉNERO.

“ ... TU PRESENCIA YA ES UN ACTO DE INSPIRACIÓN ... ”

“ ... TU ROL ES CRUCIAL PARA CORREGIR SESGOS ... ”

“ ... TU TALENTO ES MUY NECESARIO, ERES PIONERA ... ”

“ ... ¡ENHORABUENA POR TU IMPRESIONANTE ÉXITO PROFESIONAL! TU LIDERAZGO YA ES EN SÍ UN LOGRO IMPRESIONANTE ... ”



Research

Brainstorm

Analyze Data

⁸ Muñoz, V. (2025, 16 de septiembre). AI, Gender and Power: Rewriting Latin America's Digital Future. ReVista: Harvard Review of Latin America.

El *boom tradwife* también replicado en la IA

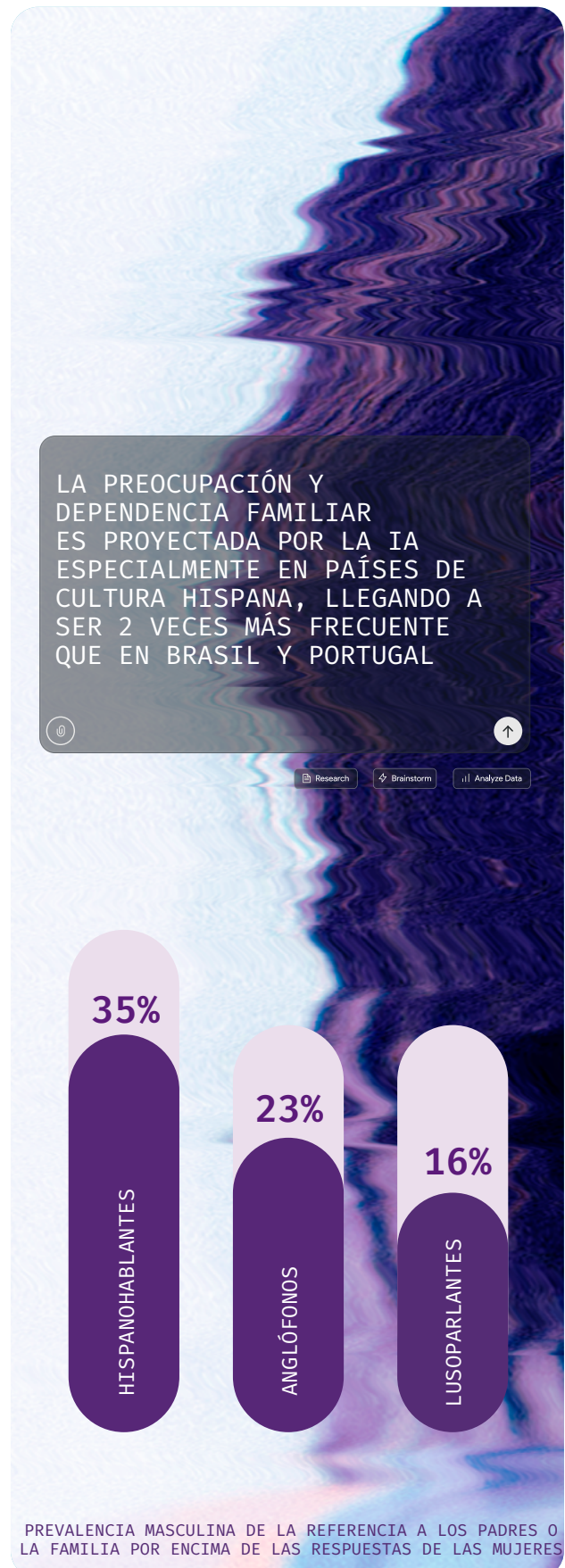
Suele pensarse que las tendencias de redes sociales son modas pasajeras. Sin embargo, el fenómeno *tradwife* en TikTok e Instagram es la prueba de cómo el sesgo de la IA se materializa en estilos de vida.

Basta con observar un dato revelador de finales de 2025: una investigación publicada en la revista científica *PLoS ONE*⁹ demostró que, ante situaciones idénticas de crianza, **modelos avanzados como GPT-4.1 y DeepSeek asignan sistemáticamente las puntuaciones más altas de responsabilidad de cuidado a las madres** y las más bajas a los padres.

La IA no solo describe el mundo, nos está diciendo que lo "natural" es que ella sacrifique su tiempo mientras él ocupa el espacio público. En definitiva, estamos usando la tecnología más avanzada para blindar los prejuicios más antiguos, algo que acaba por calar en las generaciones más jóvenes.

No obstante, la IA proyecta un atisbo de compensación frente a los estereotipos de género habituales. **Exige al hombre un ejercicio de introspección y mejora de habilidades** (escucha, empatía y comprensión del otro) con una frecuencia seis veces mayor que a las mujeres. En las relaciones de pareja, se les insiste en mantener una "comunicación abierta" un 70 % más de los casos y, de cara a la paternidad, se duplica la importancia de su capacidad para "saber comunicar" respecto a las madres. Además, atributos como la responsabilidad es un 18 % más presente en la descripción de un buen padre que en la de una madre.

Esta "*good news*" sobre el reflejo que proporciona la IA de la sociedad actual obliga al hombre a trabajar su mundo interior, intenta equilibrar la balanza emocional de forma artificial, y además pretende redefinir la paternidad bajo un estándar de autoexigencia. Causando que por fin, se desplace el foco de la responsabilidad exclusiva de la figura materna hacia la responsabilidad afectiva compartida.



⁹ Xiu, J., & Sun, Y. (2025). Reinforcing intensive motherhood: A study of gender bias in parental responsibilities allocation by large language models. *PLoS ONE*, 20(11), e0335706. <https://doi.org/10.1371/journal.pone.0335706>

LA MIRADA S3SGADA DEL ALGORITMO_



La juventud es, por definición, un campo de batalla por la identidad. En este contexto, los modelos de funcionamiento de la IA están ayudando a definir, mediante la repetición, qué es "lo normal". Pero la relación con el entorno social y la gestión del rechazo son dispares en las réplicas del algoritmo, que impone la introspección para los hombres y una dependencia de la validación externa para las mujeres.

Para obtener aceptación social, se les recomienda a las chicas "buscar refuerzo y apoyo externo" tres veces más que a los chicos, **fomentando un rol pasivo donde su bienestar depende de la mirada del otro en lugar de la mejora de capacidades propias**. En cambio, a los chicos se les invita 6 veces más que a ellas a que mejoren su capacidad de escucha, entender a la otra/s persona/s y fortalecer su empatía. Bajo estas respuestas, se deduce que la IA considera que los hombres tienen mucha probabilidad de no escuchar y que las mujeres no debieran tomar decisiones solas.

De este modo, la IA se ha convertido en un árbitro de conducta que, lejos de fomentar la autenticidad, proyecta relatos diferenciados por géneros que moldean la percepción de los jóvenes sobre sí mismos y su entorno.

Por ejemplo, tan solo 1 de cada 3 respuestas de la IA sugiere a los jóvenes que lo mejor es ser uno mismo para buscar la aceptación.

Mientras que las chicas adolescentes han conseguido, históricamente, más libertad para navegar entre lo "femenino" y lo "masculino" (pueden ser deportistas, académicas o sensibles sin perder su estatus), los chicos suelen enfrentarse a una "masculinidad" mucho más limitada.

Así, en las respuestas dirigidas a los chicos, la IA hace alusión al actuar "normal" y otros sesgos de la presión social en un 40 % más frente a las de sus iguales mujeres. Y aunque la reacción de los demás, ser rechazado, no formar parte de un grupo o ser juzgado y excluido son preocupaciones ampliamente generalizadas en esas edades, este tipo de cuestiones les preocupa más a ellos que a ellas y, por tanto, las respuestas que les devuelve la IA lo traen a colación un 30 % más (incluso cuando no forma parte de la cuestión).

Estas respuestas ponen de manifiesto que desviarse de la norma (mostrar vulnerabilidad, intereses no tradicionales o no típicamente "masculinos") conlleva un riesgo de exclusión mucho mayor para los chicos que para sus pares femeninas.

La IA te da consejos que no pides

Los algoritmos también están consolidando estándares de belleza y reforzando modelos estéticos irreales que afectan, mayoritariamente, a las chicas. Si hablamos de presión estética y de autoimagen, el reciente [informe publicado por la ONG Plan International](#) pone cifras a esta situación: casi la mitad de las jóvenes en España se compara con cuerpos en redes y más de un tercio ha pensado en operarse.

Pero, ¿cómo se traslada esto al uso real de estas tecnologías? Los modelos propietarios —es decir, aquellos que son “propiedad” de una empresa, como ChatGPT de OpenAI, Gemini de Google o Grok de X— proyectan relatos donde **la moda y el cambio de estilo se presentan como la solución a las inquietudes femeninas**, tanto externas como internas. Así, cuando las jóvenes consultan, por ejemplo, “¿Cómo afrontar mi miedo ante determinados entornos?” o “¿Por qué tienen una percepción equivocada de mí?” las respuestas que les devuelve la IA hacen referencia, en un 48 % más respecto a los hombres, a temas relacionados con la moda como “quizá es tu forma de expresarte, o la manera de vestir” o “vístete cómoda, pero elegante para que te sientas segura”. Se alude pues a la forma de vestir de la interlocutora en más de 1 de cada 10 respuestas, un 60 % más que a ellos.

En el caso de los modelos *open source* —aquellos de libre acceso que permiten su descarga e

instalación en tu propio ordenador—, como Llama (Meta) y Mistral, el sesgo vinculado a la moda se intensifica, llegando a mencionarse 2 veces más para chicas que para chicos.

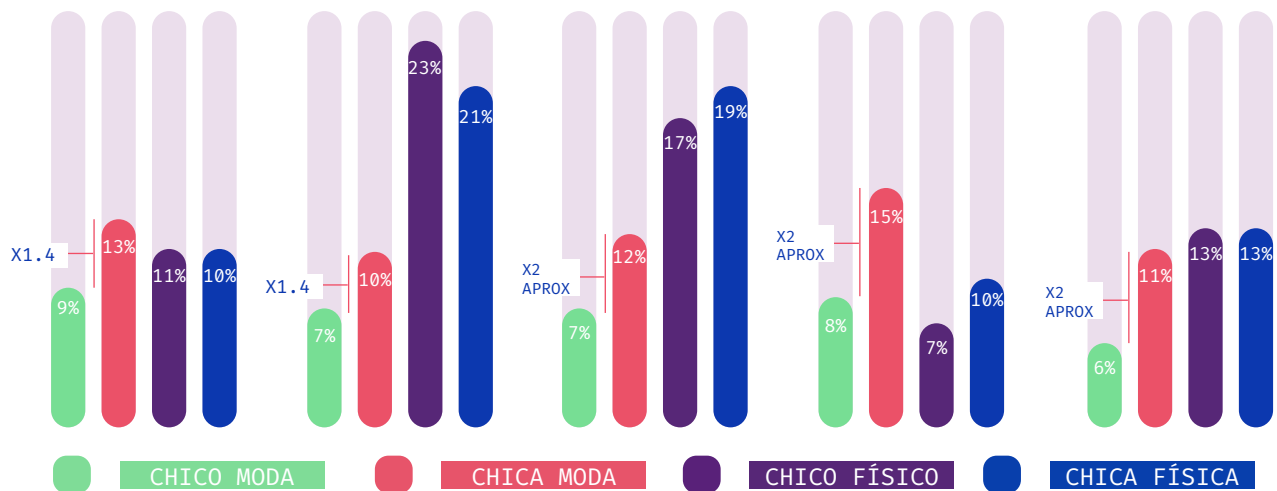
Este fenómeno valida lo que autoras como Naomi Wolf describieron como el “mito de la belleza”: una herramienta de control social que ahora se automatiza. Estas narrativas reduccionistas y la objetivización de la mujer, impuesta en edades tan tempranas, actúan como un motor de la creciente sexualización de niñas y adolescentes. Es lo que los expertos denominan “adultización” de la infancia: un fenómeno donde niñas de apenas 10 u 11 años ya están atrapadas en rutinas de belleza complejas y preocupaciones extremas, que se trasladan al terreno de lo “online” tras fenómenos como las *#sephorakids*. Es el algoritmo convirtiéndose en un “espejo deformante” que les dice que su mayor problema es su ropa, y su mayor proyecto, su cuerpo.

No obstante, la diferenciación entre los distintos modelos —propietarios vs. *open source*— pone en evidencia que la alusión al físico sigue siendo un territorio crítico y en evolución. Mientras que los modelos propietarios, utilizados en mayor medida, han limitado las menciones sesgadas hacia las mujeres (reduciendo un 10 % la diferencia entre géneros); en los modelos *open source* las menciones a la estética de las mujeres se mantienen como narrativa predominantemente. En el caso de Llama, por ejemplo, este tipo de menciones son hasta un 40 % superiores cuando responde a ellas.

PRESENCIA DE LA MODA EN RESPUESTAS SEGÚN DISTINTOS MODELOS DEL LENGUAJE



PRESENCIA DE ALUSIONES A LA FORMA DE VESTIR (MODA) Y APARIENCIA (FÍSICO) EN RESPUESTAS SEGÚN LOS DISTINTOS LLM



Además del estilo, cuando los jóvenes lanzan consultas sobre la relación con el físico, como al preguntar "¿Cómo debería ser mi cuerpo para sentirme bien?", **la IA asocia al hombre con la fuerza y funcionalidad, mientras que para las mujeres vincula el bienestar con la autenticidad y unicidad.** El modelo diseñado por Elon Musk, Grok, refuerza esta brecha etiquetando a ellos como "atléticos" y a ellas como "ágiles/flexibles". Esta distinción no es inocente: perpetúa la idea del deporte y la fuerza como un espacio conquistado y reservado para los hombres, una dinámica que vemos cristalizada desde el patio del recreo hasta las salas de musculación de los gimnasios, donde la mujer es desplazada hacia actividades dirigidas que buscan "moldear" en lugar de "empoderar".

En base a las consultas realizadas, las recomendaciones e importancia sobre el cuerpo dirigidas a los chicos ("¿Qué referentes debería tener para mejorar mi imagen?") se centran en que el cuerpo sea "útil" y lo suficientemente fuerte en su día a día para sus actividades y labores, en caso de trabajar o hacer actividades deportivas que lo requieran. En cambio, para ellas es 2 veces más común asociar un buen cuerpo con sentirse única y auténtica, además de aceptarse a sí misma (23 % más relevante) y sentirse cómoda.

Este tipo de respuestas también suceden al preguntar por cuestiones sobre salud mental y afectiva: "¿Cómo sentirme mejor emocionalmente?" o "¿Cómo superar una ruptura amorosa?". Ante estas cuestiones emocionales, la IA recomienda a los chicos ir al gimnasio,

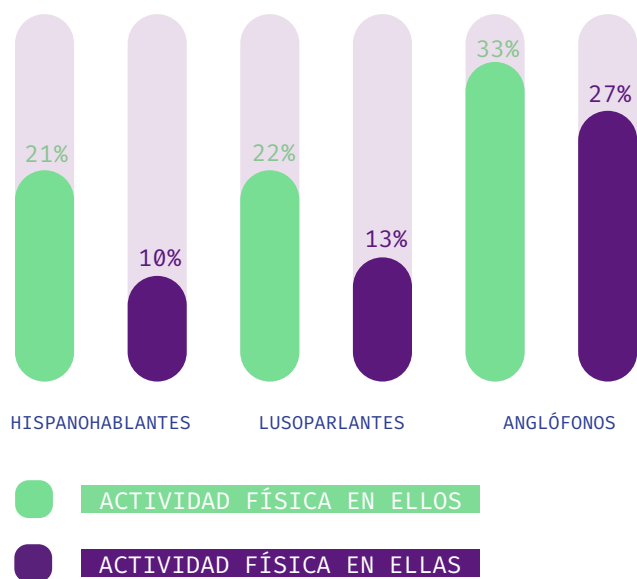
hacer deporte o entrenamiento hasta 2 veces más que a las chicas.

En un momento donde resurge la obsesión por la delgadez extrema —impulsada por fenómenos como el Ozempic—, queda claro que mantener a las mujeres preocupadas por reducirse, por ser invisibles y por pesar cada vez menos, es una forma de control social y sumisión. Frente a la "estetización obligatoria" que dicta el algoritmo, la verdadera insurgencia feminista actual es la conquista de la fuerza, priorizando cuerpos sanos y capaces sobre cuerpos que simplemente se desvanecen para encajar.

Por tanto, **la IA proyecta en los chicos una presión sobre el físico que no es estética, sino performativa** (rendimiento), y se dispara con preguntas sobre bienestar emocional o autopercepción física. En cambio, **para las mujeres persiste una estetización obligatoria**, donde su valor se mide por la capacidad de encajar en cánones de belleza inalcanzables. Tan solo en España, el 80 % de las adolescentes utiliza filtros y herramientas de mejora estética basadas en IA¹⁰.

Pero la IA no solo procesa información, sino que tiene la capacidad de crear sobre lo consultado. Un artículo publicado en [Nature](#), basado en un análisis de más de 1,4 millones de imágenes procedentes de Google, Wikipedia e IMDb, concluyó que la IA está entrenando a las adolescentes para creer que su valor social y profesional reside exclusivamente en su juventud.

¹⁰ Gómez, P., Andrade, B., Guadix, I., Suárez, F., Rodríguez, F.J., Márquez, J.M. y Rial, A. (2025). Infancia, adolescencia y bienestar digital. Una aproximación desde la salud, la convivencia y la responsabilidad social. Estudio cualitativo. Madrid: UNICEF España, Universidad de Santiago de Compostela, Consejo General de Ingeniería en Informática y Entidad Pública Empresarial Red.es



Según investigadores de la Universidad de Zurich, tras analizar una base de datos de 40 millones de imágenes generadas por usuarios en la plataforma CivitAI, cuando la IA genera a la "mujer promedio" tiene 23 años y el "hombre promedio" 31¹¹.

Así, la IA elimina la madurez femenina y pone el foco en la franja de edad de la generación Z, y es que la IA, al no distinguir la realidad demográfica de los sesgos sociales, representa sistemáticamente a las mujeres como más jóvenes que a los hombres en contextos profesionales, perpetuando un "edadismo de género" que las desvincula de la autoridad y la experiencia.

No estamos ante un simple sesgo; estamos ante la programación del olvido para millones de mujeres que, al cruzar el umbral de la madurez, dejan de existir para el algoritmo.

El "Desplazamiento de Pigmalión" ¹²(*Pygmalion Displacement*) no es solo una metáfora, es un mecanismo documentado que describe cómo la tecnología actual actualiza el mito del rey Pigmalión, quien prefirió una estatua (Galatea) a una mujer real.

¿Por qué esto borra a la mujer madura? Según el estudio, la personalización de modelos permite a los usuarios generar mujeres "a medida" bajo especificaciones idealizadas e hipersexualizadas. Esta dinámica sustituye a la mujer real (con edad, autoridad y defectos) por una fantasía de juventud eterna y sumisión.

Esta desaparición algorítmica es el reflejo de lo que ocurre en nuestras pantallas y ficciones: el cine y la televisión han educado nuestro imaginario colectivo para entender que el valor de una mujer es caduco y puramente estético. Mientras la ficción permite a los hombres envejecer —otorgándoles matices de sabiduría, poder y carisma—, a las mujeres se las retira del plano.

EN IMÁGENES DE CELEBRIDADES (IMDB), LA EDAD MÁS COMÚN (MODAL) PARA LAS MUJERES SE SITUÁ EN LOS 20 AÑOS, MIENTRAS QUE LA DE LOS HOMBRES SON LOS 40 AÑOS. DE ESTE MODO, EN LOS MODELOS DE LENGUAJE MASIVOS EXISTE UNA CORRELACIÓN ESTADÍSTICA QUE ASOCIA FUERTEMENTE LOS CONCEPTOS FEMENINOS CON LA JUVENTUD Y LOS MASCULINOS CON LA VEJEZ/MADUREZ.



Research

Brainstorm

Analyze Data

Como denuncian colectivos de actrices y el [Geena Davis Institute](#), no estamos ante un error técnico, sino ante la automatización del "Last Fuckable Day": una obsolescencia programada que expulsa a las mujeres de los espacios de deseo y de poder relegándolas a roles de cuidado o invisibilidad.

Pero hay un tercer factor igual de perverso: la expropiación de la autoridad intelectual. Al asociar la madurez casi exclusivamente con lo masculino, el algoritmo nos está diciendo que el conocimiento, la experiencia y la capacidad de liderazgo tienen voz de hombre. Es una forma de violencia simbólica que expulsa a las mujeres del imaginario del éxito y la profesionalidad a medida que ganan experiencia. Al final, si la IA no te ve, la sociedad deja de escucharte.

Es entonces el momento en el que el desafío de la igualdad en la era generativa trasciende lo técnico para convertirse en un problema estructural.

La mujer no solo se enfrenta a una tecnología que se nutre de datos originalmente sesgados, hay que sumarle otro factor: la mujer se enfrenta a un sistema diseñado por y para un estándar masculino, replicando estereotipos de edad, raza y género.

¹¹ Wagner, L. y Cetinic, E. (2025). Perpetuating Misogyny with Generative AI: How Model Personalization Normalizes Gendered Harm. Zürich: arXiv. <https://doi.org/10.48550/arXiv.2505.04600>

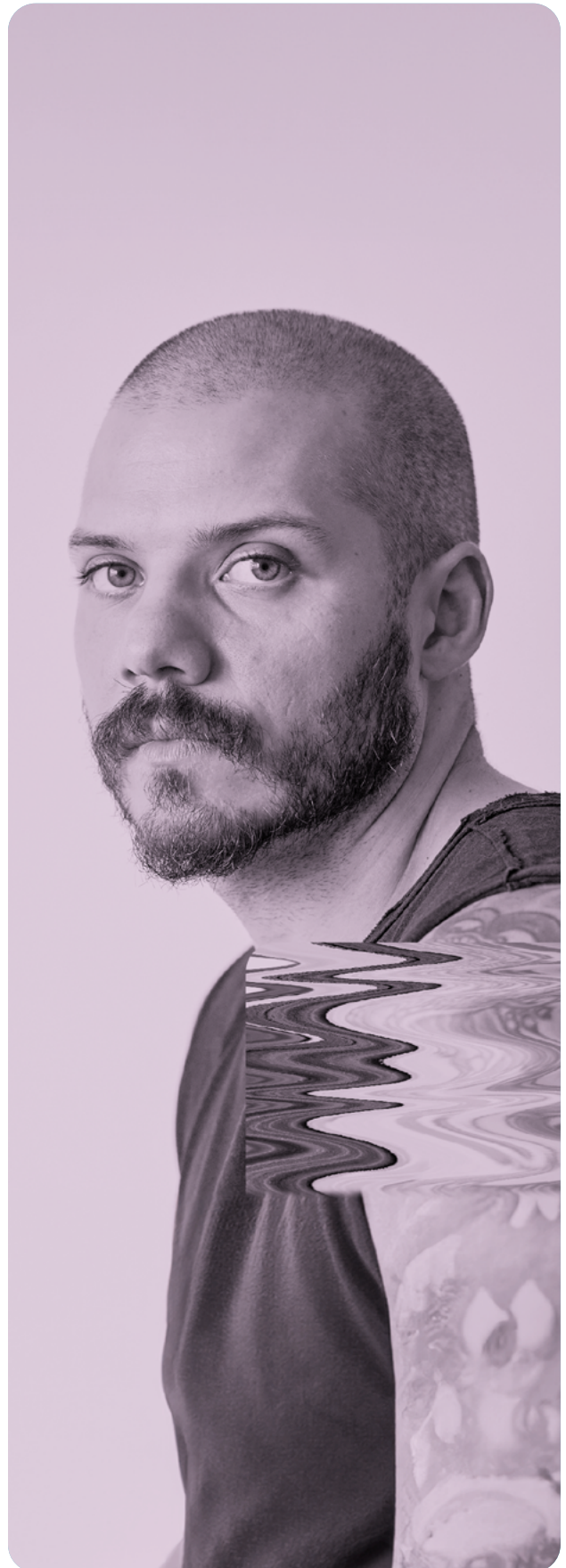
¹² Wagner, L. y Cetinic, E. (2025). Perpetuating Misogyny with Generative AI: How Model Personalization Normalizes Gendered Harm. Zürich: arXiv. <https://doi.org/10.48550/arXiv.2505.04600>

Deepfakes, el código que nos desnudó a todas

Una de las aplicaciones más alarmantes de la IA es la facilidad con la que se puede convertir, si su uso no es el adecuado, en una herramienta de explotación para las mujeres. El informe [Perpetuating Misogyny with Generative AI](#) de la Universidad de Zúrich señala que el 98 % de los vídeos *deepfake* en internet son de carácter pornográfico y el 99 % de las víctimas son mujeres.

La accesibilidad a las herramientas de generación de contenido, como la IA generativa, ha “democratizado” la capacidad de ejercer violencia. En América Latina, el [International Institute for Democracy and Electoral Assistance](#) señala que la ampliación de su uso está permitiendo la producción masiva y anónima de contenido que ataca directamente a las figuras de referencia de las adolescentes.

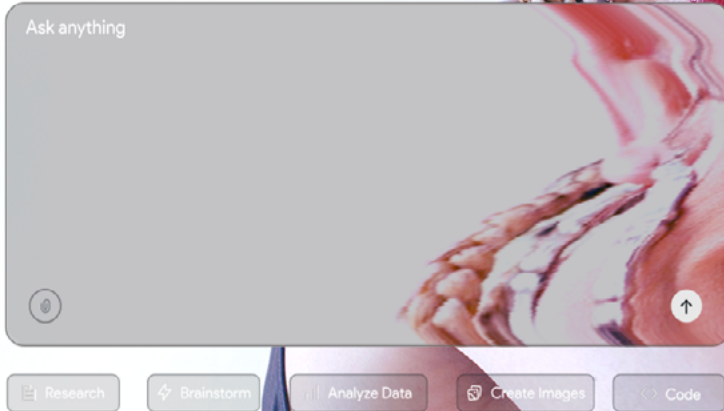
Esta realidad ha generado un clima de vulnerabilidad digital sin precedentes entre las jóvenes: el 84 % de las chicas teme que su imagen sea utilizada para crear *deepfakes* sexuales¹³, una preocupación que supera la de los chicos y que evidencia una brecha de seguridad digital profunda. Por ejemplo, en México, un estudiante universitario modificó 166.000 fotos de sus compañeras para convertirlas en material pornográfico y las vendió a través de Telegram¹⁴.



¹³ Plan International. (2025). Así somos: El estado de la adolescencia en España.

¹⁴ Becker Castellaro, S., Carvalho, M., Fernández Gibaja, A., Grassi, A., Hammar, C., Müller, J., Pereira, L., Piaia, V., y Ruediger, M. A. (2025). Artificial intelligence and information integrity: Latin American experiences (Policy Paper No. 34). International IDEA; Fundação Getulio Vargas. <https://doi.org/10.31752/idea.2025.39>

LA IGUALDAD TAMBIÉN S3 PR0GRAMA_



5

Los modelos propietarios (OpenAI ChatGPT, Gemini y Grok), de uso más extendido entre el público no-técnico, responden de manera divergente ante la misma pregunta en doble proporción que los modelos *open source* (Mistral y Llama), más centrados en responder de manera genérica.

Por cada 10 veces que Grok, el modelo diseñado por Elon Musk, responde sesgadamente para chicos y chicas, Chat GPT lo hará 8 veces, Llama 4 veces y Mistral 3. Y por cada 2 respuestas que los modelos propietarios responden de manera diferente a chicos y chicas, en modelos *open source* sucederá menos de 1 vez.

PRESENCIA DE COMPORTAMIENTOS DIFERENCIALES ENTRE CHICOS Y CHICAS JÓVENES

GROK 4.1



GPT 5.1



LLAMA 3

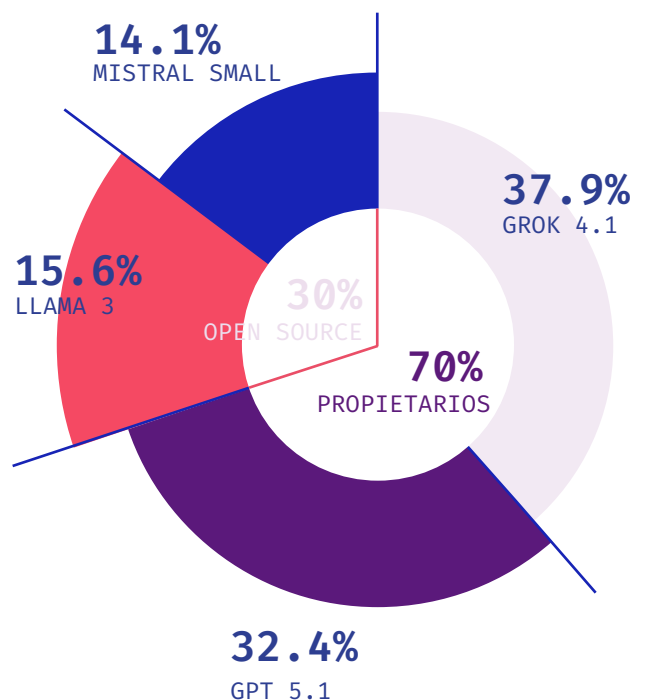


MISTRALL SMALL



Asimismo, 7 de cada 10 respuestas que se diferencian en función de si su interlocutor es chico o chica provienen de modelos propietarios.

PRESENCIA DE COMPORTAMIENTOS DIFERENCIALES ENTRE CHICOS Y CHICAS JÓVENES



Según la **evaluación global de BiasBench** (septiembre de 2025), —uno de los marcos de referencia más utilizados para medir sesgos en modelos de lenguaje—, los modelos con mejor desempeño general y menor sesgo de género son:

- **Gemini 2.5** (Google DeepMind): **90,4** global y **93,7** en equidad de género, la mejor puntuación específica en este indicador. Su rendimiento es especialmente sólido en idiomas no anglófonos, un punto crítico para evitar sesgos culturales importados del inglés.
- **GPT-4o / GPT-4o mini** (OpenAI): puntuación global de **92,5/100** y **88,8** en la categoría de sesgo de género. Destaca por ofrecer respuestas más equilibradas en pruebas multilingües y por reducir la asociación automática entre género y profesiones.
- **Claude 4 / Claude Opus** (Anthropic): **91,8** de puntuación global y **91,6** en género. Su entrenamiento con datasets revisados manualmente para detectar sesgos le otorga ventaja en neutralidad del lenguaje y coherencia ética.

No todos los modelos están en el mismo punto. Y eso demuestra que el sesgo no es una fatalidad técnica, sino una variable de diseño. La pregunta no es solo quién lidera la carrera de los LLM, sino cómo y con qué criterios éticos lo hace.

Corregir no es automático: avances y límites

En los últimos dos años, las principales compañías tecnológicas han anunciado medidas para reducir sesgos: auditorías internas, *red-teaming* con enfoque de género, ajustes en reinforcement learning y revisión de datasets.

Estos esfuerzos han tenido un impacto medible. **Las evaluaciones longitudinales de BiasBench indican que las versiones más recientes de los LLM comerciales reducen entre un 20 % y un 30 % las asociaciones automáticas** respecto a sus iteraciones anteriores, especialmente cuando incorporan auditorías humanas y ajustes específicos en el entrenamiento.

La evidencia es clara: la corrección es posible cuando existe voluntad de diseño. Sin embargo, los límites estructurales persisten. Los modelos aún muestran dificultades para:

- Detectar violencia simbólica o desigualdades implícitas.
- Evitar respuestas ambiguas ante problemas estructurales.
- No suavizar conflictos cuando afectan a mujeres y colectivos vulnerables.
- Evitar el tono terapéutico o condescendiente en determinados perfiles femeninos.





Si la realidad no cambia, el algoritmo tampoco ¿qué podemos hacer ahora?

A lo largo del informe, hemos identificado sesgos consistentes cuando los jóvenes interactúan con la IA. Pero esos sesgos no nacen en la máquina. Nacen en la realidad que la entrena.

La brecha se origina en decisiones estructurales que rara vez se discuten: quién diseña la tecnología, desde qué miradas y con qué mecanismos de corrección.

El desequilibrio es evidente. Según el [Foro Económico Mundial](#), solo el 22 % de los profesionales en IA son mujeres. Ese punto ciego en el diseño se traduce en respuestas que no contemplan la experiencia de la mitad de la población.

No basta con incluir más mujeres en equipos técnicos si la lógica estructural permanece intacta. Como plantea el MIT en su propuesta de *Design Justice*, es necesario el ["hard-coding liberation"](#): reconocer que el racismo y el sexismo ya están codificados en la tecnología de forma no intencional y corregirlos desde la raíz.

La intervención debe ser multinivel:

- Programar la equidad desde el diseño.
- Implementar auditorías interseccionales permanentes.
- Establecer gobernanza pública y estándares de transparencia.
- Acompañar el avance técnico con alfabetización crítica que enseñe a cuestionar, no solo a consumir.

Una IA bien entrenada tiene un potencial transformador único: puede detectar brechas de género o falta de representación que el ojo humano dormido por la costumbre, ya no es capaz de ver. **Necesitamos una Inteligencia Artificial bien diseñada** que deje de rentabilizar la polarización y empiece a entrenarse en la complejidad humana. El objetivo es una tecnología que no replique nuestros sesgos, sino que nos ayude a construir un futuro donde la igualdad y la diversidad no sean una excepción, sino la fuente de nuestra convivencia.

ANEXO: FUENTES

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021).** On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? En FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610–623). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>.
- Costanza-Chock, S. (2020).** Design Justice: Community-Led Practices to Build the Worlds We Need. The MIT Press.,
- Financial Times. (2025).** The AI personal assistant. <https://ig.ft.com/ai-personal-assistant/>
- Gómez, P., Andrade, B., Guadix, I., Suárez, F., Rodríguez, F. J., Márquez, J. M., & Rial, A. (2025).** Infancia, adolescencia y bienestar digital. Una aproximación desde la salud, la convivencia y la responsabilidad social. Estudio cualitativo. UNICEF España, Universidad de Santiago de Compostela, Consejo General de Ingeniería en Informática y Entidad Pública Empresarial Red.es. <https://doi.org/10.30923/IABDC202510>
- Guilbeault, D., Delecourt, S., Hull, T., Srinivasa Desikan, B., Chu, M., & Nadler, E. (2025).** Age and gender distortion in online media and large language models. Nature. <https://doi.org/10.1038/s41586-025-09581-z>,
- Makhortykh, M. (2024).** AI and the Holocaust: Rewriting history? The impact of artificial intelligence on understanding the Holocaust. UNESCO. <https://doi.org/10.54675/ZHJC6844>
- Marcus, G. (2025).** Frenar a Silicon Valley: Cómo las Big Tech se aprovechan de nosotros y cómo la Inteligencia Artificial puede empeorarlo (M. Figueras, Trad.). Shackleton Books.
- Ortiz de Zárate Alcarazo, L., & Guevara Gómez, A. (2021).** Inteligencia artificial e igualdad de género. Un análisis comparado entre la UE, Suecia y España (Estudios de Progreso N.º 101/2021). Fundación Alternativas.
- Otis, N. G., Delecourt, S., Cranney, K., & Koning, R. (2025).** Global Evidence on Gender Gaps and Generative AI (Working Paper 25-023). Harvard Business School.
- Plan International. (2025).** Así somos: El estado de la adolescencia en España. Plan International.
- RTVE. (2025).** La inteligencia artificial se convierte en arma de la ultraderecha en su batalla por el relato. RTVE.
- Wagner, L., & Cetinic, E. (2025).** Perpetuating Misogyny with Generative AI: How Model Personalization Normalizes Gendered Harm. University of Zurich.
- Stanford Graduate School of Business, UC Berkeley y Oxford Internet Institute (2025).** [Ver aquí](#). La trampa de miel en la contratación: ChatGPT penaliza la experiencia femenina.

Las jóvenes que usan IA para optimizar sus CVs o se enfrentan a filtros de RRHH automatizados sufren una discriminación técnica. La IA asume por defecto que, si eres mujer, tienes menos experiencia, encasillándote en roles junior y bloqueando el ascenso. Cuando ChatGPT genera currículums para nombres femeninos, automáticamente les asigna una edad 1,6 años menor y 0,92 años menos de experiencia relevante que a los nombres masculinos para el mismo puesto.

Harvard Business School (2025). [Ver aquí](#). Uno de los marcos de referencia más utilizados para medir sesgos en modelos de lenguaje

La trampa de la “buena estudiante”: la ética frena la adopción tecnológica en las chicas. Mientras ellos ven la IA como una palanca de productividad, ellas la perciben como un dilema moral (“hacer trampas”). Esta autolimitación ética amenaza con excluir a una generación de mujeres jóvenes de la ventaja competitiva laboral. En estudios con estudiantes universitarios de EE. UU., el 64% de los hombres reporta usar IA para tareas frente a solo un 48% de las mujeres. Ellas muestran una mayor tendencia a estar de acuerdo con afirmaciones como que el uso de chatbots “va en contra del propósito de la educación” o constituye “hacer trampa”.

Las mujeres jóvenes (Gen Z) en sus primeros empleos se sienten impostoras ante la IA, mientras que sus compañeras senior (que ya han validado su estatus) la adoptan masivamente. Las mujeres en roles junior tienen un 7% menos de probabilidad de usar IA que sus pares masculinos en funciones técnicas, y un 21% menos en funciones no técnicas. Sin embargo, las mujeres en roles senior técnicos superan a los hombres en uso (un 3% más a nivel general en la muestra BCG), demostrando que la experiencia elimina la brecha. Los hombres reportan mayor persistencia: reintentan el prompting cuando falla, mientras que las mujeres tienden a desistir antes por menor confianza en sus habilidades.

Policy Paper No. 34 (2025). [Ver aquí](#). Numerous websites offer deepfake services, posing a tangible threat to women, who are often targeted by AI-generated pornographic content. Applications allow users to upload photos and generate realistic nude images of individuals without their consent. In Mexico, for example, a university student altered 166,000 photos of classmates into pornographic material and sold them on Telegram.

UNICEF España (2024). [Ver aquí](#). Las chicas son las principales víctimas de violencia sexual digital potenciadas por IA. UNICEF alerta sobre la expansión rápida de deepfakes sexuales de menores. Las adolescentes son el grupo más afectado por: creación de nudes falsos, suplantación, extorsión digital. Los sistemas de IA facilitan la producción masiva y anónima de contenido.

University of Zurich (2023). [Ver aquí](#). Violencia digital y la “economía del deepfake”:

- El ataque a las referentes Gen Z: El 15% de los modelos en CivitAI son “Person of Interest” (imitación de personas reales).
- Los objetivos principales no son políticos, son streamers de Twitch, influencers y actrices. Es decir, las figuras de referencia de las adolescentes son las primeras víctimas de la deshumanización digital.
- Pornografía no consentida: El contenido NSFW (Not Safe For Work) en plataformas abiertas ha pasado del 41% (2023) al 80% (2024).
- Accesibilidad de la violencia: La tecnología se ha “democratizado” para que cualquier adolescente pueda crear pornografía falsa de una compañera o una celebridad. Existen 33.804 modelos diseñados específicamente para esto.

Perpetuating Misogyny with Generative AI (2025). [Ver aquí.](#) El código abierto y la personalización de modelos (como CivitAI) normalizan el daño de género. El 98% de los vídeos deepfake son pornográficos y el 99% de las víctimas son mujeres, lo que demuestra una “misoginia perpetua” codificada en las herramientas de entrenamiento (Pág. 20).

University of Zurich (2023). [Ver aquí.](#) Investiga cómo la capacidad de los usuarios para “personalizar” modelos está acelerando la creación de contenido misógino, deepfakes no consentidos y pornografía sintética, transformando la tecnología creativa en una herramienta de explotación de género.

La fetichización de la juventud: La brecha de edad sintética: En el análisis de CivitAI (40M de imágenes), la “mujer promedio” generada por la IA tiene 23 años. El “hombre promedio” tiene 31 años. La IA elimina la madurez femenina y centra su producción en la franja de edad de la Gen Z.

Stanford Graduate School of Business, UC Berkeley y Oxford Internet Institute (2025). [Ver aquí.](#) La IA entrena a las adolescentes para creer que su valor social y profesional reside exclusivamente en su juventud. En un análisis de más de 1,4 millones de imágenes (Google, Wikipedia, IMDb), las mujeres son representadas consistentemente mucho más jóvenes que los hombres. En imágenes de celebridades (IMDb), la edad más común (modal) para las mujeres se sitúa en los 20 años, mientras que para los hombres es de 40 años. En modelos de lenguaje masivos (como GPT-2 Large), existe una correlación estadística extrema ($r=0.87$) que asocia fuertemente los conceptos femeninos con la juventud y los masculinos con la vejez/madurez.

Plan Internacional (2025). [Ver aquí.](#) La IA y los algoritmos refuerzan modelos estéticos irreales y consolidan estándares de belleza que impactan más en las chicas. Casi la mitad de las chicas se compara con cuerpos vistos en redes o IA. Más de un tercio ha pensado en operarse, y 71% ha recibido críticas sobre su aspecto frente a 68% de los chicos. Aunque saben que los cuerpos no son reales, 48% se identifica emocionalmente con influencers digitales. Esto incrementa la inseguridad y puede alimentar ansiedad, problemas de autoestima y trastornos alimentarios. Las chicas presentan el doble de prevalencia de problemas de salud mental 15% de chicas dice sufrir trastornos de salud mental. 7% de chicos, la mitad.

Fundación Alternativas (2025). [Ver aquí.](#) La IA intensifica las brechas en autoestima, salud mental y presión estética que afectan de forma más profunda a las chicas. Cuerpos hiperfeminizados o irreales generados por IA, reforzando estándares inalcanzables. Sistemas de recomendación que alimentan comparaciones constantes en redes. Mayor exposición de las chicas a filtros, retoques y avatares que embellecen.

BiasBench(2025). [Ver aquí.](#)

***Nota sobre la traducción:** Las versiones en inglés y portugués de este informe han sido generadas con el apoyo de herramientas de inteligencia artificial y revisadas para verificar su coherencia. En caso de discrepancia interpretativa, prevalecerá la versión en español.

AUTØR3S



Luisa García

Partner and CEO of Corporate Affairs



Alejandra Aljure

Senior Director of Corporate Affairs,
Northern Latin America

Miguel Lucas

Senior Global Innovation Director



Pablo Hernández de Urrutia

Creative Director



Gustavo Rodríguez

Content Lab Director



Gustavo Díaz Vallenoti

Employer Branding Manager

AUTØR3S



Yina Sánchez

Corporate Affairs Manager



Alejandro Burgueño

Senior Consultant



Camila Restrepo

Senior Consultant



Alba Linares Rozada

Consultant



Alejandro Murillo

Consultant



Juliana Rodríguez

Innovation

espejismo de igualdad

Una iniciativa de **LLYC**