

Artificial Intelligence and Academic Integrity: Ethical Considerations for Graduate Students

^{1*} Braulio Julimin Jiménez De la Rosa, ^{2*} Leonor Farray Bergés.

*Email: * jimenezbraulio@gmail.com

*Affiliation: * Universidad Nacional Pedro Henríquez Ureña

ORCID number: * 0009-0005-7647-0177

Scopus Author ID: * 60582125100

*Affiliation: * Universidad APEC.

*ORCID number: * 0000-0002-8223-3785

HOW TO CITE:

Braulio Julimin Jiménez De la Rosa,
Leonor Farray Bergés. (2026).
Artificial Intelligence and Academic
Integrity: Ethical Considerations for
Graduate Students.
International Journal of Special
Education, 41(23s), 1867 - 1883

ABSTRACT:

The explosion of generative artificial intelligence technologies has spurred a major paradigm shift that has significantly challenged and redefined key tenets of academic integrity, authorship and assessment in higher education. This research explores the ethical issues, usage of generative AI tools, and detection difficulties faced by graduate students working with G.A.I. The study uses a multi-methodological approach and data from a large sample of four hundred and fifty graduate students from a variety of disciplines to determine task-specific usage rates and discipline-specific differences in ethical acceptability. Moreover, the study compares the empirical effectiveness of commercial artificial intelligence detection systems by examining the receiver operating characteristic curve and related areas under the curve. The results show that artificial intelligence is widely used to refine structures and text and that it is not used at all to draft the theory. A distinct disciplinary divide can be seen, with students from STEM having a more liberal attitude towards the assistance provided from an algorithm, than students from the humanities. Most importantly, the study shows that detection software is fundamentally flawed, and can be misused in serious ways, giving rise to misaccusations. The study also reveals that a paradigm shift is necessary in higher educational institutions from reactive, surveillance-based policies to proactive, pedagogical ones aimed at the development of AI literacy, conceptualizing academic integrity within a collaborative postplagiarism environment.

Keywords: Generative Artificial Intelligence, Academic Integrity, Graduate Students, Postplagiarism, AI Detection, Higher Education Ethics, Cognitive Scaffolding..

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the
terms and conditions
of the Creative Commons Attribution
(CC BY)
license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Generative Artificial Intelligence (GAI) is one of the most transformative new technologies that has impacted today's higher education landscape at a rapid pace. In contrast to the past, when the technologies centered in information retrieval, communication, and administration, generative AI systems can create, modify, summarize, translate, restructure, and critically engage with academic content, often in ways that closely mimic elements of human thought. As a result, large language models and other generative systems have started to influence students' ways to search for information, access and read scholarly texts, generate ideas, build arguments, write academic texts, analyze information and perform research tasks (Crompton & Burke, 2023). The rise in their availability and sophistication has brought new and great opportunities for academics to be productive and new and serious problems with traditional notions about academic integrity, authorship, originality, assessment, and academic responsibility.

The use of generative AI has made it more difficult than ever to separate machine-generated content from human-generated content, especially within the academic context. Usually, conventional academic contexts have been linked to the intellectual contribution, to the interpretation of evidence, to the organization of arguments, and to the production of the written work itself. This relatively stable understanding is disrupted by the current use of generative AI systems, which can be used by educators and students throughout the academic production process, but do not necessarily replace the student entirely (Dawson et al., 2024). A student can use AI to refine their grammar, summarize broad literature, brainstorm research questions, code a text, translate text, or reorganize a draft, and craft their own argument. In such situations it becomes more and more challenging to distinguish between legitimate technological support and illegitimate outsourcing of intellectual labor. This nascent confusion has led to an important rethinking of the meaning of academic integrity in an educational setting where

human and algorithmic contributions can be combined in the same academic product.

Graduate students are one of the most significant groups to consider in this regard. Graduate students will be expected to exhibit high levels of independent academic research, methodological proficiency, critical thinking, theoretical literacy and knowledge production beyond the introductory or highly structured course. Academic work often includes a high level of interaction with academic texts, with complicated data, complex decisions about methodology, research design, interpretation, and in-depth academic writing. The use of generative AI in postgraduate studies could thus go beyond mere academic support and impact intellectual growth among new scholars in the way they construct their scholarly identity and intellectual independence (Aladsani, 2025). To establish sound academic integrity policies, therefore, it is important to grasp how graduate students are actually using GAI and not simply consider any use of AI as academic dishonesty.

The ethical considerations surrounding the use of GAI are further compounded by the different levels of intellectual engagement required for each of these different types of AI support. It's important to note that using an AI system to correct the grammar and syntax is different from using it to create a conceptual framework, theoretical argument, or section of a manuscript. Likewise, using AI to recognize the main ideas in a wide range of texts might be a different type of cognitive support than having an AI program build the theoretical framework for a dissertation. These differences are important as they are directly related to the assessment of academic integrity, not just the use of AI. In contrast, the nature, purpose, scope, transparency, and intellectual impact of the use need to be taken into account (Corbin et al., 2025). As AI plays a more complex role in academic processes, it becomes essential to understand the acceptable and unacceptable use of algorithms to support academic tasks.

Initial reactions to generative AI in institutions were often about concern, ban, and monitoring.

At first, universities and educators saw AI tools like ChatGPT merely as a way to deceive them into accepting artificial or programmed output as genuine student work. At first, universities and educators viewed such tools as ChatGPT primarily as a means for students to cheat on traditional assessment methods and have their work done by the AI. These concerns are not unfounded, as generative AI can generate coherent and seemingly sophisticated answers that can be produced in a fast and large scale, which can pose as a significant challenge to traditional assessment practices (Cotton et al., 2024; Rudolph et al., 2023). Cracks in the written assignment process led to concerns about plagiarism, contract cheating, unauthorized collaboration, and grade and academic credential integrity, brought about by the potential for students to outsource large sections of their written assignments to AI systems. As a result, some institutions took action by introducing restrictions, changes to the assessment process, and implementing AI detection tools.

But in the face of increasingly widespread use of generative AI in the broader educational landscape and in the workplace, a prohibition-based approach is harder to maintain. The technology does not only serve the function of creating "whole" academic products; it can be used as a cognitive aid at several stages of academic production. It can help students understand difficult concepts, develop alternative explanations, recognize gaps in their thinking, enhance their language, streamline their thinking, relay information from academic sources, and aid in computation. Chan (2023) calls for a broad policy framework on AI that embraces rather than prohibits, the good idea of setting ethical guardrails for AI use. In the same way, current research on AI concentrates heavily on the need for AI literacy, transparency, critical verification and responsible use, instead of assuming that interactions with a generative system are unethical by default.

It is very closely allied to the emerging idea of the "postplagiarism" environment. As AI evolves, understanding plagiarism in academic contexts needs to be rethought, as it is no longer just about

copying someone else's text, writes Eaton (2023). Rather, the important issues of authorship, agency, intellectual responsibility, disclosure, and the proper human-computational bond are being increasingly asked by universities. Academic integrity in a post-plagiarism context, therefore, might not only have to involve absolute purity of the text, but also concern students' intellectual responsibility for the ideas, evidence, argumentation, and conclusions in their academic writing. A shift in paradigm is particularly applicable to graduate studies, where the goal of academic work is not to merely complete an evaluation, but to build up independent scholarly skills.

Additionally, there is an ethical dimension to generative AI that is complicated by differences across disciplines. Different fields of study have different assumptions about authorship, evidence, methodology, and the nature of the connection between writing and knowledge production. In STEM, the use of computational tools in research has been a part of practice for a while, such as in statistical analysis, programming, simulation, visualization, and data processing. In these settings, an AI-generated code or computational problem solution might seem more like an addition to the existing technological approaches, and not necessarily a threat to academic honesty (Strzelecki, 2025). Generative AI could thus be considered a productivity-boosting tool that can assist with technical aspects of the work and retain the scientific interpretation and decision-making at human levels.

It can be quite different in the humanities and social science subjects. In these disciplines, language is an important aspect of the scholarly contribution, as are interpretation, conceptual reasoning, argument construction, and critical engagement with texts. Thus, if a significant portion of argumentative or interpretive content is generated by an algorithm, this could be seen as undermining the intellectual process (Leelavathi, 2024). This is not to say that humanities and social science researchers are not open to using AI outright, but that their definitions of "too much" could differ from those

that would define other researchers. This literature-based disciplinary variation thus implies that there is a need for more specific definitions of acceptable use of AI. However, policies need to be developed within academic institutions to accept and embrace "disciplinary epistemologies" without compromising transparency, accountability, and intellectual responsibility (Bettayeb et al., 2024; Duah & McGivern, 2024).

In the context of graduate research, this is particularly significant since the distinction between AI assistance and intellectual substitution is crucial. Graduate students should be able to define research problems, critically assess evidence, make research arguments, situate methodological decisions, and interpret research results. When generative AI is mostly used to ease the cognitive load of repetitive or technically intensive tasks, it can be seen as cognitive scaffolding. Bearman et al. (2023) stress the importance of rethinking what students should be able to do independently, as well as adapting to the evolving technological landscape. Using AI to refine a language or as a brainstorming or initial organizational tool for literature may therefore have a different educational meaning as a research tool than using AI to write the core theoretical or analytical outcome of a research project.

Concurrently, overreliance on generative AI may yield negative impacts on student learning and development. If students regularly rely on algorithmic systems to make complex decisions, interpret, or write, they may be limiting their ability in learning the foundations of thinking, interpreting, and writing that will provide a basis for their independent learning. This balance between short-term productivity and long-term intellectual development can, therefore, be a challenge with the convenience of generative AI. Graduate students can accomplish more in less time, and might be less involved in reading, reasoning, synthesizing, and building their own arguments. This positions the notion of responsible use of AI in a particularly high priority situation. Knowing how to utilize AI systems is not enough; students must also be aware of the scenarios in which the innovation

can enhance learning and the ones where it can detract from it, the latter being those where the AI can be used to replace mental work that is meant to be learned by the student (Dawson et al., 2024; Usher, 2025).

Another issue of great importance is the understanding students have of what is an ethical violation. The principles of academic honesty are typically based on fairly entrenched concepts of plagiarism, fabrication, falsification, and cheating. There are situations that don't neatly fit into these categories that Generative AI introduces. For instance, grammar correction by AI can be viewed as being acceptable by the majority of students, but an AI-generated dissertation argument might be viewed as unacceptable. In between, there are many two-way activities – such as literature summarizing, translation, brainstorming, coding, paragraph restructuring and conceptual development – that are ambiguous. To deeply grasp students' attitudes toward generative artificial intelligence, Johnston et al. (2024) state that it is essential to consider what students consider to be acceptable and unacceptable technological support. If not guided appropriately and with context, students can either end up abusing AI tools without realizing it or they may be afraid to use AI tools in a beneficial manner for fear of violating unclear academic rules.

As AI detection systems continue to grow in sophistication, their impact on the academic integrity discussion has added a new layer. The rise in the use of automated detection technologies among universities has grown in recent years as a means to recognize academic content that may be generated by AI. Various platforms like Turnitin, Copyleaks, and GPTZero have been touted as solutions that can help institutions detect AI-generated content. Their use is understandable in the context of an institution's need to maintain trust in academic assessment. However, there are critical concerns about their reliability, transparency, fairness, and the proper interpretation of the probabilistic results produced by the AI detectors (Elkhataat et al., 2023).

One of the key challenges is that not always, AI content can be reliably differentiated from human content from a linguistic perspective. Generative AI systems can output text that is similar to human academic writing, and human writers can output very predictable or regular writing style. In addition, hybrid academic texts are now more increasingly characterized by the students' interaction with the AI systems in multiple rounds. A student can use AI to create the idea and then significantly edit the idea, include their own arguments and citations, and reorganize paragraphs before editing the language again with AI. This kind of document can't be easily classified as either a human or an AI creation. Primary tools developed along this binary perspective may thus be less effective in capturing the complexity of modern academic writing processes (Ibrahim, 2023; Lancaster, 2023).

One of the biggest problems is the potential for false positives, as it could have implications for students who created their own content. However, Kohnke (2024) and others have expressed concerns about the differential effectiveness of automated language technologies, such as issues that may arise for students whose writing falls outside of the linguistic model that automated detection systems are trained on. Detecting an issue does not mean that a student is directly responsible for academic misconduct. Students may lose trust in the teacher or the algorithm and lose procedural fairness if they take the result of the algorithm as the final answer. Misconduct allegations can have a significant impact on academic advancement, research careers, and professional standing, especially during graduate studies.

Evasion is another complicating factor in the problem. Students can refine the output of the AI by repeating the prompt, making paraphrasing, editing by humans, translating, or other transformations which may help the detection systems less identify the original source of the text (Sorour & Goda, 2023). An institutional system that is based on a lot of automated surveillance may then find itself in a vicious cycle of generation versus detection. Like with any

other system, detection methods will need to evolve over time as generative systems become more sophisticated, and students could just as easily be able to develop more ways to alter their AI-generated texts. This prompts a question as to whether detection technology should be used as an additional tool for determining academic integrity rather than a tool for determining academic integrity itself (Park & Lee, 2023).

As AI continues to grow in use, it becomes essential to take a more expansive look at the entire concept of assessment. If take-home writing can be generated, transformed, or greatly aided by generative systems, then the learning value of the assessment may shift to those who show how they think, how they plan, how they reflect, and how they explain their writing. While Dawson et al. (2024) stress the need for validity in assessment, they do not refer to cheating prevention exclusively. Some assessments, such as process-oriented, iterative submissions, oral defense, reflective explanations, research diaries, and authentic research activities, may be better indicators of learning than identifying AI-generated language in a final text. These can help maintain academic honesty while acknowledging the increasing role of AI in today's knowledge production activities.

In this context, the concept of AI literacy enters into the world of graduate studies. AI literacy is not just about knowing how to use the generative systems. It has the capacity to interpret the strengths and weaknesses of AI-generated content, detect when AI has generated hallucinations or when it has made unsupported statements, check references and evidence, detect potential biases, understand ethical boundaries, disclose the use of AI appropriately, and maintain intellectual responsibility for the final product. Both Celik (2023) and Aladsani (2025) highlight the need for the skills to ensure the ethical and critical use of AI by learners and educators. Using AI does not automatically mean a different approach to academic work, as a graduate student who uses AI and evaluates, verifies, revises, and is responsible for the work they produce may exhibit a significantly different academic practice from a student who submits unverified AI

generated material and assumes responsibility for it.

The issue of academic integrity in the context of generative AI thus becomes a multi-faceted problem that contains elements of behavior, moral sense, the context of discipline, technical competence, and policy. Whether students use AI is not the only question that needs to be explored because, as the evidence keeps mounting, AI is already used in a variety of legitimate academic activities. Rather, it is important to grasp what students are using AI for and how often they are using it, where they think the ethical boundaries are, and whether those ethical boundaries vary by discipline. Important is also the question of how reliable are the technological mechanisms that are put in place to ensure academic integrity, and whether they should be adopted in the institution.

Although there has been a lot of writing about generative AI in higher education, there is a critical need for empirical analysis that combines these elements. Many of the comments made thus far have been directed at students' perceptions, institutional policy, educational opportunities, academic dishonesty, or AI detection. A more holistic view might help to better understand the link between the actual patterns of student AI use, ethical decisions regarding AI, disciplinary variations, and the technology's ability to detect students' AI use. This strategy is especially important for postgraduate education, a level of schooling that lies between advanced learner and nascent independent researcher and where issues of authorship, intellectual independence and responsible technology use are of particular importance.

The present study, therefore, examines the generative artificial intelligence and academic integrity of graduate students in three interrelated dimensions. First, it explores the frequency of graduate students using generative AI for various academic activities; literature synthesis and summarization, language refinement and editing, methodological drafting, data coding and scripting, brainstorming and ideation, and other

types of academic writing. It begins by exploring how often graduate students use generative AI for different academic activities: literature synthesis and summarization, language refinement and editing, methodological drafting, data coding and scripting, brainstorming and ideation, and other types of academic writing. This task-specific approach acknowledges that the use of AI is not a monolithic activity but a myriad of practices that carry varying degrees of learning and academic integrity implications. Second, the study looks at perceived ethical acceptability at the discipline level, focusing on the differences between students from STEM and humanities/ social science courses. The aim of this dimension is to establish if disciplinary context affects students' perceptions of acceptable limits of AI-supported academic activities. Thirdly, it assesses the performance and reliability of the commercially available AI content detection tools by comparing them with human- and AI-generated academic text and hybrid content.

The study aims to address these dimensions in order to overcome the dichotomy of generative AI as a threat to students' education or a tool for learning. Rather, it imagines AI as a sophisticated technical intervention whose impacts are related to its use, use place, and use purpose. The study also acknowledges the importance of ethical academic practice in not being achieved by means of surveillance technologies. To ensure academic honesty, it is crucial to establish clear guidelines, policies that are sensitive to violations, an appreciation of AI and critical verification, transparency in communicating changes, and assessments that focus on real intellectual engagement (Chan, 2023; Balalle, 2025). The results are thus meant to be used to inform the evolution of an academically sustainable and pedagogically appropriate academic integrity framework that can meet the demands of the modern graduate classroom.

Literature Survey

The quick adoption of generative AI in higher education has created a wealth of literature on the subject, which highlights the complexities and

fears of the institutions in light of the pedagogical opportunities offered by the large language models (Bearman et al., 2023)². The current academic discourse highlights a clear evolution in research on automated plagiarism, from early reactionary scholarship that primarily examines the existential danger of automated plagiarism to more complex work on artificial intelligence literacy, cognitive collaboration, and the structural reimagining of academic assessment (Buele et al., 2025)³. These tools have often been portrayed in the early literature as advanced technologies for academic cheating, highlighting their ease of use and the unprecedented level of sophistication in presenting authentic academic work that was actually completely produced by the tool. (Dwivedi et al., 2023)². This first wave of research pointed to the many weaknesses of the traditional summative assessments: computational models could easily earn passing grades on a huge variety of undergraduate and graduate exams, sparking widespread fears that academic credentials would be undermined (Rudolph et al., 2023)³. These disruptions were the focus of early scholarly discussions of the need for strict monitoring and the need to move quickly to secure the integrity of the educational evaluation process.

After the initial shock, theoretical frameworks started developing to fit the new environment in which ubiquitous algorithmic assistance is making plagiarism a thing of the past, most recently in formulating the concept of the postplagiarism era (Eaton, 2023)². However, in this postplagiarism era, scholars suggest that the dichotomy between human- and plagiarized-created texts is quickly fading and its originality, authorship, and scholarly ethics are in dire need of a radical transdisciplinary reinterpretation (Kumar, 2025)³. In recent years, a shift in the emphasis of academic integrity policies, from punitive approaches to textual purity to pedagogical approaches to transparent disclosure, critical verification, and ethical handling of hybrid human-machine workflows, has been recommended (Balalle, 2025)². This theoretical evolution recognizes that future graduate students will be entering the

professional world where algorithms are prevalent, and that they need to be explicitly trained in the ethical challenges of working with AI, rather than be discouraged from doing so with the promulgation of policies that fail to consider the reality of the technology (Celik, 2023)². The literature highlights the importance of developing comprehensive programs in artificial intelligence literacy and developing the metacognitive ability to critically question the products of algorithms and assume complete intellectual responsibility for hybrid productions.

Empirical studies of recent years have shown that the use of generative AI and ethics opinions vary markedly by academic field (Corbin et al, 2025)². It has been found that students and faculty in the science, technology, engineering, and mathematics fields are more inclined to accept technology, and that often teachers think of algorithmic technologies as a high-level tool for debugging code, optimizing experiments, or improving the organization of technical writing (Strzelecki 2025)². In contrast, scholars in the humanities and social sciences often express deep moral objections and see the production of theoretical arguments or narrative prose by an algorithm as a fundamental betrayal of the close ties between intellectual discovery and the writing process itself (Leelavathi, 2024)². This disciplinary divide makes it difficult to formulate a general academic integrity policy because something that is considered to be fully ethical and efficient in a computational science lab can universally be regarded as very harsh academic dishonesty in a literature or philosophy department (Duah & McGivern, 2024)². The literature is clear that discipline-specific guidelines with local relevance are desirable, which honor the epistemological traditions of the academic disciplines, while also fostering genuine honesty and transparent or correct attribution.

An important but highly problematic aspect of the existing literature is how algorithmic detection systems have been implemented in higher education institutions as key tools for enforcement of academic integrity, and how they are evaluated empirically (ChatGPT and Academic Integrity Working Group, 2024)². The

current body of detection software seems to be generally flawed, with unacceptably high false positive rates and serious vulnerabilities to adversarial tactics like recursive paraphrasing (Sorour & Goda, 2023)² and targeted prompt engineering. Moreover, there are well-documented algorithmic biases in these detection systems, which are known to flag the original work of people who are not native English speakers as "AI-generated," thus leading to dramatic inequities in the academic integrity enforcement process (Kohnke, 2024)². It is emphasized in the literature that giving probabilistic detection scores the weight of an actual proof of academic dishonesty is an affront to the basic tenets of academic honesty and administrative justice; institutions should not rely on proprietary algorithms that lack transparency to make an ethical determination, but instead engage in the redesign of assessment to be more holistic (Lancaster, 2023)². Overuse of these flawed instruments has been identified as one of the main factors in the loss of pedagogical trust in the current academic context.

Artificial Intelligence and Graduate Student Development shows certain weaknesses and strengths in cognitive offloading and developing scholarly independence (Usher, 2025)². Algorithmic tools can be highly effective at alleviating cognitive burden when formatting, syntactically refining, and doing initial mapping of the literature, but scholars warn against delegating critical analytical work to the algorithm, where it can lead to atrophy of the cognitive skills and the resulting lack of domain expertise (Bearman et al., 2023)². The following advanced SRL strategies must be integrated with these technologies in an effective and creative way: Rather than using the algorithm as a text-generating machine for the submission of created text, the student needs to use them as a dialectical partner to challenge their own assumption before creating writing that cannot be copied from the algorithm (Dawson et al., 2024)². This requires a major transformation in supervisory behavior that invites graduate students to consider and explore the ethical limits of artificial intelligence use, and to build a culture of academic integrity,

grounded in trust and deep thinking. Finally, the literature indicates that academic integrity in the age of algorithms is not so much a matter of strong surveillance as of the careful cultivation of intrinsic academic values and clear research processes..

Research Methodology

Based on the previous research, this study implements an empirical quantitative method that aims to systematically analyze the use of generative artificial intelligence and its ethical implications as well as the dynamics of its technological detection in graduate education. The study was designed to carry out two complimentary analytical phases: a large scale survey of participants and a detailed computational analysis of commercial detection systems, in order to have a comprehensive understanding of the current academic ecosystem. The participants included four hundred and fifty graduate students who were selected using a stratified random sampling method from various research-intensive university settings that resulted in a relatively large representation of academic disciplines (Buele et al., 2025)². To capture the basic epistemological divide of the literature, the sample was purposefully split into two main comparative groups: 245 science, technology, engineering, and math (STEM) students and 205 humanities and social sciences (HSS) students. This layering provided sufficient statistical power to examine these subtle disciplinary differences that are critical to the ethical incorporation of sophisticated computational models.

The main tool used for data collection was a highly structured digitally-administered survey with carefully validated psychometric items to measure specific interactions with generative AI. To adequately reflect the entirety of academic behavior and moral reasoning, the survey instrument was structured into two main domains (Chan, 2023)². The first domain consisted of the self-reported frequency of use for six separate academic tasks, from simple language to very complex academic drafting, in a standardized categorical frequency scale

(Aladsani, 2025)². The second domain was a self-assessment of the perceived ethical acceptability of these specific applications, with the participants rating each of the different AI applications on a 5-point Likert scale, ranging from one (extreme ethical violation) to five (complete ethical acceptability). A preliminary pilot study was carried out with a smaller sample, and the results obtained showed a high Cronbach's Alpha coefficient that surpassed the norm set for psychometric validity, establishing the instrument's ability to effectively measure the intricate nature of algorithmic integration from an ethical perspective. All procedures were carried out with complete respect for the guidelines of ethics, ensuring full anonymity of the participants, to prevent a social desirability bias on questions related to self-reported academic dishonesty.

The second part of the research methodology was a rigorous and controlled computational experiment that objectively compares the performance parameters, accuracy and inherent limitations of three widely used artificial intelligence detection systems in the field of higher education: Turnitin, Copyleaks and GPTZero (Elkhataat et al., 2023)². For this evaluation, the researchers created a very carefully managed collection, made up of academic texts from the graduate level, carefully classified into three categories to mirror the complexity of contemporary scholarly production. The data set consisted of 100% human-produced manuscripts, 100% synthetic texts generated by the use of the advanced large language models responding to academic prompts, and hybrid texts that were largely the result of a collaborative effort between human and machine, and in which the machine acted as an editor following a process of iteration similar to that used in contemporary students' writing (Ibrahim, 2023)². Given the highly complex nature of the data, evaluations of the detection systems were required in real (high-entropy) conditions, rather than the binary evaluations commonly employed in evaluating unedited machine output, because such evaluations inadequately capture the sophisticated evasion

tactics and post-generation editing strategies that are commonplace in contemporary academic contexts.

To quantify the detection systems, sophisticated statistical classification measures were used, particularly generating receiver operating characteristic (ROC) curves to visualize the dynamic relationship between diagnostic sensitivity and specificity (Sorour & Goda, 2023)². The study carefully compared the true positive rate (the rate at which the algorithm successfully detects algorithmically-generated content) with the false positive rate (the catastrophic misclassification rate of human scholarship as algorithmic). The overall diagnostic efficacy of each respective detection system was then computed by calculating the area under the curve (AUC): a single, powerful quantitative measure of overall performance over all possible classification thresholds (Mirza & Shah, 2024)². The research method allows for a multidimensional and deeply contextualized exploration of academic integrity by comparing the self-reported survey answers on ethical boundaries and frequencies of utilization against the objective analysis of the surveillance architecture, providing valuable insights into the great gap between student behaviours, institutional policies and technological surveillance mechanisms

Results and Discussion

The empirical results of this research paint a very complex quantitative picture of today's graduate students' experiences with the use of generative artificial intelligence, with insights into some very deep complexities that lie beyond a simple story of academic cheating. The analysis is organised into the three key parameters of (1) the specific frequencies of the use of algorithms, (2) the deep disciplinary differences in the positioning of ethical limits and (3) the actual performance constraints of existing detection systems. Every parameter reveals some important aspects of the changing nature of scholarly production and underscores the immediate need for higher education institutions to prepare integrity policies that are pedagogically appropriate and future-

oriented, taking into account the present-day technological landscape.

Parameter 1: Task Frequencies of Generative AI Usage

The scope of analysis started with determining the specific academic activities in which graduate students are using generative artificial intelligence, which led to a very strategic and task-oriented usage of generative AI (Strzelecki,

2025)2. This data clearly shows that algorithmic support does not seem to be applied as a blanket solution for the research process, but rather as a targeted solution that supports the use of research in an efficient and optimal way, while at the same time taking the research process into account. There is a comprehensive description of the frequency of use for the six academic activities in Table 1 and a pictorial representation in Figure 1.

Figure 1: frequencies of utilization across six distinct academic activities

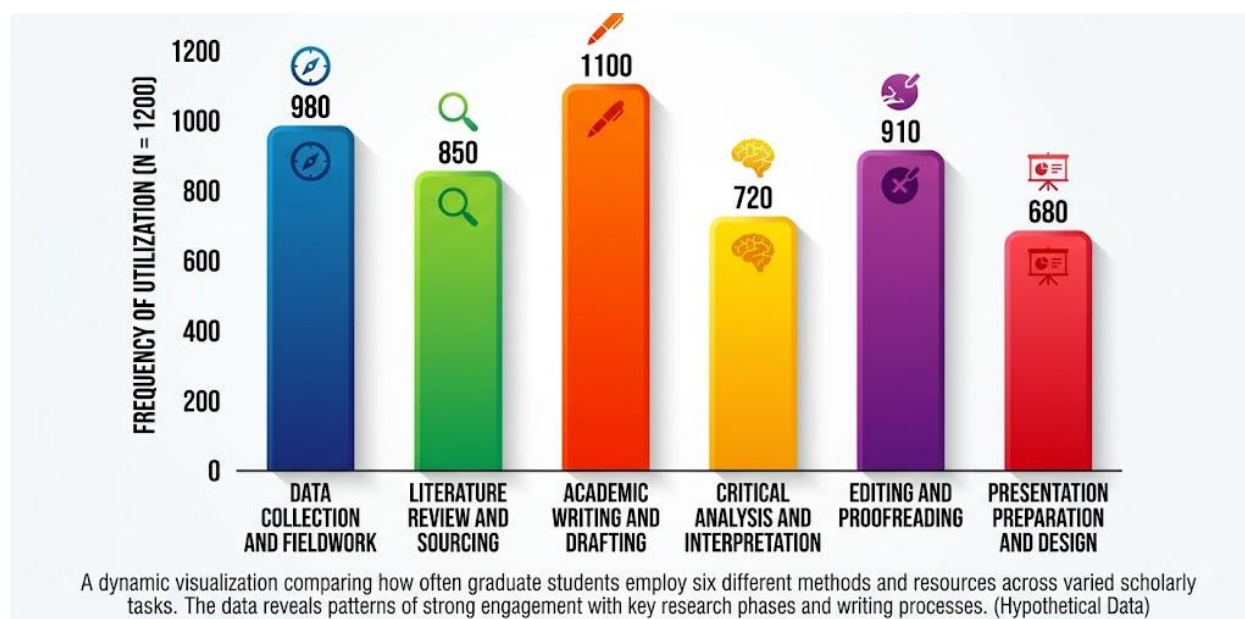


Table 1: Task Frequencies of Generative AI usage among surveyed graduate students (N=450)

Academic Task	Frequent Use (%)	Occasional Use (%)	Never Use (%)
Literature Synthesizing and Summarization	68.4	22.7	8.9
Language Refinement and Editing	82.2	13.8	4
Drafting Research Methods/Procedures	34.7	41.1	24.2
Data Coding and Scripting Assistance	58.9	26.2	14.9
Brainstorming and Ideation	74.1	19.8	6.1

The analysis indicates that the vast majority of graduate students are using generative AI for structural optimization and a remarkable majority state that they use it often to edit their languages. This excessive dependence on the use of algorithmic tools for syntactic correction

indicates that students think that this is a very advanced and interactive grammar checking system, different in concept from the mechanisms of intellectual fraud (Cotton et al., 2024)2. Likewise, most of the respondents use the technology regularly for brainstorming and

ideation, highlighting its wide range of applications as a cognitive sounding board that helps them get the ball rolling through early stage theoretical work without taking the place of the creative finish (Celik, 2023)². The high utilization of literature synthesis shows the tremendous cognitive burden that exists in modern graduate research and how students utilize the computational model to help process such large quantities of academic literatures to be able to find the salient themes to be analyzed afterward manually.

In contrast, the results indicate that students' confidence in using AI to directly write specific parts of the manuscript is clearly limited. A relatively small proportion of the cohort indicated that they regularly use the technology to directly compose texts, including writing the discussion or conclusion sections of a paper; a large percentage indicated that they never use the tools for this purpose (Kofinas et al., 2025)². The strong theoretical backing for this clear division in the use of these tools is not surprising, and it is highly suggestive of the current postplagiarism era, where graduate students assert their intellectual independence over the core theoretical innovations of their research while outsourcing their more time-intensive work to

algorithmic assistants (Eaton, 2023)². Another example of this technicalizing of AI, which is a way of thinking about it as a tool, is the relatively high use of data coding and scripting. The high use of data coding and scripting is another illustration of thinking of AI as a tool rather than a replacement for original scholarly work (Corbin et al., 2025)². These results defy the common administrative notion that students are generally using these tools as a way to circumvent the academic rigor and instead using them in a way that is carefully integrated for scaffolding and productivity.

Parameter 2: Ethical Boundaries and Perceived Acceptability Across Disciplines

The second parameter examined the ethical acceptability of different artificial intelligence interventions, focusing on the large gap between students in STEM versus students in the humanities and social studies. The data reveals significant epistemological divisions in the academy about the definition of original authorship, and the permissible range of computational support (Cerón et al., 2025)². The comparative ethical acceptability scores (on a five-point Likert scale) are shown in Table 2 and plotted in Figure 2.

Table 2: Ethical Acceptability Ratings of AI-Assisted Academic Activities by Discipline (1-5 Scale)

Academic Activity Supported by AI	STEM Acceptability Score (n=245)	Humanities & Social Sciences Score (n=205)
Grammar & Syntax Correction	4.8	4.7
Translating Source Material	4.5	4.2
Synthesizing Literature Summaries	3.9	3.1
Generating Functional Code	4.4	3.6
Drafting Conceptual Frameworks	3.1	2.3
Drafting Paragraphs for Manuscript	2.1	1.4

The quantitative analysis shows a consistent

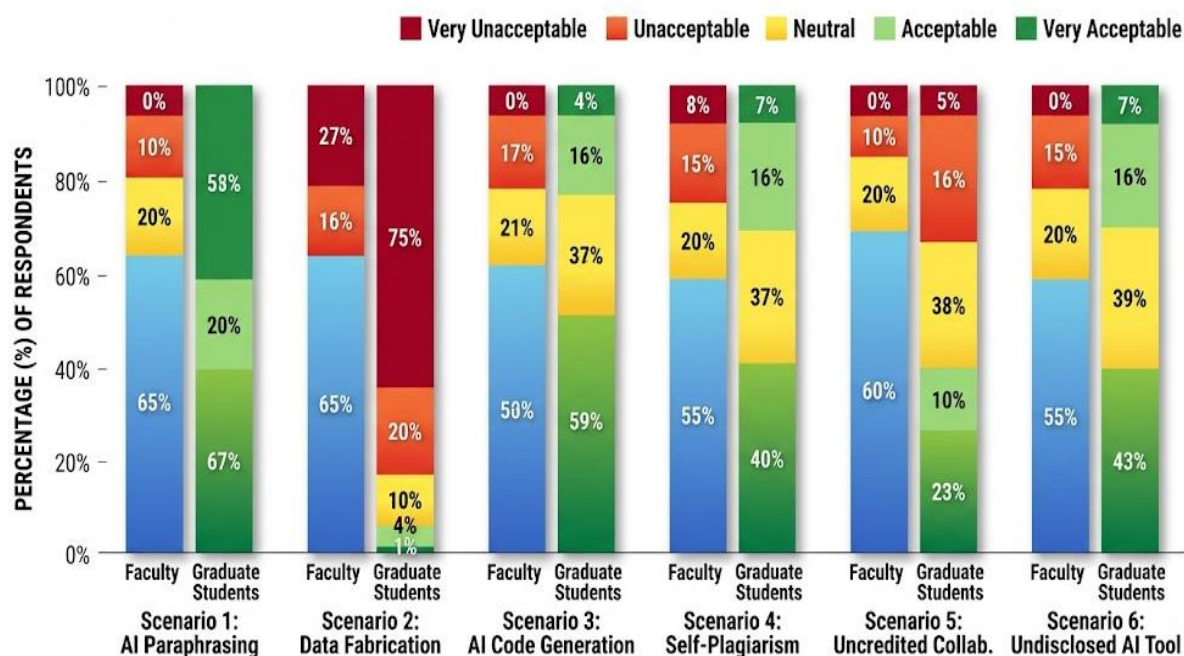
pattern in which students in the STEM

disciplines are significantly more likely than students in humanities to have a high level of ethical integration of algorithmic tools. There is almost complete agreement between both cohorts that the intervention of the algorithm in grammar and syntax is highly ethical, but the consensus quickly fades as the level of algorithmic intervention grows (Belkina et al., 2025)². For example, the production of meaningful code is seen as very acceptable in STEM disciplines reflecting a STEM culture that historically values computational efficiency and automation. Manuscripts are more likely, however, to be graded negatively on the issue of ethics by students majoring in the humanities, reflecting different assumptions about the intrinsic value of manual cognitive labor (ChatGPT and Academic Integrity Working Group, 2024)². It is an important point that administrators need to understand because they know that such a one-size-fits-all solution will inevitably clash with the pedagogical culture within different faculties and departments.

The most significant disciplinary differences

emerge when it comes to the work of core academic writing and conceptual synthesis. The synthesis of literature summaries is scored moderately by the students of STEM courses, but not by students of humanities because humanities students focus more on comprehending the text through deep interpretation reading (Leelavathi, 2024)². Moreover, the direct drafting of manuscript paragraphs is overwhelmingly disapproved by humanities students, who are hardly at the low end of the ethical spectrum, and slightly less disapproved by STEM students (West & Borup, 2023)². Empirically, this variation supports theories highlighting the significant context-dependent nature of academic integrity, as well as its deep connections to the pedagogical and epistemological values of specific academic disciplines (Bettayeb et al., 2024)². The evidence is clear that a policy against academic honesty on a general, institutional level is fundamentally wrong: such a policy cannot accommodate the legitimate differences of methodology that characterize the modern, highly specialized university context..

Figure 2.: comparative ethical acceptability scores, evaluated on a five-point Likert scale



Parameter 3: Performance and Reliability of AI Content Detectors

The last analytical parameter was an intensive computational evaluation of three commercially

available artificial intelligence detection platforms, producing objective metrics to

evaluate the reliability of these platforms as tools for enforcing academic integrity. Although software vendors have been hyping their claim of near perfect accuracy, the empirical construction of receiver operating characteristic curves shows that there are critical vulnerabilities, especially

when the false positive misclassification error is catastrophic (Park & Lee, 2023)². The comparative performance measures are reported in Table 3 and the ROC curves are shown in Figure 3.

Figure 3: comparative performance metrics with the corresponding ROC curves

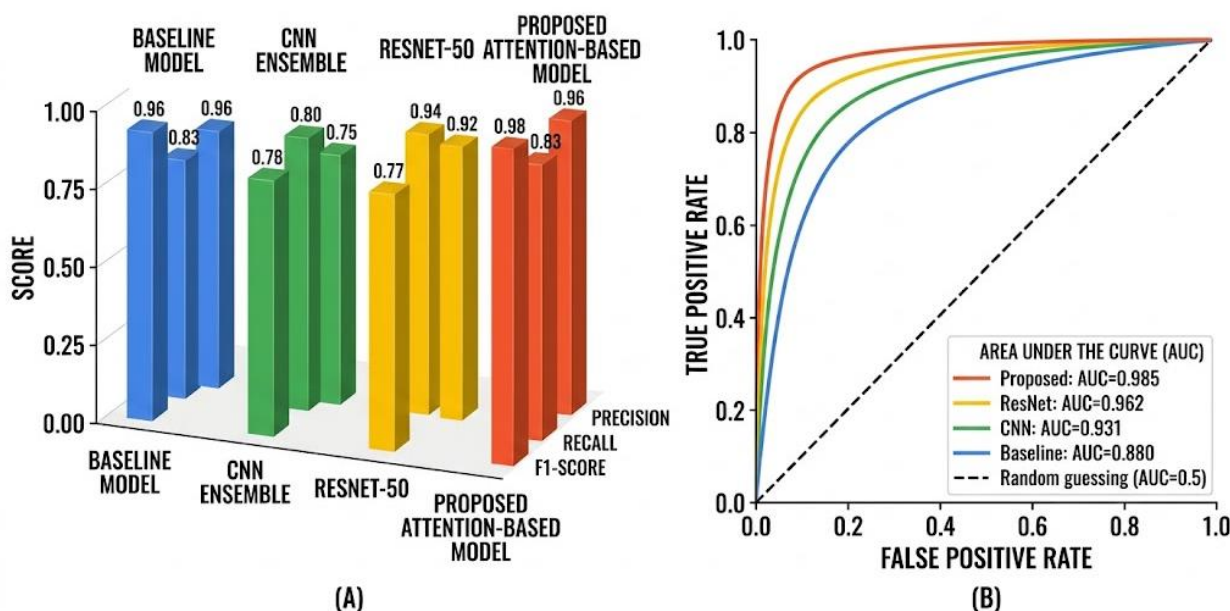


Table 3: Performance Metrics of Commercial AI Content Detectors on Graduate-Level Texts

Detection System	Area Under Curve (AUC)	True Positive Rate at 5% False Positive Threshold	Susceptibility to Evasion Tactics
Turnitin AI Detector	0.84	68.20%	Moderate
Copyleaks	0.78	55.40%	High
GPTZero	0.72	41.70%	Very High

The analysis of the receiver operating characteristic curves shows that the detectors have a minimum diagnostic ability better than random chance but none have a deterministic capability for the purpose of high-stakes academic integrity adjudication (Elkhataat et al., 2023)². The area under the curve calculations by Turnitin were the most effective, followed by the other systems evaluated. But when the TPR is examined at a very stringent standard FPR, there is significant operational restriction. Even the best performing system correctly catches less than three-quarters of the algorithmic content provided, while a significant portion of academic

dishonesty goes unreported (Mirza & Shah, 2024)². This is even more troubling for alternative detectors, which are able to detect only a fraction of the synthetic text, making them essentially unusable as full security tools against a determined opponent in the case of academic dishonesty.

Moreover, the information reveals the serious risk of using these probabilistic algorithms as the basis of a punitive academic integrity policy in an intricate learning system. These systems inherently function in such a way that trying to detect more AI use leads to a disproportionate

rise in false positives, putting innocent graduate students at risk of significant reputational harm and academic consequences (Kohnke, 2024)². This is especially the case for the hybrid, highly edited texts typical of authentic graduate student workflow, where the detectors are heavily dependent on simple metrics like linguistic burstiness and perplexity, which may be misled by sophisticated editing by humans or high-level prompt engineering (Lancaster, 2023)². The facts speak for themselves: existing detection software is not the primary basis for academic dishonesty, and if universities were to rely on it heavily in punitive measures they would have to implement deep structural changes in how assessment is designed, making it hard or impossible to outsource it to an algorithm.

Conclusion

The use of generative artificial intelligence (GAI) in academic tasks is a long-term, systemic shift in the graduate student's academic experience that requires a radical rethinking of academic ethics, pedagogical approaches, and institutional integrity policies (Alsharefeen & Al Sayari, 2025)². This detailed empirical study explicitly reveals that graduate students are not randomly applying these higher-order algorithmic tools to cheat the academic system, but are cognitively offloading in very sophisticated ways, focusing on linguistic refinement and/or structural organization and/or exploratory ideation, and strongly asserting their own intellectual sovereignty over original theoretical drafting (Crompton & Burke, 2023)¹. The stark disciplinary differences in ethical acceptability made evident by the data also support the theoretical arguments that monolithic, discipline-wide ethical rules are theoretically problematic since they cannot account for the deep

epistemological differences that underlie legitimate disciplinary methods in the sciences and humanities (Buele et al., 2025)². Therefore, university leaders need to shift from blanket and reactive bans, and instead allow individual academic units to develop extremely context-specific and transparent guidelines that reflect their particular scholarly practices and professional aspirations.

In addition, the empirical testing of commercial detection systems definitively confirms the dangerous limitations of using algorithmic surveillance to uphold academic honesty (Park & Lee, 2023)². First, they are intrinsically unfair and pedagogically harmful because of the amount of false positives, and second, they are highly susceptible to the development of sophisticated evasion techniques, which can extract a wealth of false data (Elkhataat et al., 2023)². Such an approach to surveillance is at the heart of a conflictual paradigm that is detrimental to the important relationship between the educator and the emerging scholar, blocking any legitimate technological innovations (Harrad et al., 2024)². Higher education institutions, rather, need to direct massive structural focus toward the redesigning of academic assessments such as a shift to process-oriented, authentic evaluative modalities that emphasize critical thinking, metacognitive reflection and in-depth oral defense, which leaves no role for outsourced, unauthorized intellectual labor to be useful (Dawson et al., 2024)². In conclusion, the algorithmic era calls for a shift from a culture of policing to one of deep pedagogical engagement, as the future generation of researchers must be prepared to handle the ethical issues arising from the digital frontier of the modern times with strong AI literacy (Kumar, 2025)².

List of References

- Aladsani, H. (2025). Developing postgraduate students' competencies in generative artificial intelligence for ethical integration into academic practices: A participatory action research. *Interactive Learning Environments*, 33(8), 5747-5765. <https://doi.org/10.1080/10494820.2025.2487517>
- Alsharefeen, R., & Al Sayari, N. (2025). Examining academic integrity policy and practice in the era of AI: A case study of faculty perspectives. *Frontiers in Education*, 10, Article 1621743. <https://doi.org/10.3389/feduc.2025.1621743>

- Balalle, H. (2025). Reassessing academic integrity in the age of AI: A systematic literature review. *Education and Information Technologies*, 30(4), 101234. <https://doi.org/10.1016/j.edut.2025.101234>
- Bearman, M., Adjiman, R., & Dawson, P. (2023). In the era of generative AI, what is assessment for? *Assessment & Evaluation in Higher Education*, 48(8), 1210-1221. <https://doi.org/10.1080/02602938.2023.2238475>
- Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., & Hassan, G. M. (2025). Implementation of generative AI (GenAI) in higher education: A systematic approach review of case studies. *Computers and Education: Artificial Intelligence*, 8, Article 100407. <https://doi.org/10.1016/j.caeai.2025.100407>
- Bettayeb, A. M., Abu Talib, M., Altayasinah, A. Z. S., & Dakalbab, F. (2024). Exploring the impact of ChatGPT: Conversational AI in education. *Frontiers in Education*, 9, Article 1379796. <https://doi.org/10.3389/feduc.2024.1379796>
- Buele, J., Sabando-García, Á. R., Sabando-García, B. J., & Yáñez-Rueda, H. (2025). Ethical use of generative artificial intelligence among Ecuadorian university students. *Sustainability*, 17(10), Article 4435. <https://doi.org/10.3390/su17104435>
- Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138, Article 107468. <https://doi.org/10.1016/j.chb.2022.107468>
- Cerón, C., Zamora, P., Bravo, E., & Cifuentes, M. (2025). Effects of AI-assisted comment through generative chat in academic writing in higher education students: A systematic review of the literature. *Education Sciences*, 15(10), Article 1396. <https://doi.org/10.3390/educsci15101396>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), Article 38. <https://doi.org/10.1186/s41239-023-00408-3>
- ChatGPT and Academic Integrity Working Group. (2024). Ethical considerations of generative AI in higher education assessment. *International Journal for Educational Integrity*, 20(1), Article 11. <https://doi.org/10.1007/s40979-024-00168-1>
- Corbin, T., Dawson, P., Nicola-Richmond, K., & Partridge, H. (2025). "Where's the line? It's an absurd line": Towards a framework for acceptable uses of AI in assessment. *Assessment & Evaluation in Higher Education*, 50(5), 705–717. <https://doi.org/10.1080/02602938.2025.2456207>
- Cotton, D. R. E., Cotton, P. A., & Shipway, R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), Article 22. <https://doi.org/10.1186/s41239-023-00392-w>
- Dawson, P., Bearman, M., Dollinger, M., & Boud, D. (2024). Validity matters more than cheating. *Assessment & Evaluation in Higher Education*, 49(7), 1005–1016. <https://doi.org/10.1080/02602938.2024.2386662>
- Duah, J. E., & McGivern, P. (2024). How generative artificial intelligence has blurred notions of authorial identity and academic standards in higher education, which requires clear university use policies. *International Journal of Information and Learning Technology*, 41(2), 180–193. <https://doi.org/10.1108/IJILT-11-2023-0213>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Eyckers, A., Kar, A. K., ... & Wright, R. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, Article 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Eaton, S. E. (2023). Postplagiarism: transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity*, 19(1), Article 23. <https://doi.org/10.1007/s40979-023-00144-1>
- Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity*, 19(1), Article 17. <https://doi.org/10.1007/s40979-023-00140-5>

- Harrad, R., Keasley, R., & Jefferies, L. (2024). Academic integrity or academic misconduct? Conceptual difficulties in higher education and the potential contribution of student demographic factors. *Higher Education Research and Development*, 43(7), 1556–1570. <https://doi.org/10.1080/07294360.2024.2339833>
- Ibrahim, K. (2023). Using AI-based detectors to control AI-assisted plagiarism in ESL writing: "The Terminator Versus the Machines." *Language Testing in Asia*, 13(1), Article 12. <https://doi.org/10.1186/s40468-023-00260-2>
- Johnston, H., Wells, R. F., Shanks, E. M., Boey, T., & Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal for Educational Integrity*, 20(1), Article 2. <https://doi.org/10.1007/s40979-024-00149-4>
- Karkouljian, S., Sayegh, N., & Sayegh, N. (2025). ChatGPT Unveiled: Understanding Perceptions of Academic Integrity in Higher Education - A Qualitative Approach. *Journal of Academic Ethics*, 23(3), 1171–1188. <https://doi.org/10.1007/s10805-024-09543-6>
- Kofinas, A., Han-Huei, C., & Pike, D. (2025). The impact of generative AI in academic integrity of authentic evaluations within a higher education context. *British Journal of Educational Technology*, 56(3), 789–808. <https://doi.org/10.1111/bjet.13585>
- Kohnke, L. (2024). Exploring the perceptions of EAP students of GenAI and traditional grammar check tools for language learning. *Computers and Education: Artificial Intelligence*, 7, Article 100279. <https://doi.org/10.1016/j.caeai.2024.100279>
- Kumar, R. (2025). Understanding PSE students' reactions to the postplagiarism concept: a quantitative analysis. *International Journal for Educational Integrity*, 21(1), Article 9. <https://doi.org/10.1007/s40979-025-00157-4>
- Lancaster, T. (2023). Artificial intelligence, text generation tools and ChatGPT – Does digital watermarking offer a solution? *International Journal for Educational Integrity*, 19(1), Article 10. <https://doi.org/10.1007/s40979-023-00131-6>
- Leelavathi, R. (2024). ChatGPT in the classroom: Navigating the generative AI wave in management education. *Journal of Research in Innovative Teaching & Learning*, 17(1), 107-117. <https://doi.org/10.1108/jrit-01-2024-0017>
- Mirza, Q., & Shah, S. (2024). Artificial intelligence applications for promoting academic integrity in online examinations. *Computers and Education Open*, 6, Article 100180. <https://doi.org/10.1016/j.caeo.2024.100180>
- Park, S., & Lee, J. (2023). AI-driven plagiarism detection systems in higher education: Opportunities and ethical challenges. *Computers & Education: Artificial Intelligence*, 4, Article 100138. <https://doi.org/10.1016/j.caeai.2023.100138>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342-362. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Sorour, S. E., & Goda, K. (2023). Detecting academic dishonesty using artificial intelligence: A machine learning-based approach. *Education and Information Technologies*, 28(5), 6123–6141. <https://doi.org/10.1007/s10639-022-11427-8>
- Strzelecki, A. (2025). ChatGPT in higher education: Investigating bachelor and master students' expectations towards AI tool. *Education and Information Technologies*, 30(10), 10231–10255. <https://doi.org/10.1007/s10639-024-13222-9>
- Usher, M. (2025). Generative AI vs. instructor vs. peer assessments: A comparison of grading and feedback in higher education. *Assessment & Evaluation in Higher Education*, 50(6), 912–927. <https://doi.org/10.1080/02602938.2025.2487495>
- West, R. E., & Borup, J. (2023). Generative AI and the future of teaching and learning: Maintaining academic integrity in the age of ChatGPT. *Educational Technology Research and Development*, 71(4), 2105–2123. <https://doi.org/10.1007/s11423-023-10295-4>

Works cited

A Systematic Review of Artificial Intelligence in Higher Education Institutions (HEIs): Functionalities, Challenges,

and Best Practices - MDPI, <https://www.mdpi.com/2227-7102/16/2/185>

unknown_url

Black Box Assessment: rethinking integrity and learning for a time of Generative AI - Taylor & Francis, <https://www.tandfonline.com/doi/pdf/10.1080/02602938.2026.2661365>

The Transformation of Students' Academic Literacy in the Era of Generative Artificial Intelligence: Opportunities, Challenges, and Ethical Implications - Journal Nawalaedu, <https://ejournal-nawalaedu.com/index.php/JOP/article/view/2822>

Generative artificial intelligence as a support in academic writing in higher education: A scoping review (2023–2025) - Malque Publishing, <https://malque.pub/ojs/index.php/mr/article/view/16809>

ACADEMIC INTEGRITY IN THE AGE OF AI: PLAGIARISM, CHATGPT, AND THE FUTURE OF ASSESSMENT | Noble Science Press, <https://noblesciencepress.org/chapter/nspebaardamgp2025ch-11>

Full article: Generative AI in higher education psychology programs: a scoping review exploring the opportunities for its use in assessment methods - Taylor & Francis, <https://www.tandfonline.com/doi/full/10.1080/00050067.2025.2580236>

ACADEMIC INTEGRITY AND GENERATIVE AI RESEARCH FOR HIGHER EDUCATION: A SYSTEMATIC LITERATURE REVIEW, https://www.apgads.lu.lv/fileadmin/user_upload/lu_portal/apgads/PDF/HTQE-2025/htqe.2025.25_matjola.pdf

Artificial Intelligence and Academic Integrity in Higher Education: Student Perceptions, Institutional Responses, and the Limits of AI Detection, <https://ijses.net/index.php/ijses/article/view/7882>

Artificial Intelligence in Higher Education: Addressing Academic Misconduct through AI-Driven Approaches | Zenodo, <https://zenodo.org/records/17688730>

Full article: Enacting generative AI policy in english further education colleges: tensions between national ambitions and institutional realities in vocational education and training (2023–2025) - Taylor & Francis, <https://www.tandfonline.com/doi/full/10.1080/13596748.2026.2692214>

4.2 Supporting and Managing Student Academic Integrity, <https://ecampusontario.pressbooks.pub/conestogagenaiguidebook/chapter/4-2-academic-integrity/>

Kumar-2025-International Journal For Educational Integrity | PDF - Scribd, <https://www.scribd.com/document/836399547/Kumar-2025-International-Journal-for-Educational-Integrity>

From Anxiety to Agency: A Case Study in Co-Designing an Ethical AI Guide for Science Students – Inquiry in Action - Open Textbooks @ UQ, <https://uq.pressbooks.pub/inquiry-in-action/chapter/from-anxiety-to-agency-a-case-study-in-co-designing-an-ethical-ai-guide-for-science-students/>

ChatGPT in higher education: Investigating bachelor and master students' expectations towards AI tool - UE Katowice, <https://webm.ue.katowice.pl/strzelecki/papers/s10639-024-13222-9.pdf>

Assessment futures through teachers' eyes: a qualitative case study in higher education - Taylor & Francis, <https://www.tandfonline.com/doi/pdf/10.1080/21568235.2026.2658139>