Online Learning in Educational Research
Volume 6, Issue 2, 155-167.
e_ISSN: 2798-2351
https://www.journal.foundae.com/index.php/oler/index
DOI: 10.58524/oler.v6i2.981

Deep-AI: An AI-Integrated Deep Learning Model for Advancing Scientific Literacy among Generation Z

Indrawati Wilujeng*
STKIP Al Hikmah Surabaya, INDONESIA

Trise Nurul Ain
STKIP Al Hikmah Surabaya, INDONESIA

Adhy Putri Rilianti
STKIP Al Hikmah Surabaya, INDONESIA

Faiz Hasyim
STKIP Al Hikmah Surabaya, INDONESIA

Maskanatul Fiqiyah
Universiti Pendidikan Sultan Idris, MALAYSIA

Article Info

Article history:

Received: November 29, 2026
Revised: May 7, 2026
Accepted: May 27, 2026

Keywords:

Deep-AI;
Generation Z;
Learning model;
Scientific literacy.

Abstract

The growing adoption of artificial intelligence (AI) in education has created new opportunities to strengthen scientific literacy, yet most AI-supported approaches function as instructional tools rather than integrated pedagogical models. This study investigated the preliminary effectiveness of Deep-AI, an AI-integrated deep learning model designed to enhance Generation Z students' scientific literacy through adaptive learning, scientific inquiry, and reflective learning processes. A quasi-experimental pretest–posttest design with replication across two cohorts was conducted involving 100 undergraduate students. Both cohorts participated in the same Deep-AI learning intervention, and scientific literacy was assessed before and after implementation. The findings demonstrated consistent improvements in students' scientific literacy across both cohorts, with moderate learning gains and strong educational effects. Students also reported highly positive learning experiences, indicating that the model successfully promoted engagement, conceptual understanding, and evidence-based reasoning. Comparable outcomes between the two cohorts further suggest the consistency and robustness of the model across different learning groups. This study contributes an AI-integrated instructional framework that positions artificial intelligence as a cognitive learning partner, providing a promising pedagogical approach for fostering scientific literacy and higher-order thinking in higher education within the Smart Society 5.0 era.

To cite this article: Wilujeng, I., Ain, T. N., Rilianti, A. P., Hasyim, F., & Fiqiyah, M. (2026). Deep-AI: An AI-integrated deep learning model for advancing scientific literacy among generation Z. *Online Learning in Educational Research,6*(2), 155-167. https://doi.org/10.58524/oler.v6i2.981

INTRODUCTION

The Society 5.0 era is characterized by the integration of digital technology into almost all aspects of life, which requires people to have the ability to think critically, creatively, and adaptively in facing the complexity of global problems (Wilujeng et al., 2025). Consequently, education must respond by fostering scientific literacy and integrating AI-supported learning approaches (Relmasira & Donaldson, 2025;Wang et al., 2024). One of the relevant strategies to face this era is the implementation of smart education (Dimitrova et al., 2024). Smart education emphasizes the use of technology to support an innovative and student-centered learning process (Ma'arif et al., 2025; Shakhina et al., 2023). Along with this, the world is currently focused on achieving SDG 4, namely

* **Corresponding Author**
Indrawati Wilujeng, STKIP Al Hikmah Surabaya, INDONESIA ✉ indrawati.physics@gmail.com

quality education (McCullough et al., 2022). For this reason, people must have scientific literacy to understand, use, and evaluate, not only knowledge but also technology (Istyadji & Sauqina, 2023).

In the context of the Smart Society era, scientific literacy has become a fundamental competency that includes not only understanding concepts but also the ability to evaluate, apply, and innovate with scientific knowledge. Scientific literacy is considered a crucial 21st-century skill for Generation Z, especially because they will lead the Indonesian nation towards a Golden Indonesia 2045 (Listiani, 2025; Rafida et al., 2024). In PISA 2006 science, scientific literacy refers to the ability to apply scientific knowledge, understand science as inquiry, recognize its role in shaping society, and engage with science-related issues as a reflective citizen (Bybee et al., 2009). There are 4 levels in scientific literacy with different levels of difficulty, namely nominal, conceptual, functional, and multidimensional (Istyadji & Sauqina, 2023; Rahmani et al., 2021). The majority of Generation Z, however, only possess indicators with the lowest level of difficulty (Sholahuddin et al., 2021).

Various international reports, including the Programme for International Student Assessment (PISA), increasingly confirm that Indonesia consistently shows low levels of scientific literacy (Amaditha et al., 2024; Prasasti & Anas, 2023; Yusmar & Fadilah, 2023). In 2018, Indonesia was ranked 70th out of 78 participating countries in scientific literacy (OECD, 2019). Between 2018 and 2022, scientific literacy scores showed no significant improvement (Lestari et al., 2024). Students at universities are expected to possess advanced scientific literacy skills (Berlian et al., 2021). However, the scientific literacy of 76,9% of students is still at the lowest level (Muhajir et al., 2021). Syahmani et al., (2024) reported similar findings regarding students' low scientific literacy. According to Sumanik et al. (2021), students' scientific literacy in the knowledge aspect is relatively low. Meanwhile, according to Limiansih et al. (2021), the mean scientific literacy of students is low. This is quite worrying because individuals with these scientific literacy skills will find difficulties in processing scientific information (Yang et al., 2024). This condition emphasizes the need for more effective educational interventions that align with the characteristics of contemporary university students. In this study, all participants are analytically positioned as Generation Z, referring to individuals born between 1997 and 2012 (Jayatissa, 2023), which corresponds to the current cohort of higher education students.

One way to increase scientific literacy is by using the right learning model (Maiyanti et al., 2025; Nurwati et al., 2024; Saputra & Octavia, 2024). This is in line with the results of the research of Jihannita et al (2024), which revealed that classroom learning activities can improve scientific literacy skills. However, conventional learning models are considered less interactive and fail to build long-term retention (Bitu et al., 2024). Meanwhile, although inquiry-based or discovery learning represents a broad family of student-centered approaches, their effectiveness can be influenced by contextual factors. For example, limitations in individual adaptability and digital infrastructure have been reported as challenges in certain settings (Mrosso et al., 2025). Therefore, a new learning model needs to be implemented to improve students' scientific literacy. The Deep-AI learning model was chosen to improve Generation Z's scientific literacy because it combines a deep learning approach (deep understanding through critical thinking and contextual application) with AI integration for adaptive personalization, instant feedback, and interactive simulations. These features have been associated with increased learner engagement, motivation, and improved concept retention, aligning with the preferences of Generation Z as digital natives (Deckker & Sumanasekara, 2025; Syukur et al., 2023).

As digital natives, Generation Z closely follows the latest technological developments (Yakubu et al., 2025). Currently, especially post-pandemic, AI is a rapidly developing technology. This can be utilized to enrich the learning experience and improve the overall quality of learning for Generation Z (Chen et al., 2020; Wulansari et al., 2025; Zulfiani et al., 2025). AI in education has an important role, including for personalized learning, reliable tutors, chatbots, and automated scoring systems that can save time, increase efficiency, and provide consistent feedback (Harry & Sayudin, 2023). AI can provide adaptive feedback, create personalized learning, and increase student engagement. Today, AI is a technology that plays an important role in improving the learning experience (Alamari, 2025). AI has great potential in improving scientific literacy through a personal, adaptive, and interactive approach (Jannah et al., 2025; Linh, 2025; Rajwaa & Mukti, 2025). Generation Z, who grew up with digital technology, needs to expand their interactions with AI, because most future job fields are projected to be highly dependent on this technology.

Previous research has explored various AI-supported learning approaches, including intelligent tutoring systems, adaptive learning environments, AI-assisted inquiry learning, and AI-supported scientific literacy intervention (Engka et al., 2026). While these approaches demonstrate the potential of AI for personalization and feedback, they are typically implemented as tools within existing pedagogical frameworks rather than as integrated instructional models (Létourneau et al., 2025). Therefore, a gap remains in developing a structured learning model that explicitly integrates deep learning approaches with AI to enhance scientific literacy. Recent studies suggest that artificial intelligence can support personalized learning experiences and adaptive assessment, offering opportunities to enhance students' learning outcomes and engagement (Choiriyah et al., 2025). Thus, the purpose of this study is to fill the gap by testing the effectiveness of the application of the Deep-AI learning model in improving the scientific literacy of Generation Z students. Specifically, this study aims to investigate the extent to which the model improves students' scientific literacy and to assess the consistency of these improvements across the experimental and replication classes, thus it can be recommended as an innovative alternative in supporting the achievement of scientific literacy and 21st-century skills.

METHOD

This study used a quantitative method with a quasi-experimental pretest-posttest design with replication across two classes (Nayeri et al., 2023). Two classes were used: the experimental class (Class A) and the replication class (Class B) (Fraenkel et al., 2012). Both received the same treatment, namely the Deep-AI learning model, and were administered pretest and posttest to examine the consistency of the model's effects. The participants belong to Generation Z, defined as individuals born between 1997 and 2012 (Jayatissa, 2023), which aligns with the current cohort of university students. The research subjects involved were as many as 100 students from two different study programs: Physics Education Study Program (PE) and Primary School Education Study Program (PSE). The inclusion of students from different academic backgrounds may have influenced baseline scientific literacy levels. Therefore, this study does not assume equivalence between groups but instead focuses on within-group improvement and the consistency of patterns across cohorts. Scientific literacy was considered relevant to both study programs, as it supports critical reasoning in analyzing scientific phenomena and effective science teaching practices. It also aims to test the generalizability of the Deep-AI model by replicating the results in diverse educational contexts, thereby increasing the reliability and validity of the findings.

The 100 students enrolled in the 2023 and 2024 cohorts represent both study programs (41 students from PE and 59 students from PSE). The distribution of gender is 55% female and 45% male students. The distribution of participants across classes and study programs was as follows: Class A: 21 PE students and 20 PSE students, Class B: 20 PE students and 39 PSE students. The study was conducted at STKIP Al Hikmah, providing the institutional context of a teacher education institution in Indonesia. Most participants had moderate prior experience with AI-supported learning, while their prior science achievement was categorized as low. They were selected based on an initial scientific literacy assessment adapted from PISA. Those classified in the low category, according to predetermined cut-off criteria, were included using purposive sampling. This approach may introduce selection bias and limit generalizability.

The study was conducted in several sequential stages. It began with a preliminary study to identify students' initial scientific literacy level, followed by participant selection using purposive sampling. The intervention was implemented using a one-group pretest-posttest design with replication across two classes. The Deep-AI learning model was applied over two sessions, each lasting 100 minutes, focusing on the topic of electric circuits. Pretest and posttest were administered to measure changes in students' scientific literacy. To ensure implementation fidelity, the learning activities were carried out based on structured lesson plans and consistently monitored throughout the intervention. The data were subsequently analyzed and interpreted to evaluate the effectiveness and consistency of the model, followed by drawing conclusions and formulating recommendations. The flowchart of that procedure is shown in Figure 1.

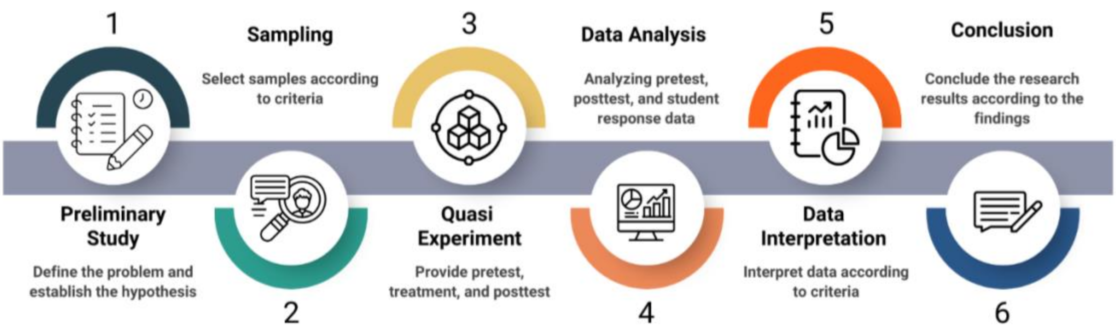

Figure 1. Research Procedure

During the preliminary study, reference studies and field studies were conducted. The instruments used were lecturer interview sheets and student questionnaires. In the quasi-experimental stage, the implementation of the Deep-AI learning model was carried out. The Deep-AI model books and model learning tools had previously been validated by experts and can be used to test their effectiveness (Wilujeng et al., 2025). The Deep-AI learning model was validated by three experts using a 4-point Likert scale. The results showed a high level of validity, with a mean score of 3.69 (very valid) and an Aiken's V coefficient of approximately 0.90. Inter-rater agreement indicated a high level of consistency among the validators. For example, the Deep-AI model development needs component received scores of 3.5, 3.5, and 2.5, reflecting strong agreement and validity across raters. The syntax of the Deep-AI learning model and the measured level of scientific literacy are shown in Table 1.

Table 1. Deep-AI Learning Model's Syntax and Measurable Level of Scientific Literacy

Learning phase	Measurable level of scientific literacy	Role of AI
Activate	Nominal	AI provides diagnostic question prompts and flash quizzes to identify students' prior knowledge
Explore	Conceptual	AI helps provide simulations as well as explanations of concepts.
Interpret	Conceptual	AI analyzes experimental results/simulation data and provides feedback on student understanding.
Create	Functional	AI supports students in designing through guided design suggestions and design error checking.
Share	Multidimensional	AI helps students prepare reports with data visualization, graphs, and communication enhancement tools.
Reflect	Multidimensional	AI provides automatic reflection in the form of an analysis of the strengths and weaknesses of student learning outcomes and recommendations for follow-up actions.

At this stage, the instruments used included: 1) a scientific literacy skills test sheet (pretest-posttest), 2) an observation sheet, and 3) a student response questionnaire sheet. The pretest-posttest instrument consisted of 8 essay items designed to assess varying levels of scientific literacy. A scoring rubric with a 4-point scale was developed, with specific descriptors provided for each item. Content and construct validity were established through expert judgement. The reliability of the instrument was assessed using Cronbach's alpha (α= 0.729), indicating acceptable internal consistency. One multidimensional question, for example, is "*How would you design a circuit to ensure that if one component fails, the others will still function?*" This item is considered multidimensional because it requires the integration of conceptual understanding, scientific reasoning, and problem-solving in a real-world context, along with justification of the proposed solution. In the next stage, the students' pretest and posttest results were analyzed. Responses were scored manually using the established rubric by two independent raters. To ensure scoring consistency, both raters were trained on the rubric before the assessment. In cases of scoring discrepancies, discussions were conducted to reach consensus. Inter-rater reliability was ensured by involving two raters and calculated using the Intraclass Correlation Coefficient (ICC) (ICC pretest=0.83, good; ICC posttest= 0.89, good), demonstrating a high level of agreement. Each student's pretest-posttest scores were analyzed using SPSS software. The results of the analysis were interpreted to draw conclusions and provide recommendations based on these results. The effectiveness of the learning model was analyzed based on the mean n-gain score, t-test, and student responses.

N-Gain

N-gain indicates the level of students' scientific literacy improvement (Sutinah & Ristiana, 2023). Comparing the difference between the pretest and the highest score with the difference between the pretest and the posttest yielded the n-gain score. Next, the n-gain score was compared using Table 2's criterion.

Table 2. Category of N-Gain (Hasyim et al., 2020)

N-Gain score	N-Gain category
0.70<N-Gain	High
0.30<N-Gain<0.70	Medium
N-Gain<0.30	Low

The n-gain was used as one of the criteria to indicate improvement in students' scientific literacy, with a value of ≥ 0.30 (medium category) considered as the minimum threshold.

Statistical Test

The paired t-test and the independent t-test were the two methods employed. The independent t-test was used to determine whether there was a statistically significant difference in the mean gain scores between the two classes, and the paired t-test was used to determine whether there was a significant increase in students' scientific literacy between pretest and posttest scores (Hasyim et al., 2020). However, both of the above tests can only be used when the data is distributed normally and homogeneous. Therefore, it is necessary to carry out a normality test and a homogeneity test first. The normality test will show whether the sample data used is distributed normally or not. It was evaluated based on the distribution of the difference scores (posttest - pretest). Meanwhile, the homogeneity test shows whether the data used is homogeneous or not.

The paired t-test is used to ascertain if an effect exists if the data is regularly distributed. However, the Wilcoxon test is applied if the data is not regularly distributed. Once an effect is known, the next step is to determine its magnitude (effect size). Effect size was calculated using Cohen's dz for paired data, calculated by dividing the standard deviation of the variations in scores by the mean difference between the pretest and posttest scores. Next, the outcomes are contrasted with Table 3's classifications.

Table 3. Effect Size Category (Hasyim et al., 2024)

Interval	Category
>1.00	Strong effect
0.51 – 1.00	Moderate effect
0.21 – 0.50	Modest effect
0 – 0.20	Weak effect

To determine whether the effect produced is consistent or not, an independent t-test is required. The independent t-test is performed when the data is distributed normally. However, if the data is not normally distributed, then the Mann-Whitney test is used.

Student Responses

Both quantitative and qualitative analyses were performed on the student answer surveys. The percentage ratio between the students' total score and the maximum possible combined grade was computed to perform the quantitative analysis. The computation's outcomes were then contrasted with Table 4's criteria.

Table 4. Student Response Percentage Criteria

Percentage (%)	Assessment criteria
81 – 100	Excellent
61 – 80	Very Good
41 – 60	Good
21 – 40	Fair
0 – 20	Poor

Three complementing criteria were used to assess the model's efficacy: 1) n-gain falls into the moderate or higher group; 2) scientific literacy significantly increases at $\alpha=0.05$ with $p<0.05$; and 3) student answers fall into the good category (Hasyim et al., 2024). These criteria were used collectively to provide a comprehensive assessment of the model's effectiveness.

RESULTS AND DISCUSSION

The Deep-AI learning model has been proven valid and practical by previous research. To evaluate the preliminary effectiveness of the model, it must meet three criteria, namely n-gain, t-test, and student response questionnaire.

N-Gain Results

The N-gain was calculated based on the pretest and posttest scores of the scientific literacy skills test sheet. The n-gain results of class A and class B are shown in Table 5.

Table 5. N-Gain of Class A and Class B

Description	Class A		Class B	
	Pretest	Posttest	Pretest	Posttest
Mean score	14.35	23.73	15.63	24.10
The mean of N-gain	0.52		0.51	
Category	Medium		Medium	
Standar Deviation (SD) N-gain	0.16		0.15	
Confidence Interval (CI) N-gain	95% (0.47 – 0.57)		95% (0.48 – 0.55)	

Table 5 shows that the mean posttest score was greater than the pretest score, both in class A and class B. The mean n-gain for class A was 0.52, while for class B it was 0.51. Both were in the medium category. To provide a more detailed analysis, N-gain scores were examined across scientific literacy levels (nominal, conceptual, functional, and multidimensional), including standard deviation and 95% confidence interval to describe variability and precision. The distribution of scores for each indicator is presented in Table 6.

Table 6. N-Gain Scores of each Scientific Literacy Level

Class	Scientific Literacy Levels	N-Gain	Standard Deviation	95% Confidence Interval
A	Nominal	0.60	0.43	(0.47, 0.74)
	Conceptual	0.58	0.23	(0.51, 0.66)
	Functional	0.52	0.19	(0.46, 0.58)
	Multidimensional	0.42	0.18	(0.36, 0.47)
B	Nominal	0.66	0.40	(0.55, 0.76)
	Conceptual	0.58	0.26	(0.51, 0.64)
	Functional	0.50	0.20	(0.45, 0.55)
	Multidimensional	0.42	0.16	(0.38, 0.46)

Statistical Test Results

On the other hand, a normality test was carried out using the Shapiro-Wilk test (Anwar et al., 2024; Gusvianti et al., 2025) and a homogeneity test using the *Levene* test (Thomann et al., 2024). Class A and class B n-gain data were utilized. After the Deep-AI learning model was put into practice, this test was utilized to see if science literacy had significantly increased. The results of the assumption testing are presented in Table 7.

Table 7. Assumption Testing Results for N-Gain Scores

Assumption Test	Data	Statistic	Sig.	Interpretation
Shapiro–Wilk	Class A N-gain	0.925	0.221	Normally distributed
Shapiro–Wilk	Class B N-gain	0.901	0.124	Normally distributed
Levene's Test	Class A vs Class B N-gain	–	0.510	Homogeneous

Before conducting the inferential analysis, assumption testing was performed on the N-gain scores. The results are presented in Table 7. The Shapiro–Wilk test indicated that the N-gain scores of both Class A ($p = 0.221$) and Class B ($p = 0.124$) were normally distributed. Furthermore, Levene's test showed that the variance of the two groups was homogeneous ($p = 0.510$). Since both normality and homogeneity assumptions were satisfied, parametric analyses using paired and independent t-tests were considered appropriate. The results of these analyses are summarized in Table 8.

Table 8. Summary of Inferential Statistical Analysis

Analysis	Class/Data	Mean Difference	SD Difference	p-value	Effect Size	Interpretation
Paired t-test	Class A (Pretest–Posttest)	9.38	3.68	< 0.001	2.55	Significant improvement; Strong effect
	Class B (Pretest–Posttest)	8.47	2.70	< 0.001	3.14	Significant improvement; Strong effect
Independent t-test	Class A vs Class B (N-gain)	0.01	–	0.840	0.006	No significant difference between groups

Note. The independent t-test comparing N-gain scores between Class A and Class B yielded $t(98) = 0.202$, $p = 0.840$, with a 95% confidence interval ranging from -0.06 to 0.07, indicating comparable outcomes across groups.

The inferential statistical analysis is summarized in Table 8. The paired t-test results revealed significant improvements in scientific literacy scores from pretest to posttest in both Class A and Class B ($p < 0.001$). The corresponding effect sizes were 2.55 and 3.14, respectively, indicating strong effects of the Deep-AI learning model. Furthermore, the independent t-test showed no significant difference in N-gain scores between the two classes ($t(98) = 0.202$, $p = 0.840$), suggesting that the model produced consistent outcomes across cohorts.

Student Responses Result

The third requirement that must be fulfilled relates to student responses. These responses refer to students' responses after they had been involved in implementing the Deep-AI learning model. These responses can include both advantages and disadvantages of the implementation. Response data was obtained from a questionnaire distributed after the lesson. The results of the student responses are shown in Figure 2.

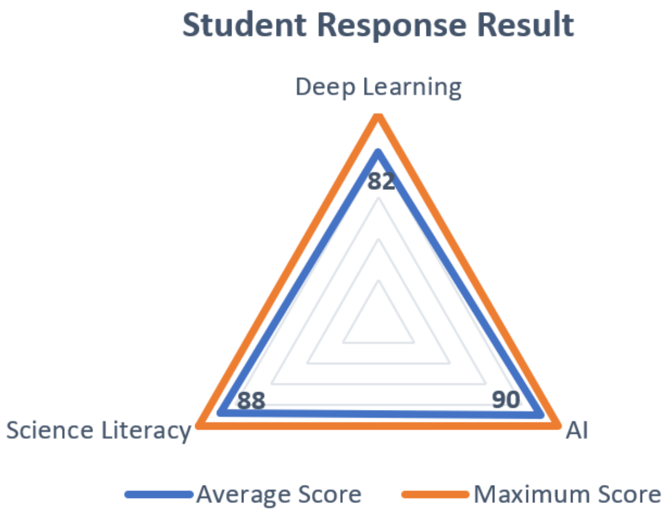

Figure 2. Student Response Results

Based on Figure 2, students responded to several aspects, including deep learning, the use of AI, and scientific literacy skills. The mean score for all three aspects reached 86,7 categorized as excellent. This shows that students responded positively to the implementation of the Deep-AI learning model.

Discussion

The implementation of the Deep-AI learning model showed how the six-phase syntax was able to increase scientific literacy indicators. In the activate phase, trigger questions such as *"What differences do you predict between current flow in series and parallel circuits?"* successfully elicited students' initial conceptions, including common misconceptions (e.g., the belief that electric current is "used up" after passing a resistor). This early elicitation became an important foundation for subsequent learning, because the misconceptions identified at the beginning resurfaced throughout the learning process and were gradually corrected as students engaged with AI-supported evidence (El Fathi et al., 2025).

In the explore and interpret phases, AI appeared to function as a digital scaffolding tool rather than as a source of direct answers, as reflected in the design of the learning activities and students' engagement patterns. Students checked their initial predictions by consulting scientific sources, visualizing phenomena, and testing values using interactive simulations and computational tools. The combination of AI-powered information retrieval and peer discussion resulted in deeper meaning-making; students did not simply replace old ideas with new information, but negotiated conceptual change through evidence-based argumentation. This aligns with the perspective that scientific understanding develops when learners continuously compare their existing mental models with empirical or simulated evidence (Hoffman et al., 2023; Létourneau et al., 2025).

The create and share phases further strengthened students' conceptual understanding by requiring them to externalize their reasoning through the design of simple circuit simulations and pedagogical explanations intended for school-level learners. Producing a circuit model using PhET Simulations combined with AI feedback encouraged precision, because students were compelled to justify the logic behind each connection rather than rely on procedural trial and error. Sharing their work with peers through the learning management system (LMS) also amplified accountability, as students were aware that their work would be visible to classmates within a structured and instructor-moderated environment. It is an aspect strongly associated with the development of scientific literacy and pedagogical readiness (Spasopoulos et al., 2025).

Finally, the reflect phase contributed to metacognitive development by encouraging students to acknowledge the evolution of their understanding. AI helped prompt reflection, not by judging learning, but by encouraging students to identify which misconception was corrected, what evidence led to the conceptual shift, and what reasoning strategies were most effective. This indicates that AI was most productive when positioned as a cognitive partner by supporting thinking processes rather than replacing them. Collectively, the six phases did not merely sequence learning activities but shaped a coherent pathway from eliciting prior knowledge to achieving conceptual, pedagogical, and metacognitive outcomes in electrical circuits. These findings suggest that the Deep-AI learning model has strong potential for teacher education contexts, where mastery of scientific concepts must be accompanied by the ability to explain them accurately and meaningfully to future students.

The purpose of developing this learning model is to improve the scientific literacy of Generation Z. Based on the explanation of each phase above, the goal is focused on the explore, interpret, and create phases. This is because in these three phases, students explore data, interpret scientific information, and produce something science-based as a result of their thought process. However, that doesn't mean the other phases aren't important. The activation phase plays an important role as a trigger for initial engagement. Meanwhile, the share and reflect phases strengthen understanding through communication and evaluation. These findings suggest that AI-facilitated exploration, interpretation, and creation processes can provide a more immersive and contextual learning experience for Generation Z.

The Deep-AI learning model is developed from the master framework of the information process model, which also serves as the foundation for Problem-Based Learning (PBL) and Project-Based Learning (PjBL) models. However, the Deep-AI learning model offers novelty by integrating a deep learning approach and the explicit use of AI at each phase. The role of AI in the Deep-AI learning model lies in its ability as a digital scaffolding that supports higher-order thinking processes. In classroom implementation, AI functioned as an intelligent learning assistant that supported students' thinking processes through adaptive prompts and feedback based on their inputs. However, this support was limited to interaction-based responses and did not involve direct real-time analysis of students' cognition. Through automated assessment features based on data analytics, AI helps students organize scientific information, visualize experimental data, and simulate scientific phenomena, allowing them to simultaneously develop critical, analytical, and creative thinking skills. Thus, AI is not merely a technical tool but acts as a cognitive partner in the classroom, strengthening the formation of deep scientific concepts through personalized learning interactions (Chen et al., 2020; Hoffman et al., 2023).

In general, there was a significant improvement in students' post-test scores after participating in the implementation of the Deep-AI learning model. Based on Table 8, the effect sizes were 2.55 for Class A and 3.14 for Class B. These values are categorized as very large effect sizes. This may have been influenced by the limited scoring scale (1-4), which resulted in low score variability. Improvements in scientific literacy were observed across all levels, with a tendency toward higher gains at the nominal and conceptual levels based on Table 6. Students generally achieve greater progress at these levels because they primarily involve recognizing scientific concepts, recalling relevant information, and understanding basic relationships between phenomena (Bybee et al., 2009; Istyadji & Sauqina, 2023). These skills are more directly supported through instruction, discussion, guided activities, and repeated exposure during learning. This conclusion was based on posttest results, which indicated a substantial increase in scores at both levels, supported by the results of observation of the learning process, which confirmed the active involvement of students in problem-solving by utilizing AI.

Conversely, the multidimensional level showed relatively lower improvement compared to the other levels, based on Table 6 (Bybee et al., 2009; Rahmani et al., 2021). This can be due to the characteristics of students who mostly do not have sufficient initial scientific literacy, leaving relatively high room for improvement from the basic level. In addition, the orientation of Deep-AI learning model development, which emphasizes complex problem-solving, need stimulates functional and multidimensional skills more than simply mastering terminology.

Implicitly, this study suggests that the integration of AI into an information process-based learning model (Deep-AI learning model) may support improvements in students' scientific literacy, especially at the nominal and conceptual levels. However, these findings should be interpreted within the context of the study and do not imply a generalizable effect across broader populations. The findings may provide preliminary insights for educators in designing learning environments that incorporate AI as a supportive tool for student learning (Alamari, 2025; Chen et al., 2020; Choiriyah et al., 2025). Furthermore, these findings may have broader relevance in supporting long-term educational goals such as SDG 4, although this study does not directly evaluate such impacts.

LIMITATIONS

This study has several limitations. First, the absence of a control or comparison group limits the ability to attribute the observed improvements solely to the intervention. Second, the application of the model is still limited to the context of certain science courses and to Generation Z subjects in one educational institution, thus generalization of results to other fields or different age groups needs to be done carefully. Third, there is AI-access inequality; the use of AI technology is highly dependent on the availability of digital infrastructure and the readiness of lecturers who teach, which, of course, varies between institutions. Fourth, although AI supports the learning process, the potential testing effects, selection bias, and limited internet access can affect the fairness of learning outcomes. Finally, novelty effects cannot be ruled out, as students' initial engagement with an AI-supported learning model may have temporarily enhanced their performance. These limitations should be considered when interpreting the findings and in designing future studies.

CONCLUSION

The Deep-AI learning model was found to be effective, as indicated by the mean N-gain score of 0,515 in the medium category. The results of the paired t-test showed a significant difference between the pretest and posttest scores in both classes. The effect size resulting from the implementation of the learning model was 2.55 for class A and 3.14 for class B, with a category of strong effect. Meanwhile, the independent t-test confirmed the consistency of improvements in scientific literacy skills across both the experimental class and the replication class. In addition, the results of the student response questionnaire reached a mean score of 86,7, categorized as excellent. The findings suggest that students' scientific literacy improved after participating in the Deep-AI learning model, although the causal claim should be interpreted cautiously due to the absence of a conventional control group. Recommendations for further research are to implement the Deep-AI model by involving more research subjects from different campuses to further validate the consistency of its effectiveness.

AUTHOR CONTRIBUTIONS

IW contributed to conceptualization, methodology, supervision, manuscript writing, and funding acquisition. TNA contributed to methodology and project administration. ADP contributed to resource preparation and data entry. FH contributed to validation and formal analysis. MF contributed to formal analysis and supervision.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Prof. Dr. Sarwanto, S.Pd., M.Si., and Ms. Nuril Munfaridah, S.Pd., M.Pd., Ph.D., for their support and contribution as validators in this research. The constructive feedback and suggestions provided have significantly improved the quality of this research. The author expresses gratitude to the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia's Directorate of Research, Technology, and Community Service (DRTPM) for funding this study under the Regular Fundamental Research scheme with SP DIPA--139.04.1.693320/2025 revision 4 dated April 30, 2025.

REFERENCES

Alamari, A. B. M. (2025). A training system based on AI applications and chatbots : Impact on developing ethical awareness among postgraduate students. *International Journal of Advanced and Applied Sciences*, *12*(9), 180–185. https://doi.org/10.21833/ijaas.2025.09.017

Amaditha, F. F., Gia, S., Fitri, S., Ekanara, B., Sultan, U., & Tirtayasa, A. (2024). Students' communication skills and scientific literacy ability in biology through guided inquiry learning. *Internaional Journal of Biology Education Towards Sustainable Development*, *4*(1), 21–31. https://doi.org/10.53889/ijbetsd.v4i1.435

Anwar, R., Elbashir, A. M., Magdy, R., Ahmad, Z., & Al-Thani, N. J. (2024). Effectiveness of STEM based workshop for deaf education: Exploratory study. *Heliyon*, *10*(e36012), 1–12. https://doi.org/10.1016/j.heliyon.2024.e36012

Berlian, M., Mujtahid, I. M., Vebrianto, R., & Thahir, M. (2021). Profil literasi sains dalam pembelajaran IPA di era Covid-19: Studi kasus di Universitas Terbuka. *Journal of Natural Science and Integration*, *4*(1), 77–84. https://doi.org/10.24014/jnsi.v4i1.11662

Bitu, Y. S., Setiawi, A. P., Bili, F. G., Iriyani, S. A., & Patty, E. N. S. (2024). Pembelajaran interaktif: Meningkatkan keterlibatan dan pemahaman siswa. *J-KIP (Jurnal Keguruan Dan Ilmu Pendidikan)*, *5*(2), 193–198. http://dx.doi.org/10.25157/j-kip.v5i2.14697

Bybee, R., McCrae, B., & Laurie, R. (2009). PISA 2006: An assessment of scientific literacy. *Journal of Research in Science Teaching*, *46*(8), 865–883. https://doi.org/10.1002/tea.20333

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Choiriyah, S., Ramadhan, S., Nugroho, A., & Putra, H. R. (2025). Artificial intelligence-driven learning assessment in faculties of education : An exploratory study. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, *6*(3), 482–495. https://doi.org/10.31538/munaddhomah.v6i3.1937

Deckker, D., & Sumanasekara, S. (2025). The role of artificial intelligence in education: Transforming learning and teaching. *EPRA International Journal of Research and Development (IJRD)*, *10*(3), 5–15. https://doi.org/10.36713/epra20429

Dimitrova, V., Nikolov, N., & Gospodinov, T. (2024). Education in the era of smart cities: Transformation and opportunities. *Environment. Technology. Resources. Rezekne, Latvia Proceedings of the 15th International Scientific and Practical Conference*, *2*, 352–357. https://doi.org/10.17770/etr2024vol2.8084

El Fathi, T., Saad, A., Larhzil, H., Lamri, D., & Al Ibrahmi, E. M. (2025). Integrating generative AI into STEM education: Enhancing conceptual understanding, addressing misconceptions, and assessing student acceptance. *Disciplinary and Interdisciplinary Science Education Research*, *7*(1). https://doi.org/10.1186/s43031-025-00125-z

Engka, E., Fajar, M., Zaman, B., Informatika, T., Novel, V., Pages, M., & Black-box, P. (2026). Pemanfaatan artificial intelligence sebagai alat bantu pengujian aplikasi visual novel mystic pages menggunakan metode black-box. *Jurnal kharisma tech*, *01*, 1–14. https://doi.org/10.55645/kharismatech.v21i1.602

Fraenkel, J. F., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill.

Gusvianti, V., Haryadi, R., Rama, Y., Muchtar, D., Program, P. E., Sultan, U., & Tirtayasa, A. (2025). The effect of PhET simulation learning media to improve science proses skills. *Jurnal Pendidikan Indonesia Gemilang, 5*(1), 15–19. https://doi.org/10.53889/jpig.v5i1.547

Harry, A., & Sayudin, S. (2023). Role of AI in education. *Interdiciplinary Journal and Hummanity (INJURITY)*, *2*(3), 260–268. https://doi.org/10.58631/injurity.v2i3.52

Hasyim, F., Prastowo, T., & Jatmiko, B. (2020). The use of Android-based PhET simulation as an effort to improve students' critical thinking skills during the Covid-19 pandemic. *International Journal of Interactive Mobile Technologies*, *14*(19), 31–41. https://doi.org/10.3991/ijim.v14i19.15701

Hasyim, F., Prastowo, T., & Jatmiko, B. (2024). Critical thinking-independent learning: A model of learning to improve students' critical thinking skills. *European Journal of Educational Research*, *13*(2), 747–762. https://doi.org/10.12973/eu-jer.13.2.747

Hoffman, R. R., Mueller, S. T., Klein, G., & Litman, J. (2023). Measures for explainable AI: Explanation goodness, user satisfaction, mental models, curiosity, trust, and human-AI performance. *Frontiers in Computer Science*, *5*, 1096257. https://doi.org/10.3389/fcomp.2023.1096257

Istyadji, M., & Sauqina. (2023). Conception of scientific literacy in the development of scientific literacy assessment tools: A systematic theoretical review. *Journal of Turkish Science Education*, *20*(2), 281–308. https://doi.org/10.36681/tused.2023.016

Jannah, D. M., Huda, S., Yasin, M., Sanatang, Fakhri, M. M., Wahid, S. N., Amukune, A. B. K., Soeharto, & Stephen. (2025). AI literacy, technical skills, and ethical awareness in predicting students' learning performance. *Online Learning in Educational Research*, *6*(1), 1–12. https://doi.org/10.58524/oler.v6i1.565

Jayatissa, K. A. D. U. (2023). Generation Z – A new lifeline : A systematic literature review. *Sri Lanka Journal of Social Sciences and Humanities*, *3*(2), 179–186. https://doi.org/10.4038/sljssh.v3i2.110

Jihannita, J., Fadly, W., Ekapti, R. F., Luthfiana, D., & Widowati, A. (2024). The development of science module integrated with ethnoscience of singo barong mask to improve scientific literacy and cultural preservation attitudes. *Journal of Innovation in Educational and Cultural Research*, *5*(2), 356–363. https://doi.org/10.46843/jiecr.v5i2.790

Lestari, D. P., Paidi, P., & Suwarjo, S. (2024). Effect of the inquiry-based nature of science argumentation instructional model in scientific literacy skills. *Journal of Education and Learning*, *18*(3), 734–744. https://doi.org/10.11591/edulearn.v18i3.21024

Létourneau, A., Martineau, M. D., Charland, P., Karran, J. A., Boasen, J., & Léger, P. M. (2025). A systematic review of AI-driven intelligent tutoring systems (ITS) in K-12 education. *Npj Science of Learning*. https://doi.org/10.1038/s41539-025-00320-7

Limiansih, K., Melani, M., & Susanti, I. (2021). Identifikasi profil literasi sains mahasiswa PGSD. *DWIJA CENDEKIA: Jurnal Riset Pedagogik* , *5*(2), 313–325. https://doi.org/10.20961/jdc.v5i2.56281

Linh, T. T. (2025). Factors influencing the use of ChatGPT in student learning in Vietnam. *International Journal of Advanced and Applied Sciences,* *12*(7), 76–86. https://doi.org/10.21833/ijaas.2025.07.007

Listiani. (2025). Scientific literacy in Indonesian secondary education: Are we ready to be scientifically literate society? *International Journal of STEM Education for Sustainability*, *5*(1), 41–52. https://doi.org/10.53889/ijses.v5i1.445

Ma'arif, S., Fathoni, A. M., Suyudi, S., Yudi, U., & El Muna, N. (2025). The interplay of organizational culture, knowledge management system, and learning evaluation in improving English language skills: A mixed-method study in Indonesian pesantrens. *Kharisma: Jurnal Administrasi dan Manajemen Pendidikan, 4*(2), 272–286. https://doi.org/10.59373/kharisma.v4i2.107

Maiyanti, A. A., Huda, S., Anggraini, A., Laili, U. F., Muniroh, L., & Umam, R. (2025). Transformation of physics learning: Integrating virtual laboratories to improve students' scientific literacy skills. *Online Learning in Educational Research (OLER),* *5*(1), 155–172. https://doi.org/10.58524/oler.v5i1.661

McCullough, B. P., Kellison, T., & Melton, E. N. (2022). The routledge handbook of sport and sustainable development. In *Routledge Taylor and Francis Group*. https://doi.org/10.4324/9781003023968

Mrosso, V., Ndibalema, P., & Ephrem, A. (2025). Enhancing inquiry-based learning skills with ICT to pre-service teachers in Tanzanian teacher colleges: Constraints and mitigating strategies. *Educational Technology Quarterly*, *2025*(2), 169–185. https://doi.org/10.55056/etq.886

Muhajir, S. N., Masara, A., & Ramdani, G. (2021). Tingkatan literasi sains mahasiswa calon guru fisika. *Jurnal Pendidikan Universitas Garut*, *15*(1), 378–384. https://doi.org/10.52434/jpif.v1i1.1259

Nayeri, N. D., Noodeh, F. A., Nia, H. S., Yaghoobzadeh, A., & Goudarzian, A. H. (2023). Statistical procedures used in pretest-posttest control group design: A review of papers in five Iranian journals. *Acta Medica Iranica*, *61*(10), 584–591. https://doi.org/10.18502/acta.v61i10.15657

Nurwati, N., Abdullah, H., & Palloan, P. (2024). Enhancing scientific literacy skills: Development of physics teaching materials based on Islamic values. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni, 13*(2). https://doi.org/10.24042/jipfalbiruni.v13i2.24119

OECD. (2019). *OECD Multilingual summaries PISA 2018 Results (Volume I) What students know and can do (Spanish)*. *I*(Volume I), 2018–2020. https://doi.org/10.1787/5f07c754-en

Prasasti, R. D., & Anas, N. (2023). Pengembangan media digital berbasis flipbook untuk meningkatkan kemampuan berpikir kritis pada peserta didik. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, *4*, 694–705. https://doi.org/10.31538/munaddhomah.v4i3.589

Rafida, T., Suwandi, S., & Ananda, R. (2024). EFL students ' perception in Indonesia and Taiwan on using artificial intelligence to enhance writing skills. *Jurnal Ilmiah Peuradeun, 12*(3), 987-1016. https://doi.org/10.26811/peuradeun.v12i3.1520

Rahmani, R., Mustadi, A., Maulidar, & Anwar Senen. (2021). The development of teaching materials based on context and creativity to increase students scientific literacy. *Jurnal Ilmiah Peuradeun*, *9*(2), 345-364. https://doi.org/10.26811/peuradeun.v9i2.506

Rajwaa, D. N. E., & Mukti, W. A. H. (2025). Inovasi pembelajaran IPA berbasis kecerdasan buatan: Meningkatkan literasi sains di sekolah. *Jurnal Studi Ilmiah*, *1*(1), 1–6.

Relmasira, S. C., & Donaldson, J. P. (2025). Deep learning pedagogies enhance AI literacy in elementary students: A five-cycle implementation study. *Online Learning in Educational Research*, *5*(1), 143–154. https://doi.org/10.58524/oler.v5i1.720

Saputra, R. S., & Octavia, B. (2024). e-module assisted by augmented reality with a discovery learning model on virus material to increase scientific literacy and learning independence for class X high school students. *Jurnal Penelitian Pendidikan IPA*, *10*(11), 8507–8518. https://doi.org/10.29303/jppipa.v10i11.8795

Shakhina, I., Podzygun, O., Petrova, A., & Gordiichuk, G. (2023). Smart education in the transformation digital society. *Modern Information Technologies and Innovation Methodologies of Education in Professional Training Methodology Theory Experience Problems*, *67*(1), 51–64. https://doi.org/10.31652/2412-1142-2023-67-51-64

Sholahuddin, A., Hayati, N., Rilia Iriani, Saadi, P., & Susilowati, E. (2021). Project-based learning on ethnoscience setting to improve students' scientific literacy. *The 4th International Conference on Mathematics and Science Education (ICoMSE) 2020*, 1–8. https://doi.org/10.1063/5.0043571

Spasopoulos, T., Sotiropoulos, D., & Kalogiannakis, M. (2025). Generative AI in pre-service science teacher education: A systematic review. *Advances in Mobile Learning Educational Research*, *5*(2), 1501–1523. https://doi.org/10.25082/amler.2025.02.007

Sumanik, N. B., Nurvitasari, E., & Siregar, L. F. (2021). Analisis profil kemampuan literasi sains mahasiswa calon guru pendidikan kimia. *QUANTUM: Jurnal Inovasi Pendidikan Sains*, *12*(1), 22–32. https://doi.org/10.20527/quantum.v12i1.10215

Sutinah, C., & Ristiana, M. G. (2023). The development of assisted worksheets differentiation learning based on learning style : How great it can help students ' mathematical understanding ability? *Jurnal Cakrawala Pendas*, *9*(2), 205–214. https://doi.org/10.31949/jcp.v9i2.4541

Syahmani, S., Iriani, R., Kusasi, M., Prasetyo, Y. D., Norhasanah, H., & Rahman, F. A. (2024). Culturally wetland responsive teaching to improve science literacy and wasaka character. *Journal of Innovation in Educational and Cultural Research*, *5*(2), 196–206. https://doi.org/10.46843/jiecr.v5i2.1075

Syukur, P. A., Fakhri, M. M., Putra, K. P., Adiba, F., & Arifiyanti, F. (2023). AI literacy meets ethics: Critical appraisal's mediating role in shaping ethical awareness in higher education. *Online Learning in Educational Research*, *5*(1), 57–71. https://doi.org/10.58524/oler.v5i1.508

Thomann, H., Zimmermann, J., & Deutscher, V. (2024). Computers & education how effective is immersive VR for vocational education ? Analyzing knowledge gains and motivational effects. *Computers & Education*, *220*(July), 105127. https://doi.org/10.1016/j.compedu.2024.105127

Wang, X., Huang, J., Yang, J., Wang, X., & Wang, F. Y. (2024). Prescriptive manufacturing in society 5.0: Human autonomous organizations and "on-demand" smart services. *IFAC-PapersOnLine*, *58*(3), 139–144. https://doi.org/10.1016/j.ifacol.2024.07.140

Wilujeng, I, Ain, T. N., Rilianti, A. P., Hasyim, F., & Fiqiyah, M. (2025). The effectiveness of H5P-asisted differentiated-independent learning model to increase low-ability students' scientific literacy. *Jurnal Pendidikan IPA Indonesia*, *14*(1), 32–41. https://doi.org/10.15294/jpii.v14i1.19560

Wilujeng, Indrawati, Ain, T. N., Rilianti, A. P., & Hasyim, F. (2025). Development of a deep-AI learning model to improve generation Z's scientific literacy through a deep learning approach and AI integration. *Jurnal Ilmiah Pendidikan Fisika*, *9*(3), 461–471. https://doi.org/10.20527/jipf.v9i3.16836

Wulansari, O. D. E., Nursiyanto, T. M., Zaini, B. B., & Zaini. (2025). Augmented reality for 3D geometric shapes and nets combination for Android. *International Journal of Electronics and Communications System, 5*(1), 69–79. https://doi.org/10.24042/ijecs.v5i1.27574

Yakubu, M. A., Sain, Z. H., Lawal, U. S., Budiman, S. A., & Permana, D. (2025). Analysis of online academic writing tools usage among postgraduate students in North-west Nigeria. *Kharisma: Jurnal Administrasi dan Manajemen Pendidikan, 4*(1), 1–13. https://doi.org/10.59373/kharisma.v4i1.69

Yang, S., Brossard, D., Li, N., & Barolo, L. (2024). Bridging gaps in COVID-19 vaccine knowledge : Effects of multimodal narratives on message elaboration and recall across science literacy levels. *Clinical Epidemiology and Global Health*, *28*(June), 101681. https://doi.org/10.1016/j.cegh.2024.101681

Yusmar, F., & Fadilah, R. E. (2023). Analisis rendahnya literasi sains peserta didik Indonesia: Hasil PISA dan faktor penyebab. *LENSA (Lentera Sains): Jurnal Pendidikan IPA*, *13*(1), 11–19. https://doi.org/10.24929/lensa.v13i1.283

Zulfiani, Z., Suwarna, I. P., Zaky, R. A., Islami, E., & Sari, I. J. (2025). Trends in SAMR research in teaching and learning from 2019 to 2024 : A systematic review. *International Journal of Advanced and Applied Sciences, 12*(4), 99–106. https://doi.org/10.21833/ijaas.2025.04.012