Proceedings of Machine Learning Research 319:37–50, 2026 IndabaX Nigeria

EDAIL-EduAI-NG: Benchmarking Educator AI Readiness for Scalable Deployment in Low-Resource Classrooms

Oluwakemi D. Olurinola OLURINOLA.OLUWAKEMI@OOUAGOIWOYE.EDU.NG
Department of Science and Technology Education,
Olabisi Onabanjo University

Sakinat O. Folorunso Sakinat SAKINAT.FOLORUNSO@OOUAGOIWOYE.EDU.NG
The Artificial Intelligence System Group,
Department of Computer Sciences,
Olabisi Onabanjo University

Patrick I. Owor PSCF25260015@OOUAGOIWOYE.EDU.NG
The Artificial Intelligence System Group,
Department of Computer Sciences,
Olabisi Onabanjo University
Ogun, Nigeria

Editor: Sakinat Folorunso, Roseline Ogundokun, and Francisca Oladipo

Abstract

Artificial Intelligence (AI) is increasingly accepted as a change agent in the transformation of education systems, but the scale-up in resource-scarce contexts is hindered by a lack of educators' readiness and the absence of longitudinal FAIR-aligned benchmarks. This paper presents EduAI-NG, an open source, FAIR-aligned longitudinal data set from the EDAIL (Educators' AI Literacy) program in Nigeria, aimed at measuring and enhancing AI capacity building among educators in resource-scarce contexts. The data set consists of 2,239 pre-training and 1,068 post-training data, along with a matched data set of 770 educators. The study constructs a composite Teacher AI Deployment Readiness Index (TADRI-lite) and provide lightweight machine learning baselines to validate the dataset's predictive utility under efficiency constraints. Results show substantial improvements in educator preparedness following the intervention: mean AI understanding increased from 2.80 to 3.93 (Cohen's = 0.852), ($p < 0.001$), while the composite readiness index demonstrated a very large effect (d 1.29) with excellent internal consistency (Cronbach's = 0.93). The gradient boosting benchmarks further confirm that a set of ethical and pedagogical indicators can reliably predict high AI readiness. The findings suggests that context-aware professional development has the potential to significantly affect educator AI readiness in the Global South.

Keywords: FAIR data, AI education, benchmark dataset, low-resource environments, Nigeria

1. Introduction

Artificial intelligence is transforming educational systems from K–12 to higher education, AI-enabled tools are increasingly viewed as catalysts for making teaching more efficient and helping students learn better (Sanusi et al., 2022; UNESCO, 2024; Saal et al., 2025). The academic study of Artificial Intelligence (AI) in education has become one of the major arenas of academic research around the globe, and several studies(Zawacki-Richter et al., 2019; Mahmoud, 2020; Ifraheem et al., 2024; Mukagiahana et al., 2024; Salimi, 2025; Olurinola, 2025) outlines its effect on education and learning practices in diverse educational systems. Cross-country evidence shows strong awareness of AI's importance across Africa, but insufficient practical engagement and inequitable access to AI learning opportunities (Aryee et al., 2025) . Despite the growing interest in AI education research globally, the field still suffers from a shortage of large-scale, longitudinal, and FAIR-compliant educator datasets, particularly within low-resource contexts. This scarcity constrains reproducibility, cross-country comparison, and evidence-based policy design.

In Nigeria, however, the adoption of AI-based educational technologies is emerging, particularly among teachers who are central to its implementation (Ayanwale et al., 2024; Olugbade et al., 2025). This spread of AI use in educational systems worldwide has raised concerns about its potential impact on the Nigerian teaching profession (Zhang et al., 2024; Awofala et al., 2025; Olurinola, 2025). The question is whether Nigerian teachers are fully aware of and ready for the rapid growth and development of artificial intelligence and its impact on education. Recent literature has reported varied levels of awareness (Ayanwale et al., 2024; Adelana et al., 2024; Femi, 2025; Obidiebube et al., 2025; Adu and Olugbade, 2026).

To this end, the EDAIL (Educators' AI Literacy) program was initiated as a national teacher capacity-building program in Nigeria. The EDAIL program offers structured exposure to AI concepts, AI pedagogy, ethical considerations, and integration strategies that can be applied in resource-constrained school environments. With a documented need to strengthen AI literacy among its educators, a large yet diverse educational system and a rapidly expanding digital ecosystem, Nigeria provides a particularly important context for works of this nature.

This paper proposes a long-term benchmark dataset called EduAI-NG that is aligned with the FAIR criteria. It assesses the impact of AI capacity-building on teacher readiness to deploy AI in the classroom and examine changes in perceived AI competence and construct a composite Teacher AI Deployment Readiness Index (TADRI-lite), and establish a lightweight machine learning baselines suitable for low-resource environments. The findings of this study show a substantial improvement in educator readiness while also revealing persistent structural bottlenecks, particularly infrastructure and pedagogical integration that continue to limit scalable AI deployment in classrooms.

This paper makes the following contributions:

- Introduce EduAI-NG, one of the first FAIR-aligned longitudinal benchmark datasets capturing educator AI readiness in an African context.
- Offer baseline lightweight machine learning benchmarks that show the effectiveness of the proposed dataset in prediction scenarios, with constrained resource assumptions.
- Quantify the impact of AI training using a validated composite readiness measure (TADRI-lite), showing a large improvement in educator readiness (Cohen's ($d = 0.94$)).
- Identify key deployment bottlenecks affecting scalable AI adoption in Nigerian classrooms and provide evidence to guide future capacity-building initiatives.

2. Methodology

2.1. The EDAIL Programme

The EduAI-NG dataset was generated through the EDAIL (Educators' AI Literacy) initiative, which is a capacity-building program aimed at improving the competencies of educators in Artificial Intelligence (AI) literacy in the context of pedagogical applications. It was specifically aimed at teachers in resource-constrained environments and was based on four main competency domains:

(i) Foundational AI literacy.
(ii) Pedagogical integration of AI tools.
(iii) Ethical and responsible AI use, and
(iv) Practical classroom deployment readiness.

The training was conducted through a structured blended learning approach encompassing pedagogical applications, live demonstrations of AI technologies, and their application in teaching through virtual workshops. The participants were drawn from all six geopolitical zones of the country, with support from the National Senior Secondary Education Commission (NSSEC), providing a representative sample in terms of the subjects' areas and levels of the sampled schools. Data was collected through structured questionnaires, which were administered in an online format using Google Forms to facilitate easy distribution and response collection. The anonymized, cleaned and preprocessed dataset was subsequently subjected to further analysis to support the study's objectives.

The study is subject to certain limitations that should be acknowledged. The voluntary nature of participation may have introduced self-selection bias, as enrollees are likely more technologically inclined than the average educator, potentially limiting generalizability. Additionally, reliance on Google Forms for data collection presupposes stable internet access and digital devices, which may have excluded educators in more remote or under-resourced settings.

2.2. EduAI-NG Dataset

2.2.1. Data Collection Design

EduAI-NG employs a "pre" and "post" approach to assess educators' readiness in the use of AI. Surveys conducted before the training focused on aspects such as the educators' readiness, awareness, and experience, while the survey conducted after the training focused on all the previously measured aspects and newer ones like what was learned, confidence, and readiness.

The dataset consists of:

- 2,239 pre-training responses.
- 1,068 post-training responses.
- 770 matched educators linked via anonymized email identifiers.
- 51–63 structured variables spanning multiple competency dimensions.

This matched cohort enables robust within-participant analysis of the impact of training. The dataset is publicly available at https://doi.org/10.17632/pmjkz4nxp6 (Olurinola et al., 2026)

2.2.2. Feature Taxonomy

For analytical purpose, the survey variables were grouped into four analytical categories:-

1. AI Foundations: prior conceptual understanding, awareness and exposure to AI.
2. AI in Pedagogy: use of AI in lesson planning, classrooms and instructional integration
3. Ethics and Responsible AI: bias, fairness and data awareness.
4. Deployment Readiness Indicators: confidence, customization ability and applied competence.

Likert-scale responses were encoded on a five-point ordinal scale ranging from Very Poor (1) to Excellent (5) and a four-point ordinal scale ranging from Strongly Disagree(1) to Strongly Agree (4), other type of responses are nominal scale, Standardized state names and ratio values.

2.3. FAIR Data Preparation

To ensure the appropriate reuse of data and the replicability of results, the EduAI-NG datasets were created in accordance with the FAIR principles (Findable, Accessible, Interoperable, Reusable). The FAIR principles for the datasets are:

1. Findability: standardization of all variables and proper documentation of the datasets.
2. Accessibility: the data is provided in a structured and easily accessible table format.
3. Interoperability: the Likert scales and vocabulary are kept consistent
4. Reusability: all data are anonymized. Robust data-cleansing protocols are included to ensure consistency, alongside user guidelines for appropriate secondary analysis.

Personally identifiable information (PII) was removed prior to analysis. The resulting dataset is supported by a comprehensive data dictionary.

2.4. Preprocessing and Cohort Matching

Data preprocessing involved:

1. normalization of identifiers for cohort matching.
2. removal of duplicate submissions entries.
3. list-wise deletion of of missing responses for model training
4. ordinal encoding of Likert responses
5. alignment of matching pre–post educator records

The matched cohort (n = 770) served as the primary population for longitudinal impact analysis.

2.5. Composite Readiness Index (TADRI-lite)

The study constructed a Teacher AI Deployment Readiness Index (TADRI-lite) as a measure to quantify multidimensional preparedness for AI integration. The index aggregates the four validated competency indicators :-

1. understanding of AI concepts
2. ability to apply AI to teaching challenges
3. knowledge of ethical AI considerations
4. ability to incorporate AI tools into teaching practice

Let:

- i = educator index, i=1, ... N
- j = readiness component index, j=1, ... 4
- x_{ij} = Likert score of the educator i on the component j
- Likert scale: $x_{ij} \in \{1, 2, 3, 4, 5\}$

The composite score was computed as the mean of normalized Likert responses. Internal consistency was evaluated using Cronbach's alpha, and longitudinal impact was assessed using paired effect size analysis.
Cronbach's alpha α measures the internal consistency reliability of a multi-item scale. It quantifies how well the items jointly measure the same latent construct (AI readiness). For a scale with K items:

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{j=1}^{K} \sigma_j^2}{\sigma_T^2}\right) \tag{1}$$

Where:

- K = number of items (4)
- σ_i^2 = variance of item j
- σ_T^2 = variance of the total composite score
- α = Cronbach's alpha reliability coefficient

The value *ranges between the values of 0 and 1. When its value is < 0.60, it is poor, >0.80 good and ≥ 0.90 is excellent*

The Teacher AI Deployment Readiness Index (TADRI-lite) was calculated as the arithmetic mean of four Likert-type competency indicators, which include conceptual understanding, pedagogical application, ethical awareness, and classroom integration ability.

Formally, for educator *i,* the index is defined as:

$$TADRI_i = \frac{1}{K}\sum_{j=1}^{K} x_{ij} \tag{2}$$

where:

- K = 4 (number of components)
- x_{ij} are ordinal-encoded Likert responses

Higher values indicate greater readiness for AI-enabled instructional deployment.

2.6. Baseline Prediction Tasks

To demonstrate the utility of EduAI-NG for downstream machine learning research, The study defined a binary classification task: High AI readiness prediction, where educators with post-training scores above the ordinal midpoint were labeled as high readiness.
This task reflects a realistic deployment screening scenario in resource-constrained education systems.

2.7. Models and Evaluation Metrics

Lightweight Predictive Modeling To evaluate EduAI-NG performance under resource-constrained conditions, The study turned to a number of tree-based ensemble models. Particularly, Gradient boosting methods stood out as the best choice being naturally suited for handling structured tabular data, and generally requiring less computation than deep neural networks;

Models of it's class are particularly suitable for low-resource educational analytics because:

- They handle heterogeneous tabular features effectively
- Feature scaling requires a minimum amount of work
- The model produces high quality results at a relatively low cost.

Gradient boosting is a powerful supervised machine learning algorithm that builds a strong predictive model by sequentially combining multiple weak learners, typically decision trees, such that each new model focuses on correcting the errors of the previous ones. Gradient Boosting can be formally defined as follows:

$$F_M(x) = \sum_{m=1}^{M} \gamma_m h_m(x), \tag{3}$$

where:

- x denotes the input feature vector,

- $h_m(x)$ is the m-th weak learner (decision tree),
- m is the corresponding learning weight, and
- M is the number of boosting iterations.
- $FM(x)$ represents the final predictive model after M iterations (or M learners)

At each step of the fitting process, which refers to the training stage where a model learns from data by adjusting its parameters to minimize error, a new learner (decision tree model) is trained on the negative gradient of the loss function. This allows the prediction error to decrease iteratively while keeping the amount of computation required tractable. In this study, the model was trained using an ensemble of 200 sequential decision trees to progressively improve predictive performance.

Classification Metrics
The performance of the model was evaluated using Accuracy, F1-score and the Receiver Operating Characteristics Area under the Curve (ROC-AUC) metrics. These metrics help provide complementary perspectives on the behaviour of the classifier particularly under class imbalance

Accuracy
Accuracy measures the proportion of correctly classified samples:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where:

- TP = true positives
- TN = true negatives
- FP = false positives
- FN = false negatives

Accuracy is useful as a broad measure of correctness, but it can paint a misleading picture when classes are imbalanced. Hence, additional metrics were also reported.

Precision and Recall
Precision and Recall can be defined as:

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN} \tag{5}$$

Precision measures the purity of predictions, while recall captures the overall sensitivity to the positive class (high-readiness educators).

F1-Score

The F1-score provides the harmonic mean of precision and recall:

$$F1\text{-}Score = \frac{Precision * Recall}{Precision + Recall} \tag{6}$$

The use of F1-score is particularly appropriate when both false positives and false negatives carry practical implications, as in readiness screening scenarios.

ROC-AUC

The Receiver Operating Characteristic (ROC) curve plots the true positive rate (TPR) against the false positive rate (FPR) across decision thresholds:

$$TPR = \frac{TP}{TP + FN}, FPR = \frac{FP}{FP + TN} \tag{7}$$

The Area under the Curve (AUC) is computed as:

$$AUC = \int_0^1 TPR(FPR)d(FPR) \tag{8}$$

and summarizes the model's ranking ability independent of any fixed threshold. Higher ROC-AUC values indicate stronger separbility between high- and low-readiness educators.

Training Impact Effect Size

Beyond predictive performance, the impact of the EDAIL intervention was quantified using Cohen's (d) on the matched cohort. For paired samples, the effect size is computed as:

$$d = \frac{\underline{\Delta}}{s_{\triangle}} \tag{9}$$

This formulation captures the standardized magnitude of the intervention effect independent of sample size. Following conventional interpretation, values of d $\geq$ 0.8 indicate a large practical effect.

where:

- $\underline{\Delta}$ represent the mean pre–post difference,
- $s_{\triangle}$ represent the standard deviation of the differences.

This formulation captures the standardized magnitude of the intervention effect independent of sample size. Following conventional interpretation values of d $\geq$ 0.8 indicate a large practical effect.

Reproducibility

All experiments for this study were implemented in Python using consistent pipelines, fixed random seeds, and carefully documented preprocessing steps. Our aim is that this approach not only enables transparent benchmarking but also makes the EduAI-NG datasets straightforward for others to build on, particularly as the field continues to explore how scalable AI tools can be applied in educational settings.

2.8. Ethical Considerations

Ethically, the research adheres to the responsible conduct of AI research and human subject research. People can voluntarily participate in the EDAIL program. Informed consent was given before the survey. Data anonymization was done during preprocessing. The released data set does not contain any personal information. The EduAI-NG resource is for research and capacity building purposes only, similar to the ethical use of AI in education.

3. Results and Discussion

3.1. Training Impact

3.1.1. Cohort Characteristics

The EduAI-NG dataset demonstrates how ready Nigerian educators are to use AI. The study consists of 2,239 pre-training responses, 1,068 post-training responses, as well as a matched group of 770 educators. This is useful for comparing the same individuals before and after training, eliminating biases that could be obtained by only studying one point in time. The responses of these educators across the country provide strong evidence of national participation, as well as enough data to judge the effectiveness of training in resource-constrained environments.

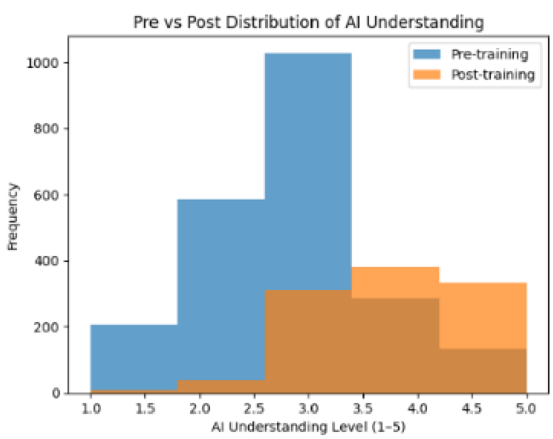

Figure 1: Demonstrates how much people's knowledge of AI changes after training, illustrating a clear rightward shift in educator competence.

3.1.2. Change in AI Understanding

Figure 2 presents the distributional shift in educators' self-reported understanding of artificial intelligence before and after the intervention. The mean AI understanding score increased from 2.80 (pre training) to 3.93 (post training) on the five-point Likert scale, representing an absolute gain of +1.13 points.

The histogram demonstrates that the distribution has shifted substantially to the right after the training, with a notable decrease in the number of responses in the lower bands (1-2) and a notable increase in the higher bands (4-5). This pattern suggests that the effects of the EDAIL intervention were broadly distributed across participants, rather than concentrated within a small, high-performing subgroup.

The results show that the training was successful in improving the level of understanding of AI for most teachers, not just the top performers. The results are supported by the large effect size, Cohen's d = 0.852, and the results of the paired tests, showing that the results are highly significant ($p < 0.001$).

3.1.3. Composite AI Deployment Readiness (TADRI-lite)

To examine the multifaceted and multidimensional aspects of preparedness for classroom Integration of AI, The study employed the Teacher AI Deployment Readiness Index (TADRI-lite). As illustrated in Figure 2, distinct improvement emerged between pre- and post-training, the composite score rose from 2.61 (pre) to 3.90 (post), reflecting an average increase of 1.29. Reliability analysis reinforced the validity of the readiness framework, with a Cronbach's α of 0.93, demonstrating strong internal consistency of the readiness construct.

Notably, the greater improvement observed in the composite index when compared to the single-item understanding measure suggests that the EDAIL initiative enhanced not just conceptual awareness but also applied, ethical, and pedagogical dimensions of AI readiness. This trend toward multidimensional improvement and comprehensive advancement holds particular significance for real-world classroom implementation, where operational competence frequently lags behind conceptual familiarity.

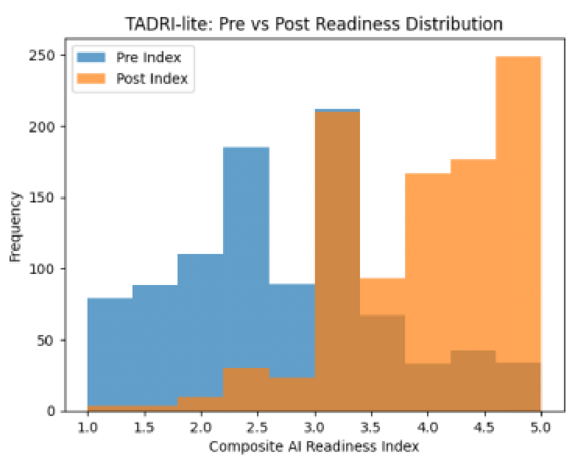

Figure 2: TADRI-lite composite readiness distribution comparing pre and post training educators, indicating substantial gains in multidimensional AI deployment readiness.

3.2. Predictive Benchmark Performance

To assess the analytical utility of EduAI-NG, the study implemented a lightweight machine learning benchmark for high AI readiness classification. Using three core readiness indicators: ability to apply AI concepts to solve teaching challenges, knowledge of ethical considerations in AI use, and methods of incorporating AI tools into teaching practice, the LightGBM model achieved strong performance on the held-out test set:

- **Accuracy:** 0.8897
- **F1-score:** 0.9184
- **ROC-AUC:** 0.9174

In figure 3, the ROC curve shows fairly consistent discriminative performance over different thresholds. Additionally, the balanced precision-recall profile implies the model works well in identifying educators across various readiness levels, picking up both higher- and lower-readiness subjects reliably.

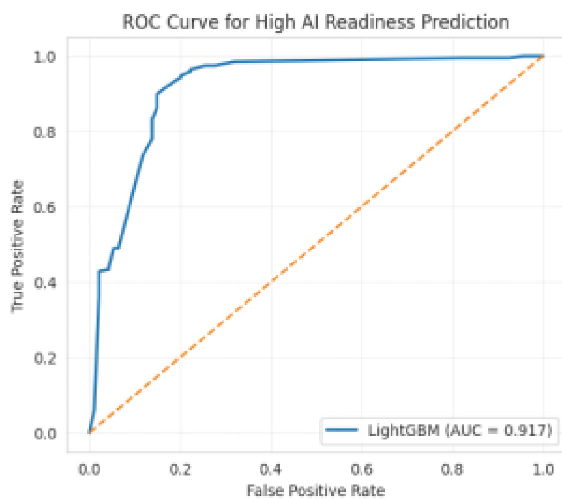

Figure 3: Receiver operating characteristic (ROC) curve for high AI readiness prediction using the lightweight LightGBM model.

Notably, this performance was achieved using a compact feature set and an efficiency-oriented model configuration. The results, therefore, confirm that EduAI-NG contains meaningful predictive signals suitable for lightweight deployment scenarios, consistent with the IndabaX emphasis on scalable AI under resource constraints.

3.3. Discussion

3.3.1. Alignment with Prior Work

These large gains in EduAI-NG are consistent with the growing literature that suggests that effective professional development can lead to positive gains in teachers' AI/digital skills.

For instance, a recent study by Sanusi et al. (2022) found that effective AI education plans can increase teachers' readiness. Regarding the African context, Ayanwale et al. (2024) found that more pre-service teachers in Nigeria were aware of AI. However, the issue of readiness and how to measure it remained a gap. This current study adds to the African context by demonstrating the long-term large-scale data that shows the positive impact on readiness. Additionally, the effect size found in the current study was relatively high at Cohen's d = 0.852. In the technology integration literature in education, the average effect size is considered to be moderate.

3.3.2. Implications for Scalable AI in Low-Resource Settings

The results have some implications for the use of AI in schools in the Global South. The substantial pre-post improvements suggest that, human capacity constraints commonly identified as a major impediment to AI adoption may be more manageable than thought when training is well-targeted and practical. Also, the high internal reliability of TADRI-lite lends support to the practicability of standardized readiness measurement tools for educators. Such tools would be essential for tracking national AI capacity development initiatives. Furthermore, the success of small, simple predictive models also suggests that readiness analytics might actually be useful in practice without the need for strong computer infrastructure. Such strong infrastructure is not available in low- and middle-income countries for learning analytics.

3.3.3. Contribution to the Field

Overall, EduAI-NG advances the AI-in-education literature in three important ways:

1. It provides one of the first FAIR-aligned longitudinal educator AI readiness datasets from an African context.

2. It demonstrates that AI capacity building can yield large, measurable improvements at scale.

3. It establishes that educator readiness is predictively tractable using lightweight models, supporting efficient deployment in resource-constrained environments.

These contributions directly support the theme of building scalable AI systems that function effectively beyond high-resource settings.

4. Conclusion and Future Work

EduAI-NG is a long-term FAIR-aligned dataset from the EDAIL project demonstrating the applicability of the FAIR datasets in the advancement of scalable and culturally grounded AI adoption in education by bridging the gap between capacity building and deployment analytics. This work not only supports policy-informed AI capacity planning but also facilitates more rigorous meta-analysis and cross-context analysis. This dataset is intended to serve as a foundation for future research into the deployment of responsible AI in education, the utility of machine learning in a low-tech environment, and teacher readiness Modelling.

The findings presented in this study are strong evidence that professionally informed context-specific training can improve teachers' readiness for AI integration, even within resource constrained educational environment. These findings support the emerging tenet that educators' AI readiness is not purely an Individualistic competency issue, but rather a multidimensional ecosystem-dependent phenomenon. Future work envisages multiple directions in this line of research. Such as, the introduction of context information, like school infrastructure, internet availability, etc., could present a better insight into models for analyzing deployment issues. Lastly, performing this research in different countries with similar FAIR approaches might help researchers compare readiness levels for AI implementation in the different educational systems across Africa.

Acknowledgments

The Google Academic Research Award (GARA) funded this work under the EDAIL initiative.

References

O. P. Adelana, M. A. Ayanwale, M. I. Adeoba, D. O. Oyeniran, N. Matsie, and D. Olugbade. Machine learning algorithm for predicting pre-service teachers' readiness to use brain-computer interfaces in inclusive classrooms. In *Proceedings of the 2024 IEEE 5th International Conference on Electro-Computing Technologies for Humanity (NIGERCON)*, pages 1–10. IEEE, 2024. doi: 10.1109/NIGERCON62786.2024.10927281.

I. E. Adu and D. Olugbade. Students data security metrics and the ethical implication for artificial intelligence deployment. In B. Edwards, H. Abuhassna, D. Olugbade, O. Ojo, and W. J. Wan Yahaya, editors, *AI in Education, Governance, and Leadership: Adoption, Impact, and Ethics*, pages 189–220. IGI Global Scientific Publishing, 2026. doi: 10.4018/979-8-3373-5550-4.ch008.

J. N. Aryee, P. Davies, G. A. Torsah, M. M. Apaw, C. D. Boateng, S. M. Mwando, and L. K. Amekudzi. Building capacity for artificial intelligence in africa: A cross-country survey of challenges and governance pathways, 2025. URL https://arxiv.org/abs/2512.05432.

A. Awofala, M. Bazza, O. Ojo, A. Oladipo, O. Olabiyi, and A. Arigbabu. Structural equation modeling of nigerian science, technology and mathematics teachers' adoption of educational artificial intelligence tools. *Digital Education Review*, 46:51–64, 2025. doi: 10.1344/der.2025.46.51-64.

M. A. Ayanwale, O. P. Adelana, R. R. Molefi, O. Adeeko, and A. M. Ishola. Examining artificial intelligence literacy among pre-service teachers for future classrooms. *Computers and Education Open*, 6:100179, 2024. doi: 10.1016/j.caeai.2024.100179.

I. Femi. Pre-service elementary teachers' attitudes and utilisation of artificial intelligence in academic activities. *Profesi Pendidikan Dasar*, 12(2):147–156, 2025. doi: 10.23917/ppd.v12i2.10233.

S. Ifraheem, M. Rasheed, and A. Siddiqui. Transforming education through artificial intelligence: Personalisation, engagement and predictive analytics. *Journal of Asian Development Studies*, 13(2):250–266, 2024. doi: 10.62345/jads.2024.13.2.22.

A. Mahmoud. Artificial intelligence applications: An introduction to education development in the light of covid-19 challenges. *International Journal of Research in Educational Sciences*, 3(4):171–224, 2020.

J. Mukagiahana, A. Sibomana, and J. Ndiritu. Teachers' understanding of gender responsive pedagogy and its application in teaching process. *Journal of Pedagogical Research*, 2024. doi: 10.33902/jpr.202423067.

J. Obidiebube, A. Ikwelle, B. Iwuagwu, J. Okwuba, and E. Ikana. Prospects and challenges of ai integration into nigerian educational systems. *International Journal of Latest Science and Engineering Research*, 2025. doi: 10.70382/caijlser.v7i8.017.

D. Olugbade, O. D. Olurinola, E. A. Tolorunleke, and A. F. Afolayan. SDG 4 in nigeria: Gender-related impacts of UNESCO MGIEP digital teacher training course. *Discover Education*, 4(1):547, 2025. doi: 10.1007/s44217-025-00961-1.

O. D. Olurinola. A systematic review of AI integration frameworks and the emergence of the EDAIL framework for teachers in LMICs. *Journal of Robotics and Automation Research*, 6(4):01–09, 2025. doi: 10.33140/JRAR.06.04.09.

O. D. Olurinola, S. O. Folorunso, and P. I. Owor. Eduai-ng: A fair longitudinal dataset of ai readiness among nigerian educators (edail), May 2026.

P. E. Saal, K. Chetty, N. Ntshayintshayi, T. Moosa, and N. Masuku. A scoping review of the integration of artificial intelligence in primary and secondary schools (2020–2024). *Journal of Education (University of KwaZulu-Natal)*, 98:62–85, 2025.

F. Salimi. Aligning policy and practice: The World Bank's approach to edtech in sub-saharan africa. *Policy Futures in Education*, 23(6):1134–1156, 2025. doi: 10.1177/14782103251324275.

I. T. Sanusi, S. A. Olaleye, S. S. Oyelere, and R. A. Dixon. Investigating learners' competencies for artificial intelligence education in an african K-12 setting. *Computers and Education Open*, 3:100083, 2022. doi: 10.1016/j.caeo.2022.100083.

UNESCO. AI competency framework for teachers. Technical report, UNESCO, 2024.

O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur. Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(1), 2019. doi: 10.1186/s41239-019-0171-0.

N. Zhang, F. Ke, C. Dai, S. A. Southerland, and X. Yuan. Leveraging AI-supported virtual simulation to support preservice teachers' responsive teaching. *British Journal of Educational Technology*, 56(3):1148–1169, 2024. doi: 10.1111/bjet.13522.