



Programa BridgeAI

Modulul 4: Etică, diversitate și drepturile omului în IA

Exercițiul interactiv 2:

Diversitatea în seturile de date, reprezentarea și modul în care apare prejudecata

Noiembrie 2025

De: Paradisskolan 4-9, Trollhättan, Suedia



**Co-funded by
the European Union**

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.



Exercițiul interactiv 2 – Prejudecată și reprezentare în IA

Temă: Diversitatea în seturile de date, reprezentarea și modul în care apare prejudecata

Durată: 25–30 de minute

Mărimea grupului: grupuri mici (3–4 elevi)

Obiectiv

Elevii vor înțelege că reprezentarea insuficientă în datele de antrenare ale IA conduce la prejudecăți și vor observa că echitatea depinde de includerea unor grupuri diverse.

Materiale necesare

- Scenario Cards (AI tools with missing representation)
- Mini-fișă de lucru: „Identifică lipsa de reprezentare”
- Markers & paper OR digital notes
- Timer

Rezultat așteptat

Fiecare grup realizează:

- A completed **mini worksheet** showing the representation gap
- A clear explanation of the **bias** that results
- One realistic idea to fix the problem

Elevii observă legătura:

Missing representation → unfair AI (bias)

More representation → fairer AI



Instrucțiuni

Pasul 1: Pregătire

Profesorul introduce exercițiul și explică:

„IA învață din exemple. Dacă exemplele prezintă doar anumite tipuri de persoane sau situații, IA va avea dificultăți cu celelalte.

Aceasta se numește reprezentare insuficientă și provoacă prejudecăți - rezultate inechitabile.”

Profesorul oferă un exemplu:

O IA vocală antrenată în principal pe vocile adulților vorbitori de engleză nu înțelege accentele somaleze, siriene sau suedeze ale adolescenților.

Profesorul explică obiectivul activității:

„Sarcina voastră este să identificați unde lipsește reprezentarea în fiecare scenariu și cum generează acest lucru prejudecăți.”

Pasul 2: Activitate

Elevii lucrează în grupuri și:

Fiecare grup primește un card cu un scenariu (cardurile sunt prezentate mai jos).

Completează o mini-fișă de lucru răspunzând:

Mini-fișă de lucru: „Identifică lipsa de reprezentare”

1. **Who is well-represented in the AI training data?**
2. **Who is missing or underrepresented?**
3. **What kind of bias does this create?**
4. **Who might be treated unfairly?**
5. **One way to fix it (add representation)?**

Apoi, grupurile se rotesc și prezintă în grupuri mai mici sau își prezintă pe scurt concluziile în fața întregii clase.



Pasul 3: Discuție finală

Discuție cu întreaga clasă, coordonată de profesor:

1. **Which groups were missing most often?**
(copii, femei, persoane care nu vorbesc engleza, persoane cu tonuri mai închise ale pielii etc.)
2. **How did missing representation lead to bias in each scenario?**
3. **Why is it important to include people from Sweden, Somalia, Syria, Europe, etc. in datasets?**
→ To make AI fair for your own classroom and community.
4. **How can AI teams reduce bias?**
→ Include more diverse data, test with different groups, ask for feedback.
5. **What happens in society when AI is unfair?**
→ People lose trust, can't access services, get wrongly judged.

Carduri cu scenarii (pentru exercițiul 2)

Mai jos sunt 5 carduri cu scenarii axate în mod specific pe reprezentare și prejudecăți:

Cardul de scenariu A – Problema tonului pielii în emoji

O IA sugerează emoji-uri utilizatorilor.

A fost antrenată în principal pe imagini cu persoane cu pielea deschisă.

Result: The AI rarely suggests dark-skinned emojis.

→ Who is underrepresented? Who is affected?

Cardul de scenariu B – Camera „inteligentă” din clasă

A camera counts how many students raise their hands.

It works well on lighter skin, but struggles to detect darker skin tones.

→ Who is left out? What bias happens?

Cardul de scenariu C – Asistentul vocal

The voice assistant was trained on adult English-speaking voices.

It mishears teenagers, Swedish accents, Somali accents, and Arabic accents.

→ Who is underrepresented? What problem does it cause?

Cardul de scenariu D – Aplicația sportivă



A fitness AI was trained mostly on data from adult men.
It gives wrong health advice for girls and younger teens.
→ Who does the AI understand best? Who does it misunderstand?

Cardul de scenariu E – Instrumentul de traducere

An AI translator learned mostly from German, Spanish, and French texts.
It makes big mistakes with Somali and Arabic sentences.
→ Which languages are represented? Which are missing?