

# Exploring the bias of generative AI



I.T.T. Panetti-Pitagora – Bari (BA)  
Docente referente: Giuseppe Mastrandrea





Il progetto si è svolto con una classe quarta nell'ambito delle attività curricolari di TPSIT, informatica ed educazione civica, con l'obiettivo di far riflettere gli studenti sulle implicazioni etiche dell'AI, partendo da un'attività pratica.

**Fase 1 - Formazione e pianificazione** (3 settimane). Il docente referente ha introdotto i concetti basilari di ML, modelli generativi, bias algoritmico, stereotipi sociali e rappresentazione. Gli studenti hanno studiato casi di discriminazione algoritmica e appreso le basi della classificazione dei toni di pelle (scala di Fitzpatrick) e delle categorie etniche. È stata poi definita la metodologia di lavoro: selezione delle 15 professioni da analizzare (7 ad alto reddito, 8 a basso reddito) e dei 5 assi di valutazione.

**Fase 2 - Generazione** (4 mesi). La classe è stata divisa in 5 gruppi, ciascuno assegnato a un modello AI (ChatGPT/DALL-E, Gemini, Leonardo AI, Canva, Bing). Ogni gruppo ha generato immagini per tutte le 15 professioni utilizzando prompt standardizzati in lingua inglese per evitare influenze linguistiche legate al genere. Sono state prodotte circa 40 immagini per ogni combinazione professione-modello, per un totale di oltre 3000 immagini.

**Fase 3 - Etichettatura** (4 mesi). Ogni studente ha analizzato e classificato le immagini generate, annotando su fogli di calcolo condivisi i valori per ciascun asse: sesso (maschile, femminile), etnia (occidentale, orientale, estremo orientale, afrodiscendente, africano), tonalità della pelle (scala di Fitzpatrick) età (bambino/a, ragazzo/a, adulto/a, anziano/a), espressione facciale (felice, neutro, triste). Per garantire affidabilità, un campione di immagini è stato etichettato da più studenti per verificare la concordanza delle valutazioni.

**Fase 4 - Analisi dei dati** (2 settimane). Gli studenti hanno aggregato i dati, calcolato le percentuali per ogni categoria e prodotto visualizzazioni grafiche. È emerso come i modelli riproducano stereotipi di genere (CEO e giudici prevalentemente maschili, casalinghi prevalentemente femminili), etnici (professioni ad alto reddito associate a etnia occidentale) e di età.

**Fase 5 - Divulgazione.** Il lavoro è stato pubblicato sul sito della scuola e presentato alla DevFest 2025: gli studenti hanno esposto metodologia, risultati e riflessioni davanti a una platea di professionisti.

**Strumenti utilizzati:** Modelli AI, fogli di calcolo, Python, PHP/HTML, Git/GitHub.

L'esperienza ha sviluppato competenze trasversali: pensiero critico, metodologia della ricerca empirica, lavoro collaborativo, analisi statistica, public speaking e consapevolezza etica. Inoltre come hard skill: uso di strumenti di generative AI via API, sviluppo web e deploy di un prodotto web.