



BIAIXOS DE GÈNERE EN ELS MODELS DE LLENGUATGE EXTERNS

Nota didàctica simplificada

Objectius del recurs

Aquesta infografia té per objectiu sensibilitzar el personal bibliotecari, el personal docent i l'estudiantat sobre la presència de biaixos de gènere en els sistemes d'intel·ligència artificial (IA), en concret en els models de llenguatge externs (LLM).

Resultats d'aprenentatge previstos

- L'estudiantat comprèn què són els biaixos en els sistemes d'IA i com aquests deriven de les dades, de les decisions algorítmiques i dels contextos de desplegament.
- Identifica àmbits concrets en què la IA perpetua desigualtats de gènere.
- Analitza críticament el paper dels LLM en la reproducció d'estereotips de gènere.
- Reconeix la importància d'incorporar mesures preventives en les fases inicials del desenvolupament de sistemes d'IA.

Preguntes per a la reflexió

- Com afecta la manca de diversitat — especialment la de gènere — en els equips de desenvolupament d'IA als resultats i a les decisions que prenen aquests sistemes?
- Quins riscos socials comporta l'ús massiu de LLM esbiaixats en contextos com l'educació o la comunicació pública?
- En quina mesura la IA pot contribuir tant a prevenir com a facilitar formes de violència de gènere, especialment en entorns digitals?
- Quines estratègies reguladores, ètiques o tècniques serien necessàries per reduir els biaixos en IA abans del seu desplegament?
- Quines diferències es poden observar en la manera com els sistemes d'IA representen el gènere segons el context cultural o geogràfic?

Ús en la docència

Aquest material educatiu sobre el biaix de gènere es pot utilitzar com a activitat introductòria en seminaris de formació per a l'estudiantat o en activitats de formació contínua i desenvolupament professional del personal bibliotecari.

