

AI Use and Thesis Quality among University Students: An Empirical Analysis from Liberia

Stephen Sunday Browne

Department of Library Services, African Methodist Episcopal University 34 Camp Johnson Road,
Monrovia, Liberia

DOI: <https://doi.org/10.51584/IJRIAS.2026.11060194>

Received: 14 June 2026; Accepted: 19 June 2026; Published: 08 July 2026

ABSTRACT

This study examines the relationship between AI usage and thesis quality among university students in Liberia, conducted between May and December 2025. The research utilizes institutional submission report data from 1,059 student documents, which were analyzed using an academic integrity platform to assess three main quality indicators: AI-generated content scores, textual similarity scores (reflecting plagiarism), and grammatical quality scores.

The findings indicate that while AI usage remained relatively low for much of the study period, averaging 20.8%, a statistically significant increase occurred in the final quarter of 2025, resulting in a mean AI score of 52.7% in December. A significant inverse correlation ($r = -0.100$) between AI scores and similarity scores suggests that students utilizing AI tools tend to produce textually original work. However, the negative correlation between AI usage and grammar scores ($r = -0.223$) raises important concerns about writing quality and student autonomy.

Undergraduate and graduate students showed a higher dependence on AI compared to the general student population. The article discusses the implications for academic policy, faculty oversight, and the prudent integration of AI methodologies in higher education in Liberia.

Keywords: Artificial Intelligence, Academic Integrity, Thesis Quality, Plagiarism Detection, Higher Education, Liberia, AME University, AI Score, Grammar Quality

INTRODUCTION

The rapid proliferation of large language models (LLMs), notably OpenAI's ChatGPT, Google's Gemini, and Anthropic's Claude, has fundamentally transformed the landscape of academic writing on a global scale. In higher education institutions across Sub-Saharan Africa, particularly in Liberia, these technologies have emerged at a time characterized by limited resources, increasing student enrollment, and mounting pressure to produce research-ready graduates. Consequently, the intersection of AI accessibility and academic integrity presents an urgent and empirically under-explored challenge for institutions such as African Methodist Episcopal (AME) University.

AME University, one of Liberia's leading private universities, has actively sought to uphold the integrity of student academic output through an institutional submission and plagiarism-detection system. This study examines a dataset of 1,059 student document submissions from May to December 2025, providing a unique empirical perspective on patterns of AI-generated content, textual similarity, and writing quality across theses, dissertations, research proposals, and assignments.

Two central questions drive this research:

1. To what extent are students at AME University utilizing AI tools in the production of their academic work, particularly theses and dissertations?

2. How does the level of AI usage correlate with other dimensions of academic quality, specifically originality (as indicated by similarity/plagiarism scores) and grammatical proficiency?

By addressing these questions through quantitative analysis of institutional submission data, this paper contributes original insights to the growing body of literature on AI in African higher education and offers actionable recommendations for institutional policy.

LITERATURE REVIEW

Since the public introduction of ChatGPT in late 2022, research on AI-assisted academic writing has surged significantly. Cotton et al. (2023) revealed that students across various disciplines are utilizing large language models (LLMs) for drafting, paraphrasing, and synthesizing literature—tasks that are essential for thesis development. Dehouche (2021) documented that AI-generated text typically demonstrates low lexical similarity to human-written sources, a finding that directly informs the inverse relationship seen in the current study between AI scores and similarity scores.

Farrokhnia et al. (2023) conducted a SWOT analysis of ChatGPT's impact on education and concluded that while there are notable productivity gains, significant threats to critical thinking development and academic integrity also exist. This concern is particularly acute within higher education contexts in developing countries, where digital literacy varies widely. The uncritical adoption of AI tools in these settings may undermine the very epistemic skills that graduate education aims to foster.

AI Detection Tools and Their Limitations

The emergence of AI detection tools, such as those integrated into platforms like Drillbit, Turnitin, and iThenticate, has introduced a new dimension to monitoring academic integrity. Heather and Flint (2023) observe that these tools assign probabilistic scores rather than making definitive determinations, raising concerns about false positives, especially for non-native English speakers who may produce writing that statistically resembles AI outputs. In the context of Liberian education, where English serves as the official medium of instruction but is seldom the first language, this limitation is particularly significant.

Liang et al. (2023) found that the performance of detection tools declines considerably when students employ paraphrasing strategies or combine AI-generated content with human writing. This phenomenon may partially account for the observation that several submissions with high AI scores in the current dataset also achieved relatively low similarity scores.

AI in African Higher Education

Empirical research on the use of AI in African higher education remains limited. Aderibigbe and Mosia (2024) conducted a survey among university students in South Africa and discovered that, although there was a high level of awareness regarding AI tools, the institutional policies governing their use were either nonexistent or poorly communicated. Similarly, Ndukwe and Daniel (2020) highlighted that resource limitations in West African universities, such as insufficient library databases and inadequate research infrastructure, might make AI tools more appealing as alternatives to thorough literature engagement.

In Liberia, AME University's initiatives to implement systematic plagiarism detection and AI monitoring represent a groundbreaking advancement in quality assurance. This study offers the first empirical analysis of the data produced through these efforts.

METHODOLOGY

This study employs a quantitative, descriptive-analytical research design. The primary data source is an institutional submission report obtained from AME University's academic integrity monitoring system, specifically the Submission Reports dataset. This dataset includes all submissions processed from May 1, 2025, to December 17, 2025, covering approximately seven and a half months across two academic semesters.

The total sample consists of 1,059 documents submitted by students at various academic levels, including undergraduate, graduate, and doctoral candidates. A significant majority of these submissions, 1,012 (95.6%), are classified as thesis documents, while smaller numbers include dissertations (11), research papers (2), assignments (31), and miscellaneous categories. All submissions are in English and originate from AME University in Monrovia, Liberia.

Key Variables

The study's primary variables consist of three quantitative indicators derived from the institutional detection system:

1. ****AI Score (%):**** This metric indicates the likelihood that a specific document contains AI-generated text, as determined by the detection platform. A higher score reflects a greater estimated proportion of AI-generated content.
2. ****Similarity Score (%):**** This percentage represents the extent to which the text in a submission aligns with content from the platform's reference corpus. It serves as a proxy for identifying potential plagiarism or inadequate paraphrasing.
3. ****Grammar Score (0–100):**** This composite measure assesses grammatical accuracy and writing quality, as evaluated by the detection system. Higher values signify better grammatical performance.

Analytical Approach

Descriptive statistics, including means, medians, standard deviations, and frequency distributions, were calculated for all three variables across the entire dataset, as well as across relevant subgroups: document type, student level, and submission month. To explore the bivariate relationships between AI scores, similarity scores, and grammar scores, Pearson correlation coefficients were computed. A threshold-based classification approach was employed to categorize documents into risk levels (low, moderate, and high) based on specified AI score ranges. All analyses were performed using Python with the standard library and were verified for internal consistency.

RESULTS

Descriptive Statistics: Overall Sample

Table 1 summarizes the descriptive statistics for all three quality indicators across the full sample of 1,059 submissions.

Table 1: Descriptive Statistics for Key Quality Indicators (N = 1,059)

Metric	N	Mean (%)	Median (%)	Std. Dev.
AI Score (All Documents)	1,059	20.8	19.0	13.4
AI Score (Theses only)	1,012	19.9	18.0	—
Similarity / Plagiarism Score	1,059	22.2	21.0	10.8
Grammar Score	1,059	85.7	89.0	—

Source: AME University Institutional Submission Reports, May–December 2025.

The average AI score across all submissions stood at 20.8% (median = 19.0%, SD = 13.4), indicating that, on average, roughly one-fifth of the submitted text was estimated to be generated by AI. The mean similarity score was 22.2% (median = 21.0%, SD = 10.8), suggesting a moderate level of textual overlap with existing

sources. Grammar performance was generally strong, with an average score of 85.7 out of 100 (median = 89.0), reflecting competent English-language writing overall.

When focusing specifically on thesis and dissertation documents ($n = 1,023$), the mean AI score was slightly lower at 19.9%, with a median of 18.0%. This aligns with the expectation that formally supervised thesis work may involve more interaction between students and faculty, leading to a reduced reliance on AI compared to less supervised assignments.

AI Use by Student Level

A disaggregation of AI scores by student academic level reveals a counterintuitive but theoretically coherent pattern: students at higher academic levels exhibit greater AI use (Table 2).

Table 2: AI Score, Similarity, and Grammar by Student Level

Student Level	N	Mean AI Score (%)	Mean Similarity (%)	Mean Grammar Score
Undergraduate	15	47.4	15.3	83.1
Graduate Student	108	31.8	19.0	82.3
Student (General/Unspecified)	905+	20.1	22.6	85.9
Overall Sample	1,059	20.8	22.2	85.7

Source: AME University Institutional Submission Reports, May–December 2025

Undergraduates recorded the highest mean AI score at 47.4%, followed by graduate students at 31.8% and general/unspecified students at 20.1%. The undergraduate group is the smallest in terms of absolute size ($n = 15$), and their elevated mean AI score may indicate focused usage in specific courses or during the thesis-completion phase, rather than reflecting a broader behavioral trend. Nonetheless, the differences observed across educational levels merit attention. In contrast, similarity scores exhibit a downward trend, decreasing from general students (22.6%) to graduate students (19.0%) and finally to undergraduates (15.3%). This inverse relationship between AI usage and textual similarity will be explored further in Section 4.4.

Temporal Trends in AI Score and Submission Volume

Monthly analysis reveals a striking temporal pattern in AI score means across the study period (Table 3).

Table 3: Monthly Mean AI Score and Submission Volume (May – December 2025)

Month	N	Mean AI Score (%)	Mean Similarity (%)	Trend / Note
May 2025	194	24.0	22.1	Baseline period
June 2025	671	18.0	22.7	Peak volume; AI drops
July 2025	41	19.5	19.8	Low volume
August 2025	50	19.2	23.9	Stable
September 2025	71	30.8	19.9	AI begins rising
October 2025	18	38.9	18.6	Sharp AI increases

November 2025	2	23.0	4.5	Very low volume
December 2025	11	52.7	15.7	Highest AI mean — alarming

Source: AME University Institutional Submission Reports, May–December 2025

June 2025 marked the highest volume of submissions, with 671 entries accounting for 63.4% of the annual total. This spike coincided with a notably low average AI score of 18.0%, suggesting that during peak thesis submission periods, students may produce more original work under the close supervision of faculty. In contrast, as submission volume dwindled in the latter part of the year, AI scores increased significantly. In September, the AI score reached 30.8%, followed by 38.9% in October, and then 52.7% in December. The December mean is particularly significant, as it represents the first instance in which the average AI score exceeded the 50% mark—indicating that, on average, over half of the text in submissions for that month was estimated to be AI-generated. This temporal pattern, where high submission volumes correlate with lower AI usage and low volumes with higher usage, may reflect the moderating influence of institutional oversight. During peak periods, faculty advisors and submission checkpoints engage more actively, likely discouraging reliance on AI. Conversely, in quieter periods, oversight may diminish, creating an environment more conducive to AI-assisted writing.

Risk Classification by AI Score Band

Using threshold-based categorization consistent with emerging AI-integrity benchmarks, submissions were classified into three bands:

Table 4: Distribution of Submissions by AI Score Risk Band

AI Score Band	Range (%)	No. of Docs	% of Total	Risk Designation
Low AI Use	0–24	739	69.8%	Acceptable
Moderate AI Use	25–49	287	27.1%	Review Required
High AI Use	≥50	32	3.0%	Academic Integrity Concern

Source: AME University Institutional Submission Reports, May–December 2025

A significant majority of documents (739; 69.8%) are classified within the Low AI Use category, indicating that most AME University students do not heavily depend on AI tools. Conversely, 287 documents (27.1%) fall into the Moderate band, where the AI's contribution to the text is substantial enough to merit faculty review. A total of 32 documents (3.0%) surpass the 50% AI usage threshold, categorizing them as high use and raising direct concerns regarding academic integrity. Notably, of these 32 high-AI documents, the majority, approximately 20, were assignments rather than theses. This suggests that assignment work may be perceived as carrying lower stakes or facing less stringent oversight than formally submitted thesis documents.

Correlation Analysis

Two Pearson correlations were calculated to explore the relationships among the three quality indicators:

1. ****AI Score and Similarity Score****: The correlation coefficient was $r = -0.100$, indicating a weak negative correlation. As AI scores increase, similarity scores tend to decrease slightly. This aligns with the hypothesis that AI-generated text is stylistically and lexically distinct from the indexed corpus. AI tools do not replicate existing sources; rather, they generate statistically probable text, which may evade similarity detection, potentially representing a form of academic dishonesty.

2. ****AI Score and Grammar Score****: Here, the correlation was $r = -0.223$, demonstrating a moderate negative correlation. A higher use of AI is associated with lower grammar scores. This finding may seem counterintuitive, as one might anticipate that AI tools would enhance grammatical quality. However, it may indicate that students who heavily rely on AI tools engage less thoroughly in the editing and revision process, submitting AI-generated drafts with limited personal refinement. Furthermore, the AI-generated content, particularly in the context of Liberian English, may exhibit grammatical irregularities that are not recognized by human evaluators but are flagged by automated grammar analysis.

This negative correlation between AI use and grammar scores is one of the most significant findings of this study from a policy perspective, suggesting that extensive use of AI does not necessarily replace writing skills. In some cases, it may even undermine detectable writing quality while increasing the volume of text produced.

Grammar Quality Distribution

Grammar scores were classified into four quality bands. Out of the 1,059 submissions, 442 (41.7%) achieved scores in the Excellent range (90–100); 405 (38.2%) received good scores (80–89); 173 (16.3%) were rated Fair (70–79); and 32 (3.0%) were marked Poor (below 70). The high percentage of Excellent and Outstanding scores (a combined total of 79.9%) indicates a relatively strong level of written English proficiency among AME University students. However, the negative correlation with AI usage suggests that heavier users of AI tend to cluster in the lower grammar bands.

Plagiarism (Similarity) Patterns

Out of twenty-one submissions (2.0%) that exhibited similarity scores of 50% or higher, two extreme cases were noted at 98% and 100%. Importantly, none of these high-similarity cases had elevated AI scores, all falling below 30%. This observation highlights that conventional plagiarism, defined as the direct copying of existing text, and AI-generated writing primarily represent distinct integrity risks that necessitate separate detection methods. Institutions that rely solely on similarity detection overlook the AI aspect, while those focused only on AI detection may miss instances of traditional plagiarism.

The statistical analysis reveals a negative correlation of $r = -0.100$ between artificial intelligence (AI) and similarity scores. This finding reinforces the conclusion that the two risks are largely independent from one another. Consequently, it underscores the necessity for a dual-detection policy framework that addresses each risk separately and effectively.

DISCUSSIONS

The most striking revelation from this study is not merely the current moderate level of artificial intelligence (AI) utilization among students; rather, it is the observable trajectory of this usage. Over a mere six-month period, AI engagement surged dramatically from a modest mean score of 18.0% in June 2025 to an impressive 52.7% by December 2025. This trajectory suggests the potential for a profound transformation in how students at AME University approach the creation of academic text. The implications extend beyond disciplinary concerns, signaling a significant structural change in the interaction dynamics between students and the process of scholarly knowledge production.

An intriguing temporal pattern emerges from the data: as the frequency of supervision decreases, the reliance on AI increases. This observation aligns with the principles of opportunity theory as discussed in academic dishonesty research (McCabe et al., 2012), indicating that instances of misconduct tend to rise when individuals perceive a reduced risk of detection and lower institutional accountability. The conclusion is unmistakable: the current strategy of passive monitoring is inadequate. To address these challenges effectively, AME University must implement proactive, process-oriented interventions that are strategically integrated throughout the academic calendar.

Perhaps the most pivotal theoretical insight is the almost complete independence of AI usage scores from similarity scores, evidenced by a correlation coefficient of $r = -0.100$. This finding contradicts the common

assumption among administrators that a single integrity platform adequately captures the full spectrum of academic dishonesty. There is a fundamental distinction between AI-generated writing and plagiarized text: the former relies on probabilistic language modeling to craft original-seeming content, while the latter replicates existing text either verbatim or with minimal alterations.

For AME University, this distinction carries major repercussions. A student who achieves a low similarity score cannot automatically be deemed to have engaged genuinely with their research. For instance, a student whose thesis comprises up to 80% AI-generated content may present a similarity score between 5% and 10%, since AI-generated text is not only distinct in style but also rich in lexical variety. It is essential for the university to recognize that a favorable similarity report does not equate to true intellectual engagement or independence.

From an educational standpoint, the negative correlation between AI usage and grammar scores ($r = -0.223$) warrants careful consideration. In theory, AI writing tools are designed to enhance grammatical quality; however, the real-world data imply otherwise. Several plausible explanations could account for this surprising trend:

Students might be leveraging AI for generating initial drafts, which they then submit without rigorous editing. This practice introduces grammatical structures typical of AI-generated text, diverging from the standards expected in academic writing. Heavy reliance on AI could diminish the active practice of writing that is critical to developing robust grammatical skills. In essence, using AI as a substitute may inhibit essential skill development. Students with weaker writing abilities may gravitate toward AI tools out of necessity due to their challenges, resulting in a selection effect where low grammar scores and high AI usage co-occur. Disentangling these influences will necessitate longitudinal data; however, the currently observed correlation is concerning enough to merit institutional scrutiny regarding AI's impact on authentic writing development.

These findings must be contextualized within Liberia's unique higher education landscape. Universities in Liberia are grappling with compounded challenges that stem from the aftermath of the civil conflict from 1989 to 2003 and the Ebola epidemic of 2014 to 2016. As they face rapid enrollment growth amidst limited faculty resources and inadequate infrastructure, the integration of technology into educational frameworks becomes increasingly complex, particularly when digital access remains inconsistent. For graduate students crafting theses on intricate topics related to economics and governance, many of which were reflected in the AME University submissions, AI tools might present an enticing shortcut in research endeavors, especially when access to library resources, research methodology training, and sufficient engagement with advisors is lacking.

Additionally, the district's remarkably high volume of thesis submissions in June 2025, totaling 671 documents within a single month, raises critical questions regarding the feasibility of effective thesis supervision on such a scale. When faculty members are tasked with guiding dozens of students simultaneously through the pressures of looming submission deadlines, the capacity to provide in-depth formative feedback that could mitigate reliance on AI shortcuts becomes severely limited.

CONCLUSION

This study presents the first systematic analysis of AI usage in academic submissions at the African Methodist Episcopal University, Liberia. Analyzing 1,059 student documents submitted between May and December 2025, it finds that while the average AI score is moderate at 20.8%, there is a significant upward trend, culminating in a December mean of 52.7%. The low correlation between AI and similarity scores ($r = -0.100$) indicates that traditional plagiarism detection methods are insufficient for addressing AI integrity risks. Additionally, the negative correlation between AI use and grammar ($r = -0.223$) raises concerns about writing development in an AI-dominated landscape. The findings reveal patterns of higher AI usage during low-supervision periods and among graduate and undergraduates, pointing to institutional factors over individual choices. Addressing this issue requires systemic responses, including clear policies, dual-detection standards, ongoing supervisory engagement, and robust AI literacy education.

AME University's commitment to integrity monitoring is a proactive step, with the data providing useful insights for evidence-based governance. The challenge now is to respond to these findings before the rising trend of AI use fundamentally transforms the nature of academic theses at the university.

REFERENCES

1. Aderibigbe, S. A., & Mosia, P. A. (2024). University students' perceptions of artificial intelligence in academic writing. *Journal of Applied Research in Higher Education*, 16(2), 45–60.
2. Cotton, D. R. E., Cotton, P. A., & Reuber, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239.
3. Dehouche, N. (2021). Plagiarism in the age of massive generative pre-trained transformers (GPT-3). *Ethics in Science and Environmental Politics*, 21, 17–23.
4. Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 1–15.
5. Heather, J., & Flint, A. (2023). Are AI detectors reliable? Exploring the accuracy of AI-generated text detection tools. *Journal of Learning Development in Higher Education*, 27, 1–14.
6. Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), 100779.
7. McCabe, D. L., Butterfield, K. D., & Treviño, L. K. (2012). *Cheating in College: Why Students Do It and What Educators Can Do about It*. Johns Hopkins University Press.
8. Ndukwe, I. G., & Daniel, B. K. (2020). Teaching analytics, value, and tools for teacher data literacy. *International Journal of Educational Technology in Higher Education*, 17(1), 1–20.