

# Informe

Informe

Grupos focales  
sobre sesgos en la  
**Inteligencia Artificial**

2025



## Detección participativa de los sesgos de la Inteligencia Artificial

### OBSERVATORIO DE LOS CONTENIDOS AUDIOVISUALES UNIVERSIDAD DE SALAMANCA

Facultad de Ciencias Sociales.

Campus Miguel de Unamuno. Edificio F.E.S.

Paseo de Francisco Tomás y Valiente, s/n. 37007.

Salamanca. España

[https://www.ocausal.es/investigacion/proyectos/deteccion-participativa-de-los-sesgos-de-la-inteligencia-artificial-ia-bias/ia\\_bias/](https://www.ocausal.es/investigacion/proyectos/deteccion-participativa-de-los-sesgos-de-la-inteligencia-artificial-ia-bias/ia_bias/)

Julio 2025

Las imágenes empleadas en esta publicación han sido diseñadas por colaboradores de Unsplash y se han descargado a través de su plataforma [www.unsplash.com](http://www.unsplash.com) bajo los términos de uso vigentes.



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2023-2024 [FCT-23-19454]

# Equipo de trabajo

## Equipo de trabajo

**Investigador principal:**

Carlos Arcila Calderón

**Investigadores:**

Valeriano Piñeiro Naval

Patricia Sánchez Holgado

Félix Ortega Mohedano

Laura Rodríguez Contreras

M. José Higuera

Maximiliano Frías Vázquez

Germán Rodríguez Wilches



# Contenido

---

- 01** Introducción
- 02** Objetivos
- 03** Metodología
- 04** Resultados
- 05** Global de grupos focales
- 06** Conclusiones generales
- 07** Referencias

# Introducción

## Introducción

En la actualidad, una parte sustancial de nuestras decisiones cotidianas está mediada, directa o indirectamente, por sistemas de inteligencia artificial (IA). Desde los buscadores que utilizamos hasta los algoritmos que seleccionan contenidos en redes sociales, pasando por asistentes virtuales, sistemas de navegación o herramientas de diagnóstico médico, la IA se ha integrado de manera progresiva en los entornos domésticos, laborales y educativos. Sin embargo, esta expansión tecnológica no está exenta de riesgos.





La Unión Europea define la IA como la capacidad de una máquina para imitar funciones humanas como el razonamiento, el aprendizaje, la planificación o la creatividad. Pero, como han señalado numerosas investigaciones, estos sistemas no son imparciales ni objetivos por sí mismos: están diseñados, entrenados y supervisados por personas, y por tanto reflejan —consciente o inconscientemente— los sesgos, estereotipos y desigualdades presentes en la sociedad (Broussard, 2018).

En este contexto, el proyecto “Detección participativa de los sesgos de la Inteligencia Artificial” (IA-BIAS), financiado por la FECYT y desarrollado por la Universidad de Salamanca, se propone involucrar activamente a la ciudadanía española en la reflexión y la detección de sesgos algorítmicos. Frente a una narrativa tecnológica que tiende a presentar la IA como infalible o neutral, este proyecto apuesta por una ciencia ciudadana crítica, en la que las personas puedan identificar, comprender y, en la medida de lo posible, corregir los mecanismos que producen resultados discriminatorios o injustos.



A diferencia de estudios previos que se han limitado a explorar las percepciones generales hacia la IA (Holder et al., 2018; Albarrán et al., 2020; Arcila et al., 2021), este informe ofrece una doble aportación: por un lado, recoge cómo distintos grupos sociales entienden y utilizan estas tecnologías; por otro, examina si pueden detectar los sesgos presentes en ellas, con qué criterios y con qué nivel de conciencia crítica.

La relevancia de esta tarea se entiende mejor si se considera que los sesgos algorítmicos no surgen de fallos técnicos aislados, sino de una cadena compleja de decisiones humanas. Como advierte Belloso Martín (2022), pueden originarse en el diseño del modelo, en la selección de los datos de entrenamiento, o en la forma en que los resultados son interpretados. Estos elementos demuestran que la IA no solo hereda las desigualdades históricas que existen en los datos, sino que puede amplificarlas si no se gestiona con responsabilidad ética y transparencia.

Con el propósito de explorar esta problemática desde una perspectiva social y generacional, se llevaron a cabo tres grupos focales diferenciados por edad: jóvenes (18-30 años), adultos (31-55) y mayores (56+). Este enfoque permitió comparar percepciones, usos, preocupaciones y propuestas de acción frente a la IA y sus sesgos.



# Objetivos

## Objetivos

Identificar la experiencia de los ciudadanos con herramientas y plataformas de IA con el fin de identificar los posibles sesgos que encuentran los españoles al utilizar estas tecnologías.

# Metodología

## Metodología

Los tres grupos focales fueron desarrollados el 27 de marzo de 2025 con una duración de 60 minutos cada uno. Se utilizó un guion semiestructurado con bloques temáticos sobre definición de IA, usos cotidianos, percepción de sesgos, impactos sociales, regulación y propuestas de mitigación. Las sesiones fueron grabadas, transcritas y analizadas cualitativamente con enfoque temático e inductivo hasta alcanzar saturación.

**Técnica empleada:**

Grupo focal semiestructurado.

**Total de participantes:**

30 (10 por grupo de edad).

**Criterios de segmentación:**

Edad (18-30, 31-55, 56+)

# 04

## Resultados

### Resultados

#### 4.1 Grupo de edad: 18 – 30 años

##### 4.1.1 Concepción de la IA

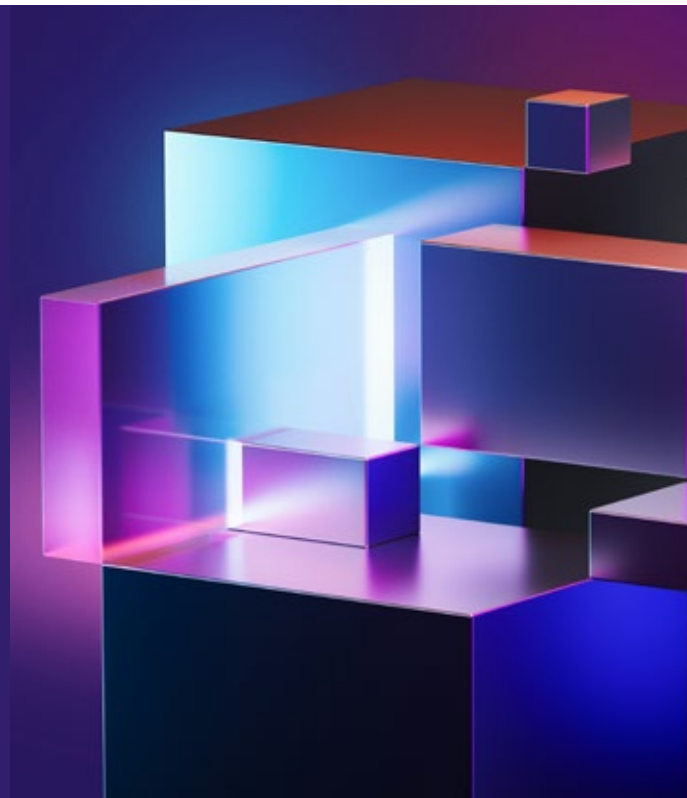
La conversación desmitifica la idea de «inteligencia». Varios participantes coinciden en que se trata de un modelo de predicción o de «un Google bastante sofisticado». Se la define como herramienta estadística más que como mente autónoma; algunos participantes coinciden en calificarla como «término de marketing». El consenso refleja alfabetización digital y escepticismo frente a la retórica comercial.

## 4.1.2 Uso cotidiano y herramientas

ChatGPT domina el repertorio para texto, acompañado de Copilot, DeepL y generadores de imagen. Mientras algunas personas enumeran hasta media docena de servicios, otras afirman «prefiero hacerlo yo mismo con Photoshop» y declaran no usar IA. Esta polaridad muestra adopción intensiva y rechazo principista dentro del mismo rango etario.

## 4.1.3 Fuentes de información y confianza

Surge desconfianza al emplear IA como fuente sobre sí misma. Se advierte que cada modelo «refleja lo que piensa la compañía que lo gestiona», por lo que, para obtener una mirada imparcial, «hay que ir por lo tradicional y leer». Otra opinión señala que los resultados suelen ser «un poco erróneos» y que la IA solo sirve para contrastar datos.



## 4.1.4 Uso cotidiano y herramientas

Se valora la rapidez para traducir o redactar, pero afloran miedos a la pereza cognitiva y a la pérdida de autoría creativa. Una voz considera «un error garrafal delegar la primera idea a una máquina» porque «estamos perdiendo la capacidad humana de controlar todo el proceso creativo». Desde la enseñanza se observa que el alumnado «está súper perezoso... se pierde el pensamiento crítico».

## 4.1.5 Percepción y ejemplos de sesgos

Existe acuerdo en que la IA reproduce prejuicios humanos. Se mencionan sesgos raciales en reconocimiento facial («reconoce mejor rostros asiáticos») y la necesidad de entrenar modelos adicionales para texturas afro. Se alerta de que entrenar con foros tóxicos produce «una IA ultramisógina». También se critica que ciertas herramientas lingüísticas privilegien corrientes tradicionales e ignoren variantes recientes. Una minoría declara no haber percibido sesgos en su experiencia.



## 4.1.6 Impacto social percibido

El riesgo más repetido es la pereza intelectual: «nos conduce a un mundo como el que imaginó Huxley en Mundo feliz... dejamos de pensar». Se añaden temores sobre intrusismo laboral, pérdida de la «esencia del alma humana», dependencia emocional - personas que «humanizan a ChatGPT y le consultan todo»— y la difusión cotidiana de contenido misógino «red-pill». El ámbito educativo aparece como zona de impacto y oportunidad.

## 4.1.7 Regulación, responsabilidad y mitigación

Se percibe urgencia de un marco regulador. Se propone crear «un comité de ética que valide las bases de datos, igual que con animales» y se insiste en que la responsabilidad abarque «toda la cadena», desde quienes diseñan y entrenan los modelos hasta las instituciones que los implementan. También se plantea la alfabetización digital desde la infancia para fomentar un uso crítico y evitar dependencias.



## 4.1.8 Conclusiones

Los y las jóvenes describen la inteligencia artificial como un sofisticado sistema de predicción y búsqueda, no como una mente autónoma. Este posicionamiento invita a emplear un lenguaje realista en futuras acciones de alfabetización. Reconocen la utilidad práctica de la IA —sobre todo para traducción y redacción—, pero expresan inquietud por el empobrecimiento del pensamiento creativo y la creciente pereza intelectual, especialmente entre estudiantes que recurren a la herramienta para cualquier tarea.

El riesgo más repetido es la pereza intelectual: «nos conduce a un mundo como el que imaginó Huxley en Mundo feliz... dejamos de pensar». Se añaden temores sobre intrusismo laboral, pérdida de la «esencia del alma humana», dependencia emocional - personas que «humanizan a ChatGPT y le consultan todo»— y la difusión cotidiana de contenido misógino «red-pill». El ámbito educativo aparece como zona de impacto y oportunidad.



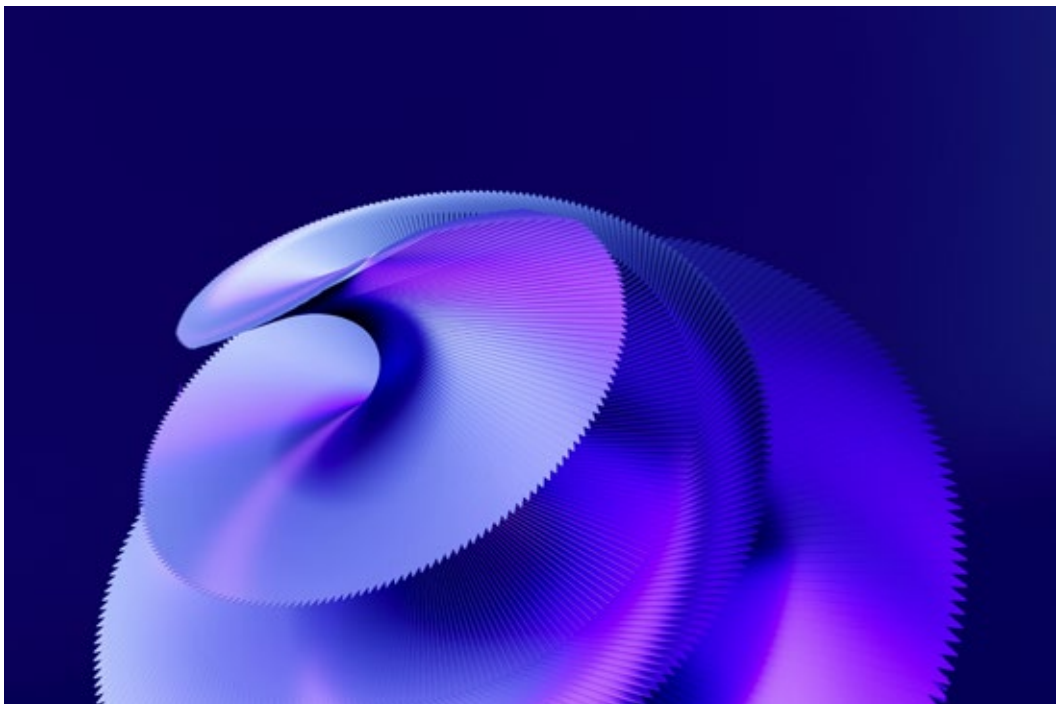
# Resultados

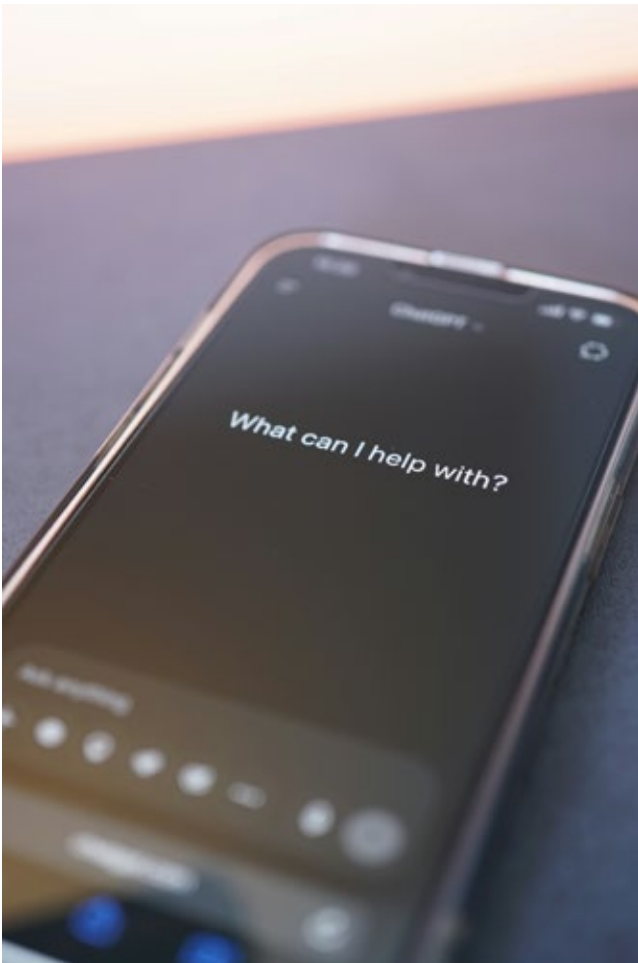
## Resultados

### 4.2 Grupo de edad: 31 – 55 años

#### 4.2.1 Concepción de la IA

Las primeras intervenciones describen la IA como «la capacidad de las máquinas para procesar información de manera similar a la de un humano», «un sistema que aprende de la interacción» o «un software que aprende». Más adelante se subraya que los algoritmos «replican los datos de los que se alimentan» y, por tanto, heredan sus prejuicios. La definición técnica convive con una conciencia temprana de que la IA nunca es neutral.





## 4.2.2 Fuentes de información y confianza

Quienes se informan activamente acuden a «artículos científicos en revistas especializadas» y a «páginas especializadas». La prensa generalista sirve para un seguimiento cotidiano, pero se valora la literatura académica cuando se busca rigor. El grupo coincide en que la propia IA es una fuente poco fiable si no se contrasta con lecturas “tradicionales”.

## 4.2.3 Usos cotidianos, ventajas y límites

Entre los participantes se encontraban: Docentes, sanitarios, ingenieros y doctorandos quienes contaron los usos intensivos de ChatGPT, Copilot, DeepL, generadores de imágenes y asistentes de voz. Las ventajas más repetidas son «ahorro de tiempo», «aporte creativo» y «estructurar ideas cuando me bloqueo». Al mismo tiempo se señalan límites: «empieza a sonar todo con el mismo tono», «inventa fuentes», «da información básica cuando sales de lo estándar» y «la versión gratuita se queda corta frente a la de pago». Varias voces resaltan que la calidad depende del prompt y que siempre hace falta «verificar y pulir» la respuesta. Dos personas dicen no emplear IA profesionalmente; otra la usa solo «para echar unas risas».

## 4.2.4 Percepción de sesgos

La evidencia de sesgo aparece con claridad en la IA generativa de imágenes: «pides personas exitosas y te muestra hombres caucásicos en un entorno corporativo». También se denuncia que los asistentes de voz vengan «por defecto con voces de mujer» y que reconozcan mejor un timbre masculino blanco. Se observa censura preventiva («pides una imagen de una pareja lesbiana y la bloquea») y sesgos ideológicos en las respuestas textuales sobre política o historia. El grupo concluye que la imparcialidad es imposible porque «las bases de datos arrastran racismo, machismo, etc.».

## 4.2.5 Decisiones injustas y ética

Emerge un debate sobre si la IA “decide” o solo ejecuta instrucciones humanas. Para algunos, culpar a la IA es «un error de base: la IA es una herramienta». Otros advierten de la pérdida de ética al delegar: «un coche autónomo deberá elegir a quién atropella» y «un hospital podría negar un tratamiento carísimo a un niño si solo evalúa rentabilidad». El consenso final es que la ausencia de valores humanos en la automatización puede derivar en resultados socialmente injustos.

## 4.2.6 Impacto social, desigualdad y dependencia

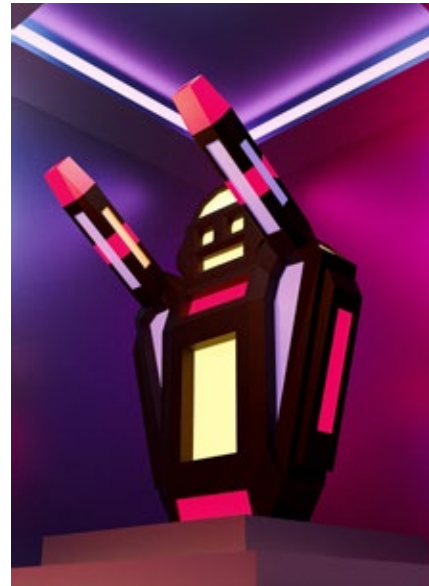
Se identifican dos grandes riesgos. El primero es la desigualdad: «Hay zonas sin internet y la brecha educativa crecerá»; «la representación global dependerá de quién alimenta los modelos». El segundo es la dependencia y el deterioro del pensamiento crítico: «mis alumnos copian y pegan todo»; «cada vez dictamos preguntas al móvil sin reflexionar». A ello se suma la preocupación por la expansión de discursos de odio, la explotación de datos personales y la consolidación de «un capitalismo de la vigilancia».

## 4.2.7 Regulación y responsabilidad

Todos aceptan la necesidad de normas, pero dudan de su alcance. Se teme que «Europa se obsesiona con prohibir y frena la innovación», mientras que EEUU innova y China avanza con su propia IA. Algunos piden un marco global; otros creen imposible regular inteligencias controladas por gobiernos autoritarios. La responsabilidad, coinciden, es compartida: diseñadores, empresas, gobiernos y usuarios «por acción u omisión».

## 4.2.8 Utilidad percibida y temores

La herramienta se considera «útil, con un gran asterisco». En ámbitos artísticos provoca rechazo: «no quiero ver la IA sustituyendo actores o voces». Surgen miedos a la suplantación de identidad –deepfakes–, a la invasión de la privacidad y a que «se implante una visión del mundo sin que lo notemos».



## 4.2.9 Propuestas de mitigación

La receta dominante es educativa: «formar en pensamiento crítico», «incluir IA en los planes de estudio» y «mostrar ejemplos concretos de sesgo». Se sugiere un repositorio público donde la ciudadanía denuncie sesgos y una línea de investigación sobre ética que acompañe cualquier despliegue tecnológico.



### 4.2.10 Conclusiones

El grupo adulto percibe la inteligencia artificial como una herramienta poderosa pero inseparable de los datos y los intereses que la alimentan. Valoran su capacidad para ahorrar tiempo, estructurar información y ampliar la creatividad, pero alertan de la homogeneización del estilo, las alucinaciones y la dependencia cognitiva. Reconocen sesgos de género, raza e ideología tanto en textos como en imágenes y voces, y temen que esos prejuicios aumenten la desigualdad, erosionen la ética en la toma de decisiones y profundicen el modelo capitalista de extracción de datos.

El debate sobre la regulación muestra una tensión entre la protección de derechos y el riesgo de frenar la innovación. La responsabilidad se concibe como un entramado que incluye a diseñadores, empresas, legisladores y usuarios. La vía de salida propuesta combina educación crítica, transparencia de datos y espacios para denunciar y corregir sesgos.

# Resultados

## Resultados

### 4.3 Grupo de edad: 56+

#### 4.3.1 Concepción de la IA

El grupo describe la IA como «una herramienta que procesa como la mente humana» y «un invento nuevo y desconocido que puede escaparse de control». Se reconoce su capacidad de aprendizaje rápido, pero se advierte que «ya no controlamos los datos» y que el poder reside en quienes manejan los servidores. La IA se percibe útil pero potencialmente peligrosa si carece de ética.

### 4.3.2 Fuentes de información y confianza

Para informarse recurren a periódicos de distintas tendencias, revistas especializadas y televisión. Coinciden en que «todo está sesgado, hay que leer varios medios». La búsqueda de datos suele ser «inconsciente» –aparece integrada en Google o en el GPS–, lo que incrementa la sensación de falta de control.



### 4.3.3 Usos cotidianos, ventajas y límites

El empleo más habitual es Google Maps, asistentes de voz, recomendaciones de Netflix y consultas rápidas en buscadores. Se valora la «inmediatez» y la «rapidez» frente a las antiguas enciclopedias. No obstante, se señala que «no aprendes nada si solo copias» y que «todo suena igual» cuando se abusa de la IA. En el ámbito sanitario se cita como «muy útil» para agilizar gestiones, mientras que en tareas creativas provoca desconfianza.

### 4.3.4. Uso cotidiano y herramientas

Se identifican sesgos en los resultados de búsqueda («lo que lees en el periódico no coincide con lo que sale en internet») y en sistemas de selección laboral que «filtran currículos contra mujeres o físicos poco saludables». También mencionan la censura de contenidos LGBT y la diferencia de resultados según país o historial. Para el grupo, la imparcialidad es imposible porque «depende de quién mete las premisas».

### 4.3.5 Decisiones injustas y ética

Surge preocupación por la automatización de decisiones críticas: «peligrosísimo que la IA agilice juicios» o niegue tratamientos médicos caros por pura rentabilidad. Se debate si la IA decide o solo ejecuta; el consenso final es que sin valores humanos incorporados «puede generar resultados socialmente injustos».

### 4.3.6 Impacto social percibido

Dos grandes temores concentran la discusión. El primero es la brecha digital: zonas sin internet y colectivos sin competencias digitales quedarían rezagados. El segundo es la dependencia cognitiva: «adolescentes vulnerables a estereotipos de belleza» y adultos que delegan la memoria y la ortografía. Se añade el riesgo de suplantación de identidad mediante voz o imagen y la extensión de un «capitalismo de vigilancia».



### 4.3.7. Regulación, responsabilidad y mitigación

La regulación se considera «vital». Se propone un «comité de sabios imparcial» y se desconfía de los gobiernos por su volatilidad ideológica. También se teme que normas estrictas en Europa frenen la innovación frente a EE. UU. y China. La responsabilidad se define como «compartida», aunque se exige supervisión de entes neutrales y renovables. Como prevención reclaman alfabetización digital desde la infancia, formación docente específica y un repositorio público donde la ciudadanía denuncie sesgos reales.



### 4.3.8 Conclusiones

Las personas mayores perciben la inteligencia artificial como una herramienta poderosa cuya utilidad práctica convive con riesgos sustanciales de manipulación y pérdida de autonomía. Reconocen sus ventajas –rapidez, comodidad y apoyo en ámbitos científicos o sanitarios–, pero insisten en que la IA heredará y amplificará los prejuicios humanos mientras los datos y los algoritmos permanezcan en manos de actores poco transparentes.

La conversación subraya dos ejes de preocupación: por un lado, la ampliación de la brecha digital y la desigualdad global; por otro, la erosión del pensamiento crítico y la dependencia emocional que podría derivarse de confiar decisiones y recuerdos a sistemas automatizados. Ante ello, reclaman una regulación firme pero equilibrada, guiada por comités éticos imparciales, y una educación crítica que empiece en la escuela y se extienda a toda la ciudadanía.

# Global de grupos focales

grupos focales

## 5.1 Concepción de la IA



Los tres grupos, coinciden en que la inteligencia artificial no debe entenderse como una mente autónoma, sino como una herramienta que opera mediante modelos estadísticos. Sin embargo, la forma de conceptualizarla presenta matices generacionales.

El grupo de 18 a 30 años la describe como «un Google bastante sofisticado» o «un término de marketing» que infla su capacidad. Esta visión refleja una actitud crítica hacia el discurso tecnocrático y una familiaridad práctica con su funcionamiento. En cambio, el grupo adulto (31-55 años) la concibe como «software que aprende» o como «un sistema que replica los datos que lo alimentan», incorporando una visión más funcional pero también más reflexiva. Las personas mayores (56+ años) la definen como «una herramienta que procesa como la mente humana», aunque advierten que «puede escaparse de control» si no se supervisa adecuadamente. En todos los casos, se reconocen tanto el potencial como el riesgo de un uso acrítico o desinformado.

## 5.2 Fuentes de información y confianza

Existe consenso entre los tres grupos, en que la IA no debe considerarse una fuente autónoma de conocimiento. Tanto jóvenes como adultos mencionan que es necesario «contrastar siempre con otras fuentes», incluyendo medios tradicionales, artículos académicos o experiencias propias.

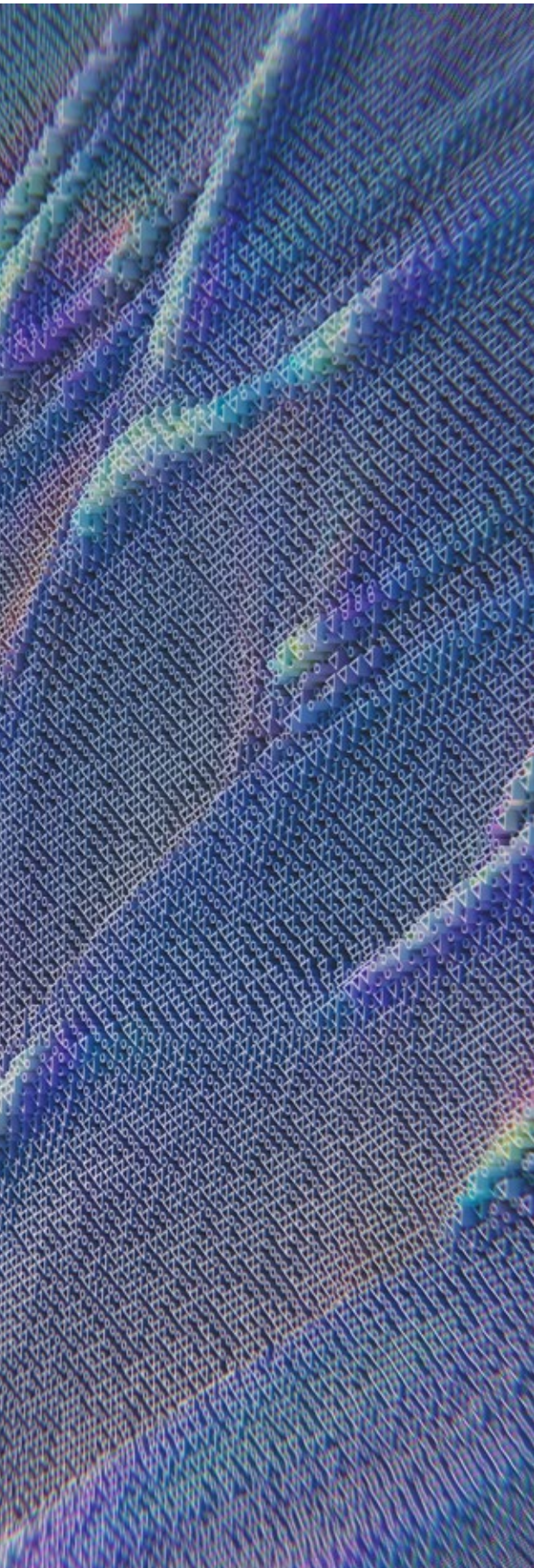


Las personas mayores, por su parte, enfatizan la necesidad de consultar medios con diferentes orientaciones ideológicas, aludiendo a la parcialidad de muchas plataformas digitales. En general, se percibe una creciente desconfianza hacia la información generada por IA si no se acompaña de criterios humanos de veracidad y contexto.

## 5.3 Usos cotidianos, ventajas y límites

El uso de herramientas basadas en IA varía según el grupo. Los jóvenes (18-30) emplean ChatGPT, Copilot, DeepL y generadores de imagen de manera intensiva, aunque algunos también expresan rechazo por motivos éticos o creativos. Los adultos (31-55) integran estas herramientas en contextos profesionales, con un uso avanzado que incluye prompts complejos, evaluación crítica de resultados y ajustes sobre la marcha. Por su parte, el grupo de mayores (56+) se concentra en usos funcionales como navegadores, asistentes de voz o recomendaciones automatizadas en plataformas de contenido.

Entre las ventajas señaladas se incluyen la inmediatez, la eficiencia y el apoyo a la organización de ideas. Sin embargo, también se destacan límites importantes: «todo suena igual», «inventa fuentes», «la versión gratuita se queda corta» y «no aprendes nada si solo copias». Se insiste en que la calidad del resultado depende del criterio del usuario y que sin pensamiento crítico, el uso de IA tiende a la dependencia y la homogeneización.



## 5.4 Percepción de sesgos

Todos los grupos reconocen que la IA reproduce prejuicios preexistentes en los datos y en quienes diseñan los modelos.

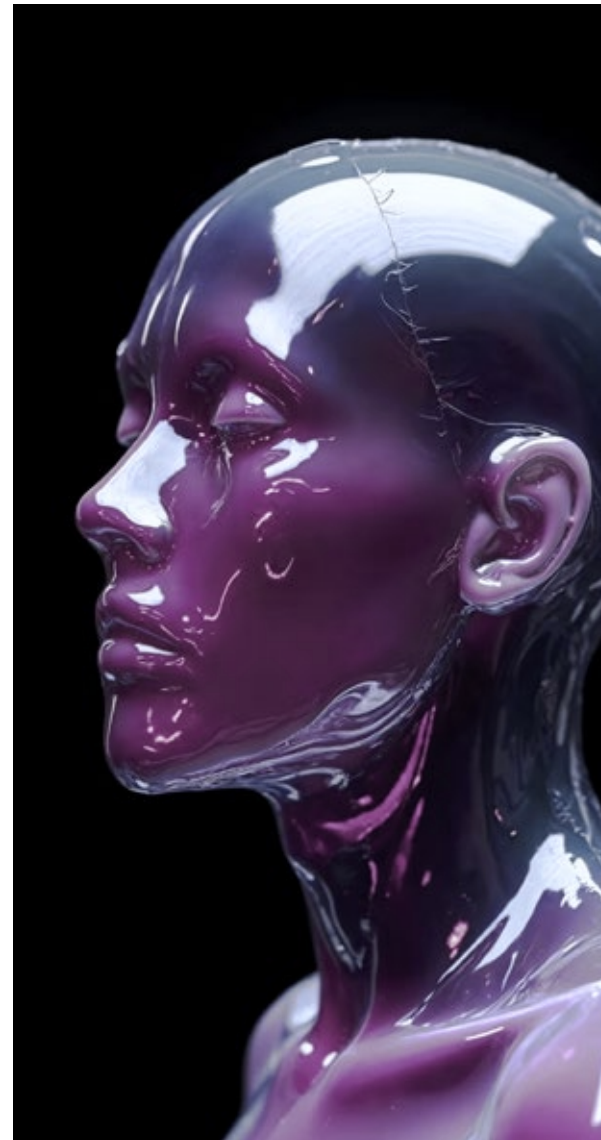
El grupo joven denuncia, entre otros, sesgos raciales en el reconocimiento facial y en la representación visual (cabellos afro, rasgos no caucásicos), así como contenido misógino derivado del entrenamiento con foros problemáticos. El grupo adulto añade ejemplos de sesgos de género en asistentes de voz (predeterminados con voz femenina), imágenes estereotipadas de éxito (hombres blancos en entornos corporativos) y censura de contenidos LGBT. Las personas mayores, por su parte, mencionan filtros en procesos de selección laboral que penalizan a mujeres o personas con determinados físicos, además de diferencias en los resultados de búsqueda según historial, país o servidor.

En todos los casos, la imparcialidad algorítmica se considera una ilusión si no se evalúan críticamente los datos y la lógica que sustentan cada sistema.

## 5.5 Decisiones injustas y ética

Los tres grupos coinciden en que la automatización de decisiones sin intervención ética puede derivar en consecuencias socialmente injustas. Los jóvenes señalan la falta de creatividad y la dependencia de los chatbots como amenazas a la autonomía intelectual. Los adultos problematizan el uso de IA en contextos de salud o justicia, donde «un coche autónomo podría decidir a quién atropella» o un sistema hospitalario podría rechazar tratamientos por rentabilidad. Las personas mayores manifiestan preocupación por la automatización de juicios o por la posibilidad de que «la IA decida sin valores humanos».

La discusión sobre si la IA “decide” o “ejecuta” remite, en todos los casos, a la necesidad de mantener el control humano como garante de la justicia social.



## 5.6 Impacto social percibido

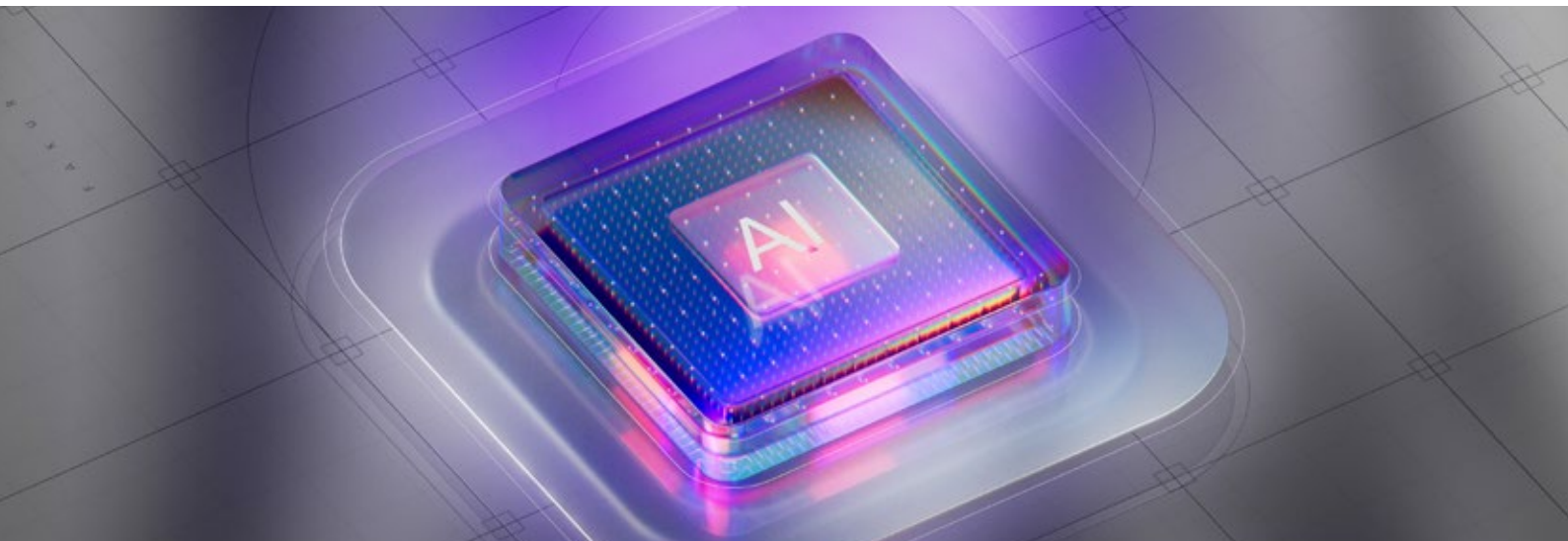
Las participantes, muestran preocupaciones específicas respecto al impacto de la IA sobre la sociedad. Los jóvenes enfatizan la pereza cognitiva, la delegación excesiva de tareas creativas y la dependencia emocional de los asistentes virtuales. Los adultos subrayan la expansión del capitalismo de vigilancia, el debilitamiento del pensamiento crítico y el riesgo de aumentar las desigualdades estructurales. Las personas mayores expresan su inquietud por la brecha digital, la polarización informativa, la suplantación de identidad y la pérdida de habilidades básicas como la memoria o la escritura.

En conjunto, se teme que el uso indiscriminado de IA sustituya funciones humanas clave, debilite la autonomía ciudadana y consolide sistemas opacos de control algorítmico.

## 5.7 Regulación, responsabilidad y mitigación

Hay consenso en que la IA debe regularse de manera urgente y con perspectiva ética. Los tres grupos proponen la existencia de comités independientes: los jóvenes hablan de “comités de datos”, los adultos de “repositorios públicos de sesgos” acompañados por marcos normativos equilibrados, y las personas mayores proponen un “comité de sabios imparcial” con supervisión renovable.

En todos los casos, la responsabilidad se define como compartida: diseñadores, empresas, gobiernos y usuarios deben asumir sus funciones de manera activa. También se subraya que las normas deben ser globales para evitar desequilibrios entre continentes y proteger los derechos de la ciudadanía en entornos cada vez más digitalizados.



## 5.8 Estrategias de mitigación

La estrategia más mencionada para prevenir los riesgos de sesgo y desinformación es la educación crítica.

El grupo joven defiende la alfabetización digital desde la infancia, mientras que los adultos piden integrarla en la formación académica y profesional. El grupo de mayores propone formación intergeneracional y específica para docentes. Todos valoran la creación de repositorios públicos de sesgos como herramienta participativa, así como ejemplos reales que permitan entrenar la mirada crítica sobre los algoritmos.

# 06

## Conclusiones Generales

Los grupos focales del proyecto IA-BIAS revelan que la ciudadanía, independientemente de su edad, tiene una visión cada vez más crítica, informada y participativa sobre la inteligencia artificial. La IA es vista como una herramienta útil pero potencialmente peligrosa si no se controla, regula ni educa en torno a su uso.

En suma, los grupos focales han permitido validar la urgencia de una intervención educativa y normativa orientada a prevenir los riesgos de una IA sin ética y potenciar su dimensión social como bien común. El enfoque intergeneracional y participativo del proyecto IA-BIAS se ve así reforzado como una vía legítima y necesaria para democratizar la tecnología.

# Referencias

## Referencias

Albarrán, I., Molina, J. M. y Gijón, C. (2020). Perception of Artificial Intelligence in Spain. International Conference Telecommunications Society (ITS), 1-18.

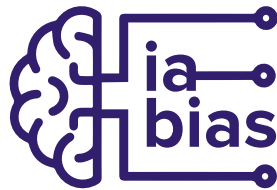
Arcila et al, 2021. Disponible en:

[https://www.ocausal.es/wp-content/uploads/2021/01/A1\\_Conocimiento-pu%CC%81blico-y-comprensio%CC%81n-de-la-Ciencia-de-Datos.pdf](https://www.ocausal.es/wp-content/uploads/2021/01/A1_Conocimiento-pu%CC%81blico-y-comprensio%CC%81n-de-la-Ciencia-de-Datos.pdf)

Belloso Martín, N. (2022). La problemática de los sesgos algorítmicos (con especial referencia a los de género). ¿Hacia un derecho a la protección contra los sesgos?. En J. Garrido Martín y R. Valdavia Jiménez (Coords.), Inteligencia artificial y filosofía del derecho (pp. 45-79). Laborum Ediciones.

Broussard, M. (2018). Artificial unintelligence: How computers misunderstand the world. mit Press.

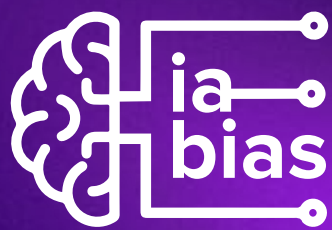
Holder, C., Khurana, V. y Watts, M. (2018). Artificial Intelligence: Public Perception, Attitude and Trust. Bristow, 1– 23.



**Detección participativa de los sesgos de la Inteligencia Artificial**



Proyecto financiado por la Fundación Española para la Ciencia y la Tecnología (FECYT) en la Convocatoria de ayudas para el fomento de la cultura científica, tecnológica y de la innovación 2023-2024 [FCT-23-19454].



2025