AI-Generated Content in Education

Opportunities and Academic Integrity
CSCI 3250 Computers and Society — Group Project Report

Group 30

Haonan JIAO	1155211226	Introduction
Chenyan XU	1155173847	Technical Introduction
Ruixuan XU	1155211213	Case Study 1
Xiangxiang WENG	1155211173	Case Study 2
Yuanpeng WU	1155193279	Discussion
Yihao TAN	1155211107	Conclusion

April 20, 2026

Abstract

The rapid proliferation of large language models (LLMs) has created both significant opportunities and serious challenges for educational institutions worldwide. This report examines the dual nature of AI-generated content in education through an analysis of the underlying technology, two real-world institutional cases, and the ethical tensions that make this issue difficult to resolve. We argue that sustainable responses require not prohibition but the development of new assessment practices, AI literacy frameworks, and institutional policies that distinguish legitimate AI assistance from academic dishonesty.

Contents

Abstract 1

1 Introduction: AI Enters the Classroom 2

1.1 What Is AI-Generated Content? 2

1.2 Why This Topic Matters 2

2 Understanding the Technology: Large Language Models 2

2.1 How LLMs Work 2

2.2 Key Capabilities and Applications in Education 3

3 Real-World Cases 3

3.1 Case 1: New York City Public Schools and the ChatGPT Ban 3

3.1.1 Background and the Initial Ban 3

3.1.2 The Reversal and Its Significance 3

3.1.3 Lessons from the NYC Case 4

3.2 Case 2: The Unreliability of AI Detection Tools 4

3.2.1 The Promise and the Problem 4

3.2.2 False Positives and Their Consequences 4

3.2.3 Implications for Institutional Policy 5

4 Discussion: What Makes This Issue Complicated 6

4.1 The Problem of Fake Mastery 6

4.2 Using versus Delegating 6

4.3 Values in Tension 6

5 What Should Happen Next 7

5.1 Proposed Solutions 7

5.2 Roles for Different Stakeholders 7

6 Conclusion 8

AI Usage Declaration 8

Team Contributions 8

References 9

1 Introduction: AI Enters the Classroom

1.1 What Is AI-Generated Content?

Artificial intelligence has moved from research laboratories into everyday life at an unprecedented pace. Among the most visible manifestations of this shift is AI-generated content—text, images, code, and other media produced by machine learning models rather than human authors. In educational settings, tools such as ChatGPT, Gemini, and Copilot are now routinely used by students to draft essays, solve problem sets, write code, and summarise readings.

The question is no longer whether AI is present in classrooms—it already is. As the CSCI 3250 course itself acknowledges, generative AI tools are in use across virtually every discipline. The real question is how educators, institutions, and students should respond.

1.2 Why This Topic Matters

Academic integrity is a cornerstone of higher education. Degrees and credentials carry value precisely because they certify that the holder has demonstrated mastery of a body of knowledge. When AI can produce plausible, high-quality work on demand, the meaning of that certification is called into question.

At the same time, blanket prohibition is neither practical nor necessarily desirable. AI tools can democratise access to high-quality tutoring, assist students with language barriers, and accelerate exploration of complex topics. The tension between these two realities makes this one of the most pressing issues in contemporary education, and one that demands careful, evidence-based analysis rather than reactive policy.

2 Understanding the Technology: Large Language Models

2.1 How LLMs Work

Large Language Models (LLMs) are neural networks trained on massive corpora of text data. The core mechanism is deceptively simple: given a sequence of words, the model learns to predict the most probable next word. Repeated billions of times across trillions of tokens of text, this training process produces a model that can generate coherent, contextually appropriate language across an enormous range of topics [1].

The architecture underlying modern LLMs is the Transformer, introduced by Vaswani et al. in 2017 [2]. Transformers use a mechanism called self-attention to weigh the relevance of every word in a context window against every other word, allowing the model to capture long-range dependencies

in language far more effectively than earlier recurrent architectures [3]. It is this architectural innovation that enabled the leap from narrow language tools to general-purpose systems capable of reasoning across domains.

2.2 Key Capabilities and Applications in Education

Modern LLMs can generate fluent text, answer questions across disciplines, write and debug code, summarise documents, and translate between languages. These capabilities position them as potentially powerful study aids: a student struggling with a calculus concept can request a step-by-step explanation tailored to their level, while a non-native English speaker can use an LLM to refine their writing without altering their underlying argument.

However, the same capabilities that make LLMs useful also make them easy to misuse. Because an LLM can produce a complete, well-structured essay on virtually any topic within seconds, the barrier to submitting work that does not reflect a student's own understanding has been dramatically lowered. This is the central tension that educational institutions are now grappling with.

3 Real-World Cases

3.1 Case 1: New York City Public Schools and the ChatGPT Ban

3.1.1 Background and the Initial Ban

In January 2023, the New York City Department of Education (NYCDOE) announced a ban on ChatGPT across all school devices and networks. The decision reflected widespread concern among educators that students would use the tool to complete assignments without engaging with the underlying material. Specifically, there were deep fears that the instant generation of essays would bypass the productive struggle necessary for developing critical thinking and fundamental writing skills. The ban was one of the first large-scale institutional responses to generative AI in education and attracted significant media attention. It represented a precautionary approach: when faced with an uncertain and potentially disruptive technology, restrict access until the risks are better understood.

Figure 1: New York City Public Schools Chancellor David Banks, pictured June 27 in New York, wrote in an op-ed published Thursday that "students are participating in and will work in a world where understanding generative AI is crucial."

3.1.2 The Reversal and Its Significance

The ban was short-lived. By May 2023—just four months later—the NYCDOE reversed course and lifted the restriction [5]. The reversal was accompanied by a statement from David Banks, Chancellor of New York City Public Schools, who noted that students are participating in and will work in a world where understanding generative AI is crucial. Instead of fearing the technology, the district began encouraging educators to explore how AI could be utilized as a supportive teaching aid and a subject of critical discussion in classrooms. This shift in position reflects a broader recognition that prohibition is not a sustainable strategy. Students who graduate without exposure to AI tools may be at a disadvantage in a workforce that increasingly relies on them.

3.1.3 Lessons from the NYC Case

The episode illustrates several important tensions inherent in institutional responses to rapidly evolving technology. There is a fundamental conflict between short-term academic integrity and long-term workforce preparation. Top-down bans are also difficult to enforce in practice, since students can access ChatGPT on personal devices regardless of school network restrictions. Furthermore, strict network bans can create unintended equity issues, disproportionately affecting lower-income students who rely solely on school-provided devices for internet access. Most importantly, the rapid reversal suggests that the appropriate institutional response is not prohibition but guided integration with clear ethical boundaries—a conclusion that has since been echoed by educational bodies worldwide.

3.2 Case 2: The Unreliability of AI Detection Tools

3.2.1 The Promise and the Problem

Figure 2: VeriGuide, one of the AI detection tools widely used in Hong Kong universities.

As AI-generated content became more prevalent, a market emerged for tools designed to detect it. Products such as VeriGuide, Turnitin's AI detector, and GPTZero promised to identify AI-written text, offering educators a technological solution to the integrity problem. However, these tools have proven deeply unreliable. Their core limitation is fundamental: they attempt to distinguish AI-generated text from human-written text based on statistical patterns, but those patterns are neither stable nor unique to AI [6].

3.2.2 False Positives and Their Consequences

The most damaging failure mode is the false positive—flagging human-written text as AI-generated. This has real consequences for students who are wrongly accused of academic dishonesty. A widely cited illustration of this problem is the case in which an AI detection tool flagged the United States Declaration of Independence as AI-generated [7].

The document, written by Thomas Jefferson in 1776, obviously predates any AI system. The result underscores the fundamental unreliability of these tools: they respond to surface-level statistical features of text, not to any genuine understanding of authorship.

The problem is not limited to historical texts. Research has shown that non-native English speakers are disproportionately flagged by AI detectors, because their writing patterns—shorter sentences, simpler vocabulary, more predictable structure—statistically resemble AI output. This creates a discriminatory dynamic in which students who are already disadvantaged by language barriers face

additional and unjust scrutiny.

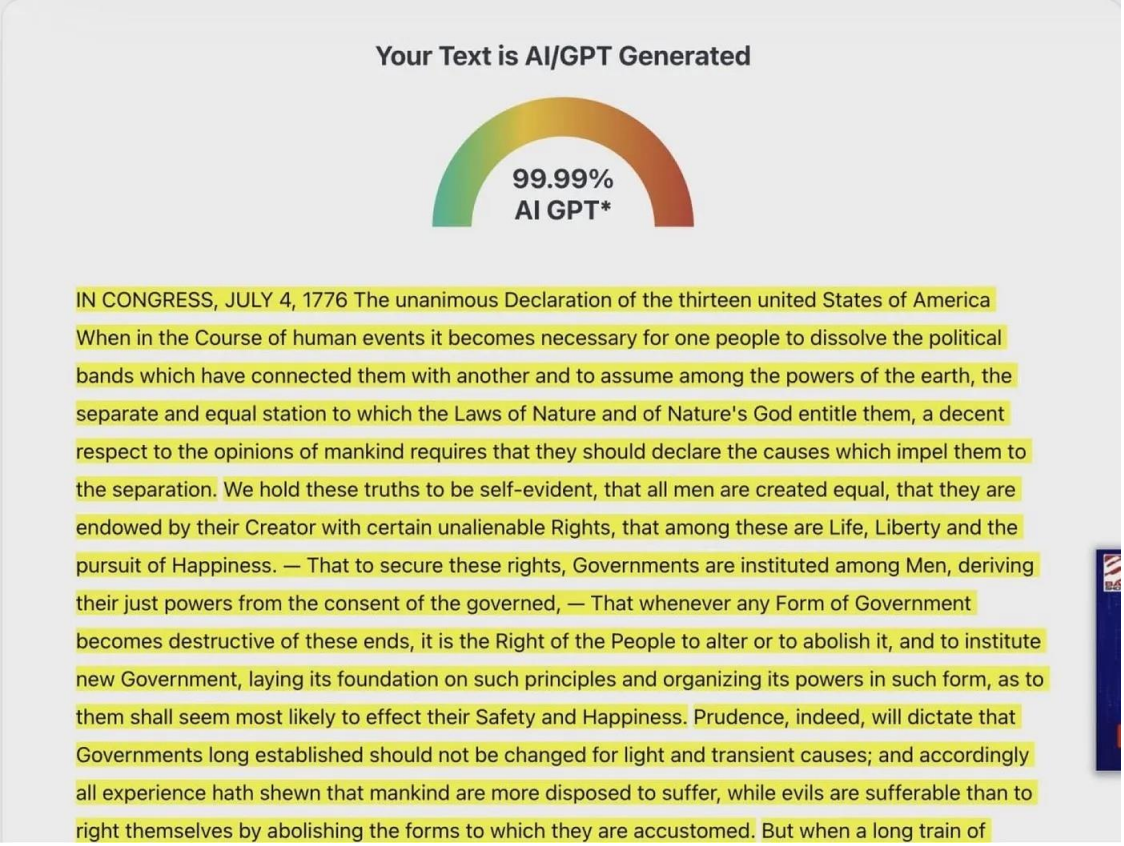

Figure 3: The U.S. Declaration of Independence (1776), famously flagged as AI-generated by an automated detection tool.

3.2.3 Implications for Institutional Policy

The unreliability of AI detection tools means that institutions cannot rely on them as a primary enforcement mechanism. Using flawed detectors as the basis for academic misconduct proceedings risks serious injustice. The technological arms race between AI generation and AI detection is unlikely to produce a reliable solution in the near term, and institutions that invest heavily in detection-based enforcement strategies may find themselves both ineffective and exposed to legal challenge.

4 Discussion: What Makes This Issue Complicated

4.1 The Problem of Fake Mastery

One of the deepest concerns about AI in education is the phenomenon of fake mastery: the appearance of competence without the underlying understanding. When a student submits an AI-generated essay, they may receive a good grade without having engaged with the material, developed their analytical skills, or practised the craft of writing. The credential they earn may not reflect genuine capability.

The analogy to calculators is instructive but limited. Calculators are now universally accepted in mathematics education, but their introduction required a rethinking of what skills students need to develop. We no longer expect students to perform long division by hand, but we do expect them to understand what division means and when to apply it. The challenge with AI is that it can perform not just computation but reasoning, synthesis, and communication—the very skills that higher education is designed to develop. If students outsource these processes entirely, the educational mission is undermined in a way that calculator use never threatened.

4.2 Using versus Delegating

A crucial distinction exists between using AI as a tool and delegating thinking to AI. Using AI might mean asking it to explain a concept, check grammar, or suggest alternative phrasings—activities where the student remains the author of their own thinking. Delegating means asking AI to generate the argument, structure the essay, and produce the analysis, with the student serving merely as a conduit between the AI and the submission portal.

The key variable is intellectual presence: whether the student is genuinely engaged in the work, or has absented themselves and allowed the AI to take over. This distinction is philosophically clear but practically difficult to enforce. The boundary between assistance and delegation is not always obvious, and it varies by context, discipline, and assignment type. A student who uses AI to brainstorm ideas and then develops them independently occupies a very different ethical position from one who submits lightly edited AI output as their own work.

4.3 Values in Tension

The AI-in-education debate involves genuine conflicts between important values that cannot be re-solved by technical means alone. Restricting AI preserves traditional assessment validity but may leave students unprepared for an AI-integrated workplace. AI tools can level the playing field for students with fewer resources, yet unreliable detection disproportionately harms non-native speakers. The productive struggle of doing difficult work is often where learning happens, yet AI can

accelerate task completion in ways that short-circuit that struggle. Rapid AI adoption may improve certain outcomes but destabilises established assessment frameworks before adequate replacements are ready. Any policy response involves trade-offs, and different stakeholders—students, educa-tors, employers, policymakers—will weigh these values differently depending on their priorities and contexts.

5 What Should Happen Next

5.1 Proposed Solutions

Given the limitations of both prohibition and unreliable detection, more constructive approaches are needed. The most promising direction is a shift in assessment design from evaluating prod-ucts to evaluating processes. Rather than relying solely on the final submitted work, assessments should incorporate process evidence such as drafts, revision histories, oral defences, and in-class components. If a student can explain and defend their work in real time, the question of whether AI assisted in its production becomes substantially less critical.

A complementary approach is to improve auditability through interaction logs. Institutions could require students to submit records of their AI interactions alongside their work, making the nature and extent of AI assistance transparent. This shifts the institutional focus from detection to disclosure, which is both more honest and more practically achievable.

Beyond assessment reform, institutions should invest in active AI literacy education. Rather than leaving students to navigate AI use without guidance, curricula should explicitly address the ethical dimensions of co-creation with AI. Students need to understand not just how to use AI tools, but when it is and is not appropriate to do so, and why the distinction matters for their own intellectual development. Finally, the NYC case demonstrates that bans are not sustainable. Institutions need coherent ethical frameworks that distinguish legitimate AI assistance from academic dishonesty, developed collaboratively with students rather than imposed from above.

5.2 Roles for Different Stakeholders

Governments and regulators bear responsibility for developing national guidelines for AI use in education, funding research into reliable assessment methods, and ensuring that AI literacy is incorporated into curricula at all levels. Educational institutions must revise academic integrity policies to address AI explicitly, invest in faculty development, and pilot new assessment formats. Technology companies should be transparent about the capabilities and limitations of their tools, and should not market AI detectors as reliable when the evidence clearly shows they are not. Indi-vidual students, for their part, should engage honestly with AI tools, disclose their use as required, and take responsibility for ensuring that their submitted work reflects their own understanding and

intellectual effort.

6 Conclusion

The arrival of powerful AI tools in education represents a genuine paradigm shift—one that cannot be reversed and should not simply be resisted. The question is not whether AI will be part of education, but how its integration can be managed in ways that preserve the educational mission while preparing students for an AI-integrated world.

Several conclusions emerge from this analysis. The concept of original work needs to be redefined: in an AI-integrated world, originality lies in intent, critical synthesis, and intellectual presence, not in the absence of technological assistance. AI is best understood as a cognitive lever—a tool that amplifies human capability when used well, but that substitutes for human thinking when misused. The evidence from the NYC case and from the failure of AI detection tools suggests that sustainable responses must be built on transparency, process-based assessment, and genuine AI literacy rather than on prohibition or surveillance. The future belongs to those who learn to orchestrate with AI—directing it, evaluating its outputs, and integrating it into their own thinking—rather than those who simply delegate to it. Building the educational frameworks that make this distinction meaningful is the central challenge facing institutions today.

AI Usage Declaration

In this report, we utilized the AI platform Claude Code to access the Claude Opus 4.7 model. This tool assisted us in brainstorming during the planning phase, generating the initial draft of the abstract, refining our language, and structural organisation of the report. We have reviewed all AI-generated output for accuracy and take full responsibility for the final content submitted.

Team Contributions

Name	Student ID	Contribution
Haonan JIAO	1155211226	Introduction (Section 1)
Chenyan XU	1155173847	Technical Introduction (Section 2)
Ruixuan XU	1155211213	Case Study 1 — NYC ChatGPT ban (Section 3.1)
Xiangxiang WENG	1155211173	Case Study 2 — AI detection tools (Section 3.2)
Yuanpeng WU	1155193279	Discussion (Section 4)
Yihao TAN	1155211107	Solutions and Conclusion (Sections 5–6)

References

[1] University of Arizona Libraries, "What is a large language model (LLM)?" 2023. [Online]. Available: https://ask.library.arizona.edu/faq/407985. [Accessed: Apr. 20, 2026].

[2] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30, 2017. [Online]. Available: https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf. [Accessed: Apr. 20, 2026].

[3] 3Blue1Brown, "Transformers, the tech behind LLMs — Deep Learning Chapter 5," *YouTube*, Apr. 1, 2024. [Online]. Available: https://www.youtube.com/watch?v=wjZofJX0v4M. [Accessed: Apr. 20, 2026].

[4] NBC News Now, "New York City public schools ban ChatGPT," *NBC News*, Jan. 2023. [Online]. Available: https://www.nbcnews.com/now/video/new-york-city-public-schools-ban-chatgpt-159425093696. [Accessed: Apr. 20, 2026].

[5] K. Rosenblatt, "New York City public schools remove ChatGPT ban," *NBC News*, May 19, 2023. [Online]. Available: https://www.nbcnews.com/tech/chatgpt-ban-dropped-new-york-city-public-schools-rcna85089. [Accessed: Apr. 20, 2026].

[6] T. G. Mathewson, "AI detection tools falsely accuse international students of cheating," *The Markup*, Aug.14, 2023. [Online]. Available: https://themarkup.org/machine-learning/2023/08/14/ai-detection-tools-falsely-accuse-international-students-of-cheating. [Accessed: Apr. 20, 2026].

[7] T. Rao, "The AI Said the U.S. Declaration of Independence Was AI-Generated," *Medium / Illumination*, Jun. 3, 2025. [Online]. Available: https://medium.com/illumination/the-ai-said-the-u-s-declaration-of-independence-was-ai-generated-85b9b6b853df. [Accessed: Apr. 20, 2026].