



GENDER BIASES IN LARGE LANGUAGE MODELS

Simplified Teaching Note

Objectives of the resource

This infographic aims to raise awareness among library staff, teaching staff, and students about the presence of gender bias in artificial intelligence (AI) systems, specifically in large language models (LLM).

Intended Learning Outcomes

- Students understand what biases are in AI systems and how they arise from data, algorithmic decisions, and deployment contexts.
- They identify specific areas where AI perpetuates gender inequalities.
- They critically analyze the role of LLMs in reproducing gender stereotypes.
- They recognize the importance of incorporating preventative measures in the early stages of AI system development.

Questions for Reflection

- How does a lack of diversity (especially gender diversity) in AI development teams affect the results and decisions of AI systems?
- What social risks does the widespread use of biased LLMs pose in contexts such as education or public communication?
- To what extent can AI contribute to both preventing and facilitating forms of gender-based violence, especially in digital environments?
- What regulatory, ethical, or technical strategies would be necessary to reduce bias in AI before its deployment?
- What differences can be observed in how AI systems represent gender depending on the cultural or geographical context?

Usage in Teaching

This educational material on gender bias can be used as an introductory activity in training seminars for students or in continuing education and professional development activities for library staff.

