

# Módulo 4. Aspectos éticos



☰ 1. Sesgos y alineación de la IA

☰ Referencias

# 1. Sesgos y alineación de la IA

---

¿Qué crees que pasaría si le pedimos a una IA que genere la imagen de una mujer exitosa? ¿Qué hay de su ropa, sus expresiones faciales y su apariencia en general?

**Figura 1. Ejemplo de sesgo en la generación de imágenes por IA: representación de una mujer exitosa**

PROMPT

Una **mu**jer **exitosa**  
Tómate un momento para imaginarlo...



RESULT



Fuente: elaboración propia.

Veamos algunas de las características de la imagen que acabamos de observar.

- La mujer es **rubia, delgada**, parece una **actriz** o una **modelo**.

- Viste de **traje**.
- Parece ser una **superheroína**.
- Está rodeada de **títulos**.

Nos lleva a preguntarnos: ¿realmente las mujeres exitosas son así? ¿Cómo y quién define que el éxito se plasme de esta forma?

Ahora, ¿qué pasa si le pedimos a una IA que haga una imagen de Mick Jagger y otra de Thalía? ¿A quién crees que hará mejor?

**Figura 2. Imagen generada por IA de Mick Jagger**

**PROMPT**

**Mick Jagger bailando y cantando**  
Tómate un segundo para visualizarlo.



**RESULT**



Fuente: elaboración propia.

**Figura 3. Imagen generada por IA de Thalía**

PROMPT

**Thalía**  
Tómate un momento **bailando** y para **cantando** imaginarlo...



RESULT



Fuente: elaboración propia.

---

¡Ups! La IA no pudo reconocer a Thalía y no la representó correctamente.

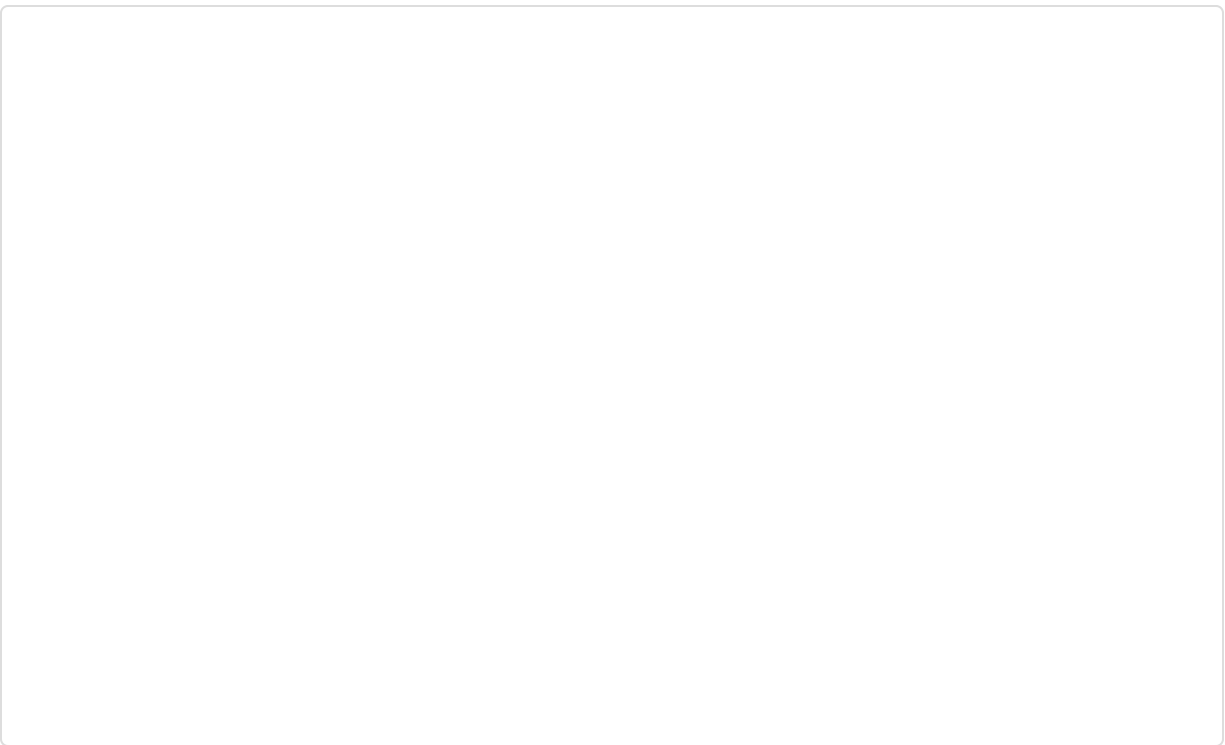
**Estos ejemplos nos dejan entrever que los estereotipos y los sesgos están presentes, como en los humanos, también en la Inteligencia Artificial.**

¿Por qué? Porque las inteligencias artificiales aprenden de las personas. Y, como parte de ese aprendizaje, heredan prejuicios y sesgos que se trasladan a la herramienta. ¿El resultado? La IA puede generar respuestas que, sin quererlo, sean discriminatorias, injustas o hasta ofensivas. Las máquinas no solo aprenden de lo que ven, sino también de lo que no ven. Si les faltan datos diversos o equilibrados, seguirán reproduciendo las mismas limitaciones.

La inteligencia artificial ha sido entrenada con datos de internet, por lo que es importante prestar atención a los posibles sesgos que pueden reproducirse a partir de esa información.

Por ejemplo, imagina que entrenamos un algoritmo para que genere imágenes de médicos, pero la mayoría de las fotos que le damos son de hombres. ¿Qué crees que pasará?

Cuando la IA cree sus propias imágenes de médicos, es muy probable que la mayoría sean hombres, aunque en la realidad haya muchas médicas mujeres también. **Si los datos que recibe representan solo un tipo específico de persona, la IA replicará esa visión limitada, perpetuando prejuicios sin darnos cuenta.**



### **Pero, ¿qué son exactamente los sesgos?**

**Los sesgos en la IA ocurren cuando los algoritmos producen resultados parciales o discriminatorios debido a los datos con los que fueron entrenados o a errores en su diseño.** Como las máquinas aprenden a partir de datos

recopilados del mundo real, pueden heredar los mismos prejuicios y desigualdades presentes en la sociedad.

Si quieres continuar explorando y profundizando sobre los sesgos en la tecnología te recomendamos explorar estos recursos:

02:27



**Fuente:** MIFF [MIF], (s.f.). *Coded Bias | Trailer*  
[video]. YouTube.

<https://www.youtube.com/watch?v=jZl55PsfZJQ>

13:19



**Fuente: Bloomberg Live [Bloomberg Live],**  
(s.f.). *The Coded Gaze: Bias in Artificial Intelligence | Equality Summit* [video].  
YouTube. <https://www.youtube.com/watch?v=eRUEVYndh9c>

## ¿Cómo evitar la reproducción de sesgos?

Existen pautas, reglas y procedimientos a seguir para ayudar a las máquinas a aprender a ser más justas, equitativas y plurales. Este proceso se llama alineación. También es controversial porque no somos todos iguales, no pensamos de la misma manera y es difícil establecer una ética de la IA.

### Ética y responsabilidad: buenas prácticas

Para establecer nuevos estándares digitales de representación, la marca Dove junto con expertos en IA, crearon el Manual de Indicaciones para la Belleza Real, con orientaciones fáciles sobre cómo crear imágenes más representativas a nivel general con herramientas de IA generativa.



**Fuente: Dove,** (s.f.). *Manual de la belleza real.*  
<https://www.dove.com/ar/stories/campaigns/keep-beauty-real.html>

Esta campaña pone de manifiesto los sesgos que tienen las inteligencias artificiales en la representación de las mujeres y

la marca hizo lo siguiente para mostrar la belleza real y no la belleza hegemónica:

01:55



**Fuente: Dove España [Dove España], (s.f.).** *El código | Un corto de Dove | Dove Proyecto Autoestima* [video]. YouTube.  
<https://www.youtube.com/watch?v=6ZERsGrWUHk>

**La IA es una herramienta muy poderosa, con infinitas posibilidades para la creación de contenidos. Por eso, en un contexto de rápida difusión de contenido creado por múltiples usuarios, es más importante que nunca recordar que debemos ejercer un uso responsable de la IA.**

**CONTINUAR**

# Referencias

---

**Bloomberg Live [Bloomberg Live]**, (s.f.). *The Coded Gaze: Bias in Artificial Intelligence | Equality Summit* [video]. YouTube. <https://www.youtube.com/watch?v=eRUEVYndh9c>

**Dove**, (s.f.). *Manual de la belleza real*. <https://www.dove.com/ar/stories/campaigns/keep-beauty-real.html>

**MIFF [MIF]**, (s.f.). *Coded Bias | Trailer* [video]. YouTube. <https://www.youtube.com/watch?v=jZl55PsfZJQ>

CONTINUAR