LEARNING LETTERS

Year 2026
Volume 6
Article 57
https://doi.org/10.20851/ll.v6.57

Beyond the algorithm: A longitudinal study of academic integrity trends in the context of artificial intelligence in STEM education

Michael Edward Ulpen[1], Manvi Gandhi[2], and Junaiz Rehmen[3]

[1] Adelaide University, Australia

[2] Adelaide University and Eynesbury College, Australia

[3] Adelaide University and Eynesbury College, Australia

The rise of generative artificial intelligence (GenAI) raises critical concerns about academic integrity and the authenticity of student learning. This paper draws on academic integrity records from the South Australian Institute of Business and Technology to explore trends in GenAI use, detection, and institutional response following the public release of ChatGPT. Using a quantitative data analysis approach, we examined patterns across multiple study periods and disciplines to highlight key differences in how GenAI is being addressed in STEM and non-STEM courses. The findings reveal a rising number of misconduct cases, especially in essay-style and computer programming assessments. These shifts raise concerns about the effectiveness of current assessment models, and the long-term credibility of STEM qualifications. The paper argues for a proactive and pedagogically informed response, emphasising authentic in-class assessments, policy reform, and inclusive learning environments, to ensure STEM graduates are equipped not only with technical skills but also with the ethical and analytical capacities required for work in industry.

*Keywords: a*cademic integrity, AI humaniser, generative artificial intelligence, Turnitin, STEM

Corresponding author: Michael Edward Ulpen, Michael.Ulpen@adelaide.edu.au

Recommended citation: Ulpen, M. E., Gandhi, M., Rehmen, J. (2026). Beyond the algorithm: A longitudinal study of academic integrity trends in the context of artificial intelligence in STEM education. *Learning Letters, 6,* Article 57. https://doi.org/10.20851/ll.v6.57

Introduction

The public availability of generative artificial intelligence (GenAI) technologies has disrupted the higher education sector (Hutson et al., 2022). GenAI is changing the ways in which students interact with their learning resources and assessment tasks (Borah et al., 2024). GenAI tools are now being used by students to generate essays, computer programs, diagrams, or even multimedia content (Zhai & Krajcik, 2024). While there are numerous advantages associated with the educational use of GenAI, it is important to understand that the availability of GenAI poses a serious threat to academic integrity (Yusuf et al., 2024; Zawacki-Richter et al., 2019). Students who use GenAI may avoid engaging with the fundamental problem-solving tasks required to form critical thinking, ethical reasoning, and reflection skills (Andrae, 2025). There is growing concern that students may graduate without demonstrating meaningful engagement with the core learning outcomes of their programs, raising questions about their readiness for industry or further academic study (Karacan-

Ozdemir et al., 2024). Therefore, this study aims to examine how the introduction of GenAI has influenced academic integrity behaviours within higher education, with particular attention to the ways in which these trends emerge across STEM disciplines.

GenAI technologies make it increasingly difficult for educators to distinguish between unethical use and appropriate academic support (Ayub et al., 2024; Lancaster et al., 2019). Various institutions have adopted Turnitin text-matching software to detect content generated by AI (Artificial Intelligence), but its use has raised some serious concerns (Kaktiņš, 2019). Widespread use of translation and grammar correction software complicates the ability of educators to determine if writing is AI generated, which raises questions about equity and inclusion for international students (Roe et al., 2023). Meanwhile, AI "humanisers" have been expressly developed to prevent AI-generated text from being detected by Turnitin and other AI detection software (Ayub et al., 2024). Given these challenges, it is incumbent upon higher education institutions to understand how they are interacting with the shifting GenAI landscape and adapt their policies and practices accordingly.

This paper will specifically consider a case study of academic integrity investigation records at the South Australian Institute of Business and Technology (SAIBT) across 13 study periods. The study will draw upon all academic integrity records associated with formal student submissions before and after the public launch of ChatGPT. The intention of this paper is to understand emerging trends in academic integrity and the use of GenAI. The findings are contextualised for STEM courses, where there is a growing need to adapt learning spaces and assessments to ensure genuine academic integrity and provide authentic learning outcomes.

This study identifies emerging patterns in GenAI-related integrity cases and examines differences between STEM and non-STEM courses in terms of case characteristics, detection, and outcomes. By providing empirical evidence of how GenAI-enabled misconduct is evolving, this study contributes practical insights for academic integrity officers, course coordinators, and policy leaders. The findings support informed decision-making regarding assessment redesign, academic integrity education, and institution-wide strategies for promoting authentic learning in a GenAI-enabled environment.

Methods

This study uses a quantitative research approach structured around a longitudinal, comparative analysis of academic integrity cases across 13 study periods from 2021 to 2025. The goal is to identify discipline-specific trends in the use of AI detection technologies and to understand how GenAI and assessment practices may influence the frequency and handling of academic integrity concerns.

The dataset was segmented into two major disciplinary cohorts:

- STEM: Information technology, engineering, and health science courses
- Non-STEM: Arts and business courses

This division was based on differences in assessment types. STEM courses typically involve tests/exams or practical tasks such as coding, problem-solving, and mathematics, where file formats and task design may not align with Turnitin's detection capabilities. Conversely, non-STEM study predominantly involves essay-style assessments, which are more likely to be submitted through Turnitin for originality and AI content checks.

Academic integrity records (n=1291) were collected from Pipefy, SAIBT's academic integrity case management system. All submitted Turnitin Similarity Reports (n=365) were extracted from these records for further analysis.

Metrics and data analysis

The following metrics were calculated per study period:

- Total student enrolments
- Number of academic integrity investigations
- Number of penalised academic integrity cases (excluding minor warnings and non-issues)
- Number of Turnitin Similarity Reports submitted
- Unique student offences compared to total offences
- Investigation and penalty rates relative to cohort size to account for cohort size fluctuations

Statistical analysis

Descriptive statistical analysis was conducted to identify trends in academic integrity cases over time. Comparisons were drawn between disciplinary cohorts to assess variations in reporting behavior, reliance on detection tools, and potential influence of assessment design on investigation rates. Turnitin Similarity Reports were examined to calculate the average percentage of text flagged as AI-generated in each study period.

Linear regression analyses were performed using Microsoft Excel and Minitab to identify trends and calculate coefficients. Residual plots in Minitab were reviewed to ensure that the conditions for linear regression were satisfied.

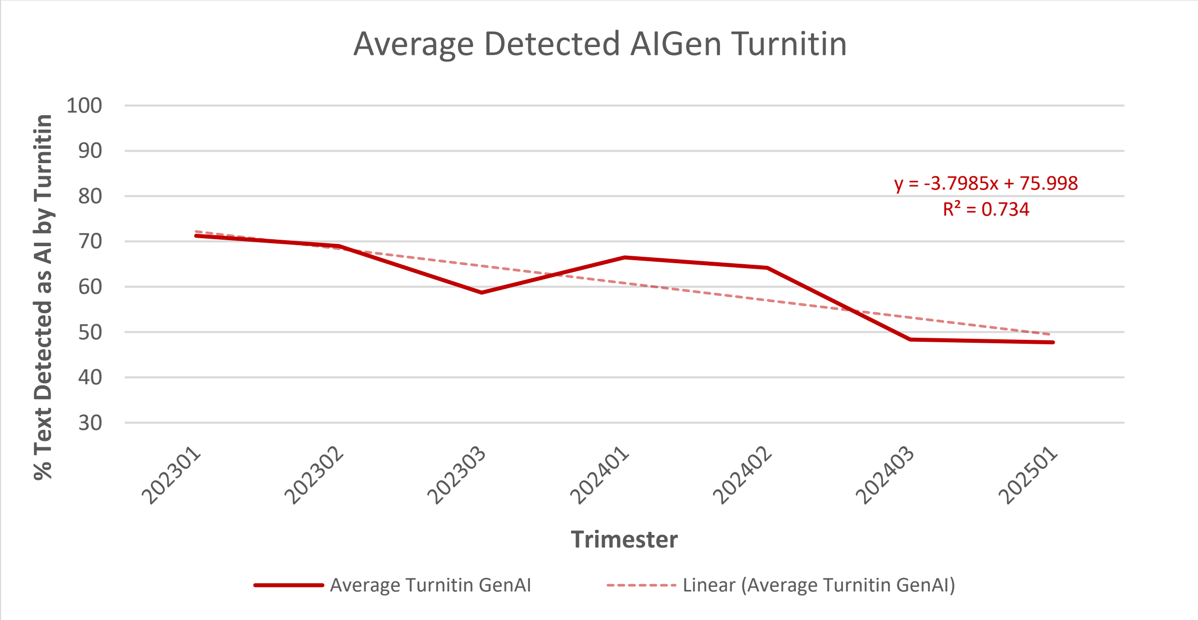
The dataset was further filtered to determine unique student offences compared to total student offences per study period. Investigation rates and penalty rates were calculated as a percentage relative to cohort size in each discipline in each study period. This was done to normalise the rate of academic integrity cases regardless of cohort size.

Findings

Detector use

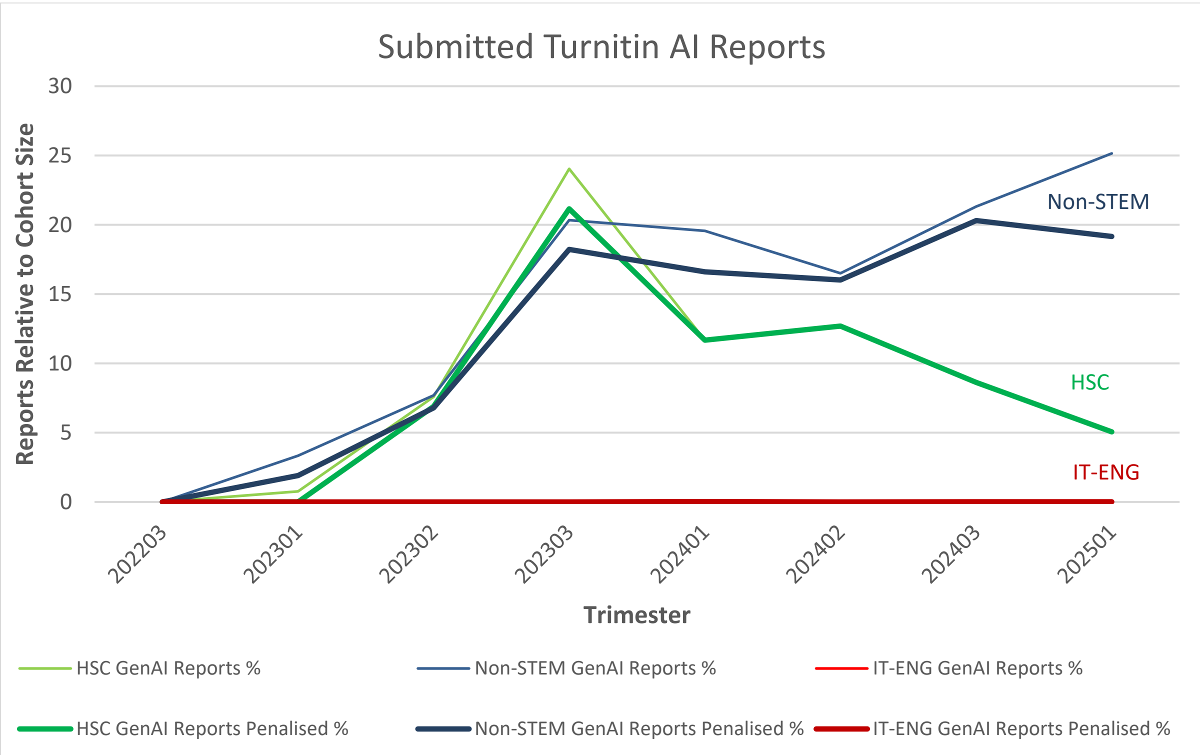
The Turnitin Similarity reports submitted between study periods 202301 and 202501 were analysed to calculate the average proportion of text that was detected as GenAI within each submission. The analysis revealed a consistent downward trend (Figure 1). From the Turnitin Similarity reports submitted in study period 202301, Turnitin indicated an average of 71.25% of text was likely AI-generated. By study period 202501, this average had decreased to 47.73% (Figure 1). During this period, various AI humanisers have been released and Turnitin has gone through several updates so this may not be an accurate measure of how much submitted text was actually AI generated.

Figure 1: Average percentage of text detected as GenAI across all Turnitin artificial intelligence reports submitted for academic integrity investigations each trimester.



Separating into IT/engineering, health science, and non-STEM cohorts revealed that teachers in IT and engineering rarely use Turnitin to detect GenAI (Figure 2). Despite this, the academic integrity officers listed the reason for the offence as “Use of artificial intelligence software” in 37.76% of the IT/engineering cases (54 cases out of 143).

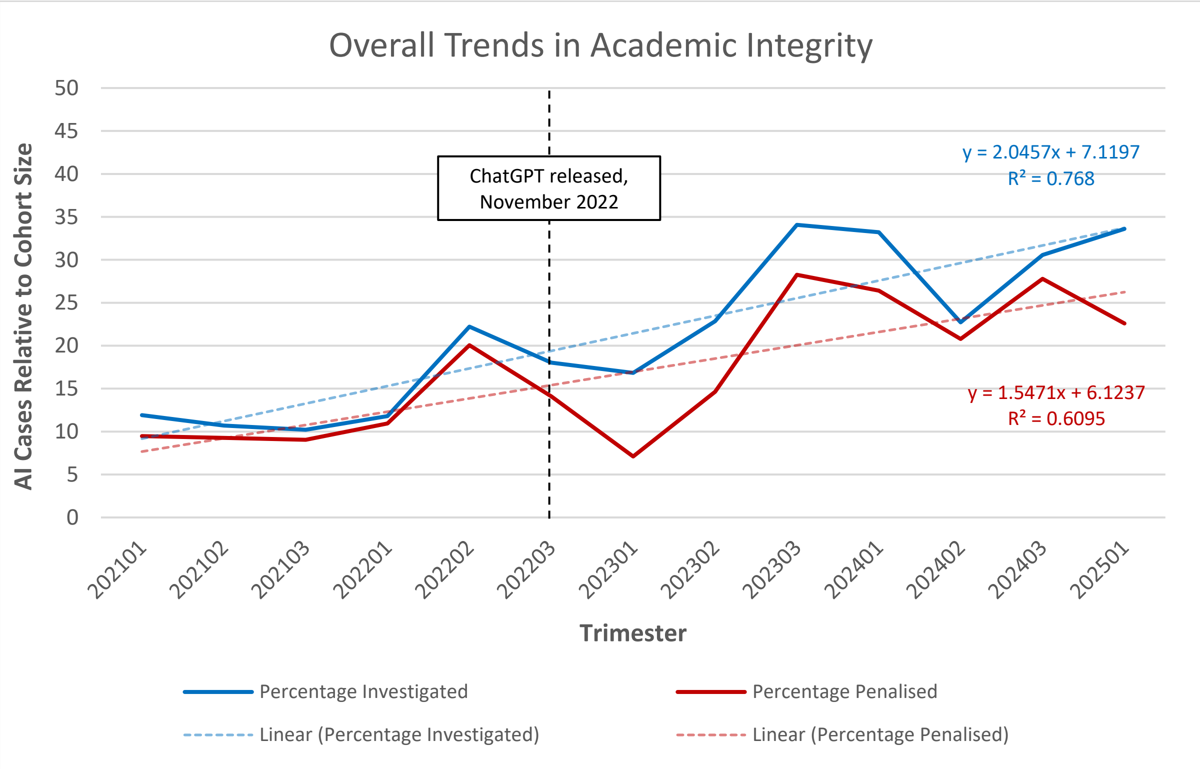
Figure 2: Number of Turnitin artificial intelligence reports submitted as additional evidence for academic integrity concerns measured relative to cohort size.



Rates and denominators

The scale of academic integrity issues was examined by dividing the total number of academic integrity cases and penalties by the number of students enrolled in each study period (Figure 3). The graph showed an upward trend in both the number of AI investigations and the number of penalties issued over time. It appears that the difference between percent investigated and percent penalised was narrower and more consistent prior to the release of ChatGPT in November 2022. The difference between investigations and penalties appears to become wider and more erratic thereafter. In 202501 the number of AI cases investigated was approximately 33%, and the number of AI cases penalised was approximately 22.5% relative to total enrolments.

Figure 3: Percentage of academic integrity concerns and penalisations relative to total enrolments each trimester.



According to simple linear regression, the number of AI cases appears to be increasing over time (Figure 3). The trendline for AI cases investigated showed a linear increase of around 2%, and the trendline for AI cases penalised showed a linear increase of around 1.5%. We did not witness any statistically significant curvature in the graph.

According to Figure 4, both STEM and non-STEM courses have seen increases in investigation and penalty rates over time. The figure demonstrates that there were stronger investigation and penalty growth rates in non-STEM courses compared to STEM courses (see Appendix A for full regression outputs).

Figure 5 represents the number of AI cases penalised and the proportion of unique student offenders, each relative to cohort size for both STEM and non-STEM courses. Some students were involved in multiple investigations or received more than one penalty within a single study period. Notably, the non-STEM cohort shows a more pronounced gap between penalised cases and unique offenders from 202303 to 202501.

Figure 4: Percentage of academic integrity investigations compared to percentage of confirmed academic integrity breaches relative to cohort size each trimester and their associated linear average trends.

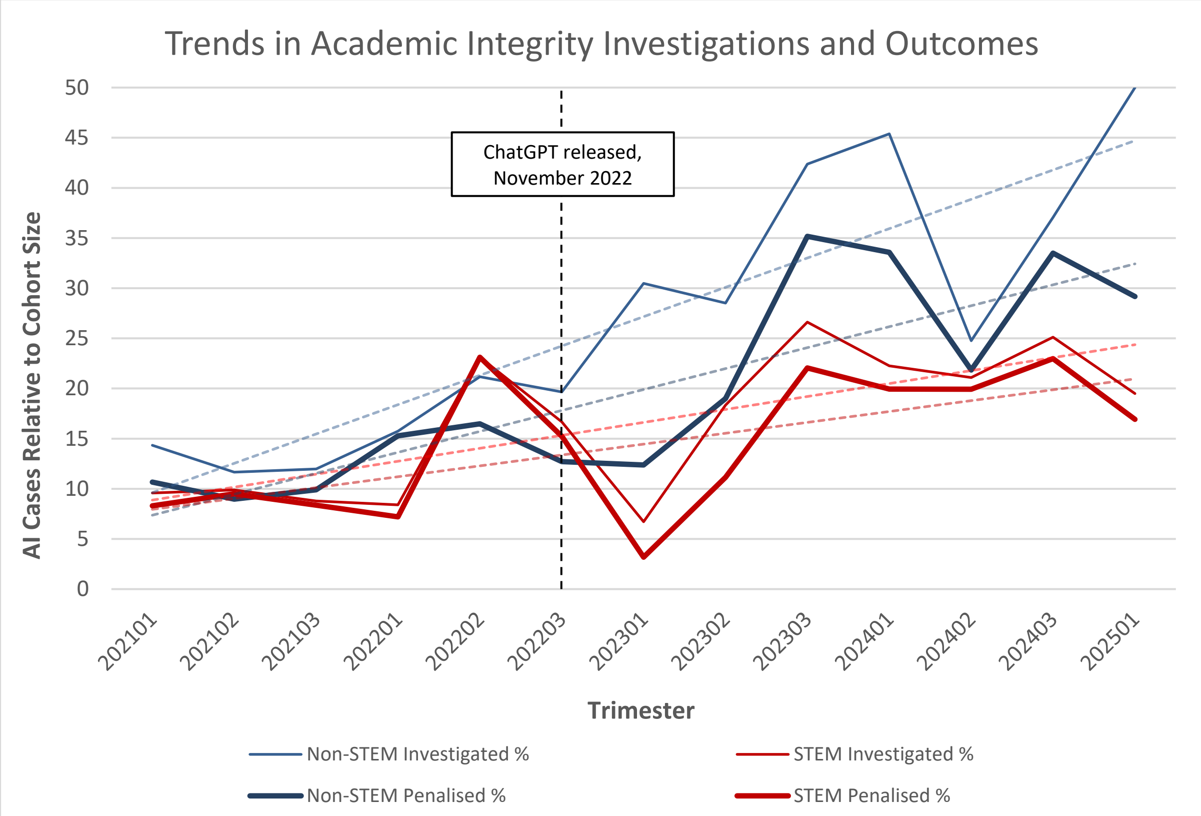
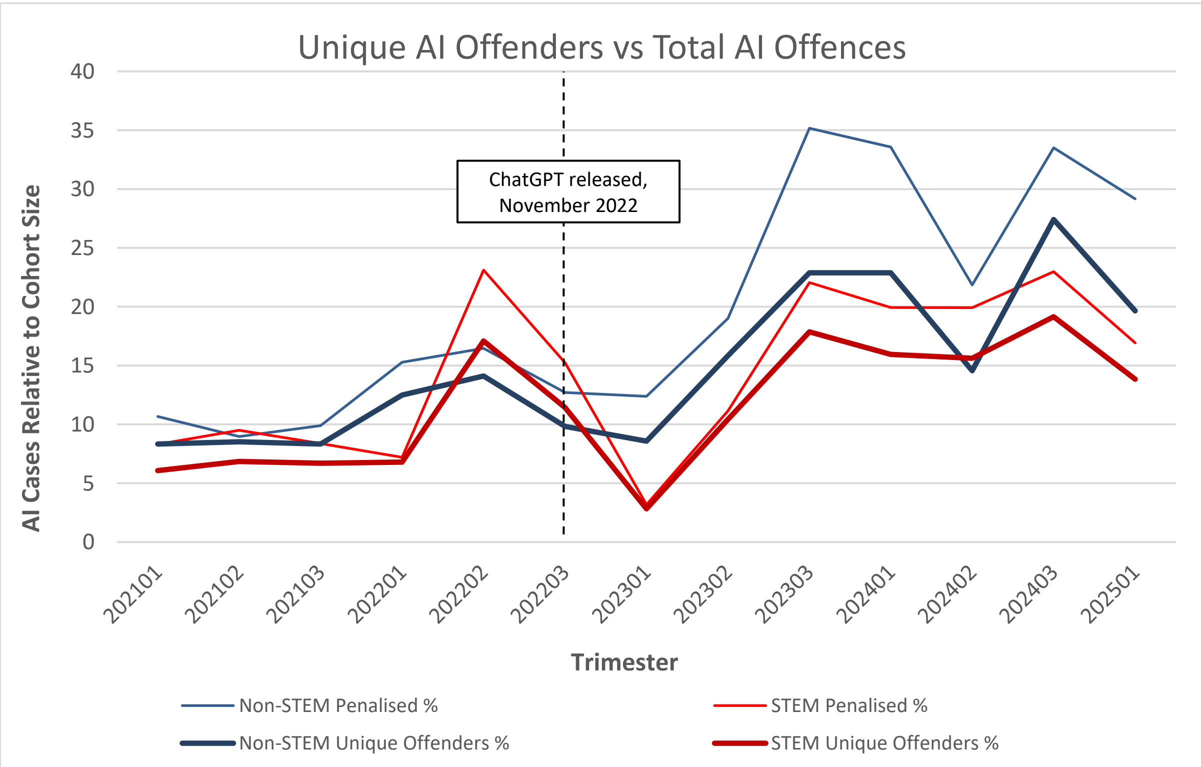


Figure 5: Unique AI offenders compared to all AI offences in each study period



Assessment differences

Academic integrity cases from each discipline were separated into assessment types: essay, multimedia, programming, math, presentation, and exam/test (refer to Appendix C for full breakdown). In IT and engineering, the highest number of academic integrity cases were raised in relation to programming assessments, and these cases were raised more consistently each study period than other assessment types. In health science, business, and arts, most academic integrity offences were related to essay style assessments. Additionally, exam/test assessments in health science appear to be less resilient to academic misconduct compared to exams/tests in other disciplines. Notably, the number of cases in the arts was 93.75% relative to the cohort size in 202501, which far exceeds the highest recorded AI cases in a study period for other disciplines, 35.25%.

Discussion

While the number of academic integrity investigations is increasing across all disciplines, the rise is more pronounced in courses with essay-style assessments. Furthermore, AI humanisers, such as aiundetect.com and quillbot.com, have been developed to avoid detection by Turnitin and other GenAI detection algorithms (Kar et al., 2024). The widespread availability of these technologies may be a contributing factor to the downward trend seen in the percentage of text determined as GenAI by Turnitin Similarity reports (Figure 1).

These concurrent trends raise critical concerns about the reliability of detection tools and the growing administrative burden of academic integrity investigations. As Dawson (2020) argues, overreliance on technological solutions risks giving institutions a “false sense of security”. The downward trend in Turnitin Similarity reports reinforces Dawson’s call to shift the focus from detection alone and towards assessment practices that are resistant to the use of GenAI.

Turnitin is unable to detect GenAI in the filetypes commonly submitted for programming assessments. Despite this, 49.48% of academic integrity cases in programming assessments between 202301 and 202501 were penalised for “use of artificial intelligence software” (48 cases out of 97). Non-programming assessments in IT and engineering had a much lower proportion of GenAI offences and fewer cases overall (19 GenAI offences out of 77 cases, 24.68%). This was far below the average of the other disciplines (327 GenAI offences out of 634, 51.58%). This suggests that non-programming assessments in STEM may be more resilient to the misuse of GenAI compared to essay style assessments in other disciplines.

Limitations and future research

The release of ChatGPT had a major impact on academic integrity, but it was not the only factor. The fluctuations in the rate of investigations and penalties seen in Figures 3 and 4 have been caused by a combination of changes to technology, policy, assessments, students, teachers and attitudes. Future research could conduct a historical analysis to reveal the reasons for these fluctuations with more precision.

Future research could measure academic integrity concerns and penalties made for assessments where Turnitin is enabled compared to similar assessments where Turnitin is disabled. This analysis could investigate to what extent evidence from Turnitin is responsible for revealing AI offences that human assessors could not otherwise identify or to what extent evidence from Turnitin is being used for initiating unreasonable AI investigations.

This study was focused on revealing emerging trends in academic integrity, however it was only able to analyse records from one dataset, so the trends may be specific to SAIBT. Performing a meta-analysis that includes other higher education institutions could improve the reliability of the trendlines.

Recommendations

To uphold academic integrity in the age of GenAI, educational institutions must shift from a reactive detection-based model to a proactive, pedagogically driven approach. Several methods have emerged from STEM courses at SAIBT that attempt to protect against the misuse of GenAI. The most effective include the use of multimedia, dispersed or scaffolded assessment instructions; assessing the output of specialised software; compulsory face-to-face assessments; and paper examinations. Requiring the student to write with reference to their personal context or with reference to in-class activities also seem to be effective deterrents to the use of GenAI. We recommend maximising the use of these strategies across all courses.

Wecks et al. (2024) found that students who use GenAI performed on average 6.71 points out of 100 worse in exams than students who did not use GenAI during their studies. Therefore, it is imperative to ensure academic integrity and genuine engagement from students throughout their studies. Investment in academic integrity processes and assessment redesign could help higher education institutions to more effectively ensure genuine academic integrity without compromising student and industry trust. A balanced approach, integrating technology, policy and pedagogy, could help safeguard the institution's credibility and ensure students graduate with the genuine competencies required for ethical and effective practice in STEM.

Conclusion

Unregulated access to GenAI is complicating the process of identifying and preventing academic misconduct while exposing critical gaps in assessment design and algorithmic detection (Troy, 2025). Employers need confidence that STEM graduates embody not only technical ability but also the ethical and analytical capacities to apply that expertise in industry (Rayner & Papakonstantinou, 2015). If higher education institutions are unable to ensure the genuine academic integrity of their graduates, then it raises dire concerns for public safety and for the long-term relevance of STEM degrees (Roy & Edwards, 2023).

Generative artificial intelligence poses a profound challenge for education. Our findings reveal that GenAI is contributing to an increase in academic integrity concerns and breaches, leading to an increased workload for teachers and academic integrity officers. Institutions that fail to adapt to the trends revealed in this report risk awarding qualifications to students who have not met the required learning outcomes for their field, which in turn could undermine public trust in STEM education and in the relevance of their qualifications in industry. Therefore, it is essential that assessments and learning environments are adapted to mitigate this threat and to ensure the authentic academic integrity of our graduates.

Acknowledgements

The authors would like to acknowledge SAIBT's Academic Board for their valuable insights. We also thank SAIBT's Academic Director, *Mona Umapathy*, for her assistance in analysing academic integrity trendlines, and Discipline Leads *Alice Dodd* and *Mahsa Karimpour* for their support with academic integrity data analysis and statistical analysis using Minitab.

Funding and ethics

No funding was received for the conduct of this research.

This study received ethics exemption (Exemption ID: 207498) from the University of South Australia.

Disclosure of conflicts of interest

The authors report no potential conflict of interest.

Disclosure of the use of AI-assisted technologies during writing

No AI-assisted technologies were used during the writing process.

About the authors

Michael Edward Ulpen completed his Bachelor of Information Technology in 2013 with a focus on games and entertainment design. He is currently an academic teacher at Adelaide University specialising in software engineering and IT systems. He has over ten years of experience in teaching in undergraduate programs. His current research interests include academic integrity and artificial intelligence.

ORCID: https://orcid.org/0009-0004-8134-2914

Manvi Gandhi completed her Master's degree in environmental science at the University of South Australia and is currently completing her PhD in artificial intelligence in education. Her doctoral research explores student perceptions and use of large language models (LLMs), the training they receive, and their understanding of ethical AI practices, alongside investigating educators' confidence in teaching ethical AI use and integrating AI into assessment design. Manvi is an Academic Lecturer in the School of Pharmacy and Biomedical Sciences at Adelaide University contributing to teaching, assessment, and curriculum development across health and biomedical sciences programs. She also serves as Program Manager for health sciences and arts programs at Eynesbury College (formerly the South Australian Institute of Business and Technology), providing academic leadership across multiple programs. With over seven years of teaching experience across higher education and pathway programs, Manvi has worked with diverse student cohorts, including international and underrepresented learners. Her teaching philosophy centres on ethical practice, inclusivity, critical thinking, and real-world relevance, with a strong commitment to preparing students for an AI-enabled future with confidence and integrity.

ORCID: https://orcid.org/0009-0001-2440-0469

Junaiz Rehmen completed his PhD at the Future Industries Institute, University of South Australia, with a research focus in energy and manufacturing. He is currently a teaching academic at Adelaide University, specialising in engineering education. He has over ten years of experience in teaching and research and has delivered a wide range of engineering courses across undergraduate and postgraduate programs. His teaching practice spans face-to-face, hybrid, and online modalities, with a strong emphasis on student engagement and effective learning design. His current research interests centre on engineering education, including online learning, assessment design, and technology-enhanced pedagogical practices.

ORCID: https://orcid.org/0000-0002-4889-2816

References

Andrae, S. (2025). Critical thinking and GenAI: A guide to action. In M. Kessinger, G. Nelson, L. Lennex, & K. Nettleton (Eds.), *Enhancing classroom instruction and student skills with AI* (pp. 377-412). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-4576-5.ch011

Ayub, T., Ahmad Malla, R., Khan, M. Y., & Ganaie, S.A. (2024). The art of deception: Humanizing AI to outsmart detection. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/GKMC-03-2024-0133

Borah, A. R., Nischith, T. N., & Gupta, S. (2024). Improved learning based on GenAI. *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, 1527–1532. https://doi.org/10.1109/IDCIoT59759.2024.10467943

Dawson, P. (2020). *Defending assessment security in a digital world: Preventing e-cheating and supporting academic integrity in higher education*. Routledge. https://doi.org/10.4324/9780429324178

Hutson, J., Jeevanjee, T., Vander Graaf, V., Lively, J., Weber, J., Weir, G., Arnone, K., Carnes, G., Vosevich, K., Plate, D., & Leary, M. (2022). Artificial intelligence and the disruption of higher education: Strategies for integrations across disciplines. *Creative Education*, 13(12). https://doi.org/10.4236/ce.2022.1312253

Kar, S.K., Bansal, T., Modi, S., & Singh, A., (2024). How sensitive are the free AI-detector tools in detecting AI-generated texts? A comparison of popular AI-detector tools. *Indian Journal of Psychological Medicine, 47*(3), 275-278. https://doi.org/10.1177/02537176241247934

Kaktiņš, L. (2019). Does Turnitin support the development of international students' academic integrity? *Ethics and Education*, *14*(4), 430–448. https://doi.org/10.1080/17449642.2019.1660946

Karacan-Ozdemir, N., Park, C., & Solberg, S. (2024). Exploring STEM career competencies with the assistance of generative AI. In *2024 IEEE Integrated STEM Education Conference (ISEC)*, 1–3. http://doi.org/10.1109/ISEC61299.2024.10665084

Lancaster, T., Robins, A. V., & Fincher, S. A. (2019). Assessment and plagiarism. In S. A. Fincher & A. V. Robins (Eds.), *The Cambridge handbook of computing education research* (pp. 414–444). Cambridge University Press. https://doi.org/10.1017/9781108654555.015

Rayner, G. M. & Papakonstantinou, T. (2015). Employer perspectives of the current and future value of STEM graduate skills and attributes: An Australian study. *Journal of Teaching and Learning for Graduate Employability, 6*(1), 110–125. https://doi.org/10.21153/jtlge2015vol6no1art576

Roe, J., Renandya, W. A., & Jacobs, G. M. (2023). A review of AI-powered writing tools and their implications for academic integrity in the language classroom. *Journal of English and Applied Linguistics, 2*(1), 23-30. https://doi.org/10.59588/2961-3094.1035

Roy, S. & Edwards, M. A. (2023). NSF Fellows' perceptions about incentives, research misconduct, and scientific integrity in STEM academia. *Scientific Reports, 13*(1), 5701. https://doi.org/10.1038/s41598-023-32445-3

Troy, D. (2025). GenAI in academia: The case for transparency and accountability, *Community College Enterprise, 31*(1), 47–59. https://research.ebsco.com/c/hwkao2/search/details/wlospazisn/related-items?db=esu

Wecks, J. O., Voshaar, J., Plate, B. J., & Zimmermann, J. (2024). Generative AI usage and exam performance. *SSRN Electronic Journal*. http://dx.doi.org/10.2139/ssrn.4812513

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education, 21*(1), 21. https://doi.org/10.1186/s41239-024-00453-6

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education, 16(*1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X. & Krajcik, J. (2024). Artificial intelligence-based STEM Education. *Uses of Artificial Intelligence in STEM Education*. https://doi.org/10.1093/oso/9780198882077.003.0001

Appendix A

Table A1: Coefficients for linear regression on the average percentage of text detected as GenAI across all Turnitin artificial intelligence reports submitted for academic integrity investigations each trimester.

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	98.8	10.4	9.47	0.000	
Time	-3.80	1.02	-3.71	0.014	1.00

Table A2: Coefficients for linear regression on percentage of cohort investigated and penalised each trimester.

			Coef	SE Coef	T-Value	P-Value	VIF
All courses	Percentage investigated	Constant	7.12	2.69	2.65	0.023	
		Time	2.046	0.339	6.03	0.000	1.00
	Percentage penalised	Constant	6.12	2.96	2.07	0.063	
		Time	1.547	0.373	4.14	0.002	1.00
non-STEM	Percentage investigated	Constant	6.69	3.98	1.68	0.121	
		Time	2.925	0.501	5.84	0.000	1.00
	Percentage penalised	Constant	5.28	3.31	1.59	0.139	
		Time	2.088	0.417	5.01	0.000	1.00
STEM	Percentage investigated	Constant	7.58	3.05	2.49	0.030	
		Time	1.291	0.384	3.36	0.006	1.00
	Percentage penalised	Constant	6.86	3.31	2.07	0.062	
		Time	1.084	0.417	2.60	0.025	1.00

Appendix B

Table B1: Rates of academic integrity breaches, expressed as the number of investigations and penalties per 1,000 enrolments across STEM and non-STEM courses from trimester 202101 to 202501.

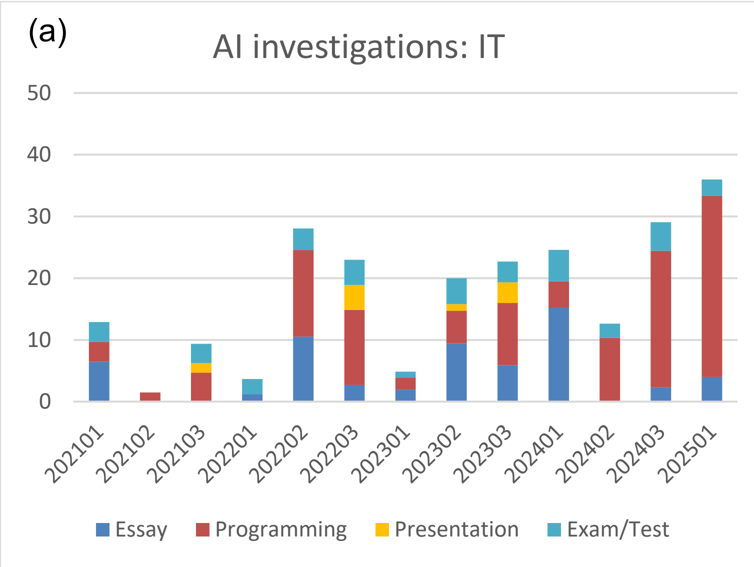
	STEM		non-STEM	
Trimester	**Investigated**	**Penalised**	**Investigated**	**Penalised**
202101	96	83	143	107
202102	99	95	117	90
202103	88	84	120	99
202201	84	72	157	153
202202	231	231	212	165
202203	167	152	197	127
202301	67	32	305	124
202302	183	112	285	190
202303	266	221	424	352
202401	223	199	454	336
202402	211	199	248	218
202403	251	230	371	335
202501	195	169	500	292

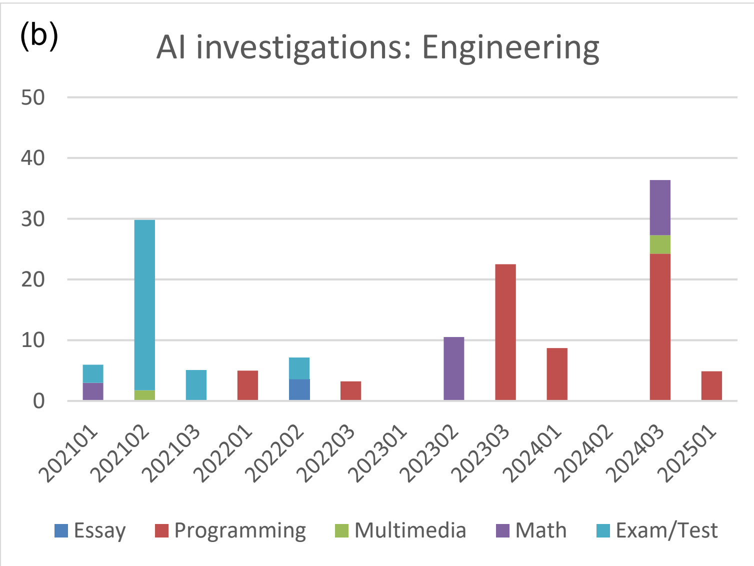
Table B2: Rates of academic integrity breaches, expressed as the number of penalised events and unique offenders per 1,000 enrolments, across STEM and non-STEM courses from Trimester 202101 to 202501.

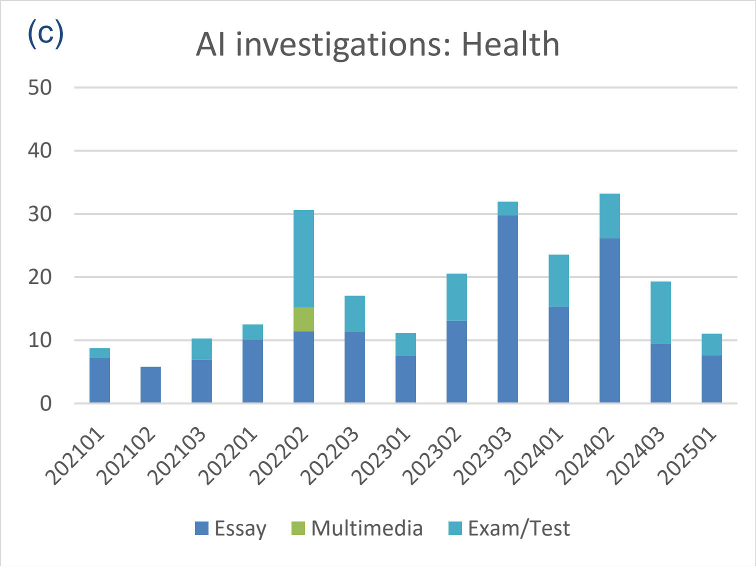
	STEM		non-STEM	
Trimester	**Penalised**	**Unique offenders**	**Penalised**	**Unique offenders**
202101	83	61	107	83
202102	95	68	90	85
202103	84	67	99	83
202201	72	68	153	125
202202	231	171	165	141
202203	152	114	127	98
202301	32	28	124	86
202302	112	104	190	158
202303	221	179	352	229
202401	199	159	336	229
202402	199	156	218	146
202403	230	191	335	274
202501	169	138	292	196

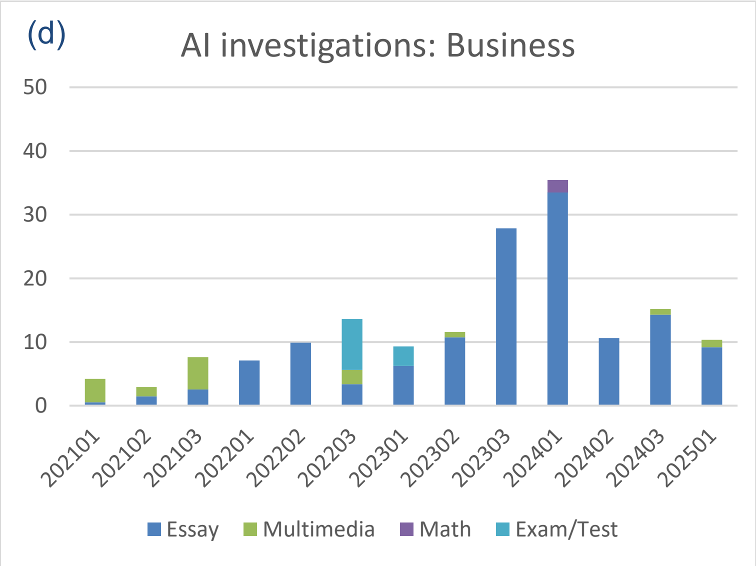
Appendix C

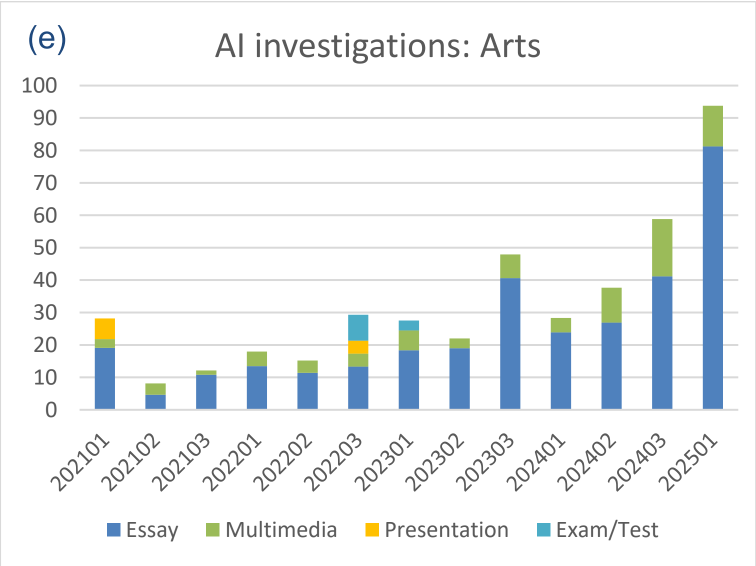
Figure C: Academic integrity investigations categorised into assessment types and reported as a percentage relative to cohort size in each discipline: (a) IT, (b) engineering, (c) health science, (d) business, (e) arts.






(c)
AI investigations: Health
50
40
30
20
10
0
202101
202102
202103
202201
202202
202203
202301
202302
202303
202401
202402
202403
202501
Essay
Multimedia
Exam/Test


(d)
AI investigations: Business
50
40
30
20
10
0
202101
202102
202103
202201
202202
202203
202301
202302
202303
202401
202402
202403
202501
Essay
Multimedia
Math
Exam/Test


(e)
AI investigations: Arts
100
90
80
70
60
50
40
30
20
10
0
202101
202102
202103
202201
202202
202203
202301
202302
202303
202401
202402
202403
202501
Essay
Multimedia
Presentation
Exam/Test