Fairness in Automated Scoring

🛑 Stop and think

Unbiased scoring is critical for ethical, especially when outcomes affect people's opportunities. So in the automated system [1] , this links to the significance of fairness.

🚩 Clarify goals

The primary goal is to create an automated scoring system that is fair and does not exhibit bias against any demographic groups. And also the automated scoring algorithm should not prejudice towards questions, also other personal characteristics.

💡 Determine known and unknown facts

Known facts	Unknown facts
Automatic scoring algorithm may have biases. But developing without checking may be easier.	Potential new biases that could arise from changes in the system, how different groups might interact with changes, and how biases may affect long-term outcomes.

🌈 Develop options

Option 1: Introduce algorithms that detect and correct or report the biases.

Option 2: Make user report in the system and correct the biases in time when have reports.

Option 3: Ignore the biases.

↪️ Consider foreseeable results of the options

Option 1: May need less future work as team does not need to keep monitoring the biases, but the development process should have challenges, as this require ongoing updates as demographics and definitions of fairness evolve.

Option 2: Need less effort in development but require a commitment to continuous review, but could effectively address subtle biases.

Option 3: This may affect the student life and also future, reduce the trust of the whole system.

📋 Refer to a code of ethics for guidance of areas to be mindful of

Regarding to EA ethics and guidelines 1.1, we should act on the basis of a well-informed conscience. [2]

ℹ️ Consult with respected staff or outside professionals

Our team does not have outside professionals for this part so we check papers on google scholar, and we got the idea of the following actions.

✅ Decide the course of action and take it

Design algorithm under biases consideration.

Provide a mechanism for users to report perceived unfairness(may not contain)

Periodically audit the system by independent evaluators.

Adjust algorithms based on feedback and bias monitoring.

Relevant data

[1] Madnani, N., Loukina, A., Von Davier, A., Burstein, J., & Cahill, A. (2017, April). Building better open-source tools to support fairness in automated scoring. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing* (pp. 41-52).

[2] Engineers Australia. (2022). Code of Ethics and Guidelines on Professional Conduct. https://www.engineersaustralia.org.au/sites/default/files/2022-08/code-ethics-guidelines-professional-conduct-2022.pdf