

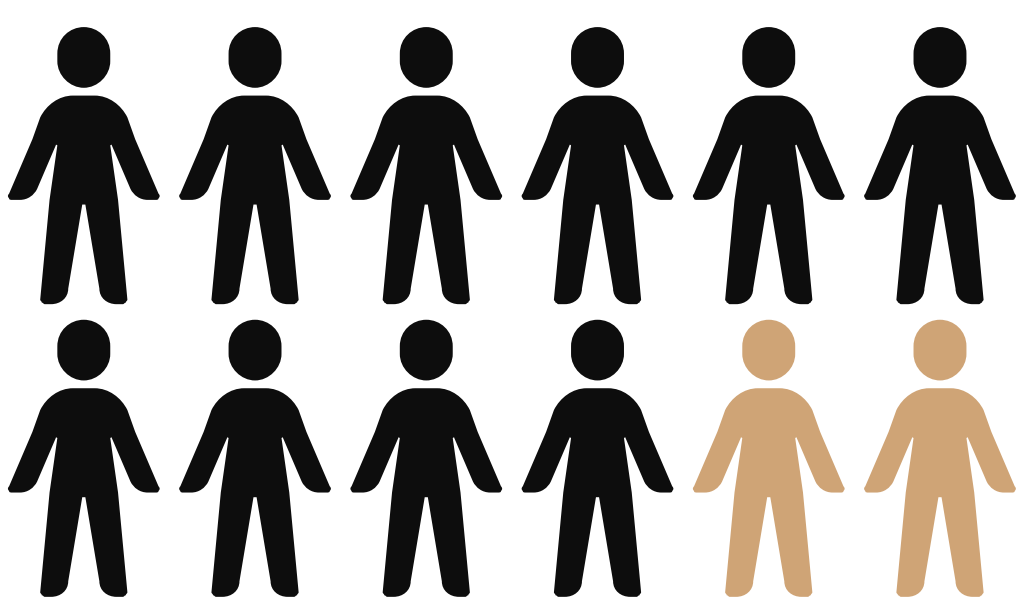
Developing a Behavioural and System-Level Tool to Detect AI Generated Content in Academic Writing

Rachel Ranjith, Trinity College Dublin

Supervised by Dr. Hitesh Tewari, School of Computer Science and Statistics,
and Dr. David Lewis, School of Computer Science and Statistics.

AI tools like ChatGPT are reshaping how students study, write, and learn.

10 in 12 students in higher education have used some form of AI in academic contexts.¹



Introduction and Objectives

Universities are split between **policing AI use** and finding ways to **integrate it** responsibly.² My project explores this tension through a dual approach: **developing a technical system** to detect AI-assisted writing, and critically reflecting on the ethical, pedagogical, and policy **implications of AI in higher education**. The primary objectives are:

- Review current AI use in education
- Develop prototype for AI detection
- Test accuracy of keystroke tracking
- Identify limitations and trade-offs
- Evaluate ethical and pedagogical impacts
- Outline strategies for responsible AI use

Methodology

REVIEW

Survey current state-of-the-art, existing tools, policies, and research

PROTOTYPE

Develop platform with keystroke tracking and chatbot

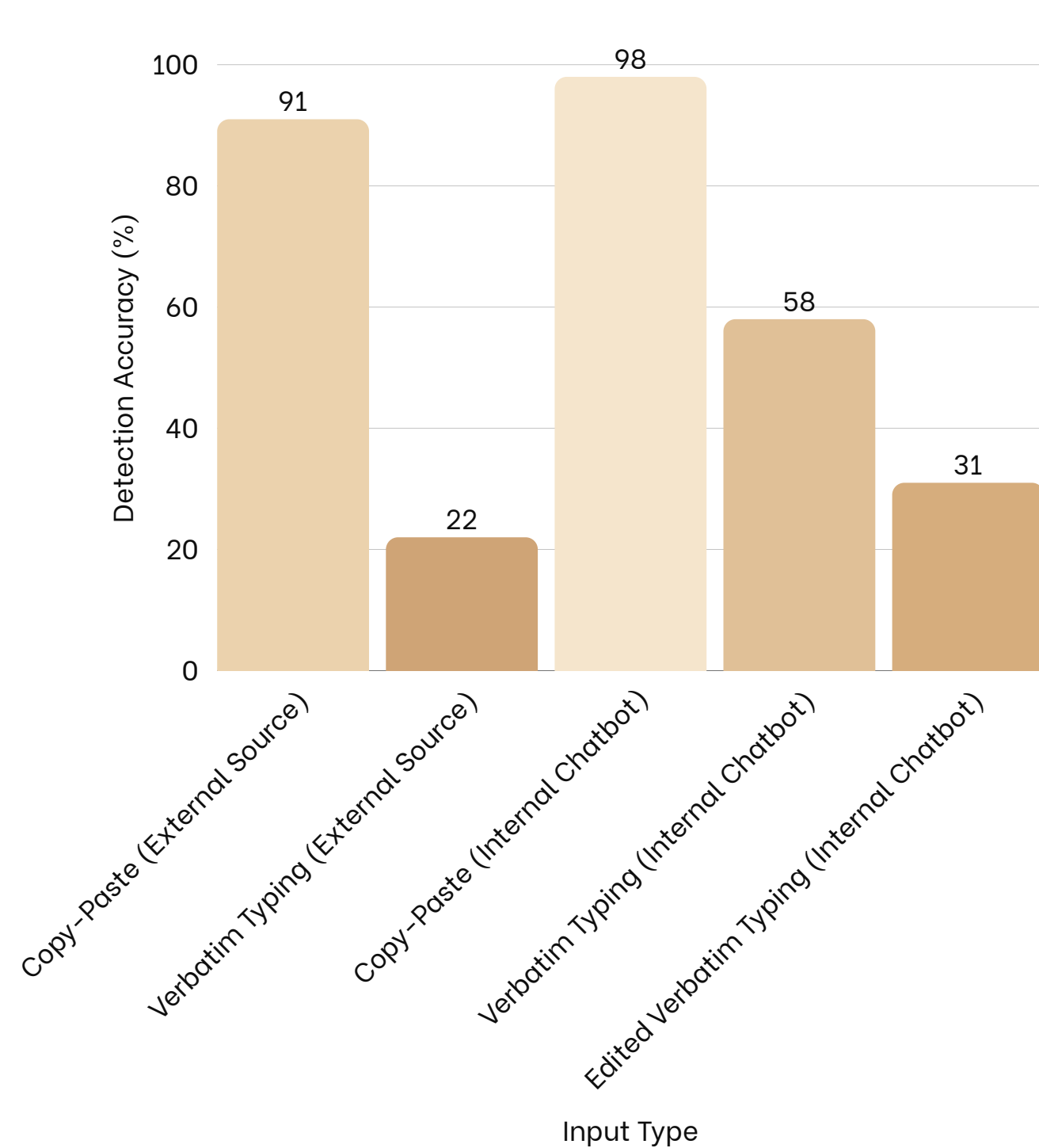
ANALYZE

Assess performance and identify limitations, run tests to benchmark

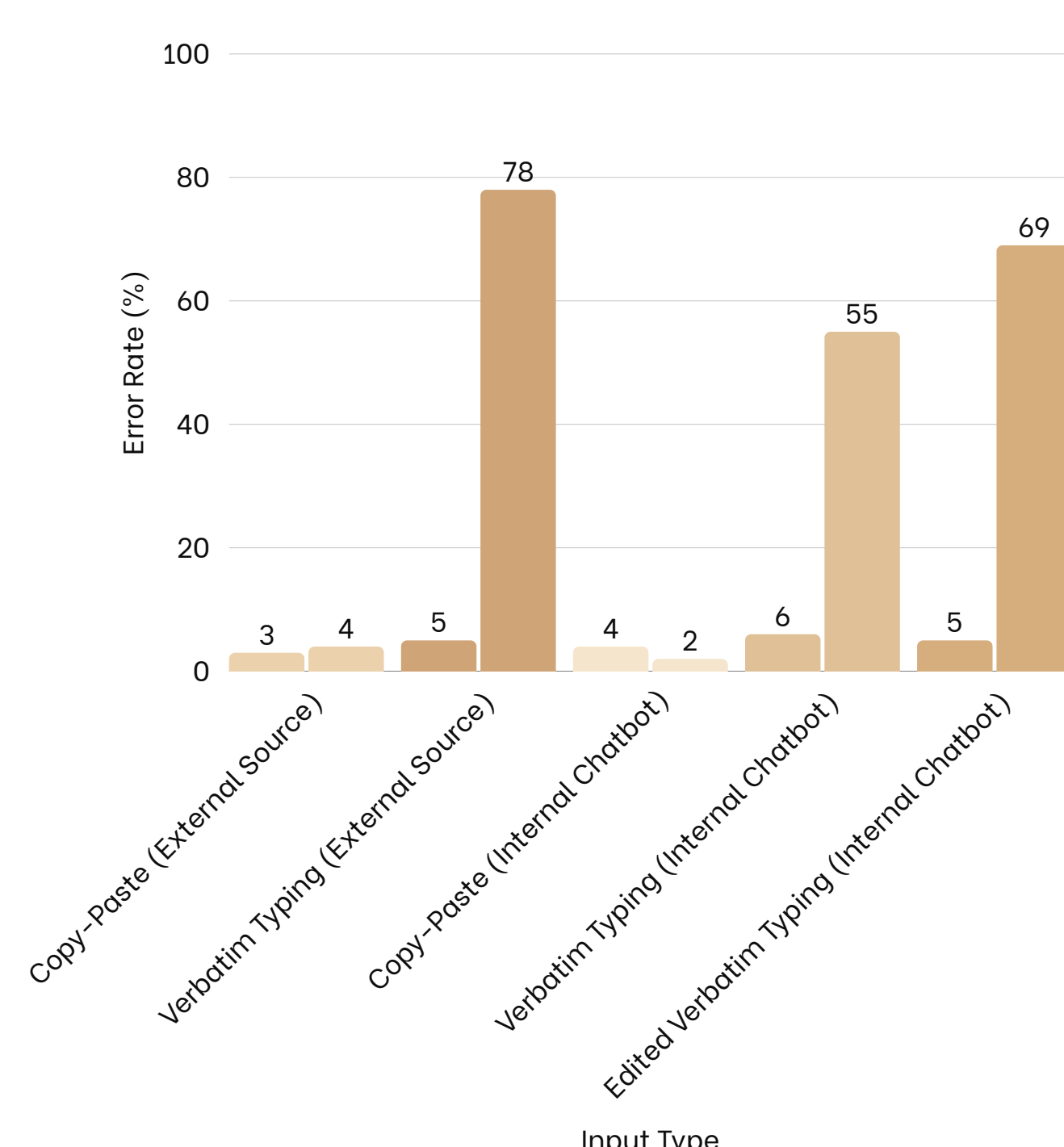
REFLECT

Examine AI ethical, pedagogical, and policy implications

Prototype Results



The system³ achieved **high accuracy** in detecting pasted text, with near-perfect identification rates. However, detection **dropped significantly** when text was retyped verbatim, and **further declined** when students used external sources. These results highlight the limits in key-stroke tracking.



This grouped bar chart **highlights errors** across different input types. Most errors are **false negatives**, showing that the system often fails to catch AI or external inputs. Over-tightening the system could **increase false positives**, wrongly questioning **student credibility**, which is ethically serious.

Conclusion

- Keystroke tracking has **technical limits** and **privacy risks**.
- Strict enforcement risks a **surveillance “witch-hunt”**; too much leniency **erodes learning**.
- Detection alone cannot solve issues like **misinformation** or **skill gaps**.
- The challenge: balance policies and pedagogy to **integrate AI responsibly** and foster **critical thinking**.⁴

Next Steps

- Explore **hybrid detection approaches** combining technical and pedagogical strategies
- Integrate **AI literacy** programs to cultivate critical engagement and deeper understanding of the risks of AI
- Investigate **policy frameworks** balancing integrity, trust, and learning outcomes
- Conduct user studies to **refine system usability** and ethical safeguards

Acknowledgement

This project was made possible through the support of the Laidlaw Scholars Leadership and Research Programme and Trinity College Dublin. I am especially grateful to my supervisors, Dr. Hitesh Tewari and Dr. David Lewis, for their invaluable guidance and feedback throughout the research process. I would also like to acknowledge the Laidlaw community and my peers, whose discussions and encouragement helped shape the direction of this work.

References

- [1] - <https://www.aiprm.com/ai-in-education-statistics/>
- [2] - <https://www.universityworldnews.com/post.php?story=20250806191842970>
- [3] - <https://github.com/Rachel-R16/Laidlaw-Research-Academic-Integrity-Platform>
- [4] - <https://www.forbes.com/councils/forbestechcouncil/2025/03/18/ai-in-academics-what-it-takes-to-build-a-sustainable-future/>