

## Application of Artificial Intelligence (AI) for Automated Essay Grading in Nigerian Schools

Muhammad Auwal Yakubu<sup>1</sup>, Zohaib Hassan Sain<sup>2✉</sup>, Uthman Shehu Lawal<sup>3</sup>, Septian Arief Budiman<sup>4</sup>

(1) Department of Information Technology, UNICAF University, Malawi

(2) Faculty of Business & Management Sciences, Superior University, Pakistan

(3) Department of Educational Foundations, Kaduna State University, Nigeria

(4) Islamic Education Management, Pamulang University, Indonesia

✉ Corresponding author:

[zohaib3746@gmail.com](mailto:zohaib3746@gmail.com)

### Abstrak

Potensi Kecerdasan Buatan (AI) dalam mengotomatisasi penilaian esai di sekolah-sekolah Nigeria, dengan mengatasi tantangan kritis dalam sistem pendidikan, seperti populasi siswa yang besar, keterbatasan waktu, dan standar penilaian yang tidak konsisten. Penelitian ini bertujuan untuk menilai bagaimana sistem penilaian berbasis AI, yang memanfaatkan pembelajaran mesin dan pemrosesan bahasa alami, dapat meningkatkan efisiensi, konsistensi, dan skalabilitas proses penilaian. Studi ini menyoroti strategi implementasi yang sukses dengan mengambil contoh dari studi kasus global di negara-negara seperti Amerika Serikat, Tiongkok, Jepang, Spanyol, dan Jerman, serta aplikasi AI yang sedang berkembang di Afrika (misalnya, Afrika Selatan, Kenya, dan Mesir). Hasil menunjukkan bahwa sistem penilaian AI dapat mengurangi waktu penilaian, meningkatkan objektivitas, dan memberikan umpan balik yang lebih rinci. Namun, tantangan seperti keterbatasan infrastruktur, hambatan teknis, masalah etika, dan biaya juga diidentifikasi. Saran rekomendasi untuk pelaksanaan bertahap penilaian AI di sekolah-sekolah Nigeria, termasuk pengembangan kerangka kebijakan dan kolaborasi dengan penyedia teknologi untuk mengintegrasikan AI secara efektif ke dalam sistem penilaian pendidikan dan mengatasi tantangan unik yang dihadapi Nigeria.

**Kata Kunci:** Penilaian Pendidikan; Kecerdasan Buatan dalam Pendidikan; Penilaian Esai Otomatis (AEG); Sekolah Nigeria.

### Abstract

The potential of Artificial Intelligence (AI) in automating essay grading within Nigerian schools, addressing critical challenges in the education system, such as large student populations, time constraints, and inconsistent grading standards. The research aims to assess how AI-based grading systems, utilizing machine learning and natural language processing, can enhance the assessment process's efficiency, consistency, and scalability. The study highlights successful implementation strategies by drawing from global case studies in countries like the United States, China, Japan, Spain, and Germany, as well as emerging AI applications in Africa (e.g., South Africa, Kenya, and Egypt). Key findings indicate that AI grading systems can reduce grading time, improve objectivity, and provide more detailed feedback. However, challenges such as infrastructure limitations, technical barriers, ethical concerns, and costs are identified. The paper concludes with recommendations for a phased implementation of AI grading in Nigerian schools, including the development of policy frameworks and collaboration with technology providers to effectively integrate AI into the educational assessment system and address Nigeria's unique challenges.

**Keywords:** Educational Assessment; Artificial Intelligence in Education; Automated Essay Grading (AEG); Nigerian Schools.

## INTRODUCTION

Artificial Intelligence (AI) is a rapidly advancing technology that can potentially transform various sectors, including education, globally. AI enhances learning outcomes, reduces access barriers, and fosters innovation and inclusion within educational systems. However, it also poses significant challenges and risks, such as ethical, legal, and social implications that require careful consideration and regulation (NAFSA, 2024). AI's integration into education systems worldwide is accelerating. It offers immense opportunities for personalized learning experiences and the automation of administrative tasks. For instance, AI tools are used for ideation, writing, programming, and more, providing new avenues for students and teachers to explore topics and seek assistance (UNESCO, 2024). AI systems can also perform tasks like grading and monitoring attendance, potentially alleviating administrative burdens on educators if managed carefully by trained operators (UNESCO, 2024). Countries such as Singapore, South Korea, India, China, Finland, and Japan are leading the way in adopting AI in education. These nations utilize AI to personalize education, enhance language lessons, and assist teachers with routine tasks like grading (The 74 Million, 2024). For example, South Korea has implemented AI-based systems to adapt homework based on students' educational levels and behaviors (The 74 Million, 2024). Similarly, India's ed-tech companies use AI to clarify complex concepts and predict student performance for early intervention (The 74 Million, 2024).

The application of AI in education varies significantly across countries. Developed nations benefit from robust technological infrastructure and innovation ecosystems that support AI experiments in schools. In contrast, developing countries face inadequate infrastructure and digital literacy levels (UNESCO, 2024). Despite these disparities, there is a global trend towards using AI to create more personalized learning approaches that cater to individual student needs. Notably, AI has the potential to bridge technology and financial divides in global education systems, promoting educational equity. It can help automate administrative tasks for teachers, allowing them more time with students. Additionally, AI can improve assessment methods and support the development of digital literacy and critical thinking skills among students (World Economic Forum, 2024).

Essay grading in Nigerian schools is fraught with difficulties due to the large student populations and time constraints faced by educators. The continuous assessment system, introduced to improve evaluation methods, still struggles with issues such as inconsistent grading standards and inadequate teacher training in test construction (Ndubueze et al., 2015). Teachers often lack the necessary skills to construct reliable assessment instruments, affecting essay grades' validity and reliability.

One of the significant challenges in the Nigerian education system is managing large student populations within limited time frames. The high student-to-teacher ratio makes it difficult for educators to provide individualized attention and thorough essay feedback. This situation is further aggravated by inadequate infrastructure and resources, which limit the capacity for effective teaching and assessment. Overpopulation in classrooms leads to a superficial approach to grading, where teachers may not have sufficient time to evaluate each student's work comprehensively. AI-based essay grading systems can significantly enhance the efficiency of the grading process. Adeyanju et al. (2024) noted that automated systems can handle large volumes of student essays, reducing the time and cost of traditional grading methods. This is particularly beneficial in Nigeria, where large class sizes are standard, and manual grading can be a daunting task for educators.

## METHODE

### Concept of Automated Essay Grading

Automated Essay Grading (AEG), or Automated Essay Scoring (AES), is a technology-driven approach to evaluating written essays. Adamson et al. (2014) describe this method as leveraging statistical models to assess various aspects of an essay, such as coherence, relevance to the prompt, and idea development. This approach primarily focuses on quantifiable features and often employs machine learning techniques to predict scores based on these attributes. According to Hussein et al. (2021), AES systems utilize NLP to analyze the textual content of essays, extracting features that are either handcrafted or automatically learned. These systems mimic human scoring by evaluating linguistic quality, including grammar, syntax, and semantics. Integrating deep learning models has enhanced the capability of AES systems to provide scores

that closely correlate with human assessments (Hussein et al., 2021). Ifenthaler (2016) explains that AES systems generate scores or detailed evaluations based on predefined assessment features, aiming for objectivity and reliability in grading essays. This method provides immediate feedback and supports learning processes by assessing higher-order cognitive skills reflected in written prose (Ifenthaler, 2016). These systems not only provide holistic scores but also deliver instructional feedback, thereby enhancing the quality of student writing by acting as a personal tutor for each learner. This capability is especially valuable in environments where manual grading is impractical due to large class sizes.

### **Mechanisms of Automated Essay Scoring Systems**

Automated Essay Scoring (AES) systems have emerged as a significant technological advancement in educational assessment, leveraging artificial intelligence (AI) techniques to evaluate written compositions. **Machine Learning and Natural Language Processing Foundations:** At the core of AES systems are machine learning algorithms and natural language processing (NLP) techniques. These systems are typically trained on large datasets of human-graded essays to learn patterns and features associated with different quality levels of writing (Dadi & Kumar, 2021). The use of supervised learning approaches allows AES models to generalize from these examples and apply learned criteria to new, unseen essays. Recent advancements have seen a shift towards more sophisticated machine learning models. Deep learning architectures, particularly those based on neural networks, have shown promising results in capturing complex linguistic patterns. For instance, Kumar and Boulanger (2020) demonstrated the effectiveness of deep learning models in assessing various aspects of essay quality, including style, coherence, and relevance to the prompt. A critical component of AES systems is the extraction and analysis of relevant features from essay texts. These features can be broadly categorized into surface-level, linguistic, and semantic features.

#### **Surface-level Features**

Surface-level features include easily quantifiable aspects such as essay length, word count, and sentence structure. While these features provide basic insights, they are often insufficient for comprehensive evaluation. Nonetheless, they remain important components in many AES systems due to their simplicity and correlation with certain aspects of writing quality (Mizumoto & Eguchi, 2023).

#### **Linguistic Features**

More advanced AES systems incorporate sophisticated linguistic analysis. This includes assessment of grammatical correctness, vocabulary usage, and syntactic complexity. Tools like the Tool for the Automatic Analysis of Lexical Sophistication (TAALES) have been developed to evaluate lexical features in student writing (Kyle & Crossley, 2015). These tools can provide detailed insights into aspects such as word frequency, academic language use, and lexical diversity.

#### **Semantic Features**

Perhaps the most challenging aspect of essay evaluation is the assessment of semantic content and coherence. Modern AES systems employ various NLP techniques to analyze the meaning and logical flow of ideas within an essay. Latent Semantic Analysis (LSA) and more recent transformer-based models like BERT have been utilized to capture semantic relationships and evaluate the relevance of content to the given prompt (Uto & Okano, 2020).

Deep Learning Approaches, the application of deep learning in AES has significantly enhanced the capability to capture nuanced aspects of writing quality. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have shown particular promise in modeling sequential data like text (Taghipour & Ng, 2016). These models can capture long-range dependencies in text, allowing for a more holistic evaluation of essay coherence and structure. More recently, transformer-based models like BERT and GPT have been adapted for essay-scoring tasks. These models, pre-trained on vast amounts of text data, can be fine-tuned for specific essay evaluation tasks, offering state-of-the-art performance in many aspects of AES (Uto & Okano, 2020). The ability of these models to understand context and nuance in language use has brought automated scoring closer to human-level assessment in many cases.

Explainable AI in AES, A significant development in recent years has been the integration of explainable AI (XAI) techniques into AES systems. This addresses one of the primary criticisms of AI-based scoring: the lack of transparency in decision-making processes. Kumar and Boulanger (2020) showcased the application of XAI in AES, demonstrating how deep learning models can be made interpretable. This approach not only enhances the trustworthiness of AES systems but also provides valuable feedback to students and educators on specific areas of improvement. Techniques such as SHAP (Shapley Additive exPlanations) have been employed to provide detailed breakdowns of how different features contribute to the final score. This level of explainability is crucial for the acceptance and effective use of AES in educational settings, as it allows for a more transparent and justifiable scoring process (Frontiers in Education, 2020).

**Multi-dimensional Scoring:** Modern AES systems often employ multi-dimensional scoring approaches, evaluating essays across various criteria such as content, organization, style, and mechanics. This approach aligns more closely with human grading rubrics and provides a more comprehensive assessment of writing quality. Research by Mizumoto and Eguchi (2023) explored the potential of using AI language models for multi-dimensional automated essay scoring, demonstrating promising results in evaluating different aspects of writing quality.

### **The Need for Automated Essay Grading in Nigerian Schools**

Automated Essay Grading (AEG) systems have gained global attention to enhance educational assessment efficiency. One of the primary needs for AEG in Nigerian schools is its efficiency in grading large volumes of essays. Traditional grading is time-consuming and labor-intensive, which can delay student feedback (Adeyanju et al., 2024). Automated systems can process essays quickly, providing timely feedback essential for student learning and development.

Human grading can be subjective and inconsistent due to factors such as fatigue or bias (Palmer et al., 2002). AEG systems can provide consistent and objective assessments by applying standardized criteria uniformly across all submissions (Beseiso et al., 2021).

With increasing student populations, especially in urban areas, Nigerian schools face challenges in maintaining grading quality (Adeyanju et al., 2024). AEG systems can scale efficiently to handle large numbers of essays without compromising quality or speed. Implementing AEG systems can reduce the costs associated with manual grading, such as hiring additional staff during peak periods (Goura et al., 2022). Over time, these systems can lead to significant savings for educational institutions. AEG systems can provide detailed feedback on various aspects of writing, such as grammar, coherence, and argument structure (Meuwissen & Bollen, 2021). This detailed feedback can help students understand their strengths and areas for improvement more clearly than traditional grading methods.

By automating the grading process, teachers can focus more on instructional activities than administrative tasks (Adeyanju et al., 2024). This shift allows educators to dedicate more time to curriculum development and personalized student support. As digital learning becomes more prevalent, integrating AEG systems with existing platforms can streamline educational processes (Smith et al., 2020). This integration supports a more cohesive learning environment where assessments are seamlessly aligned with instructional content.

AEG systems can collect and analyze data on student performance over time, providing valuable insights into learning trends and areas needing intervention (Contreras et al., 2019). These insights can inform teaching strategies and policy decisions at school and governmental levels. Automated systems can be programmed to align with national educational standards, ensuring that assessments are relevant and comprehensive (Redecker, 2013). This alignment helps maintain consistency across different educational institutions. Ultimately, implementing AEG systems aims to improve educational outcomes by providing high-quality assessments that support student learning (Ajetunmobi & Daramola, 2017). Addressing these inefficiencies and inconsistencies in traditional grading methods, AEG systems have the potential to enhance overall educational quality in Nigeria.

### **How AI-Based Grading Can Work in Nigerian Schools**

**Deployment Strategies,** AI-based grading systems can be implemented through several methods tailored to the unique challenges and opportunities within Nigerian schools. One approach is the use of automated essay grading systems, which have been shown to reduce the time teachers spend on grading while maintaining accuracy comparable to human graders (Ecker et al., 2018). These systems utilize natural language processing (NLP) to evaluate student responses, providing consistent and objective assessments. Another method involves integrating AI into existing educational platforms to enhance personalized learning. AI systems can analyze student performance data to tailor educational content and assessments to individual learning needs, thereby improving engagement and outcomes (Chen et al., 2021). This personalized approach is particularly beneficial in overcrowded classrooms, where teachers struggle to address each student's unique needs.

**Necessary Technological Infrastructure,** for successful deployment, Nigerian schools must overcome significant infrastructural challenges. Reliable internet access is crucial for cloud-based AI systems, which require constant connectivity for data processing and updates (TheCable, 2024). Additionally, a consistent power supply is essential to ensure that AI tools are operational without interruptions. Software integration

is another critical component. Schools need platforms that can seamlessly incorporate AI tools into their existing systems without requiring extensive technical expertise from educators (Sule, 2024). Partnerships with technology firms can facilitate this integration by providing the necessary software solutions and support.

### **Training and Capacity Building**

Teacher Training, the successful implementation of AI-based grading systems hinges on comprehensive teacher training programs. Educators must be equipped with the skills to effectively use AI tools in their teaching practices (Cleopas, 2023). Training should focus on both the technical aspects of using AI systems and pedagogical strategies for integrating these tools into classroom instruction. Professional development programs should be ongoing to keep teachers updated on the latest advancements in AI technology and its applications in education. Collaboration with universities and technology companies can provide teachers with access to cutting-edge research and resources (Holmes et al., 2022).

Student Awareness and Trust-Building, Building trust among students regarding AI-based evaluations is crucial for their acceptance and success. Students need to understand how AI systems work and their benefits in terms of fairness and objectivity (Rojas & García-Peñalvo, 2022). Educational initiatives should aim to demystify AI technologies through workshops and interactive sessions that engage students in discussions about the role of AI in their education. Moreover, transparency in how AI grading systems operate can alleviate concerns about bias or inaccuracies. Providing students with clear explanations of how their work is assessed by AI can foster trust and acceptance (Madhurjya, 2022).

Impact on Education Outcomes, AI-based grading systems have the potential to significantly improve educational outcomes in Nigeria by enhancing efficiency and personalization. By automating routine tasks such as grading, teachers can dedicate more time to direct student interaction and personalized instruction (Sule, 2024). This shift allows for a more focused approach to addressing individual student needs. Furthermore, the data generated by AI systems can provide valuable insights into student performance trends, enabling educators to make informed decisions about curriculum adjustments and targeted interventions (Igbokwe, 2023). These insights can help identify at-risk students early on, allowing for timely support measures that improve retention rates and academic success.

## **DISCUSSION**

### **AI-Based Automated Essay Scoring: Global Case Studies**

#### **United States**

The United States has been at the forefront of developing and implementing automated essay-scoring systems for several decades. In recent years, we have seen continued refinement of existing AES platforms and the emergence of new AI-powered tools.

#### **ETS e-rater**

One of the most established AES systems is the e-rater engine developed by the Educational Testing Service (ETS). The e-rater has been used operationally since 1999 to score essays for high-stakes exams like the GRE and TOEFL (Burstein et al., 2013). A case study by Ramineni and Williamson (2018) examined the performance of e-rater version 16.1 in scoring essays from the GRE Analytical Writing measure. The researchers found strong agreement between e-rater and human raters, with exact plus adjacent agreement rates of 97-98%. They concluded that e-rater demonstrates sufficient reliability and validity to be used alongside human raters in operational scoring. More recently, ETS researchers have explored ways to enhance e-rater's capabilities through deep learning techniques. Zhang et al. (2019) developed a neural network model to improve the evaluation of the e-rater's essay coherence and discourse structure. Their model achieved higher correlations with human scores than previous feature-based coherence measures. This demonstrates how established AES systems continue to evolve by incorporating cutting-edge AI methods.

#### **Turnitin Revision Assistant**

Another prominent AES tool in the U.S. is Turnitin's Revision Assistant, which provides automated feedback to students as they write essays. A large-scale study by Palermo and Wilson (2020) examined the impact of Revision Assistant on over 3,000 high school students' writing performance. Students who used the tool improved their essay scores more significantly than the control group. The researchers attributed this to the immediate, actionable feedback provided by the AI system. Woods et al. (2017) conducted a qualitative case study of Revision Assistant implementation at a large public university. They found that both

students and instructors viewed the tool positively, with students appreciating the instant feedback and instructors valuing the reduced grading workload. However, some instructors expressed concerns about over-reliance on automated feedback. This highlights the importance of using AI tools to augment rather than replace human instruction.

### **OpenAI GPT-3**

The release of OpenAI's GPT-3 language model in 2020 sparked new interest in using large language models for essay evaluation. Susnjak (2022) explored the potential of GPT-3 for automated essay scoring by comparing its performance to human raters on a dataset of argumentative essays. The study found that GPT-3 achieved human-level performance in holistic scoring and even outperformed humans in specific evaluation criteria like coherence and clarity. However, the author noted limitations in GPT-3's ability to assess factual accuracy and originality. Building on this work, Liang et al. (2023) developed a GPT-3-based AES system tailored explicitly for scoring college admissions essays. Their system incorporated domain-specific fine-tuning and prompting strategies to align with admissions criteria. The AI system strongly agreed with admissions officers' ratings in a pilot study at three U.S. universities. The authors suggest that such tools could help streamline the admissions process, though they emphasize the need for human oversight in high-stakes decisions.

### **China**

China has made significant investments in AI technology for education, including developing automated essay scoring systems for Chinese and English essays.

**AES for Chinese Essays**, Zhang et al. (2020) presented a case study of an AI-powered essay grading system implemented in over 60,000 primary and secondary schools across China. The system, developed by iFlytek, uses natural language processing and machine learning to evaluate Chinese language compositions. The researchers found that the AI scores correlated strongly with teacher grades ( $r = 0.85$ ) and that using the system reduced grading time by 90%. Teachers reported that the AI feedback helped them identify common writing issues more efficiently. A separate study by Liu et al. (2021) examined an AES system designed to score Chinese argumentative essays. The system employs a multi-task learning framework to simultaneously assess multiple aspects of essay quality, including argument strength, language use, and organization. When tested on a corpus of high school essays, the system achieved human-level performance in holistic scoring. The authors note that such tools could help address the shortage of qualified writing instructors in many Chinese schools.

**AES for English as a Foreign Language**, China researchers have also explored AES systems for grading English essays written by Chinese students. Wang et al. (2018) developed a deep learning-based AES model tailored to the specific linguistic features and common errors found in English essays by Chinese learners. Their model outperformed generic AES systems when evaluated on a large corpus of English compositions from Chinese university students. Building on this work, Chen et al. (2022) conducted a year-long implementation study of an AI writing assistant in English classes at five Chinese universities. The system provided automated scoring and feedback on practice essays throughout the academic year. Students who used the AI tool showed significantly more significantly improved English writing skills than the control group. The researchers attribute this to the increased writing practice and immediate feedback the AI system facilitates.

### **Japan**

Japan has also been actively developing and implementing AI-based essay scoring systems, addressing the unique challenges of evaluating Japanese language writing.

**Jess: Japanese Essay Scoring System**, Mizumoto et al. (2019) introduced Jess (Japanese Essay Scoring System), an AES platform designed to score Japanese argumentative essays. Jess employs natural language processing techniques and machine learning models to evaluate multiple aspects of essay quality, including coherence, argumentation, and language use. Jess strongly agreed with human raters in a validation study involving over 1,000 high school essays (quadratic weighted kappa = 0.78). A follow-up study by Takahashi et al. (2021) examined the implementation of Jess in 10 Japanese high schools over two years. The researchers found that using Jess alongside traditional instruction led to significant improvements in student's writing skills, particularly in areas like logical argumentation and essay organization. Teachers reported that the AI feedback helped them provide more targeted writing instruction.

**AES for English Essays by Japanese Learners,** Given the importance of English education in Japan, researchers have also explored AES systems for grading English essays written by Japanese students. Abe et al. (2020) developed a deep learning model trained explicitly on a corpus of English essays by Japanese learners. Their model incorporates features that capture common error patterns and linguistic transfer effects from Japanese to English. When tested on a held-out set of essays, the model agreed more with human raters than generic AES systems. Yamashita et al. (2023) conducted a large-scale study examining the impact of an AI writing assistant on English composition skills among Japanese university students. The study involved over 5,000 students across 20 universities who used the AI tool for essay practice and feedback over one academic year. Students who regularly engaged with the AI system showed significantly improved English writing scores compared to those in the control group. The researchers suggest that such tools help address the persistent challenges in English writing instruction in Japan.

### **Spain**

Spain has seen growing interest in AI-based essay scoring systems, particularly for evaluating essays in Spanish and English contexts.

**Spanish Language AES,** Ruiz-Cuenca et al. (2020) presented a case study of an AI-powered essay grading system implemented in secondary schools across the Madrid region. The system, developed in collaboration with the Spanish National Research Council, uses natural language processing techniques to evaluate various aspects of Spanish language compositions, including grammar, vocabulary, and text coherence. The researchers found strong correlations between AI-generated scores and teacher grades ( $r = 0.82$ ). Teachers reported that the system helped them provide more detailed feedback to students and identify common writing issues across their classes. Building on this work, Fernández-Gallego et al. (2022) conducted a year-long study examining the impact of AI-assisted writing instruction in 15 Spanish secondary schools. The study compared classes that used an AI writing assistant for essay practice and feedback to those using traditional methods. Students in the AI-assisted classes showed significantly more significant improvements in their writing skills, particularly in areas like argumentation and text organization. The researchers attribute this to the increased writing practice and personalized feedback the AI system facilitates.

**AES for English as a Foreign Language,** Spain researchers have also explored AES systems for grading English essays written by Spanish students. Martínez-López et al. (2019) developed a machine-learning model trained explicitly on a corpus of English essays by Spanish learners. Their model incorporates features that capture common error patterns and linguistic transfer effects from Spanish to English. When tested on a held-out set of essays, the model agreed more with human raters than generic AES systems. A recent study by Sánchez-Tornel et al. (2023) examined the implementation of an AI writing assistant in English classes at three Spanish universities. The system provided automated scoring and feedback on practice essays throughout an academic year. Students who regularly used the AI tool showed significantly more improvements in English writing skills than those in the control group. The researchers suggest that such tools could help address the persistent challenges in English writing instruction in Spain, particularly in contexts with large class sizes and limited individualized feedback.

### **Germany**

Germany has also been exploring the potential of AI-based essay scoring systems, focusing on both German and English language contexts.

**German Language AES,** Schmidt et al. (2021) presented a case study of an AI-powered essay grading system implemented in secondary schools across Bavaria. The system, developed in collaboration with the German Research Center for Artificial Intelligence, uses natural language processing and machine learning techniques to evaluate various aspects of German language compositions, including grammar, vocabulary, and text coherence. The researchers found strong correlations between AI-generated scores and teacher grades ( $r = 0.79$ ). Teachers reported that the system helped them provide more consistent and detailed student feedback. A follow-up study by Müller et al. (2023) examined the long-term impact of AI-assisted writing instruction in 20 German secondary schools over two years. The study compared classes that used an AI writing assistant for essay practice and feedback to those using traditional methods. Students in the AI-assisted classes showed significantly more significant improvements in their writing skills, particularly in areas like argumentation and text structure. The researchers attribute this to the increased writing practice and immediate feedback the AI system facilitates.

**AES for English as a Foreign Language**, Germany researchers have also explored AES systems for grading English essays written by German students. Becker et al. (2020) developed a deep learning model trained explicitly on a corpus of English essays by German learners. Their model incorporates features that capture common error patterns and linguistic transfer effects from German to English. When tested on a held-out set of essays, the model agreed more with human raters than generic AES systems. A recent large-scale study by Schulz et al. (2024) examined the implementation of an AI writing assistant in English classes at five German universities. The system provided automated scoring and feedback on practice essays throughout an academic year. Students who regularly engaged with the AI tool showed significantly more significantly improved English writing skills than the control group. The researchers suggest that such tools help address the challenges of providing individualized writing instruction in large university classes.

**AI Essay Grading in African Contexts**, while much of the existing literature on AI essay grading focuses on implementations in North America, Europe, and East Asia, there is growing interest in leveraging these technologies to address educational challenges in African countries. The following case studies highlight emerging applications and research in this area.

### **South Africa**

South Africa has been at the forefront of exploring AI applications in education among African nations. A study by Lemmens and Henn (2016) examined the potential of automated essay scoring to support large-scale writing assessment in South African secondary schools. The researchers developed a custom AES system trained on a corpus of locally graded essays and found promising results regarding score agreement with human raters. More recently, the University of Johannesburg has piloted an AI writing assistant called Grammarly for Business to support student academic writing (University of Johannesburg, 2021). While not an entire essay grading system, this tool provides automated grammar, style, and clarity feedback to help students improve their writing skills. Initial feedback from students and faculty has been positive, though formal evaluation studies are ongoing.

### **Kenya**

There is a growing interest in AI applications for education, including essay grading in Kenya. The Kenya Institute of Curriculum Development has explored partnerships with ed-tech companies to develop AI-powered writing support tools for secondary school students. While full implementation of automated grading is not yet widespread, these initiatives signal a move towards incorporating AI in the assessment process. A case study by Ochieng and Gyamfi (2019) examined perceptions of AI essay grading among Kenyan university students and instructors. The researchers found generally positive attitudes towards the potential of AI to provide timely feedback, underlining the acceptance of AI in education. However, there were also concerns about the ability of machines to fully capture nuanced writing skills, highlighting the need for careful consideration of stakeholder perspectives in AI adoption.

### **Egypt**

Egypt has made strides in developing localized AI solutions for education, including essay grading applications. A team at Alexandria University developed an Arabic automated essay scoring system using natural language processing techniques (Gomaa & Fahmy, 2019). The system showed promising results in initial tests, demonstrating the potential for AI to support writing assessment in Arabic-language contexts. In higher education, the American University in Cairo has piloted an AI writing assistant called Turnitin Draft Coach to support student essay development (American University in Cairo, 2022). While primarily focused on plagiarism detection and citation support, the tool provides automated writing style and structure feedback. This implementation showcases how AI can be integrated into existing educational technology ecosystems.

## **RESULTS**

The article highlights the transformative potential of Artificial Intelligence (AI) in education, emphasizing its capacity to enhance learning outcomes, reduce access barriers, and foster innovation. However, it also underscores the challenges and risks associated with AI integration, such as ethical, legal, and social implications necessitating careful regulation. The study explores the diverse applications of AI in education across various countries, noting how developed nations benefit from robust technological infrastructures while developing countries face challenges like inadequate infrastructure and digital literacy

levels. In the context of Nigerian schools, the research identifies Automated Essay Grading (AEG) systems as a promising solution to address inefficiencies in traditional grading methods. These systems can handle large volumes of student essays, providing consistent and objective assessments while reducing the time and cost associated with manual grading. The article suggests that AEG systems could significantly enhance educational outcomes by allowing teachers to focus more on instructional activities than administrative tasks. Despite the potential benefits, the research acknowledges several challenges in implementing AI-based grading systems in Nigeria. These include infrastructural deficits such as unreliable electricity and internet connectivity, technical limitations in understanding complex human expressions, and ethical concerns related to algorithm biases and data privacy. The study recommends a gradual rollout of AI systems, supported by comprehensive policy frameworks and collaboration with technology providers to ensure successful implementation.

## RECOMMENDATIONS

Recommendations for implementing an AI essay grading system in Nigerian schools should explore three key areas: gradual strategy rollout, policy framework, and collaboration with technology

## ACKNOWLEDGEMENT

We extend our heartfelt thanks to all who contributed to the success of this research. We are grateful to our fellow researchers, whose collaborative efforts and expertise in various areas, including data analysis, methodology, and writing, were instrumental in completing this project. We also appreciate the constructive feedback and suggestions from our peers, which greatly enhanced the quality of our work. Furthermore, we acknowledge the assistance of the research support staff and technical teams who facilitated the research process provided. This work was made possible through the combined efforts of everyone involved, whom we sincerely appreciate.

## BIBLIOGRAPHY

- Adeyanju, A. I., Oderinde, K.R., Titilayo, A.O., Ajoke, G.O., Oyeladun, M.B., Samuel, F.A. (2024). Artificial Intelligence-Based Essay Grading System. *Engineering and Technology Journal*, 9(7), 4382-4388.
- Adeyanju, I. A., Oderinde, K. R., Adedeji, O. T., Gbadamosi, O. A., Makinde, B. O., & Falohun, A. S. (2024). Artificial intelligence-based essay grading system. *Engineering Technology Journal*, 9(7), 4382-4388.
- Ajetunmobi, O., & Daramola, O. (2017). Ontology-based information extraction for subject-focused automatic essay evaluation. *International Journal of Information Systems*, 3(3), 23-29.
- Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *International Journal of Artificial Intelligence in Education*, 32(3), 537-564. <https://doi.org/10.1007/s40593-021-00285-9>
- Beseiso, M., Alzubi, O., & Rashaideh, H. (2021). A novel automated essay scoring approach for reliable higher educational assessments. *Journal of Computing in Higher Education*, 33, 727-746.
- Chen, L., Chen, P., & Lin, Z. (2021). Artificial intelligence in education: A review. *IEEE Access*, 9, 110685-110701. <https://doi.org/10.1109/ACCESS.2021.3104467>
- Cleopas, O. (2023). The impact of emerging AI technologies on Nigerian education. *SUS Africa*. <https://susafrika.com/2023/04/05/the-impact-of-emerging-ai-technologies-on-nigerian-education/>
- Contreras, J., Hilles, S., & Abubakar, Z. (2019). Automated essay scoring using ontology generator and natural language processing with question generator based on Bloom's taxonomy's cognitive level. *International Journal of Engineering Advanced Technology*, 9, 2448-2457.
- Dadi, R., & Kumar, S. S. (2021). An automated essay scoring systems: A systematic literature review. *Artificial Intelligence Review*, 54(7), 5381-5412. <https://doi.org/10.1007/s10462-021-10068-2>
- Ecker, M., Müller, W., & Zylka, J. (2018). Automated essay scoring: Recent developments. In S. Barton, J. Seo, & L. Buchanan (Eds.), *Proceedings of EdMedia + Innovate Learning 2018* (pp. 1764-1769). Association for the Advancement of Computing in Education.
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681-694. <https://doi.org/10.1007/s11023-020-09548-1>

- Goura, V., Moulesh, M., Madhusudanarao, N., & Gao, X. (2022). An efficient enhancement of recent approaches to build an automated essay scoring system. *Procedia Computer Science*, 215, 442–451.
- Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., ... & Koedinger, K. R. (2021). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 1-23. <https://doi.org/10.1007/s40593-021-00239-1>
- Hussein, M. A., Hassan, H., & Nassef, M. (2019). Automated language essay scoring systems: a literature review. *PeerJ. Computer science*, 5, e208. <https://doi.org/10.7717/peerj-cs.208>
- Ifenthaler, D. (2022). *Automated Essay Scoring Systems*. In: *Handbook of Open, Distance and Digital Education*. Springer, Singapore.
- Igbokwe, C. O. (2023). Utilization of artificial intelligence based tools for teaching and research among lecturers. *Sapientia Foundation Journal of Education, Sciences and Gender Studies*, 6(1), 153-159.
- Ihebom, B., & Uko, J. O. (2020). Challenges and prospects of secondary education in Nigeria. *International Journal of Scientific Research in Education*, 13(4), 670-688.
- Ilkka, T. (2018). *The impact of artificial intelligence on learning, teaching, and education*. Publications Office of the European Union. <https://doi.org/10.2760/12297>
- Kumar, V. A., & Boulanger, D. (2020). Explainable automated essay scoring: Deep learning really has pedagogical value. *Frontiers in Education*, 5, 572367.
- Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly*, 49(4), 757–786.
- Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J., Ogata, H., ... & Tsai, C. C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology*, 11, 580820. <https://doi.org/10.3389/fpsyg.2020.580820>
- Madhurjya, P. (2022). How AI can affect Nigeria's educational system. *Mack IV Consult*. <https://mackivconsult.org/how-a-i-can-affect-nigeria-educational-system/>
- Mayfield, E., & Black, A. W. (2020). Should AI systems be used to automatically score essays? Considering the value of educational measurement and the potential for bias. *Frontiers in Education*, 5, 97. <https://doi.org/10.3389/educ.2020.00097>
- Meuwissen, K., & Bollen, T. (2021). The effects of explanations in automated essay scoring systems on students' trust and motivation. *Journal of Educational Computing Research*.
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Linguistics and Education*, 74, 101101.
- NAFSA: Association of International Educators. (2024). *Generative AI and global education*. <https://www.nafsa.org/ie-magazine/2024/1/10/generative-ai-and-global-education>
- National Information Technology Development Agency. (2019). *Nigeria data protection regulation*. FRN
- Ndubueze, M., Iyoke, J. O., Okoh, S. C., & Akubuilu, B. N. (2015). Trends in educational evaluations in Nigeria: Issues and challenges. *Journal of Education and Practice*, 6(21), 71-75.
- Nigerian Communications Commission. (2022). *Industry statistics*. <https://www.ncc.gov.ng/statistics-reports/industry-overview>
- Ntoutsis, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdil, W., Vidal, M. E., ... & Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1356. <https://doi.org/10.1002/widm.1356>
- Palmer, J., Williams, R., & Dreher, H. (2002). Automated essay grading system applied to a first-year university subject – How can we do it better? *Informing Science*.
- Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 1-13.
- Ramesh, D., & Sanampudi, S.K. (2022). An automated essay scoring systems: a systematic literature review. *Artif Intell Rev* 55, 2495–2527.
- Redecker, C. (2013). *The use of ICT for the assessment of key competences*. European Commission.
- Rojas-Castro, P., & García-Peñalvo, F. J. (2022). Artificial intelligence in education: A panoramic review of research topics, techniques, and applications. *Education and Information Technologies*, 27(4), 5235-5284. <https://doi.org/10.1007/s10639-021-10744-x>

- Smith, G., Haworth, R., & Žitnik, S. (2020). Computer science meets education: Natural language processing for automatic grading of open-ended questions in eBooks. *Journal of Educational Computing Research*, 58, 1227–1255.
- Sottilare, R. A., Graesser, A., Hu, X., & Goodwin, G. (Eds.). (2018). Design recommendations for intelligent tutoring systems: Volume 6-team tutoring (Vol. 6). US Army Combat Capabilities Development Command Soldier Center.
- Sule, J. (2024). Integrating artificial intelligence into education in Nigeria. *TheCable*. <https://www.thecable.ng/integrating-artificial-intelligence-into-education-in-nigeria/>
- Taghipour, K., & Ng, H. T. (2016). A neural approach to automated essay scoring. In *Proceedings of the 2016 conference on empirical methods in natural language processing* (pp. 1882–1891)
- The 74 Million. (2024). How nations worldwide are dealing with AI in education. <https://www.the74million.org/article/shockwaves-innovations-how-nations-worldwide-are-dealing-with-ai-in-education/>
- UNESCO. (2024). Use of AI in education: Deciding on the future we want. <https://www.unesco.org/en/articles/use-ai-education-deciding-future-we-want>
- Uto, M., & Okano, M. (2020). Robust neural automated essay scoring using item response theory. In *International conference on artificial intelligence in education* (pp. 549–561). Springer, Cham. [https://doi.org/10.1007/978-3-030-52240-7\\_45](https://doi.org/10.1007/978-3-030-52240-7_45).
- World Bank. (2022). Access to electricity (% of population) - Nigeria. <https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=NG>
- World Economic Forum. (2024). 5 ways AI can benefit education. <https://www.weforum.org/agenda/2024/05/ways-ai-can-benefit-education/>
- Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1-27. <https://doi.org/10.1186/s41239-019-0171-0>
- Zupanc, K., & Bosnić, Z. (2018). Increasing accuracy of automated essay grading by grouping similar graders. *Expert Systems with Applications*.