TA'DIB JOURNAL, 28 (2), 2025, (July-December)

ISSN: 1410-8208 (*Print*) 2580-2771 (*Online*)
Available online at https://ejournal.uinmybatusangkar.ac.id/ojs/index.php/takdib/index

Assessing Artificial Intelligence Plagiarism Risk: ChatGPT vs Scite Among Islamic Education Students

Received: 13-12-2025; Revised: 20-12-2025; Accepted: 31-12-2025

Ismail *)
Universitas Islam Negeri Walisongo Semarang , Indonesia
E-mail: ismail@walisongo.ac.id

Suja'i
Universitas Islam Negeri Walisongo Semarang , Indonesia
E-mail: sujai@walisongo.ac.id

Silviatul Hasanah
Universitas Islam Negeri Walisongo Semarang, Indonesia
E-mail: silviatulh@walisongo.ac.id

*) *Corresponding Author*

Abstract: This study examines and compares the effects of ChatGPT and Scite_AI, on plagiarism tendencies among students of Islamic Religious Education in Indonesia. Adopting a quantitative research design, the study employed multiple linear regression analysis to evaluate both the partial and simultaneous influences of these tools on academic plagiarism. Prior to regression analysis, classical assumption tests—including normality (Prior to the regression analysis, classical assumption tests—including normality (Kolmogorov-Smirnov p = 0.088), multicollinearity (VIF < 10), multicollinearity (VIF < 10), heteroscedasticity (Breusch-Pagan test), and linearity (scatterplot of residuals)—were rigorously conducted to ensure model validity. The results reveal that both AI tools significantly contribute to increased plagiarism tendencies; however, ChatGPT demonstrates a markedly stronger effect (β = 0.4941; p < 0.001) compared to Scite (β = 0.1042; p < 0.001). The overall regression model is statistically significant (F = 87.32, p = 0.000) and satisfies all classical assumptions, confirming its reliability. Theoretically, this research enriches academic integrity literature by positioning AI tool typology—particularly the distinction between generative and verification tools—as a critical predictor of plagiarism behavior. Practically, it calls for differentiated AI literacy strategies in Islamic higher education, advocating for the integration of adab al-'ilmu (ethics of knowledge) into digital literacy curricula to foster moral discernment and responsible technology use among future religious educators.

Abstrak: *Studi ini meneliti dan membandingkan pengaruh ChatGPT dan Scite_AI terhadap kecenderungan plagiarisme di kalangan mahasiswa Pendidikan Agama Islam di Indonesia. Dengan mengadopsi desain penelitian kuantitatif, studi ini menggunakan analisis regresi linier berganda untuk mengevaluasi pengaruh parsial dan simultan dari alat-alat tersebut terhadap plagiarisme akademik. Sebelum analisis regresi, uji asumsi klasik—termasuk normalitas (Kolmogorov-Smirnov p = 0,088), multikolinearitas (VIF < 10), heteroskedastisitas (uji Breusch-Pagan), dan linearitas (scatterplot residual)—dilakukan secara ketat untuk memastikan validitas model. Hasil penelitian menunjukkan bahwa kedua alat AI tersebut secara signifikan berkontribusi pada peningkatan kecenderungan plagiarisme; namun, ChatGPT menunjukkan efek yang jauh lebih kuat (β = 0,4941; p < 0,001) dibandingkan dengan Scite (β = 0,1042; p < 0,001). Model regresi secara keseluruhan*

signifikan secara statistik (F = 87,32, p = 0,000) dan memenuhi semua asumsi klasik, yang menegaskan keandalannya. Secara teoritis, penelitian ini memperkaya integritas akademik. Literatur ini menyoroti tipologi alat AI—khususnya perbedaan antara alat generatif dan verifikasi—sebagai prediktor penting perilaku plagiarisme. Secara praktis, hal ini menyerukan strategi literasi AI yang berbeda dalam pendidikan tinggi Islam, serta menganjurkan integrasi adab al-'ilmu (etika ilmu) ke dalam kurikulum literasi digital untuk menumbuhkan kearifan moral dan penggunaan teknologi yang bertanggung jawab di kalangan calon pendidik agama.

Keywords : chatgpt, scite_ai, plagiarism,Islamic education

INTRODUCTION

The use of artificial intelligence among university students is now a hot topic. Holoniq (2023) reported that more than 75% of college students in the United States are interested in using AI to complete coursework. Furthermore, Study.com (2023) explained that 89% of US college students admitted to having used ChatGPT, with 48% using it to write essays (Hasanein & Sobaih, 2023). In Asia, a McKinsey report (2024) noted that the use of generative AI on campuses has increased 300% since 2022, including in Indonesia, where a survey by the Ministry of Education, Culture, Research, and Technology (2024) found that 64% of college students have tried at least one AI tool for academic purposes. Amid this explosion in adoption, two tools stand out for their distinct functions: ChatGPT (OpenAI), which generates original text based on prompts, and Scite (scite.ai), designed to verify the validity of scientific citations through contextual analysis (Morris, 2018). While both are useful, the speed and ease of use offered by AI open up serious gaps in academic integrity—particularly in the form of plagiarism, which is increasingly difficult to detect due to the "unique" yet intellectually unoriginal nature of text (Williams, 2024).

Traditional plagiarism (such as copying text without citation) is relatively easy for detection systems to identify. However, generative AI creates a new form of plagiarism: paradigmatic plagiarism, where ideas, argument structures, or narratives are adopted without attribution, but the text appears original. ChatGPT, for example, can generate essays that pass plagiarism checks because they are not literally copied, but rather derive their content from training data without acknowledgment of the original intellectual source (Božić & Poola, 2023). Meanwhile, Scite_AI, which is intended to increase scientific transparency, also poses risks if used passively. Students may simply adopt recommendations for "supporting citations" without reading primary sources, resulting in technically correct but ethically inauthentic citations (Lund & Shamsi, 2023). This phenomenon reflects the illusion of academic competence: the ability to produce output without in-depth understanding. In the context of Islamic education, which emphasizes the values of scientific honesty and responsibility, this risk is crucial, as it constitutes not only a procedural violation but also a moral violation of the principle of trustworthiness in science.

Despite the increasing urgency of this issue, empirical literature examining the causal relationship between AI use and plagiarism tendencies remains very limited. A number of international studies have led the way, but the majority are qualitative or exploratory. For example, Perkins et al. (2023), in a study across 12 universities in the UK, Australia, and Canada, found a positive correlation between the frequency of ChatGPT use and decreased cognitive engagement, but did not directly measure plagiarism. Meanwhile, Liang et al. (2024) in a study conducted in China, Singapore, and Malaysia using

simulation scenarios showed that students using AI were more likely to avoid attribution. However, the study did not differentiate between the types of tools (generative vs. verifiable). On the other hand, Zhang et al. (2024) in an OECD report emphasized the lack of quantitative data comparing the impact of various AI tools on academic behavior. Specifically, no study has statistically compared the effects of ChatGPT and Scite on plagiarism tendencies, especially in the context of Islamic education in a developing country like Indonesia. This gap is filled by this study through a rigorous quantitative approach based on multiple linear regression.

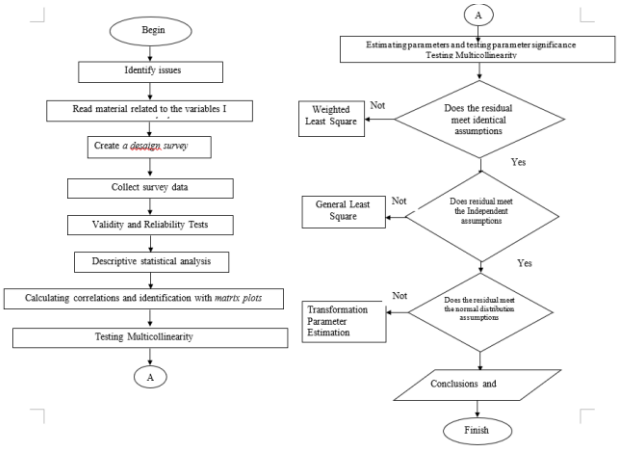
This study aims to examine and compare the effects of ChatGPT and Scite on plagiarism tendencies among Islamic Religious Education students in Indonesia. Beyond addressing a critical empirical gap—namely, the absence of comparative studies between generative AI tools like ChatGPT and AI-powered verification tools like Scite—this research contributes academically by illuminating how distinct AI functionalities shape student writing practices and ethical decision-making. By contrasting a tool that generates content with one that validates it, the study offers novel insights into the differential impact of AI affordances on academic integrity. Such a comparison is particularly significant in the Indonesian Islamic higher education context, where the integration of AI literacy must align with religious-ethical values and pedagogical goals. This article is structured systematically to ensure logical clarity and analytical depth. Following this introduction, the second section outlines the research methodology, including the survey design, the plagiarism tendency measurement instrument (validated through scale reliability testing), and the regression analysis procedure with associated assumption checks. The third section presents the main findings, including regression coefficients, partial and simultaneous significance tests, and verification of classical assumptions (normality, homoscedasticity, multicollinearity, etc.). The final section discusses the implications of these findings in both global and local contexts, emphasizing the need for a values-based, pedagogically grounded approach to AI literacy that distinguishes between high-risk generative tools and lower-risk verification tools in Islamic education.

METHOD

a) Research Design

Given that this study aims to statistically measure the simultaneous and partial effects of several independent variables on the dependent variable, a quantitative approach using multiple linear regression analysis was chosen as the most appropriate analysis strategy. This analysis process consists of several stages carried out systematically to ensure the validity and reliability of the research results. These stages include specifying the relationships between variables using matrix diagrams and Pearson product-moment correlations, testing residual data assumptions, including identical distribution, independence, and normally distributed residuals, as well as estimating parameters and testing their significance to measure the predictive impact of the independent variables on the dependent variable (Roustaei, 2024). Before proceeding with the regression analysis, testing the validity and reliability of the research instrument is necessary because data collection requires the use of a Likert scale with procedures for testing internal consistency and measurement accuracy (Özden, 2024). In summary, the research stages undertaken by the researcher are presented in the flowchart shown in Figure 1.1 below.

Figure 1.1: Research flowchart



b) Participants

Participants in this study were students at Walisongo State Islamic University in Semarang who actively use ChatGPT and Scite_Ai. This study used primary data from a three-phase survey. The first phase surveyed 1,281 students to determine the frequency of AI use, specifically ChatGPT. After screening, 330 students were classified as heavy AI users (24-35 hours per week active AI screen time, (Bozic and Pola, 2021)). The second phase involved 50 students to test the validity and reliability of the survey instrument. After the instrument was declared valid and reliable, the third phase involved multiple linear regression analysis with 330 respondents. All participants provided informed consent for the publication of their data, but maintained the confidentiality of their profiles.

c) Data Collection Technique

Data were collected using a questionnaire distributed to respondents voluntarily. This questionnaire consisted of 40 closed-ended statements divided into five Likert scales. Variables X1 and X2 were developed based on indicators from Davis' Technology Acceptance Model (TAM) theory, and variable Y1 was derived from the Fuzzy Cognitive Map (FCM) framework. The use of a Likert scale allows respondents to express objective descriptions of their experiences using a numeric score scale of 1-5 (Joshi et al., 2015). The researchers used Google Forms, a survey software. They provided interactive instructions to ensure respondents understood how to complete the questionnaire and to reduce ambiguity. This research employed convenience sampling, a non-probability sampling technique, due to the limited population studied (Golzar, Noor, & Tajik, 2022).

d) Data Analysis

In this study, the data analysis process began with testing the prerequisites for multiple linear regression analysis to ensure the validity of the statistical model. Classical assumption tests included residual normality tests using the Kolmogorov-Smirnov method, heteroscedasticity tests using the Glejser approach, autocorrelation tests using the Durbin-Watson statistic, and multicollinearity tests to evaluate the relationship between independent variables. After the assumptions were met, the researchers proceeded with inferential statistics in the form of multiple linear regression analysis at a 5% (0.05) significance level, processed using Minitab 22 software (Klein, 2024). Hypothesis testing decisions are based on the P-value: the null hypothesis (H_0) is rejected if the P-value is less than 0.05 and accepted if the P-value is greater than or equal to 0.05 (Lyu & Li, 2024). Furthermore, descriptive statistics were applied to describe respondent characteristics, including data patterns and distributions related to plagiarism and data falsification tendencies in the use of ChatGPT and Scite_AI.

e) Validity and Reliability Testing

In this study, 50 respondents were used to test the quality of the research instrument, as recommended by Arianto and Octavia (2021). Validity testing was conducted to ensure that each statement in the questionnaire was able to measure the construct it was intended to measure; a statement is considered

valid if it is relevant and on target (Fitri & Arum, 2023). In addition, a reliability test was conducted to assess the consistency of the measurement results. This means that if the questionnaire is administered to the same respondents under different conditions, the results remain stable and reliable.

1) Validity Test

The following are the results of the validity test for variable X1, with a 95% confidence level and a 5% significance level, analyzed using the Pearson product-moment correlation formula.

Table 1. Validity Test for Variable X1

Statement	*P-Value*
Statement 1	0,000
Statement 2	0,000
Statement 3	0,001
Statement 4	0,000
Statement 5	0,000
Statement 6	0,000
Statement 7	0,000
Statement 8	0,000
Statement 9	0,000
Statement 10	0,000

Table 1 shows the results indicating that all statements in the ChatGPT usage variable (X1) correlate significantly with the total statement score. Analysis of these 10 statements yielded a P-value lower than the significance level of $\alpha = 0.05$, thus rejecting the null hypothesis (H_0). This rejection of H_0 proves that all statements meet the validity criteria (Martens & Dardenne, 1998).

The following are the results of the validity test for variable X2, which was conducted subsequently.

Table 2. Validity Test for Variable X2

Statement	*P-Value*
Statement 1	0,000
Statement 2	0,000
Statement 3	0,001
Statement 4	0,000
Statement 5	0,000
Statement 6	0,000
Statement 7	0,000
Statement 8	0,000
Statement 9	0,000
Statement 10	0,000

Table 2 shows that, based on an analysis of the 10 statements regarding the Use of Scite (X2) in relation to the total statement score, all items yielded P-values below the 0.05 significance level. This resulted in the rejection of the null hypothesis (H_0) for each statement, indicating that all items were valid.

The following are the results of the validity test for variable X3. Next, a validity test was conducted for variable Y1.

Table 3. Validity Test for Variable Y1

Statement	*P-Value*
Statement 1	0,000
Statement 2	0,000
Statement 3	0,001
Statement 4	0,000
Statement 5	0,000
Statement 6	0,000
Statement 7	0,000
Statement 8	0,000
Statement 9	0,000
Statement 10	0,000

Table 3 shows the correlation significance between each statement in the tendency to plagiarize variable (Y1) and the total score for all statements. Of the ten statements tested, all yielded p-values less than the established significance level ($\alpha = 0.05$). Therefore, the null hypothesis (H_0) is rejected, meaning all statements are valid.

Next, a validity test was conducted for variable Y2. The following are the results of the validity test for variable Y2.

Table 4. Validity Test for Variable Y2

Statement	*P-Value*
Statement 1	0,000
Statement 2	0,000
Statement 3	0,001
Statement 4	0,000
Statement 5	0,000
Statement 6	0,000
Statement 7	0,000
Statement 8	0,000
Statement 9	0,000
Statement 10	0,000

Based on the analysis of the ten statements in the tendency to plagiarize (Y_1), all items showed p-values smaller than the significance level of $\alpha = 0.05$. This finding indicates that the correlation between each

statement and the total score of the variable is statistically significant. By rejecting the null hypothesis (H_0), it can be concluded that all statement items in the Data Falsification variable are declared valid and suitable for use as measurement indicators for the construct. Thus, Table 4 confirms the feasibility of the instrument used in this study in terms of item validity.

1) Reliability Test

The Cronbach's Alpha formula was used in this study to test the reliability of the data. The following are the results of the reliability test for variable X1.

Table 5. Reliability Test for Variable X1

Variable	Alfa Cronbach
X1	0,8291

Based on the analysis results in Table 5, the Cronbach's Alpha value was 0.8291. This value is above the minimum threshold commonly used in quantitative research, which is 0.600. Therefore, the null hypothesis (H_0) stating that the measurement instrument is unreliable can be rejected. Therefore, it can be concluded that the instrument used to measure variable X_1 has an adequate level of internal consistency and reliability, making it suitable for use in this study.

Next, a reliability test was conducted for variable X2. The following are the results of the reliability test for variable X2.

Table 6. Reliability Test for Variable X2

Variable	Alfa Cronbach
X2	0,8579

Based on the analysis results in Table 6, a Cronbach's Alpha value of 0.8579 was obtained. This value exceeds the minimum reliability threshold commonly used in quantitative research, which is 0.600. Therefore, it can be concluded that the measurement instrument for variable X_2, namely the use of Scite, has an adequate level of internal consistency. This indicates that the items in the instrument are relatively stable and capable of producing consistent data from one measurement to the next. Therefore, the null hypothesis (H_0) stating that the instrument is unreliable can be rejected, and the variable is suitable for use in further analysis.

Next, a reliability test was conducted for variable Y1. The following are the results of the reliability test for variable X3.

Table 7. Reliability Test for Variable Y1

Variable	Alfa Cronbach
Y1	0,8859

Based on the reliability analysis results in Table 7, a Cronbach's Alpha value of 0.8859 was obtained. This value far exceeds the minimum threshold commonly used in quantitative research, which is 0.600. Therefore, the null hypothesis (H_0) stating that the instrument is unreliable can be rejected. This rejection of H_0 indicates that the instrument used to measure the tendency to plagiarize (Y_1) has a high level of internal consistency. This means that respondents' responses to the items in the instrument are relatively stable and not due to chance, making the instrument suitable for measuring plagiarism addiction in this study.

f) Research Variables

This study aims to analyze the effect of the use of ChatGPT (X1) and Scite AI (Y1) on the tendency to plagiarize at UIN Walisongo Semarang, Indonesia, to complete course assignments. The main focus of this study is to examine how the use of these two types of AI tools impacts two indicators of academic ethics: the tendency to plagiarize (Y1) and data falsification (Y2).

RESULTS AND DISCUSSION

Results

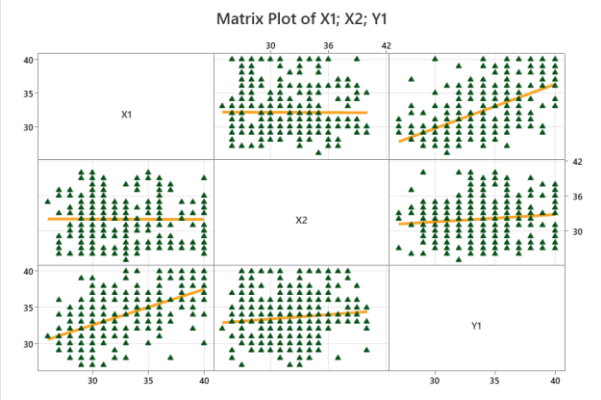
Before conducting a multiple linear regression analysis between variables X_1, X_2, and Y_1, several classical assumptions must be met to ensure the resulting model is valid and

reliable. First, it is important to examine the pattern of relationships between variables through plot analysis and a Pearson correlation matrix chart to ensure an adequate linear relationship. Second, the residuals from the regression model must be normally distributed, which can be tested using a normality test. Third, the assumption of homoscedasticity, meaning that the residual variance is expected to be constant across the range of predictor values, is important. Fourth, the residuals must be independent, which is typically tested using the Durbin-Watson statistic to avoid autocorrelation. Finally, there must be no high levels of multicollinearity between the independent variables, as this can disrupt the stability and interpretation of the regression coefficients. By meeting these five assumptions, multiple linear regression analysis can produce unbiased, efficient, and consistent estimates.

a) Pearson Correlation Plot and Chart Analysis

A matrix plot is used to determine the relationship between the dependent and independent variables. The relationship is used to visualize the shape of the linear line. The following is the result of a matrix plot between variables.

Figure 1. Plot of the matrix between variables



The relationship between the independent variables (X_1 and X_2) and the dependent variable (Y_1) shows a consistent and directional pattern. This was first revealed through the Pearson Product Moment correlation statistical test, which quantitatively confirmed a positive linear relationship between the three variables. This statistical finding was further reinforced by visual analysis of the graph, which displayed a yellow line forming an ascending pattern from the bottom left to the top right, a characteristic of a positive correlation. Thus, both from the numerical data and the graphical representation, it is clear that an increase in the values of X_1 and X_2 is followed by an increase in the value of Y_1.

Table 9. Pearson Product Moment Correlation Test

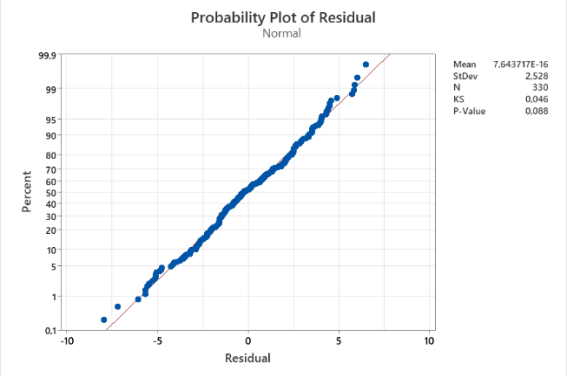
Variable	X1	X2
	P-Value	
Y1	0,000	0,040

Based on the results of the Pearson product-moment correlation test presented in Table 9, a p-value of 0.000 was obtained for the relationship between variables X_1 and Y_1, and 0.040 for the relationship between variables X_2 and Y_1. Because both p-values are less than the significance level of $\alpha = 0.05$, the null hypothesis (H_0) is rejected. This rejection of H_0 indicates that there is a significant linear relationship between each independent variable (X_1 and X_2) and the dependent variable (Y_1). Thus, it can be concluded that both X_1 and X_2 are statistically linearly correlated with Y_1.

b) Normal Distribution

The following are the results of the analysis of the normal distribution residual assumption test for data on the use of Chat GPT and Scite on plagiarism tendencies using the Kolmogorov-Smirnov method.

Figure 2. Test of the Assumption of Normal Distribution Residuals



Based on the image above, the data is visually normally distributed, as the red plot lies between the linear lines formed. The results of the normal distribution assumption test using the Kolmogorov-Smirnov test show that the p-value (0.088) is greater than the alpha value (0.05). This means that the H0 cannot be rejected. Therefore, the conclusion is that the residual data on the use of Chat GPT and Scite on the tendency to plagiarism are normally distributed.

c) Homoscedasticity

In this study, the Glacier test was used to determine the value of the identical residual assumption. The calculation results can be seen in Table 10 below.

Table 10. Identical Residual Assumption Test

Source of Variance	DF	SS	MS	F	*P-value*
Rear	1	0,822	0,8219	0,38	0,539
Residual	328	711,157	0,8219		
Whole	329	711,979			

Based on the results of the identical residuals assumption test in Table 10, a p-value of 0.539 was obtained, which is greater than the significance level of α = 0.05. Because this p-value is not small enough to reject the null hypothesis (H2), it can be concluded that there is insufficient statistical evidence to indicate a difference in residual variance between the groups. Therefore, the assumption of identical residuals is met in this regression analysis. This strengthens the validity of the model used to examine the effect of ChatGPT and Scite on plagiarism tendencies, as one of the important prerequisites in multiple linear regression, namely homogeneity of residual variance, has been met.

d) Independent

The independent residuals assumption can be tested using the Durbin-Watson test. The following are the results of the Durbin-Watson test.

Table 11. Independent Residuals Assumption Test

Durbin Watson	Dl	from	4-dL	4-dU
1,846	1,813	1,826	2,187	2,175

Based on the results of the independent residuals assumption test, as listed in Table 11, the Durbin-Watson statistic value was 1.846. This value lies between the upper limit (dU = 1.826) and the complementary upper limit (4 − dU = 2.175). Thus, this value lies within the range indicating no autocorrelation in the residuals. Consequently, the statistical decision taken is to fail to reject the null hypothesis (H_0). This means that the residuals from the regression model testing the effect of ChatGPT and Scite use on plagiarism tendencies meet the assumption of independence. In other words, the residual data are independent, making the regression model suitable for further interpretation.

e) Multicollinearity

The multicollinearity test is used to determine whether there is a correlation between the independent variables. This test can be performed by examining the Variance Inflation Factor (VIF) and the tolerance value for each independent variable. If the tolerance value is greater than 0.10 and the VIF is less than 10, it can be concluded that multicollinearity does not exist. The following are the results of multicollinearity.

Table 12. Multicollinearity

Pattern	Collinearity Statistics		Information
	Tolerance	Ligh	
Use of GPT Chat	1	1	No multicollinearity
Use of Scite_AI	1	1	No multicollinearity

Based on the analysis results in Table 12 regarding the multicollinearity test, it can be seen that the Tolerance value for each independent variable (i.e., ChatGPT Usage and Scite Usage) is 1, which is well above the minimum threshold of 0.1. Similarly, the Variance Inflation Factor (VIF) value for both variables is 1, which is below the critical threshold of 10. This condition indicates that there is no significant linear correlation between the independent variables in the regression model. Thus, it can be concluded that the regression model used is free from multicollinearity issues.

Next, a multiple linear regression analysis was conducted. The first step in this analysis was parameter estimation. The following is a multiple linear regression model obtained from the parameter estimation of the dependent variable Y1 using X1 and X2.

$$\widehat{Y_1} = 14{,}41 + 0{,}4941X_1 + 0{,}1042X_2$$

The meaning of the linear regression model equation above is that if the value of the Chat GPT usage variable increases by one unit, the data plagiarism tendency variable will increase by 0.4941, assuming the values of the other variables remain constant. If the value of the Scite usage variable increases by one unit, the plagiarism tendency variable will increase by 0.1042, assuming the values of the other variables remain constant.

The next step is to conduct a simultaneous test. A simultaneous test is used to examine the effect of independent variables simultaneously on the dependent variable. The following are the results of the simultaneous test analysis.

Table 13. A simultaneous test (ANOVA)

Source of Variance	DF	SS	MS	F	*P-value*
Residual	2	1069,66	534,83	83,19	0,000
Error	327	2102,23	6,43		
Overall	329	3171,89			

Based on the data in Table 13, which displays the results of the ANOVA test, the calculated F-value of 83.19 far exceeds the F-table value (3.023) at a significance level of α = 0.05, with a p-value of 0.000—far below the significance threshold. This condition indicates that the null hypothesis (H_0), which states there is no simultaneous influence of the tested variables, cannot be accepted. Therefore, it can be concluded that at least one of the two variables significantly influences the tendency to plagiarize among students. This finding strengthens the argument that AI-based tools, especially generative ones like ChatGPT, have the potential to become an important factor in contemporary academic dynamics, including in the context of academic integrity.

To identify which variables individually make a significant contribution, the next step is to conduct a partial test. The purpose of a partial test is to determine which independent variables specifically influence the dependent variable, in this case, the tendency to plagiarize, before proceeding to further interpret the test results.

Table 14. Partial test (t-test)

Variable	T	*P-value*	Results
X1	12,65	0,000	Reduce H0
X2	2,58	0,000	Reduce H0

Based on the data analysis, it was found that both ChatGPT and Scite usage significantly influenced the tendency to

plagiarize among students. This is indicated by a p-value of 0.000 for both variables in the partial t-test—well below the significance threshold of $\alpha = 0.05$—and calculated t-statistics (12.65 and 2.58, respectively), which exceeded the t-table value (1.97). These findings indicate that the null hypothesis (H_0) for each variable can be rejected. Furthermore, although the regression model used explained approximately 33.72% of the variation in plagiarism tendencies, the majority of the variation (66.28%) was still influenced by factors outside this model. Thus, it can be concluded that ChatGPT and Scite usage are two important factors that partially contribute to academic plagiarism.

The findings of this study reveal that both ChatGPT and Scite usage significantly contribute to an increased tendency to plagiarize, with ChatGPT demonstrating a significantly stronger influence than Scite. This is reflected in the respective regression coefficients (0.4941 for ChatGPT and 0.1042 for Scite) as well as the partial test results which show p-values close to zero for both variables. Simultaneously, the multiple linear regression model is proven valid and significant, as indicated by the calculated F-value which far exceeds the F-table and the p-value of 0.000, which means that the combination of the two independent variables is able to explain the variation in plagiarism tendencies. The validity of the model is strengthened by the fulfillment of all classical assumptions of multiple linear regression: a positive linear relationship between variables (based on matrix plots and Pearson correlation), normal distribution of residuals (Kolmogorov-Smirnov test, $p = 0.088$), homoscedasticity (Glejser test, $p = 0.539$), independence of residuals (Durbin-Watson = 1.846), and the absence of multicollinearity (tolerance = 1 and VIF = 1 for both predictors). These findings emphasize that the use of AI technology, especially generative tools like ChatGPT, needs to be accompanied by strong academic literacy and ethical awareness to prevent plagiarism, although academic tools like Scite—which are designed to enhance scientific integrity—do not appear to be completely immune to these risks.

Discussion

The findings of this study indicate that the use of artificial intelligence (AI)-based technologies, specifically ChatGPT and Scite, is significantly positively correlated with an increased likelihood of plagiarism among university students. The significantly larger regression coefficient for ChatGPT (0.4941) compared to Scite (0.1042) indicates that AI-based generative tools have a dominant influence in encouraging this unscholarly practice. This is plausible given ChatGPT's ability to generate complete text without requiring in-depth user understanding, thus facilitating the reproduction of content without adequate attribution. Furthermore, although Scite was designed as an academic verification tool to help users evaluate the quality and integrity of references, it also contributes—albeit to a lesser extent—to plagiarism. This is likely because users misinterpret Scite's function as a "validity stamp" without understanding the substance of the cited work. The statistically valid regression model supports the belief that this relationship is not merely coincidental but reflects the dynamics of AI use in contemporary academic practice.

These findings align with a number of recent studies highlighting the ethical risks of using generative AI in education. For example, Sullivan et al. (2023) found that students who frequently use ChatGPT tend to experience decreased critical thinking skills and a reliance on automated output. This finding is supported by research by Buragohain and Chaudhary (2025), which found that excessive reliance on ChatGPT can hinder students' critical thinking skills and creativity, potentially triggering passive information consumerism (Okoro, 2011). However, unlike previous research that

tends to strictly distinguish between "dangerous" tools like ChatGPT and "safe" tools like academic verification tools, this study reveals that even platforms designed to support scientific integrity—like Scite—are not entirely risk-free (Fırat, 2023). These findings extend the literature by demonstrating that the risk of plagiarism stems not only from what is used, but also from how the tool is used (Khalil & Er, 2023).

This is relevant to Zeide et al.'s (2022) argument that technology is morally neutral, but its use is strongly influenced by the user's ethical literacy and academic understanding. Thus, this study not only confirms existing concerns but also extends them to the realm of scientific tools previously assumed to be safe. In this context, educators are advised to assign assignments that go beyond the basics and encourage active engagement and critical thinking, while also explaining the limitations of ChatGPT use (Khalil & Er, 2023). Furthermore, imbalances in access to high-quality AI tools can lead to disparities in academic performance and emphasize the importance of balanced use and guidance in utilizing AI tools (Holmner, 2025).

Theoretical explanations for these findings can be analyzed through the lens of the Theory of Planned Behavior and the concept of critical digital literacy. According to this theory, intentions to engage in behavior (in this case, plagiarism) are influenced by attitudes, subjective norms, and perceived behavioral control (Segovia-Juárez & Baumgartner, 2023). Students using ChatGPT may perceive the tool as facilitating academic tasks without ethical consequences (positive attitudes), supported by group norms that view the use of AI as commonplace (subjective norms), and feel they have full control over the AI's output (Selemani et al., 2018). On the other hand, the lack of ability to evaluate, verify, and use information ethically leads to users not understanding the ethical boundaries of text reproduction (Muluk et al., 2021). In the context of Islamic education, this reflects a crisis in the internalization of the value of scientific trust, namely the moral responsibility in producing and conveying knowledge. Without reinforcing this value, any technology, even one designed to enhance integrity, can be misused.

The significant difference in the impact of ChatGPT and Scite on plagiarism reflects fundamental differences in the functions and working methods of the two tools. ChatGPT operates as a content generator that generates original text based on prompts, allowing users to copy output without adequate intellectual processing—such as paraphrasing, synthesis, or critical analysis (Sterner, 2021). This model encourages surface learning, where students simply pursue task completion without understanding the essence of the material. In contrast, Scite serves as an evidence validator that displays the context of citations in the scientific literature, supposedly leading users to a deeper understanding (Nicholson et al., 2021)

CONCLUSION

The findings indicate that both tools positively contribute to an increased risk of plagiarism, with ChatGPT having a significantly more dominant influence than Scite. This reveals a paradoxical reality in the digital academic ecosystem: even technologies designed to support scientific integrity, such as Scite, are not completely immune to misuse if not balanced with adequate ethical literacy. Meanwhile, ChatGPT's generative nature, which enables instant text production without deep cognitive processing, has the potential to encourage students to engage in surface learning and technological dependency. These results confirm that the risk of plagiarism stems not only from the technology itself, but also from how users interpret and utilize it in their daily academic practices. In the context of values-based higher education, these findings underscore the urgency of not only

adopting technological innovations but also ensuring that their use aligns with the principles of intellectual honesty, scientific responsibility, and respect for the work of others.

The primary implications of this research are twofold: theoretically, it enriches the discourse on digital literacy by demonstrating that AI literacy must encompass ethical and epistemological dimensions, not just technical skills. Practically, these findings challenge Islamic higher education institutions to reform their teaching approaches and academic policies in the face of the wave of AI adoption. The integration of values such as adab al-ilm (the principle of knowledge) and amanah ilmu (scientific trust) is an important foundation for building moral defenses against the misuse of technology. This research also makes an original contribution by being one of the first to examine the impact of Scite in the context of plagiarism, while simultaneously situating the issue of AI within the framework of Islamic academic ethics. The recommendations are clear: the use of AI in education must be guided by critical literacy training, transparent policies, and an academic culture that emphasizes authentic learning processes, not just outcomes. Without these, the technology that should be liberating could potentially bind the minds and academic morals of future generations of students.

REFERENCES

Arianto, N., & Octavia, B. (2021). The effect of service quality and distribution on purchase decisions. Journal of Business Disruption, 4(2), 98. https://doi.org/10.32493/drb.v4i2.9867

Božić, V., & Poola, I. (2023). Chat GPT and education (Preprint). https://doi.org/10.31217/osf.io/8xgj6

Buragohain, D., & Chaudhary, S. (2025). Navigating ChatGPT in ASEAN higher education: Ethical and pedagogical perspectives. Computer Applications in Engineering Education, 33(4). https://doi.org/10.1002/cae.70062

deLaubell, L. (2023). Scite. The Charleston Advisor, 24(4), 60–63. https://doi.org/10.5260/chara.24.4.60

Fırat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. Journal of Applied Learning & Teaching, 6(1). https://doi.org/10.37074/jalt.2023.6.1.22

Fitri, Y., & Arum, P. (2023). Analysis of factors that affect the human development index based on districts/cities in Central Java. J Statistika: Jurnal Ilmiah Teori Dan Aplikasi Statistika, 16(1), 424–433. https://doi.org/10.36456/jstat.vol16.no1.a7079

Golzar, J., Noor, S., & Tajik, O. (2022). Convenience sampling. International Journal of Education & Language Studies, 1(2), 72–77.

Hasanein, A., & Sobaih, A. (2023). Drivers and consequences of ChatGPT use in higher education: Key stakeholder perspectives. European Journal of Investigation in Health Psychology and Education, 13(11), 2599–2614. https://doi.org/10.3390/ejihpe13110181

Holmner, M. (2025). The future of academic writing in the age of generative AI. Proceedings of the Association for Information Science and Technology, 62(1), 1241–1243. https://doi.org/10.1002/pra2.1374

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. British Journal of Applied Science & Technology, 7(4), 396.

Khalil, M., & Er, E. (2023). Will ChatGPT get you caught? Rethinking of plagiarism detection. https://doi.org/10.35542/osf.io/fnh48

Klein, N. (2024). Distributional regression for data analysis. Annual Review of

Statistics and Its Application, 11, 321–346.

Liang, Y., Wu, C., Song, T., Wu, W., Xia, Y., Liu, Y., … Duan, N. (2024). TaskMatrix.AI: Completing tasks by connecting foundation models with millions of APIs. Intelligent Computing, 3, 0063.

Lund, B., & Shamsi, A. (2023). Examining the use of supportive and contrasting citations in different disciplines: A brief study using Scite (scite.ai) data. Scientometrics, 128(8), 4895–4900.

Lyu, B., & Li, J. (2024). Multiple-model polynomial regression and efficient algorithms for data analysis. Theoretical Computer Science, 1021, 114878.

Maalouf, M. (2011). Logistic regression in data analysis: An overview. International Journal of Data Analysis Techniques and Strategies, 3(3), 281–299.

Martens, H. A., & Dardenne, P. (1998). Validation and verification of regression in small data sets. Chemometrics and Intelligent Laboratory Systems, 44(1–2), 99–121.

Morris, E. (2018). Academic integrity matters: Five considerations for addressing contract cheating. International Journal for Educational Integrity, 14(1). https://doi.org/10.1007/s40979-018-0038-5

Muluk, S., Habiburrahim, H., & Safrul, M. (2021). EFL students' perception on plagiarism in higher education: Triggering factors and avoiding strategies. Jurnal Ilmiah Didaktika: Media Ilmiah Pendidikan Dan Pengajaran, 22(1), 20. https://doi.org/10.22373/jid.v22i1.8944

Nicholson, J., Mordaunt, M., Lopez, P., Uppala, A., Rosati, D., Rodrigues, N., … Rife, S. (2021). Scite: A smart citation index that displays the context of citations and classifies their intent using deep learning. https://doi.org/10.1101/2021.03.15.435418

Okoro, E. (2011). Academic integrity and student plagiarism: Guided instructional strategies for business communication assignments. Business Communication Quarterly, 74(2), 173–178. https://doi.org/10.1177/1080569911404064

Özden, M. G., Liu, X., Wilkinson, T. J., Üstün-Yavuz, M. S., & Morley, N. A. (2024). Predictive modelling of laser powder bed fusion of Fe-based nanocrystalline alloys based on experimental data using multiple linear regression analysis. Heliyon, 10(15).

Perkins, M., & Roe, J. (2024). Decoding academic integrity policies: A corpus linguistics investigation of AI and other technological threats. Higher Education Policy, 37(3), 633–653.

Roustaei, N. (2024). Application and interpretation of linear-regression analysis. Medical Hypothesis, Discovery and Innovation in Ophthalmology, 13(3), 151.

Segovia-Juárez, J., & Baumgartner, R. (2023). The use of artificial intelligence applications for education and scientific research. Hatun Yachay Wasi, 3(1), 98–111. https://doi.org/10.57107/hyw.v3i1.61

Selemani, A., Chawinga, W., & Dube, G. (2018). Why do postgraduate students commit plagiarism? An empirical study. International Journal for Educational Integrity, 14(1). https://doi.org/10.1007/s40979-018-0029-6

Sterner, E. (2021). Scite: Providing additional context for citation statements. Issues in Science and Technology Librarianship, (98). https://doi.org/10.29173/istl74

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity

and student learning. Journal of Applied Learning & Teaching, 6(1). https://doi.org/10.37074/jalt.2023.6.1.17

Williams, R. (2024). The ethical implications of using generative chatbots in higher education. Frontiers in Education, 8. https://doi.org/10.3389/feduc.2023.1331607

Zeide, H., Lutsenko, O., Svyrydiuk, T., & Seminikhyna, N. (2024). Empowering language learners' critical thinking: Evaluating ChatGPT's role in English course implementation. https://doi.org/10.31235/osf.io/fh6v8

Zhang, S., Zhao, X., Zhou, T., & Kim, J. H. (2024). Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. International Journal of Educational Technology in Higher Education, 21(1), 34.