Risks of discrimination in Lesson Planning with AI

Nóra Födelmesi
Keith Quille

Abstract

Artificial Intelligence (AI) is increasingly used in education, assisting in tasks such as lesson and curriculum planning for primary school students. However, AI-generated content may unintentionally reinforce discrimination based on:

- Race,
- Ethnicity,
- Religion/belief,
- Gender,
- Sexual orientation,
- Disability,
- Age, and
- Socioeconomic status.

By analyzing AI-generated lesson plans with **AI detection tools**, **academic research**, and **personal analysis**, this research explores whether AI can reliably identify discrimination or if its inherent limitations restrict its ability to do so.

Introduction

Among the many forms of discrimination AI can perpetuate, this study focuses on disability-based discrimination, particularly concerning students with **physical disabilities**.

My focus is on two subjects that present the highest risk of discrimination for students with physical disabilities: **Physical Culture** (Physical Education) and **Technology and Design**. Physical Culture inherently involves movement, coordination, and physical participation, which can create barriers for students with mobility impairments if AI-generated plans do not account for adaptive activities. Similarly, Technology and Design often requires fine motor skills, tool use, and hands-on activities that may not be accessible to all students.

Literature review

AI is widely used in education, but research shows it can reinforce bias, particularly against students with physical disabilities. Studies highlight how AI-generated lesson plans in Physical Culture and Technology and Design often assume full physical ability, excluding students with mobility impairments.

Several tools aim to identify and mitigate AI bias:

- **Perspective API** (Google Jigsaw) – Detects toxic and biased language.
- **IBM AI Fairness 360** – Evaluates fairness in AI decisions.
- **Fairlearn** (Microsoft) – Helps assess and reduce AI bias.
- **SHAP** – Provides interpretability for AI decision-making.

While these tools can flag explicit bias, they struggle with implicit discrimination, such as ableist assumptions in lesson planning.

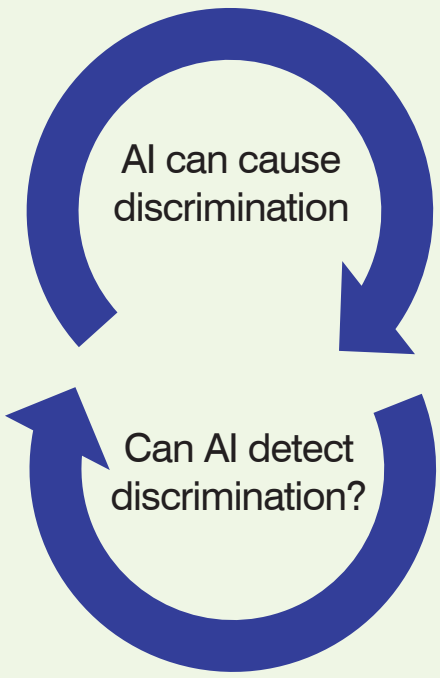
References: European Commission. (2022). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators.
Horváth, D. (2023). Artificial Intelligence in Education: Challenges and Opportunities. Budapest University of Technology and Economics.
National Center for Biotechnology Information. (2012). Bias in Artificial Intelligence: A Review of AI Discrimination in Decision-Making.
Hungarian Ministry of Education. (2020). National Curriculum Framework 2020.
Google Jigsaw. (n.d.). Perspective API: Documentation and Developer Guide.

Acknowledgements: The HCAIM (the Human-Centred AI Master's Programme) Project is Co-Financed by the Connecting Europe Facility of the European Union Under Grant № CEF-TC-2020-1 Digital Skills 2020-EU-IA-0068. This poster was created as part of the Blended Intensive Programme organized under the Erasmus + Programme of the European Union

Research Methodology

This study investigates whether AI can detect and mitigate discrimination in educational lesson planning or if it primarily reinforces bias, specifically against students with physical disabilities. The methodology consists of literature review, AI-generated lesson plan analysis, and qualitative evaluation.

AI discrimination detection tools



Conclusion

Overall, the research suggests that while AI can highlight certain biases, it is not yet capable of fully preventing discrimination in lesson planning. The results emphasize the **need for ongoing human oversight**. Future research should focus on developing AI systems that proactively generate accessible and adaptive lesson plans rather than requiring post-analysis bias correction. While I could not directly implement API-based bias detection, future research should integrate **automated tools to analyze discrimination** in AI-generated content. Further studies should also explore how AI can be trained to proactively generate inclusive lesson plans rather than relying on post-analysis bias detection.

MŰEGYETEM 1782

hcaim human centred artificial intelligence masters