ijamrsd.com
ISSN: 2583-6404
Journal of Advanced Multidisciplinary Research Studies and Development
Peer Reviewed Referred Journal | Bi-monthly | Open Access

AGE OF CHATGPT: CHALLENGES, RESPONSES, AND RELIABILITY FOR HIGHER EDUCATION

Dr. M S. Patil[1], Mr. N. J. Misal[2], Ashish R. Lahase[3]
[1] *Associate Professor, Department of Physics, Yeshwantrao Chavan College of Arts Commerce and Science Sillod. Dist. Chh. Sambhajinagar.*
Email: mspati9827@gmail.com
[2] *Assistant Professor, Department of Computer Science, Yeshwantrao Chavan College of Arts Commerce and Science. Sillod. Dist. Chh. Sambhajinagar.*
Email: misalnj@gmail.com
[3] *Department of CS&IT, Dr. Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhajinagar.*
Email: ashish.lahase@gmail

Abstract
The rapid spread of Generative Artificial Intelligence tools, especially ChatGPT, has fundamentally changed how students approach academic work, raising serious concerns about integrity, assessment validity, and equity in higher education. This paper reviews existing literature to examine three connected areas: the challenges that ChatGPT poses to academic integrity and learning outcomes; the institutional responses universities have adopted, ranging from outright bans to assessment redesign; and the reliability and fairness of AI-detection tools that have been widely deployed as an enforcement mechanism. The review finds that no single technological solution is sufficient on its own. AI detection software produces probabilistic outputs that are prone to false positives and have been shown to discriminate against non-native English writers. By contrast, authentic assessment redesign combined with structured AI literacy education shows stronger evidence of effectiveness. The paper concludes that higher education institutions need to develop clear, fair, and adaptable governance frameworks, ones that acknowledge AI as a permanent part of academic life rather than a problem to be policed away.

Keywords: Generative AI, Chatgpt, Academic Integrity, Higher Education, AI Detection, Authentic Assessment, AI Literacy, Assessment Design.

► *Corresponding Author: Dr. M S. Patil*

1. Introduction

When OpenAI released ChatGPT in November 2022, few anticipated how quickly it would reshape the academic landscape. Within just two months, the platform had attracted over 100 million active users, making it the fastest-adopted consumer technology on record [1]. For universities and colleges, this was not a gradual shift they had time to prepare for — it was an overnight disruption. Suddenly, a tool capable of writing coherent essays, generating code, and producing research summaries was freely available to any student with an internet connection.

The implications go well beyond convenience. When students can generate polished academic text in minutes, the foundational premise of most assessment systems — that the work submitted represents the student's own thinking — is called into question [2]. This is not simply about

cheating. It touches deeper questions about what education is actually trying to achieve and whether current assessment methods are still fit for purpose.

At the same time, it would be a mistake to frame generative AI solely as a threat. These tools are already embedded in professional life across nearly every field, from medicine and law to engineering and journalism. Universities that respond only by banning or policing AI use risk preparing students for a world that no longer exists [3]. The more productive question is how to integrate AI thoughtfully into higher education, in ways that preserve genuine learning while acknowledging the realities of modern professional practice.

This review addresses three interconnected questions. First, what specific challenges does widespread student access to ChatGPT create for academic integrity, learning outcomes, and equity? Second, what responses have institutions adopted, and how effective have they been? Third, how reliable are the AI detection tools that many universities now use as their primary enforcement mechanism? By examining recent literature and institutional evidence, this paper argues that lasting solutions will require a combination of authentic assessment redesign, AI literacy education, and transparent institutional policy — not technological detection alone.

2. Challenges in the Age of ChatGPT

2.1 Academic Integrity and the Boundaries of Misuse

Academic integrity has always sat at the heart of higher education. The International Center for Academic Integrity defines it through five core values — honesty, trust, fairness, respect, and responsibility — and these values form the basis for most institutional policies around student conduct [4]. What ChatGPT has done is introduce a form of academic dishonesty that these frameworks were not built to handle.

Traditional plagiarism involves copying someone else's existing work. AI-generated text does not do this — it produces original sentences that have never appeared anywhere before. This means that standard plagiarism detection tools, which work by matching text against databases of known sources, return low similarity scores for AI-generated content even when it was written entirely by a machine [2]. The technical infrastructure that universities have relied on for years is simply not designed for this kind of problem.

Adding to the difficulty is the lack of clear normative boundaries around what constitutes misuse. There is general agreement that submitting AI-generated work as one's own, without acknowledgment, is dishonest. But the picture becomes murkier when AI is used to polish grammar, brainstorm ideas, or outline a structure — activities that are difficult to distinguish from using a writing centre, a tutor, or a knowledgeable peer [5]. Without explicit guidance at the course level, both students and staff are left to make individual judgment calls, which leads to inconsistent enforcement and a pervasive sense of unfairness.

Survey research confirms that undisclosed AI use is already common. In one study, approximately 30% of students reported using ChatGPT for assessed work, with a smaller proportion doing so without any disclosure [6]. Given that self-report data typically underestimates socially undesirable behaviour, the actual figure is likely higher. The incentive structure is clear: take-home assessments with generic prompts offer high benefit and low detection risk, making them particularly vulnerable to AI misuse.

2.2 Impact on Critical Thinking and Deep Learning

One of the most frequently raised concerns about ChatGPT is what happens to learning when students routinely outsource cognitive tasks to an AI. Bloom's Taxonomy describes a hierarchy of thinking skills, with higher-order abilities such as analysis, evaluation, and synthesis sitting at the

top [7]. These are precisely the skills that most academic assignments are designed to develop. When a student asks ChatGPT to write an essay rather than wrestling with the ideas themselves, the developmental opportunity is lost.

Research on this is still in its early stages, but initial findings are consistent with this concern. Baidoo-Anu and Owusu Ansah found that while ChatGPT can support learning in structured ways — such as providing explanations or generating practice questions — unguided use tends to produce shallow engagement rather than the deeper processing associated with durable understanding [8]. Students who receive answers rather than working towards them may develop a false sense of competence, which is invisible until it matters most: in supervised exams or unfamiliar problem-solving contexts.

This risk is particularly acute in subjects where knowledge builds sequentially. In mathematics, programming, or language learning, each step depends on mastering the one before it. If students use AI to complete early assignments without actually understanding the material, they may pass courses while accumulating gaps that become critical liabilities later in their academic journey [9]. The consequences may not be visible immediately, but they are real.

2.3 Assessment Validity and the Integrity of Credentials

Beyond individual learning, widespread AI use raises questions about what academic credentials actually certify. A degree is a signal to employers, professional bodies, and society that the holder has acquired certain knowledge and skills. That signal depends on assessment processes that genuinely measure what they claim to measure. When students submit AI-generated work and receive grades for it, the signal breaks down [10].

The empirical evidence on this is striking. Rudolph et al. submitted AI-generated responses to university examination questions and found that human markers were unable to reliably identify them as machine-produced — the AI scripts received passing grades across multiple disciplines [9]. A separate blind study found the same result: AI-generated examination papers consistently fooled university markers [11]. These are not hypothetical risks. They are documented failures of assessment systems that were designed for a different technological environment.

2.4 Equity, Access, and the Ethics of AI Use

Discussions of AI and academic integrity often focus on the students who use it improperly, but there is another equity dimension that deserves equal attention: not all students have the same access to the most capable tools. The basic version of ChatGPT is free, but the more powerful models require paid subscriptions. Students from lower-income backgrounds, or those studying in countries where payment infrastructure is limited, may find themselves at a disadvantage compared to peers who can afford better AI assistance [5].

There is also the question of what AI tools do well and what they do poorly. Large language models are trained predominantly on English-language data, which means their outputs in other languages, or in academic traditions outside the Western mainstream, tend to be of lower quality. This creates an uneven playing field that largely tracks existing inequalities in access to educational resources.

Finally, there is the matter of AI hallucinations — the well-documented tendency of language models to generate plausible-sounding but factually incorrect information with apparent confidence [8]. In academic contexts, where the credibility of cited evidence matters, this is not a minor technical quirk. Students who reproduce AI-generated claims without verification risk propagating misinformation and, more immediately, undermining the quality of their own work.

3. Institutional Responses

3.1 Prohibition: Understandable but Inadequate

The first wave of institutional responses was largely reactive. Several universities moved quickly to ban ChatGPT use outright, and early in 2023, New York City public schools prohibited access to the tool on school networks. The instinct is understandable — when a new threat appears, containment seems like the obvious response. But prohibition has not held up well in practice.

The fundamental problem is enforcement. Students access AI tools on personal devices, through mobile data, and at home. Institutions have no practical mechanism to monitor or prevent this use [12]. Blanket bans therefore do not stop AI use — they simply drive it underground, making it less visible and harder to address. Australia's Tertiary Education Quality and Standards Agency explicitly acknowledged this in its 2023 guidance, noting that prohibition without assessment reform was unlikely to be effective and could be counterproductive [12].

There is also an educational argument against pure prohibition. If graduates are going to work in environments where AI is used routinely — which is already the case in most professional fields — then universities that ban AI entirely are failing to prepare students for that reality [3]. This does not mean permitting all uses of AI uncritically, but it does mean that the conversation needs to move beyond simple prohibition towards something more nuanced.

3.2 AI Detection Tools: A Flawed Safety Net

When outright bans proved difficult to sustain, many institutions turned to AI detection software as a technological substitute. Tools such as Turnitin's AI detection module, GPTZero, and Copyleaks have been integrated into academic integrity workflows at universities around the world. These tools work by analysing statistical properties of text — primarily how predictable successive word choices are (perplexity) and how much sentence length and complexity vary across a document (burstiness) — and producing a probability score intended to indicate the likelihood of AI authorship [13].

In principle, this sounds like a reasonable solution. In practice, the limitations are serious enough that many scholars have questioned whether these tools should be used as enforcement mechanisms at all [14]. The scores they produce are probabilistic, not definitive. A report stating that a document is '95% AI-generated' does not mean there is a 95% probability the document was written by AI; it means the text falls near a particular end of a statistical distribution. That distinction matters enormously when the output of the tool is used as evidence in a misconduct investigation.

3.3 Authentic Assessment Redesign

While detection tools address the symptom, assessment redesign addresses the underlying vulnerability. The core insight is straightforward: if assessments are designed so that AI cannot plausibly complete them without substantial human input, the integrity problem is largely prevented rather than just detected after the fact [15]. Authentic assessment, a concept well-established in the pedagogical literature before ChatGPT arrived, becomes a natural and effective response to the challenge.

Oral examinations are perhaps the most AI-resistant format available. A student can have AI write an essay, but they cannot have AI sit a viva on their behalf. When the assessment requires real-time articulation and defence of ideas, the underlying knowledge — or its absence — becomes immediately apparent. Staged assessments that require submission of drafts, annotated source lists, and reflective commentary are similarly resistant to wholesale AI outsourcing, because they make the intellectual process visible rather than just the final product [10].

Reflective portfolios, case-based assessments tied to specific course discussions, and tasks grounded in personal professional experience are also difficult for AI to generate convincingly, because they draw on knowledge and experience that is specific to the individual student. Villarroel et al. argue that authentic assessment is not simply an AI-deterrent but a pedagogically superior approach in its own right — one that develops more transferable competencies and produces more durable learning [15]. The arrival of ChatGPT has created a new urgency around a pedagogical reform that was already overdue.

Dawson makes the further point that authentic assessment tends to improve student engagement and motivation [10]. When students work on tasks that feel meaningful and connected to real-world contexts, they are more likely to invest genuine effort — regardless of whether AI shortcuts are available. Assessment redesign, in this sense, is not just a defensive strategy but a positive educational one.

3.4 AI Literacy and Disclosure Frameworks

A number of institutions have moved in a different direction: rather than trying to prevent AI use, they have chosen to teach students how to use it responsibly. AI literacy education covers the technical basics of how large language models work, their known limitations and failure modes, the ethical dimensions of AI use and attribution, and practical skills in critically evaluating and verifying AI-generated content [8]. The goal is to develop graduates who can engage with these tools thoughtfully and professionally, not students who have learned to hide their AI use to avoid penalties.

Several leading universities have developed explicit AI use and disclosure frameworks that reflect this approach. Harvard University's guidelines, for example, permit AI as a learning aid in many contexts while requiring students to disclose where AI has made a substantive contribution to their work [16]. The Cornell Center for Teaching Innovation has developed similar guidance emphasising transparency and critical use rather than restriction [17]. These frameworks parallel the conventions that already exist around citing secondary sources or acknowledging editorial assistance — they extend existing academic honesty principles rather than inventing new ones.

Evidence suggests that this approach has measurable benefits. Lodge et al. found that clear AI use policies, combined with explicit instruction in AI literacy, were associated with higher rates of voluntary disclosure and lower rates of covert misuse [3]. When students understand both what is expected of them and why, and when the rules are consistent and clearly communicated, the normative pressure to be honest is stronger.

4. Reliability of AI Detection Tools: A Critical Review

4.1 How Detection Tools Work — and Where They Fall Short

To understand the limitations of AI detection tools, it helps to understand what they actually measure. Most commercial tools rely on two statistical signals. Perplexity measures how predictable a sequence of words is: AI models tend to choose high-probability word sequences, resulting in lower perplexity than typical human writing. Burstiness measures variation in sentence length and complexity: humans naturally write with more variation than AI models, which tend toward more uniform sentence structures. Detection tools combine these signals using proprietary classifiers trained on labelled datasets [13].

The problem is that these signals are imperfect proxies for human versus AI authorship. Weber-Wulff et al. carried out a systematic evaluation of fourteen commercial and open-source detection tools, testing them on a controlled dataset of human-written and AI-generated texts. They found substantial variation in accuracy across tools, and — critically — no tool performed at a level that

would be considered adequate for use as primary evidence in an academic misconduct proceeding [13]. Worse, the accuracy of these tools is a moving target: as AI writing models improve, the statistical signatures that detection tools rely on become less distinctive, and the accuracy of existing detectors tends to degrade.

4.2 The Problem of False Positives

False positives — cases where genuinely human-written text is incorrectly flagged as AI-generated — are the most serious practical failure mode of detection-based enforcement. The consequences for a student who receives a false positive are severe: a formal misconduct allegation, potential suspension or expulsion, lasting reputational damage, and significant psychological harm [14]. Even when an investigation ultimately clears the student, the experience of being accused of academic dishonesty can have lasting effects on their confidence and relationship with their institution.

There is documented evidence of students facing misconduct proceedings on the basis of detection scores alone, in some cases before institutions had even published clear policies about what AI use was and was not permitted [2]. The opacity of detection algorithms means that in most such cases, students have no meaningful way to challenge the result: neither they nor their institution has access to the reasoning behind the score, making it effectively unappealable.

4.3 Bias against Non-Native English Writers

One of the most troubling findings in the detection tool literature concerns differential treatment of non-native English writers. Liang et al. conducted a controlled study comparing detection rates for essays written by verified native and non-native English speakers. They found that non-native speaker essays were flagged as AI-generated at substantially higher rates — not because they were more likely to have used AI, but because the stylistic features associated with writing in a second language overlap with the features these tools interpret as AI-generated [18]. Simpler sentence structures, more predictable vocabulary choices, and less stylistic variation are all characteristics of second-language academic writing that also resemble the output of a language model.

The implications of this finding are difficult to overstate. India, China, the Middle East, and sub-Saharan Africa together contribute millions of students to international higher education systems, the majority of whom write academic English as a second or third language. If AI detection tools are systematically more likely to flag their work as AI-generated, then these students face a higher risk of false misconduct allegations purely as a consequence of their linguistic background [18]. This is not a minor edge case — it is a structural bias embedded in the enforcement mechanism that many institutions have already adopted.

4.4 Transparency and Due Process

A final concern about AI detection tools is the opacity with which they operate. Most commercial tools do not disclose their training data, model architecture, or decision thresholds. This black-box design creates a fundamental due process problem: when an institution uses a detection score to initiate misconduct proceedings against a student, that student has no meaningful way to examine or challenge the evidence against them [14]. The basic requirements of procedural fairness — access to evidence, the opportunity to respond, a comprehensible basis for the decision — are not met.

Perkins et al. and Weber-Wulff et al. have both argued that in its current form, AI detection output does not meet the evidentiary standard required to support formal misconduct proceedings [13][14]. Their position is that detection scores can appropriately trigger a conversation or a closer look at a submission, but they should never serve as the primary or sole basis for a finding of misconduct. An institution that uses a probabilistic algorithm to discipline students without corroborating evidence is, in effect, punishing students for having a particular writing style rather than for a proven act of dishonesty.

5. Synthesis and Discussion

5.1 Teaching Students to Use AI Rather Than Hide It

Across the evidence reviewed in this paper, one theme recurs consistently: responses that treat AI as something to be concealed or defeated tend to be less effective than responses that treat it as something to be understood and governed. This has direct implications for pedagogy. Rather than building courses around the assumption that students should not use AI, the more durable approach is to build courses in which students develop critical literacy around AI — knowing when it is appropriate, how to evaluate its outputs, how to supplement its weaknesses, and how to be transparent about their use of it [8].

A graduated model of AI integration makes sense in this context: early-stage tasks where AI is permitted as a scaffold, with later assessments requiring individual demonstration of the competencies developed. This mirrors how AI is actually used in professional settings — to augment human judgment, not replace it — and prepares students more honestly for the environments they will enter [5].

5.2 Redesigning Assessment for the AI Age

The single most effective structural response to the challenge posed by GenAI is probably the redesign of assessment itself. Tasks that require individual experience, real-time performance, or demonstrated process are inherently more resistant to AI substitution than generic take-home essays [15]. But the case for redesign should not rest only on AI resistance. As Villarroel et al. argue, authentic assessment is pedagogically superior regardless of the AI context, producing deeper learning and more transferable competencies [15]. The AI challenge has simply made an already overdue reform more urgent.

Implementing this at scale requires investment in staff development. Many academic staff lack the confidence or the practical knowledge to design assessments that engage productively with AI rather than simply trying to exclude it [3]. Institutions that take this seriously will need to support their faculty with time, training, and structural incentives for curriculum development. This is a non-trivial institutional commitment, but the alternative — continuing to rely on detection tools that do not work reliably and that carry serious equity risks — is not a sustainable position.

5.3 Building Policy That Is Clear, Fair, and Adaptable

Clear policy is a precondition for fair enforcement. When students do not know what AI use is permitted and what is not, when rules vary from one course to the next without explanation, and when sanctions are applied inconsistently, the perception of arbitrary treatment is corrosive to institutional trust [4]. Universities need to establish AI use policies at the institutional level that are coherent and well-communicated, while still leaving room for course-level variation where disciplinary context genuinely requires it.

Where AI detection tools are used, their role should be clearly defined as investigative rather than evidentiary. A detection score might appropriately prompt a tutor to look more closely at a submission, or to invite a student to discuss their work. It should not, without additional corroborating evidence, form the basis of a formal misconduct allegation [13][14]. Institutions should also conduct and publish regular audits of the tools they use, including data on false positive rates and any differential rates across student groups, in order to maintain accountability for the tools they have chosen to deploy [12].

6. Conclusion

The arrival of ChatGPT in higher education has exposed structural vulnerabilities that were always present in conventional assessment design but had not previously been exploited at this scale or

speed. The challenges it creates — for academic integrity, for the development of genuine skills, for equitable treatment of students, and for the credibility of academic credentials — are real and serious [2,9]. They deserve responses that are equally serious.

This review has argued that detection-focused responses are insufficient and, in their current form, actively harmful in some respects. AI detection tools are statistically unreliable, prone to false positives that can damage students' lives, and demonstrably biased against non-native English writers [18,13]. Deploying them as a primary enforcement mechanism without acknowledging these limitations is an institutional failure of due process.

The more promising path involves investing in authentic assessment design, building AI literacy into the curriculum, and developing policies that are clear, fair, and honest about the limitations of what enforcement alone can achieve [10,15]. These are not new ideas, but the urgency is new. Institutions that respond thoughtfully to this moment have an opportunity not just to manage a disruption but to improve the quality and relevance of the education they offer.

The goal, ultimately, is not to put AI back in the box — that is not possible, and probably not desirable. It is to help students develop the judgment and integrity to use powerful tools responsibly, which is precisely what higher education has always been for [3].

References

1. K. Hu, "ChatGPT sets record for fastest-growing user base — analyst note," Reuters, Feb. 2, 2023. [Online]. Available: https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/
2. D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," Innovations in Education and Teaching International, vol. 60, no. 6, pp. 1-12, 2023. doi: 10.1080/14703297.2023.2190148
3. J. M. Lodge, K. Thompson, and L. Corrin, "Mapping out a research agenda for generative artificial intelligence in tertiary education," Journal of Applied Learning and Teaching, vol. 6, no. 1, pp. 1-10, 2023. doi: 10.37074/jalt.2023.6.1.34
4. International Center for Academic Integrity (ICAI), The Fundamental Values of Academic Integrity, 3rd ed. Clemson, SC: Clemson University, 2021. [Online]. Available: https://academicintegrity.org/
5. M. Perkins, "Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond," Journal of University Teaching & Learning Practice, vol. 20, no. 2, pp. 1-24, 2023. doi: 10.53761/1.20.02.07
6. M. Khalil and E. Er, "Will ChatGPT get you caught? Rethinking of plagiarism detection," in Proc. 15th Int. Conf. Computer Supported Education, 2023, pp. 103-112. doi: 10.5220/0011996100003470
7. L. W. Anderson and D. R. Krathwohl, Eds., A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Educational Objectives. New York, NY: Longman, 2001.
8. D. Baidoo-Anu and L. Owusu Ansah, "Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning," Journal of AI, vol. 7, no. 1, pp. 52-62, 2023. doi: 10.61969/jai.1337632
9. J. Rudolph, S. Tan, and S. Tan, "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?" Journal of Applied Learning and Teaching, vol. 6, no. 1, pp. 1-30, 2023. doi: 10.37074/jalt.2023.6.1.9
10. P. Dawson, "Authentic assessment for a digital world," Assessment & Evaluation in Higher Education, vol. 48, no. 1, pp. 1-12, 2023. doi: 10.1080/02602938.2021.1956036

11. The Guardian, "Researchers fool university markers with AI-generated exam papers," Jun. 26, 2024. [Online]. Available: https://www.theguardian.com/
12. Tertiary Education Quality and Standards Agency (TEQSA), "Guidance note: Artificial intelligence and academic integrity," Australian Government, 2023. [Online]. Available: https://www.teqsa.gov.au/
13. D. Weber-Wulff, A. Anohina-Naumeca, S. Bjelobaba, T. Foltynek, J. Guerrero-Dib, O. Popoola, P. Sigut, and L. Waddington, "Testing of detection tools for AI-generated text," International Journal for Educational Integrity, vol. 19, no. 1, pp. 1-20, 2023. doi: 10.1007/s40979-023-00146-z
14. M. Perkins, J. Roe, D. Postma, J. McGaughran, and D. Hickerson, "Game of tones: Faculty recognition of GPT-4 generated text in university assessments," arXiv preprint arXiv:2305.18226, 2023. [Online]. Available: https://arxiv.org/abs/2305.18226
15. V. Villarroel, S. Bloxham, D. Bruna, C. Bruna, and C. Herrera-Seda, "Authentic assessment: Creating a blueprint for course design," Assessment & Evaluation in Higher Education, vol. 43, no. 5, pp. 840-854, 2018. doi: 10.1080/02602938.2017.1412396
16. Harvard University, "Generative AI guidance for students," Office of Undergraduate Education, 2023. [Online]. Available: https://college.harvard.edu/
17. Cornell Center for Teaching Innovation, "Artificial intelligence and academic integrity," Cornell University, 2023. [Online]. Available: https://teaching.cornell.edu/
18. W. Liang, M. Yuksekgonul, Y. Mao, E. Wu, and J. Zou, "GPT detectors are biased against non-native English writers," Patterns, vol. 4, no. 7, p. 100779, 2023. doi: 10.1016/j.patter.2023.100779