



Programa BridgeAI

Módulo 4: Ética, diversidad y derechos humanos en la IA

Ejercicio interactivo 2:

Diversidad en los conjuntos de datos, representación y cómo se producen los sesgos

November 2025

By: Paradisskolan 4-9, Trollhättan, Sweden



**Co-funded by
the European Union**

Financiado por la Unión Europea. Las opiniones y puntos de vista expresados son únicamente los del autor o autores y no reflejan necesariamente los de la Unión Europea ni los de la Agencia Ejecutiva Europea de Educación y Cultura (EACEA). Ni la Unión Europea ni la EACEA pueden ser consideradas responsables de ellos.

Ejercicio interactivo 2 – Sesgos y representación en la IA

Tema: Diversidad en los conjuntos de datos, representación y cómo se producen los sesgos

Duración: 25-30 minutos

Tamaño del grupo: Grupos pequeños (3-4 estudiantes)

Objetivo

El alumnado comprenderá que **una representación insuficiente en los datos de entrenamiento de la IA genera sesgos**, y verá que la equidad depende de la inclusión de grupos diversos.

Materiales necesarios

- Tarjetas de situaciones (herramientas de IA con falta de representación)
- Minificha: «Detecta la brecha de representación»
- Rotuladores y papel O notas digitales
- Temporizador

Resultado esperado

Cada grupo elabora:

- Una **minificha** cumplimentada que muestra la brecha de representación
- Una explicación clara del **sesgo** resultante
- Una idea realista para solucionar el problema

El alumnado **comprende la relación:**

Falta de representación → IA injusta (sesgo)

Mayor representación → IA más justa

Instrucciones

Paso 1: Preparación

El docente presenta el ejercicio y explica:

«La IA aprende a partir de ejemplos. Si los ejemplos solo muestran determinados tipos de personas o situaciones, la IA tendrá dificultades con los demás. Esto se denomina **representación insuficiente**, y provoca **sesgo** — resultados injustos».

El docente ofrece un ejemplo:

Una IA de voz entrenada principalmente con voces de adultos en inglés no comprende los acentos somalíes, sirios o suecos de adolescentes.

Explica el objetivo de la actividad:

«Su tarea consiste en detectar dónde falta representación en cada situación y cómo esa carencia genera sesgos».

Paso 2: Actividad

El alumnado trabaja en grupos y:

Cada grupo recibe **una tarjeta de situación** (las tarjetas aparecen más abajo).

Cumplimentan una **minificha** respondiendo a las siguientes preguntas:

Minificha: «Detecta la brecha de representación»

1. ¿Quién está bien representado en los datos de entrenamiento de la IA?
2. ¿Quién falta o está infrarrepresentado?
3. ¿Qué tipo de sesgo genera esto?
4. ¿Quién podría recibir un trato injusto?
5. ¿Una forma de solucionarlo (añadir representación)?

Después, los grupos rotan y presentan sus conclusiones en grupos más pequeños o las exponen brevemente ante toda la clase.

Paso 3: Puesta en común

Debate con toda la clase dirigido por el docente:

1. **¿Qué grupos aparecían ausentes con mayor frecuencia?**
(niños y niñas, mujeres, personas que no hablan inglés, personas con tonos de piel más oscuros, etc.)
2. **¿Cómo generó sesgos la falta de representación en cada situación?**
3. **¿Por qué es importante incluir en los conjuntos de datos a personas de Suecia, Somalia, Siria, Europa, etc.?**
→ Para que la IA sea justa para su propia aula y comunidad.
4. **¿Cómo pueden los equipos de IA reducir los sesgos?**
→ Incluir datos más diversos, realizar pruebas con diferentes grupos y solicitar retroalimentación.
5. **¿Qué ocurre en la sociedad cuando la IA es injusta?**
→ Las personas pierden la confianza, no pueden acceder a servicios o son juzgadas erróneamente.

Tarjetas de situaciones (para el ejercicio 2)

A continuación se presentan **5 tarjetas de situaciones** centradas específicamente en la representación y los sesgos:

Tarjeta A – El problema del tono de piel de los emojis

Una IA sugiere emojis a los usuarios.

Fue entrenada principalmente con imágenes de personas de piel clara.

Resultado: La IA rara vez sugiere emojis de piel oscura.

→ ¿Quién está infrarrepresentado? ¿A quién afecta?

Tarjeta B – La cámara «inteligente» del aula

Una cámara cuenta cuántos estudiantes levantan la mano.

Funciona bien con pieles claras, pero tiene dificultades para detectar tonos de piel más oscuros.

→ ¿Quién queda excluido? ¿Qué sesgo se produce?

Tarjeta C – El asistente de voz

El asistente de voz fue entrenado con voces de adultos angloparlantes.

Interpreta incorrectamente a adolescentes y a personas con acento sueco, somalí o árabe.

→ ¿Quién está infrarrepresentado? ¿Qué problema provoca?



Tarjeta D – La aplicación deportiva

Una IA de actividad física fue entrenada principalmente con datos de hombres adultos. Ofrece consejos de salud incorrectos a chicas y adolescentes más jóvenes.

→ ¿A quién comprende mejor la IA? ¿A quién interpreta incorrectamente?

Tarjeta E – La herramienta de traducción

Un traductor con IA aprendió principalmente a partir de textos en alemán, español y francés.

Comete errores importantes con frases en somalí y árabe.

→ ¿Qué lenguas están representadas? ¿Cuáles faltan?



**Co-funded by
the European Union**

Financiado por la Unión Europea. Las opiniones y puntos de vista expresados son únicamente los del autor o autores y no reflejan necesariamente los de la Unión Europea ni los de la Agencia Ejecutiva Europea de Educación y Cultura (EACEA). Ni la Unión Europea ni la EACEA pueden ser consideradas responsables de ellos.