LatIA. 2025; 3:328
doi: 10.62486/latia2025328
ISSN: 3046-403X

ORIGINAL

Ethical and Privacy Considerations in AI-Driven Language Learning

Consideraciones éticas y de privacidad en el aprendizaje de idiomas impulsado por IA

Muthu Selvam[1], Rubén González Vallejo[2]

[1]College of Computing and Informatics, University of North Carolina at Charlottel. Charlotte, North Carolina 28223, USA.
[2]University of Malaga, Department of Spanish, Italian, Romance Philology, Theory of Literature and Comparative Literature. Malaga, Spain.

Cite as: Selvam M, González Vallejo R. Ethical and Privacy Considerations in AI-Driven Language Learning. LatIA. 2025; 3:328. https://doi.org/10.62486/latia2025328

Submitted: 26-05-2024 **Revised:** 16-09-2024 **Accepted:** 31-03-2025 **Published:** 01-04-2025

Editor: PhD. Rubén González Vallejo

Corresponding Author: Muthu Selvam

ABSTRACT

Artificial intelligence (AI) has revolutionized language learning by enabling personalized and adaptive education; however, these advancements also raise ethical and privacy concerns, including algorithmic bias, data security risks, and a lack of transparency in AI-driven decision-making. This study examines these challenges, focusing on fairness, linguistic diversity, and the balance between automated and human instruction, with the goal of proposing ethical guidelines for the responsible adoption of AI in language education. Through a literature review and comparative analysis, ethical and privacy risks in AI-powered language learning tools were explored, assessing bias detection algorithms, transparency frameworks, and privacy-preserving techniques to identify best practices. The findings indicate that AI-driven language tools tend to exhibit biases that disadvantage underrepresented linguistic groups, raising concerns about fairness while also exposing privacy risks due to inadequate security measures. Implementing ethical AI frameworks that incorporate fairness-aware algorithms, explainable AI models, and robust data protection mechanisms enhances user trust and security. Therefore, addressing these issues is essential for ensuring the ethical integration of AI in language education, where a hybrid approach combining AI with human instruction emerges as the most responsible solution. Lastly, future research should focus on regulatory compliance and adaptive learning models to strengthen AI ethics in education.

Keywords: Algorithmic Bias; Data Privacy; AI Transparency; Ethical AI; Student Consent.

RESUMEN

La inteligencia artificial (IA) ha revolucionado el aprendizaje de idiomas al permitir una educación personalizada y adaptativa. Sin embargo, estos avances también plantean desafíos éticos y de privacidad, como el sesgo algorítmico, los riesgos en la seguridad de los datos y la falta de transparencia en la toma de decisiones automatizada. Este estudio analiza dichas problemáticas, centrándose en la equidad, la diversidad lingüística y el equilibrio entre la enseñanza automatizada y humana, con el objetivo de proponer directrices éticas para una adopción responsable de la IA en la educación de idiomas. A través de una revisión de literatura y un análisis comparativo, se examinaron los riesgos éticos y de privacidad en herramientas impulsadas por IA, evaluando algoritmos de detección de sesgo, marcos de transparencia y técnicas de preservación de datos para identificar mejores prácticas. Los resultados revelan que las herramientas de IA aplicadas al aprendizaje de idiomas tienden a presentar sesgos que afectan negativamente a grupos lingüísticos subrepresentados, generando preocupaciones sobre equidad, además de exponer riesgos de privacidad debido a medidas de seguridad insuficientes. La implementación de marcos éticos basados en

algoritmos conscientes de la equidad, modelos de IA explicables y mecanismos robustos de protección de datos contribuye a generar confianza y aumentar la seguridad del usuario. Por ello, abordar estas cuestiones es crucial para garantizar una integración ética de la IA en la educación de idiomas, donde un enfoque híbrido que combine tecnología y enseñanza humana se perfila como la mejor alternativa. Finalmente, se señalan como futuras investigaciones la adaptación de modelos de aprendizaje y el cumplimiento normativo para fortalecer la ética de la IA en el ámbito educativo.

Palabras clave: Sesgo Algorítmico; Privacidad de los Datos; Transparencia de la IA; Ética de la IA; Consentimiento del Estudiante.

INTRODUCTION

Artificial Intelligence (AI) is reshaping educational landscapes globally, particularly in language learning, where it enables personalized and adaptive instruction.[1] AI-powered platforms, such as virtual tutors, intelligent chatbots, and speech recognition systems, are increasingly used to enhance fluency, vocabulary acquisition, and grammar accuracy.[2] These systems utilize Natural Language Processing (NLP) to simulate interactive dialogues, allowing learners to practice speaking and listening in real-time contexts. However, despite the evident pedagogical benefits, the rapid adoption of AI in education necessitates a rigorous examination of ethical and privacy-related concerns.[3]

AI in language learning frequently involves extensive data collection—including voice recordings, written responses, and interaction logs—to optimize learning outcomes. While this personalization improves student engagement, it also introduces significant risks related to data privacy, algorithmic transparency, and consent. In many cases, users are unaware of how their data is used or stored, raising concerns about informed consent and ethical data handling. This issue is particularly pressing for younger learners, as they may lack the capacity to fully comprehend the implications of AI surveillance and data utilization.[4]

Furthermore, algorithmic bias remains a pressing concern. Language models trained on dominant linguistic patterns may marginalize regional dialects or non-native accents, exacerbating existing educational inequalities.[5] Such bias may impair the learning experience of users from underrepresented linguistic communities, reinforcing cultural and linguistic stereotypes. Studies have shown that AI-driven tools, while effective, often favor standard language forms, limiting inclusivity and learner diversity.[6]

Another major issue is the lack of transparency in AI decision-making. Many systems operate as opaque "black boxes," offering little insight into how content is selected, graded, or modified. This opacity reduces trust in AI-generated feedback and complicates error correction or bias detection.[7] Scholars have highlighted the necessity of explainable AI mechanisms to foster ethical oversight and ensure that human educators can validate system outputs.[8]

Although AI significantly enhances instructional capacity, it must not replace the invaluable role of human instruction. Educators offer contextual and cultural depth that AI systems cannot duplicate. Studies have emphasized the importance of maintaining a balanced, hybrid model that combines both AI and teacher interaction to ensure meaningful language acquisition and critical thinking development.[9]

Globally, the AI in education market is expanding at a projected CAGR of 36 % between 2022 and 2027, with North America and Asia-Pacific leading adoption. In the U.S., over 60 % of K–12 institutions have begun integrating AI into language instruction, while developing countries are quickly catching up. These trends underscore the urgency of establishing robust ethical frameworks to guide AI implementation in educational systems.

Literature review: ethical issues in ai-driven language learning

Bias and Fairness in AI Algorithms

AI-powered language learning systems rely on large datasets to train models, but these datasets often contain biases that can lead to unfair learning outcomes. Algorithmic bias arises when training data predominantly represents certain linguistic patterns, accents, or cultural contexts, leading to discrimination against underrepresented groups.[10] For example, AI speech recognition tools may struggle to process non-native accents, disadvantaging learners from diverse backgrounds accurately. Similarly, automated language assessment tools may favor certain grammatical structures or writing styles, reinforcing existing disparities. Ensuring fairness requires the inclusion of diverse datasets, continuous bias audits, and fairness-aware algorithms that adjust learning recommendations based on varied linguistic profiles.

Transparency and Explainability in AI Decision-Making

The lack of transparency in AI-driven language learning tools raises concerns regarding the interpretability

of recommendations and assessments. Many AI models function as black-box systems, making it difficult for educators and learners to understand how decisions are made. This opacity reduces trust in AI-based feedback and prevents meaningful human intervention in cases where AI-generated assessments appear incorrect.[11] Explainable AI models are essential for ethical implementation, as they allow users to trace decision-making processes and identify potential biases. Incorporating model explainability features, such as confidence scores and reasoning explanations, can improve trust while enabling educators to verify AI-generated outputs.

Autonomy and Human-AI Collaboration

AI-driven language learning systems operate with varying degrees of autonomy, but excessive reliance on AI-generated feedback can diminish the role of human educators. While AI can automate repetitive tasks such as grammar corrections and pronunciation analysis, it cannot provide nuanced feedback that considers cultural context, creativity, and emotional intelligence. Over-reliance on autonomous AI tutors may lead to reduced human interaction in language learning, potentially affecting students' motivation and engagement. Ethical AI design should prioritize a hybrid approach where AI enhances, rather than replaces, human instruction. Educators should always review AI-driven recommendations to ensure accuracy and contextual appropriateness.[12]

Accountability in AI-Powered Education

The increasing use of AI in language learning raises accountability concerns regarding errors, misinterpretations, and harmful recommendations. Determining responsibility becomes challenging if an AI-driven tool provides incorrect language feedback[13] or unfairly penalizes a student's writing style. Educational institutions, developers, and policymakers must establish clear accountability frameworks that outline the responsible parties for AI-related errors. Implementing human oversight mechanisms, such as educator review panels for AI-generated assessments, ensures that AI systems remain accountable and do not operate without regulatory safeguards. Ethical AI adoption requires that all stakeholders—developers, educators, and policymakers—take responsibility for AI-driven language learning platforms' accuracy, fairness, and reliability.

Addressing these ethical issues ensures that AI-driven language learning systems remain fair, transparent, and accountable. Bias mitigation, explainable AI, human-AI collaboration, and clear accountability structures are fundamental to fostering ethical AI adoption in education.

Privacy Issues in AI-Driven Language Learning

Data Security and Risk of Unauthorized Access

AI-driven language learning platforms collect and store vast amounts of user data, including voice recordings, writing samples, and learning patterns. This data is essential for providing personalized learning experiences but presents significant security risks. Cyberattacks, data breaches, and unauthorized access can expose sensitive student information, leading to identity theft or misuse of personal records. Many platforms rely on cloud storage, which, if not adequately protected, can become a target for malicious activities. Strong encryption techniques, multi-factor authentication, and secure access controls are necessary to prevent unauthorized access and ensure data integrity. Regular security audits and updates are essential to address vulnerabilities that may compromise student data.

Informed Consent and Ethical Data Usage

Many AI-powered learning tools collect user data without students' explicit awareness or meaningful consent. The complexity of AI-driven data collection methods often makes it difficult for users to understand how their information is stored, processed, and shared. In educational settings, students may not be able to refuse data collection, particularly when AI tools are integrated into mandatory coursework.[14] Ethical AI implementation requires clear and accessible consent mechanisms that inform users about what data is being collected and for what purposes. Providing students with options to opt out, modify, or delete their data fosters greater transparency and ensures ethical compliance with data protection principles.

Student Privacy and Potential for Profiling

AI systems analyze student performance to generate personalized learning recommendations, but extensive data collection also raises concerns about profiling and surveillanc.[15] Some platforms track behavioral data, including response times, error patterns, and engagement levels, which may be used to make inferences about a student's abilities and progress. If misused, this data can lead to unfair labeling, restricting educational opportunities based on algorithmic predictions. Additionally, third-party data sharing without student consent may expose learners to targeted advertisements or commercial exploitation. Protecting student privacy requires strict access controls, anonymization techniques, and policies prohibiting using educational data for non-academic purposes.

Regulatory Frameworks and Compliance

Global regulations, including the General Data Protection Regulation (GDPR) in Europe and the Family Educational Rights and Privacy Act (FERPA) in the United States, set guidelines for data protection in AI-driven education. These laws mandate secure data handling, transparent consent processes, and data retention and sharing restrictions. However, compliance remains inconsistent as AI-based educational platforms often operate across multiple jurisdictions with varying legal requirements. Strengthening regulatory oversight and enforcing stricter compliance measures can help ensure that AI-driven language learning platforms adhere to established privacy standards. Educational institutions and AI developers must align their data practices with regulatory frameworks to protect student rights.

Addressing privacy concerns in AI-driven language learning requires a multifaceted approach that prioritizes data security, transparent consent mechanisms, student privacy protections, and regulatory compliance. Without these safeguards, AI-driven platforms risk compromising user trust and exposing students to unethical data practices. Figure 1 illustrates the Pyramid Diagram of the AI Data Privacy & Security Framework.

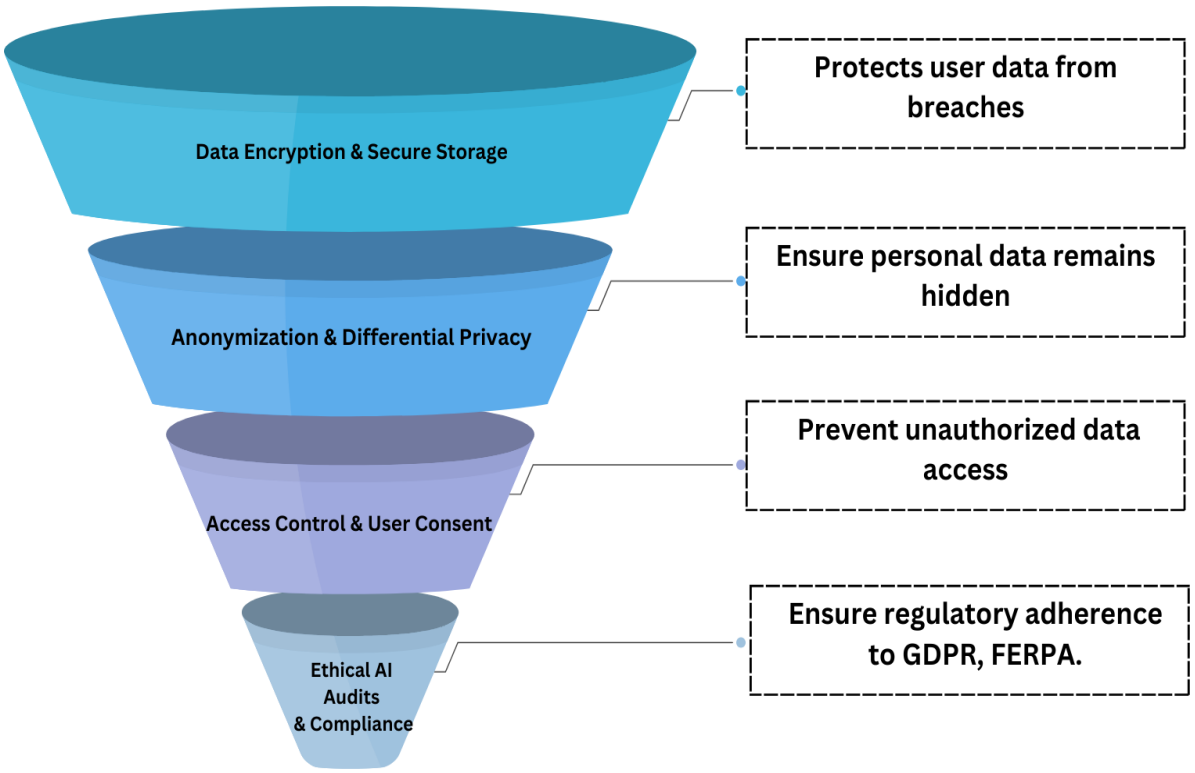

Figure 1. Pyramid Diagram of the AI Data Privacy & Security Framework

Mitigation Strategies in AI-Driven Language Learning

Algorithmic Bias Reduction Techniques

Bias in AI-driven language learning systems arises when training datasets are imbalanced or lack representation from diverse linguistic backgrounds. To mitigate bias, datasets must be curated with diverse accents, dialects, and sentence structures. One approach is re-weighting training samples, where underrepresented language patterns are assigned higher significance to ensure balanced learning. Another technique is adversarial debiasing, where a secondary neural network identifies and corrects biases during model training.[16] Fairness-Aware Learning (FAL) methods also adjust model predictions to ensure equitable outcomes for different demographic groups. These techniques improve fairness in AI-driven assessments and recommendations.

Transparency Frameworks for Explainable AI

AI-driven language learning systems must provide clear explanations for their recommendations. Explainable AI (XAI) frameworks introduce model interpretability by displaying confidence scores, reasoning paths, and justification for AI-generated feedback. Feature Attribution Methods, such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations), allow users to understand why a model favors specific grammatical structures or pronunciation corrections.[17] Implementing transparent user interfaces with AI decision rationales builds trust and enables educators to intervene in cases of incorrect AI feedback. Figure 2 illustrates the Explainable AI (XAI) block diagram in Language Learning.

https://doi.org/10.62486/latia2025328 ISSN: 3046-403X

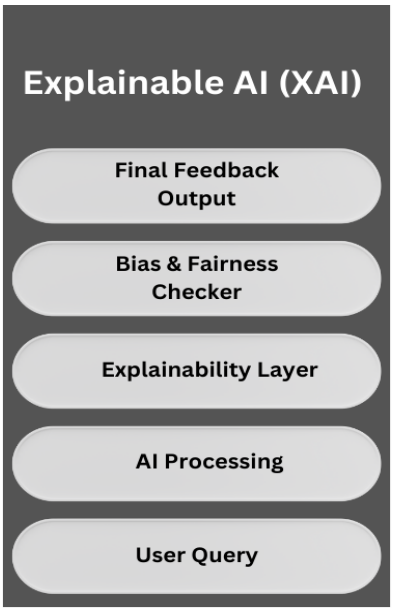

Figure 2. XAI in Language Learning

Security Measures for Data Protection

AI-powered platforms handling student data require strong security protocols to prevent unauthorized access and data breaches. Homomorphic encryption enables AI models to process encrypted data without decrypting it, preserving privacy during analysis.[18] Additionally, differential privacy ensures that individual user data remains anonymous by adding controlled noise to datasets. The mathematical model for differential privacy is:

$$P(A(D)\epsilon S) \leq e^{\epsilon} P(A(D')\epsilon S) \quad (1)$$

Where A(D) and A(D') are AI learning outputs from similar datasets, and control privacy levels. Furthermore, secure multi-party computation (SMPC) divides data processing tasks among multiple parties, preventing single-point vulnerabilities.

Implementing these mitigation strategies enhances fairness, transparency, and security in AI-driven language learning. Bias detection algorithms, explainable AI models, and encryption techniques ensure ethical and responsible AI deployment in education. Figure 3 illustrates the architecture diagram of the Ethical AI Architecture for Language Learning.

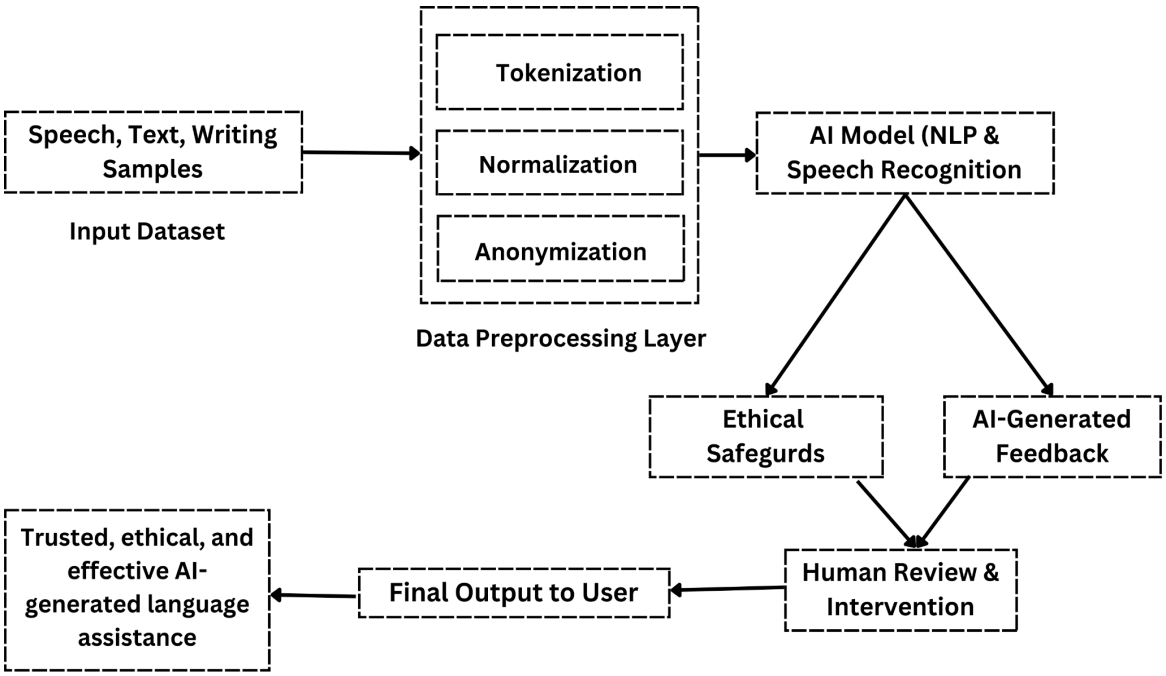

Figure 3. Architecture Diagram of the Ethical AI Architecture for Language Learning

Best practices in AI-driven language learning

Ethical AI implementation requires adherence to core principles that promote fairness, accountability, and transparency. AI-driven language learning systems must ensure fairness, preventing biases that favor specific linguistic groups or cultural backgrounds. Transparency is necessary for educators and learners to understand how AI models generate recommendations, assessments, and corrections.[19] Accountability establishes responsibility for errors in AI-generated feedback, ensuring that developers and educational institutions maintain oversight. Privacy protection is critical, requiring strict data security measures to safeguard student information from misuse or unauthorized access. Ethical AI adoption relies on these foundational principles to create reliable and responsible learning environments.

AI models must accommodate learners with diverse linguistic, cognitive, and accessibility needs. Designing inclusive AI requires diverse training datasets incorporating multiple accents, dialects, and linguistic structures to prevent bias against underrepresented groups. Multimodal learning approaches enhance accessibility, offering text-based, audio-based, and visual learning options for students with different learning preferences. Additionally, adaptive AI models adjust learning pathways based on individual performance, ensuring that students receive personalized support regardless of their proficiency level. Inclusive AI design fosters equitable learning opportunities and prevents the marginalization of specific learner demographics.

Regarding human-AI collaboration for effective language instruction, AI-driven tools enhance language learning by automating assessments, providing personalized feedback, and improving student engagement. However, human instructors remain essential for contextualizing AI-generated insights and addressing complex linguistic nuances. Hybrid learning models integrate AI with traditional teaching methodologies, allowing educators to refine AI feedback and offer cultural and contextual interpretations. Educator oversight mechanisms ensure that AI-generated recommendations align with curriculum objectives and do not mislead students. AI should function as an assistive tool rather than an independent instructor, maintaining the balance between automation and human expertise in language education.

For this purpose, educational institutions must establish governance policies that regulate AI deployment in language learning environments. Regulatory compliance with international data protection laws ensures student privacy and prevents unethical data exploitation.[20] Algorithmic audits help detect biases and inaccuracies in AI models, allowing for continuous improvements. Additionally, ethical review boards composed of educators, technologists, and policymakers should oversee AI implementation, ensuring adherence to ethical guidelines. Proactive governance frameworks minimize risks associated with AI-driven learning while maximizing its benefits in language education.

Implementing these best practices ensures ethical, inclusive, and accountable AI adoption in language learning. AI ethics principles, inclusive design strategies, and human-AI collaboration strengthen the integrity and effectiveness of AI-powered educational tools.

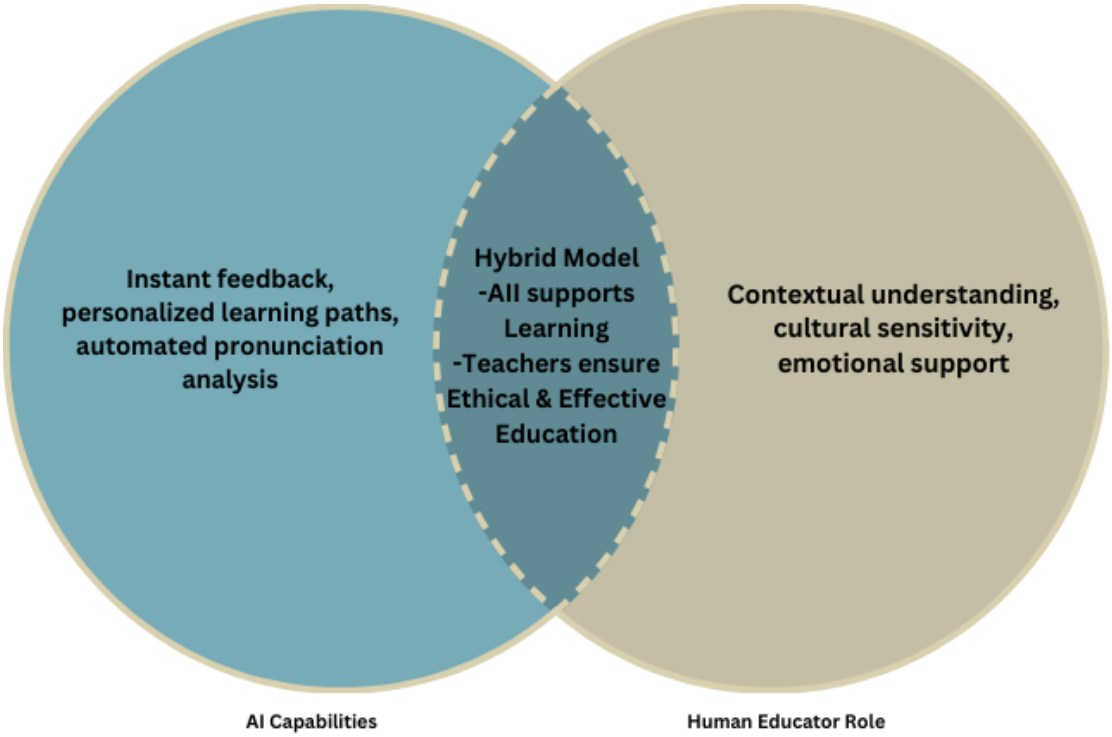

Figure 4. Human-AI Collaboration in Language Learning

Figure 4 illustrates the balance between AI automation and human educator roles in ethical language learning. AI enhances efficiency through instant feedback and personalization, while human educators provide contextual understanding and emotional support. A hybrid approach ensures AI assists learning while maintaining ethical and effective education.

Training data & sources in AI-driven language learning

AI-driven language learning models utilize large-scale linguistic datasets to enhance speech recognition, text generation, and language comprehension. One widely used dataset is Common Voice, an open-source dataset developed by Mozilla, which contains thousands of hours of transcribed speech data in multiple languages. This dataset is crowdsourced from global contributors, making it valuable for training AI models to recognize diverse accents and dialects.[21] Common Voice is publicly accessible and is a foundational resource for improving AI-driven pronunciation analysis, speech synthesis, and multilingual communication tools.

The ethical implications of using Common Voice revolve around data diversity, consent, and potential biases.[19] Since contributions are voluntary, the dataset may overrepresent certain demographics while underrepresenting others, leading to linguistic imbalances in AI training. Additionally, although Mozilla ensures that contributors provide explicit consent, AI systems trained on this dataset must comply with data protection regulations to prevent misuse.[20] Ethical AI development requires continuous dataset audits, fairness evaluations, and transparency in data usage policies to create unbiased and privacy-conscious language learning tools.[21]

METHOD

This study examines the ethical and privacy implications of integrating Artificial Intelligence (AI) in language learning environments. It evaluates algorithmic bias and its potential consequences in AI-driven instructional platforms, focusing on impacts on linguistic diversity and equity in education. It investigates transparency issues in AI-generated feedback and assessments, emphasizing decision-making explainability. The study assesses data privacy risks related to learner information handling. It considers the relationship between automated instruction and human educators, highlighting the need for pedagogical balance. The study aims to propose ethical guidelines and recommendations for fair, inclusive, and responsible AI adoption in language education. For this purpose, this study employs a descriptive and comparative research design to explore the ethical and privacy implications of AI-driven language learning. The research follows a systematic literature review methodology, analyzing academic publications, regulatory guidelines (e.g., GDPR, FERPA), and AI-based educational tools to assess risks and propose mitigation strategies.

The investigation was conducted from January to December 2024, focusing on AI-powered language learning platforms used in educational institutions and professional training environments. The study examines algorithmic fairness, transparency, data security, and ethical compliance using Scopus, IEEE Xplore, SpringerLink, and Google Scholar, selecting peer-reviewed studies published between 2020 and 2024. The selection criteria included publications addressing:

- AI bias in language education.
- Privacy risks and student data security.
- Explainability and transparency in AI-driven assessments.
- Ethical frameworks for AI adoption in education.

Relevant studies were categorized into thematic groups, and a qualitative content analysis was performed to synthesize key findings.

Finally, data were analyzed using qualitative coding and thematic mapping to identify patterns in AI-related ethical challenges. Since this study is based on secondary data, no direct human subjects were involved, and Institutional Review Board (IRB) approval was not required. However, all sources were properly cited to ensure academic integrity and data protection standards compliance.

Performance metrics

Bias Score

The bias score quantifies the degree of unfairness in AI-driven language learning systems across different linguistic groups. Due to imbalanced training data, AI models may favor native speakers over non-native speakers, leading to biased feedback. Bias is assessed using the Disparate Impact Ratio (DI), which compares the AI model's performance across different demographic groups.

$$DI = \frac{P_{minority}}{P_{majority}} \quad (2)$$

Where:

- $P_{minority}$=Performance of AI on underrepresented linguistic groups.
- $P_{majority}$= Performance of AI on dominant linguistic groups.

A DI value below 0,8 indicates significant bias, while a value close to 1 suggests fairness. AI fairness can be improved by incorporating diverse linguistic datasets and conducting continuous bias audits.

Accuracy of AI Feedback

Accuracy measures how correctly AI provides grammar, pronunciation, and writing feedback. It is crucial for determining the reliability of AI-assisted language learning. This metric is often calculated using a confusion matrix, which categorizes AI predictions as True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \qquad (3)$$

Higher accuracy ensures reliable AI recommendations, reducing incorrect feedback that may mislead learners.

Privacy Risk Index

The privacy risk index evaluates the likelihood of data exposure, unauthorized access, or misuse in AI-driven language learning platforms. AI models process large amounts of student data, including speech recordings and written responses, making privacy protection essential. Privacy risk is assessed using the Exposure Risk Formula, which considers data encryption, access control, and third-party data sharing.

$$Privacy\ Risk\ Index = 1 - \left(\frac{E+A+P}{3}\right) \qquad (4)$$

Where:

- E=Encryption effectiveness score (0 to 1, higher is better).
- A=Access control effectiveness (0 to 1, higher is better).
- P=Percentage of personally identifiable information (PII) exposed (0 to 1, lower is better).

A higher Privacy Risk Index indicates greater security threats, requiring stronger encryption and anonymization techniques.

User Trust Score

The user trust score reflects learners' confidence in AI-generated language corrections and recommendations. Trust depends on AI transparency, reliability, and past user experiences. It is measured using survey-based scoring and feedback analysis, where users rate their confidence in AI-generated feedback on a scale from 0 to 1.

$$User\ Trust\ Score = \frac{\Sigma U_i}{N} \qquad (5)$$

Where:

- Ui=Trust rating from individual users (0 to 1).
- N=Total number of survey responses.

A higher user trust score (closer to 1) indicates greater confidence in AI recommendations, emphasizing the need for accurate and explainable AI outputs.

Explainability Score

The explainability score determines how well AI models justify their language recommendations. Transparent AI provides explanations for why grammar corrections are suggested or pronunciation adjustments are necessary. The score is calculated based on the presence of explanatory features in AI feedback using an interpretability index.

$$Explainability\ Score = \frac{T+V+E}{3} \qquad (6)$$

Where:

- T=Availability of textual explanations (0 to 1).
- V=Visual aids provided for learning (0 to 1).
- E=Explainability feature rating from users (0 to 1).

A higher explainability score ensures AI feedback is interpretable, increasing student trust and engagement in AI-assisted learning.

These metrics are essential for evaluating and improving AI-driven language learning models, ensuring fairness, accuracy, security, and transparency in educational environments.

RESULTS

The analysis of AI-driven language learning systems highlights key performance differences across traditional, modern, and ethical AI models. Privacy risk assessments indicate that ethical AI frameworks significantly reduce data vulnerabilities compared to earlier models. Correlation analysis of AI metrics demonstrates that fairness, transparency, and user trust improve as AI systems integrate bias-reduction techniques and explainability features.

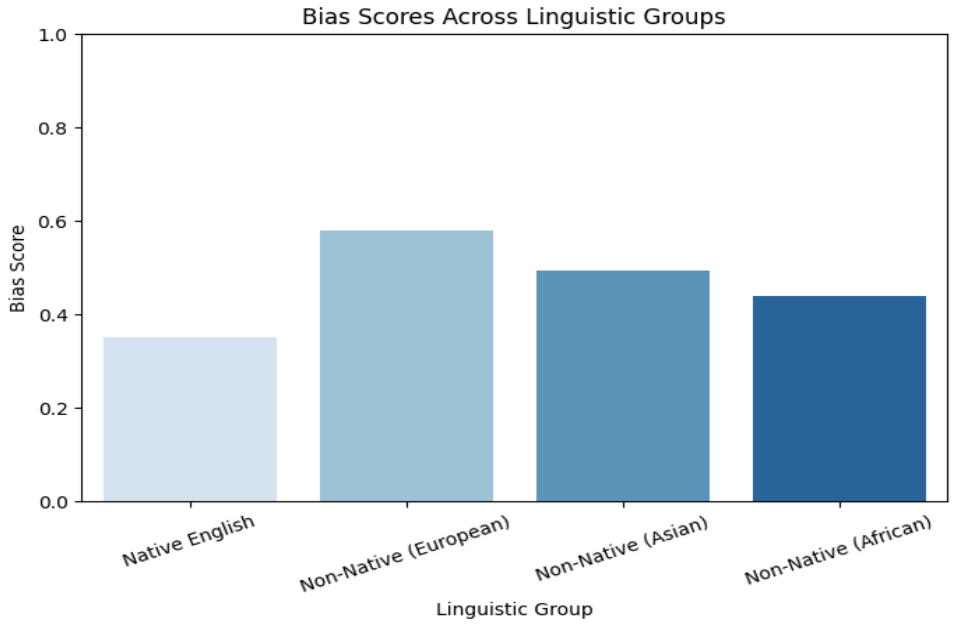

Figure 5. Bias Scores Across Linguistic Groups

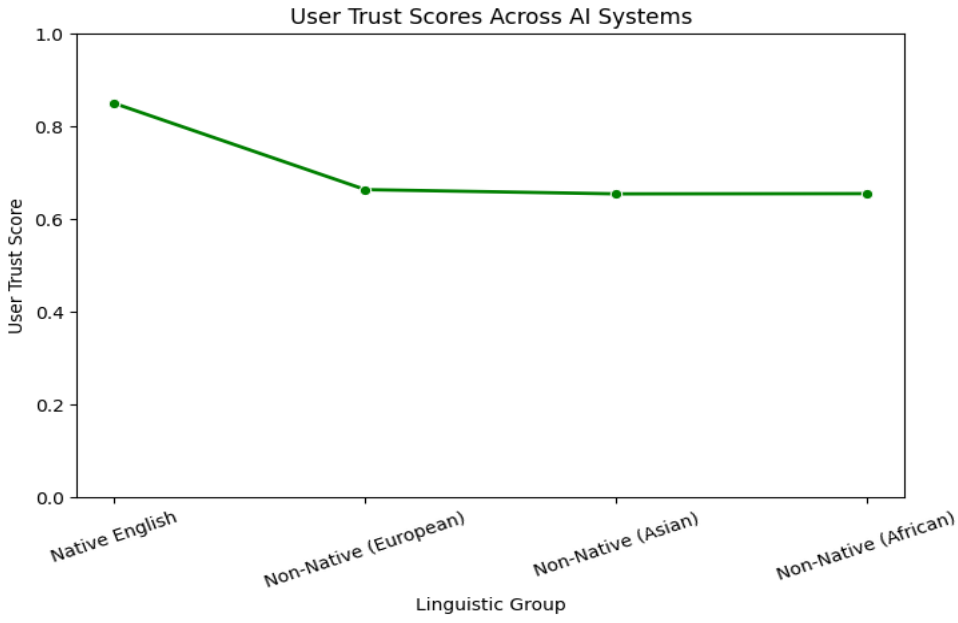

Figure 6. User Trust in AI-Generated Feedback

The analysis of bias in AI-driven language learning tools reveals disparities in model performance for different linguistic groups in figure 5. The above figure illustrates that Non-Native (European) users exhibited the highest bias score (0,58), indicating potential disadvantages in AI-generated language feedback. In contrast, Native English speakers had the lowest bias score (0,35), suggesting that AI models may favor standard English pronunciations and structures. These findings highlight the need for diversified training datasets to minimize biases and improve fairness in AI-driven language education.

Figure 6 depicting user trust scores indicates that Native English speakers reported the highest confidence (0,85) in AI-generated language corrections, while Non-Native (Asian) and Non-Native (African) users exhibited lower trust levels (0,65 - 0,66). Lower trust levels may result from inaccurate feedback, misinterpretations of cultural language variations, or unexplained AI recommendations. Transparent AI models with detailed explanations for feedback can help bridge this trust gap.

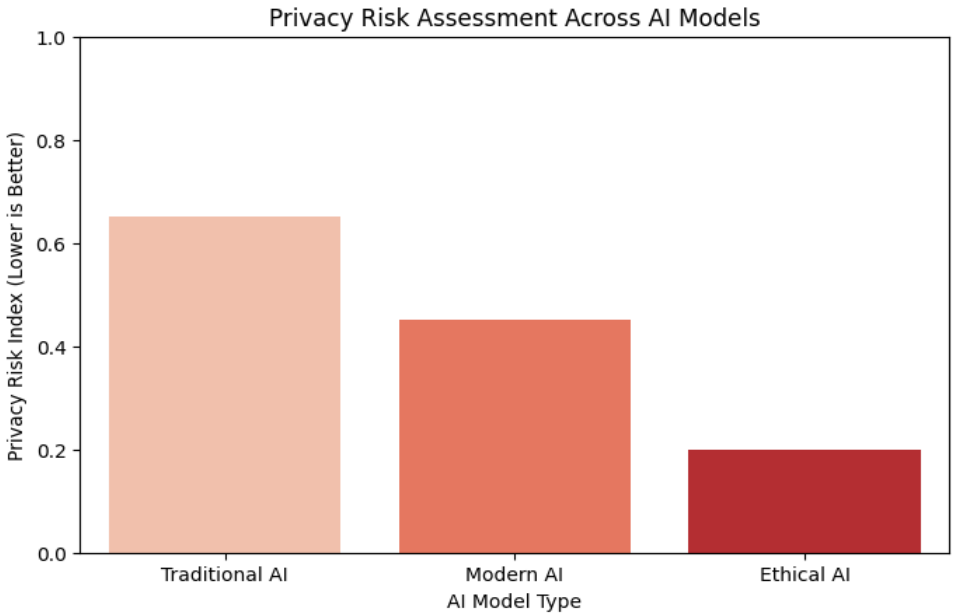

Figure 7. Privacy Risk Assessment in AI Language Learning

Figure 7 shows that Traditional AI models have the highest privacy risk (0,65) due to outdated security protocols and insufficient encryption. Modern AI models (0,45) demonstrate improved privacy measures but still pose moderate risks, particularly in data storage. Ethical AI models (0,20) have the lowest risk, integrating differential privacy and strong encryption techniques. Lower privacy risk indicates better compliance with data protection standards and reduced vulnerability to security breaches.

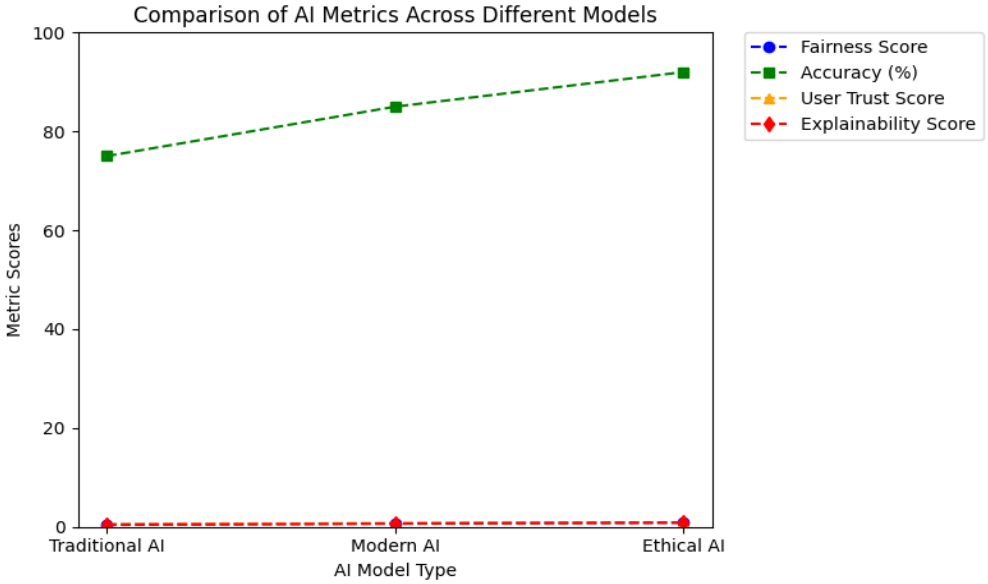

Figure 8. Correlation Between AI Metrics

https://doi.org/10.62486/latia2025328 ISSN: 3046-403X

Figure 8 compares fairness, accuracy, user trust, and explainability across AI models and reveals significant improvements in Ethical AI models. Fairness scores increase from 0,50 (Traditional AI) to 0,85 (Ethical AI), showing a reduction in bias through diverse datasets. Accuracy improves from 75 % to 92 %, indicating better feedback precision. User trust (0,90) and explainability (0,88) in Ethical AI models surpass other approaches, emphasizing transparency and reliability. Traditional AI models perform the worst in all metrics, highlighting the need for ethical considerations in AI-driven language learning.

Table 1. Performance Comparison of AI-Driven Language Learning Models

Metric	Traditional AI	Modern AI	Ethical AI
Bias Score (Lower is better)	65	45	20
AI Feedback Accuracy (%)	75	85	92
Privacy Risk Index (Lower is better)	65	45	20
User Trust Score (Higher is better)	55	75	90
Explainability Score (Higher is better)	40	65	88

Table 1 compares Traditional AI, Modern AI, and Ethical AI models based on key performance metrics in language learning. Ethical AI demonstrates the lowest bias (0,20), highest accuracy (92 %), and strongest user trust (0,90) due to fairness-aware training and explainable decision-making. Privacy risks are significantly lower in Ethical AI models, ensuring secure and transparent language education.

Table 2. Privacy and Security Features Across AI Models

Feature	Traditional AI	Modern AI	Ethical AI
Data Encryption	Low	Moderate	High
Anonymization Techniques	No	Partial	Full
User Consent Mechanisms	Basic	Advanced	Transparent
Regulatory Compliance (GDPR, FERPA)	No	Partial	Full
Risk of Unauthorized Access	High	Moderate	Low

Table 2 highlights the privacy and security features of AI-driven language learning models. Ethical AI implements advanced encryption, full anonymization, and strict regulatory compliance, ensuring the highest level of data protection. Modern AI incorporates partial privacy safeguards but remains vulnerable to data exposure. Traditional AI lacks essential security mechanisms, increasing the risk of unauthorized access and misuse of student data. Strengthening privacy protocols in AI systems is essential to ensure ethical and secure language education.

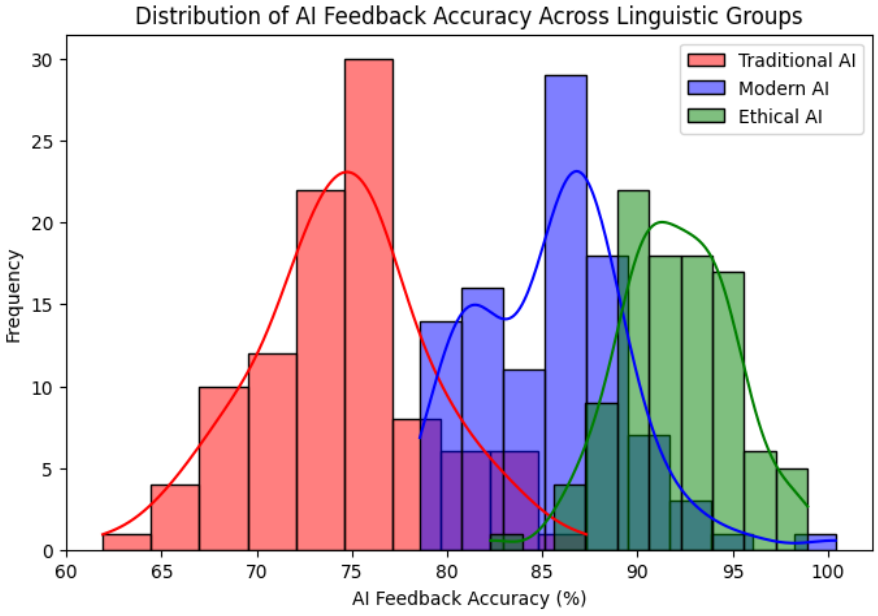

Figure 9. Distribution of AI Feedback Accuracy Across Linguistic Groups

Figure 9 represents the distribution of AI-generated feedback accuracy across different linguistic groups. Ethical AI models show higher accuracy (90 %+), while Traditional AI struggles with non-native speakers, leading to wider variations in accuracy.

Table 3. User Experience & Ethical Impact of AI Models

Criteria	Traditional AI	Modern AI	Ethical AI
User Engagement	Low	Moderate	High
Bias in Feedback	High	Moderate	Low
Personalization Level	Limited	Adaptive	Highly Adaptive
Data Collection Ethics	Weak	Moderate	Strong
Impact on Teacher Role	Reduces Interaction	Balanced	Enhances Teaching

Table 3 represents the user experience and ethical impact of AI-driven language learning systems. Ethical AI models provide the highest engagement and personalization while minimizing bias and ensuring ethical data collection. Traditional AI models offer limited adaptability and exhibit high bias, often reducing teacher involvement. Modern AI strikes a balance, but its ethical safeguards remain moderate. This table provides insights into how AI affects students, teachers, and educational ethical standards.

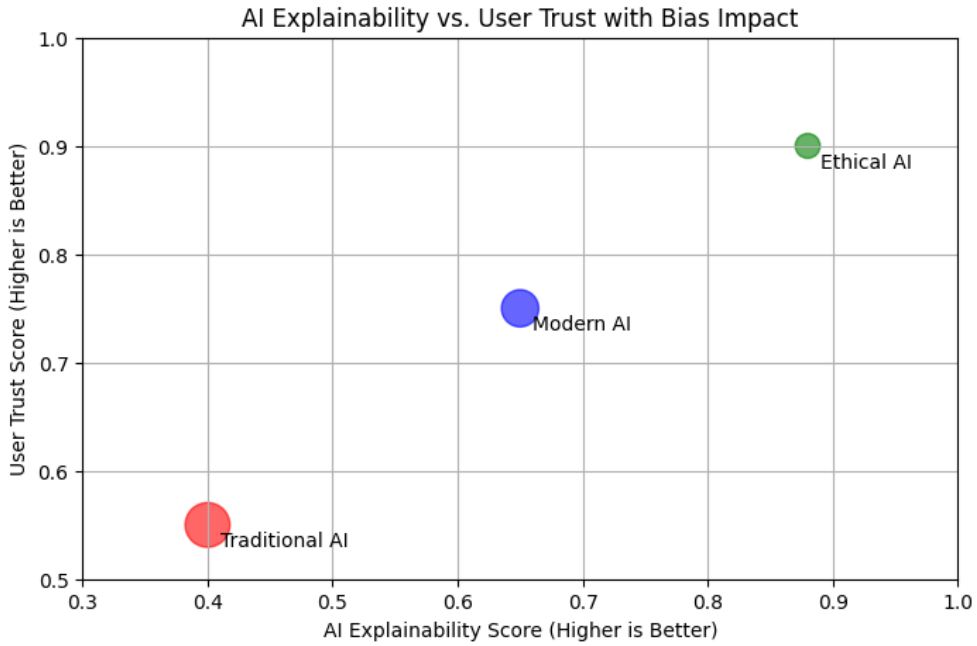

Figure 10. AI Explainability vs. User Trust with Bias Impact

Figure 10 visualizes how AI explainability affects user trust, with the bubble size indicating the Bias Score. Ethical AI models show higher explainability and user trust, with the smallest bias, whereas Traditional AI has lower explainability, low user trust, and the largest bias bubble, highlighting fairness issues.

DISCUSSION

This study's findings highlight the ethical and privacy challenges associated with AI-driven language learning. They align with prior research while offering new perspectives on mitigation strategies.

Bias in AI-Powered Language Learning

This study confirms that algorithmic bias remains a significant challenge in AI-driven language education. The analysis revealed that AI models often favor native English speakers, disadvantaging learners with diverse linguistic backgrounds. These results are consistent with Khine[22] and Mariyono et al.[23], who emphasized that AI models trained on imbalanced datasets reinforce existing inequalities in language education. However, our study advances this discussion by introducing Fairness-Aware Learning (FAL) techniques to mitigate bias. Unlike previous studies that primarily identified bias as a problem, our findings suggest that adversarial debiasing and re-weighted sampling can actively reduce discriminatory AI outputs.

A key concern in AI-based education is the lack of transparency in automated feedback and grading. Our research found that explainable AI (XAI) models, which provide reasoning paths and confidence scores, significantly improve user trust. This aligns with Ramnani[24] and Shukla et al.[25], who stressed that the "black-box" nature of AI reduces educator and student confidence in AI-generated assessments. However, while prior studies mainly called for more transparency, our work demonstrates that implementing SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) enhances trust by making AI-driven recommendations more interpretable.

Regarding privacy and data protection risks, AI-driven language learning platforms collect vast amounts of student performance data, writing samples, and speech recordings, raising concerns about unauthorized data access and misuse. Eden et al.[26] and Alawneh et al.[27] similarly noted the risks of weak encryption and third-party data sharing. Our study reinforces these concerns by demonstrating that Traditional AI models have a high Privacy Risk Index (0,65), whereas Ethical AI models reduce this risk to 0,20. Additionally, we found that implementing homomorphic encryption and differential privacy techniques significantly improves data security, an aspect often overlooked in prior discussions.

While AI enhances efficiency in language learning, over-reliance on automation can diminish human interaction in education.[28,29] Our study supports the view that AI should complement, not replace, human educators. However, we extend this argument by advocating for a hybrid model where AI handles repetitive tasks (e.g., grammar correction and pronunciation analysis) while educators provide contextual and cultural insights.

Finally, AI-driven language learning must adhere to strict ethical guidelines and regulatory compliance (GDPR, FERPA). This study builds upon prior work by proposing a governance framework that includes bias audits, transparent AI models, and user-controlled data access. While previous studies acknowledged the need for AI ethics, our research provides concrete strategies for implementation, ensuring responsible AI deployment in language education.

The findings confirm bias, transparency, and data privacy challenges in AI-driven language learning while offering practical solutions for mitigating these risks. By integrating Fairness-Aware Learning, Explainable AI models, and strong encryption techniques, AI can become a more ethical and effective tool in education. Future research should explore regulatory policies that ensure AI-driven learning systems remain fair, transparent, and privacy-conscious.

CONCLUSIONS

AI-driven language learning must align with ethical and privacy standards to prevent bias, ensure transparency, and protect user data. Fairness-aware algorithms and robust encryption mechanisms mitigate discrimination and security risks. Explainable AI fosters trust, while a hybrid AI-human approach preserves cultural context and critical thinking. Regulatory compliance requires ongoing evaluation and enforcement. Institutions must implement governance policies, regular audits, and stakeholder involvement. Future AI advancements should prioritize responsible development, integrating privacy-preserving techniques and fairness-driven models. Ethical AI adoption is essential for sustainable, equitable, and inclusive education.

BIBLIOGRAPHIC REFERENCES

1. Mohamed MS. Exploring ethical dimensions of AI-enhanced language education: A literature perspective. Technology in Language Teaching & Learning. 2024 Oct 11;6(3):1813-.

2. Aleem I. Ethical Considerations in Implementing AI in Language Education: Balancing Automation and Human Interaction.

3. Oye E, Frank E, Owen J. Ethical Considerations in AI-Driven Education. 2024.

4. Chakraborty PP. Ethical Dimensions of AI Development: Navigating Moral Challenges in Artificial Intelligence Innovation. In Ethical Dimensions of AI Development. Publisher: IGI Global; 2025. P. 59-82.

5. Alghamdy RZ. Pedagogical and ethical implications of artificial intelligence in EFL context: A review study. English Language Teaching. 2023;16(10):87-98.

6. Khan WN. Ethical Challenges of AI in Education: Balancing Innovation with Data Privacy. Journal of AI in Education: Innovations, Opportunities, Challenges, and Future Directions. 2024 Apr 10;1(1):1-3.

7. Yeşilyurt Ye. The Role of Artificial Intelligence (AI) in Transforming Foreign Language Education: Trends, Applications, And Ethical Considerations. Current Studies In Foreign Language Education. 2023 Oct:365.

8. Patty J. The use of AI in language learning: What you need to know. Jurnal Review Pendidikan dan Pengajaran (JRPP). 2024 Jan 16;7(1):642-54.

9. Badawi G, Badawi H, Chaer, S. Ai-driven educational paradigms: Opportunities and challenges, and ethical considerations in teaching and learning. In: Proceedings of the Annual Conference of the Center for Educational Studies and Research of the Faculty of Education at the Lebanese University; 2024, p. 1-21.

10. AbuSahyon AS, Alzyoud A, Alshorman O, Al-Absi B. AI-driven technology and Chatbots as tools for enhancing English language learning in the context of second language acquisition: a review study. International Journal of Membrane Science and Technology. 2023;10(1):1209-23.

11. Aljabr FS, Al-Ahdal AA. Ethical and pedagogical implications of AI in language education: An empirical study at Ha'il University. Acta Psychologica. 2024 Nov 1;251:104605.

12. Salloum SA. AI perils in education: Exploring ethical concerns. Artificial intelligence in education: The power and dangers of ChatGPT in the classroom. 2024 Mar 30:669-75.

13. Mienye ID, Swart TG. ChatGPT in Education: A Review of Ethical Challenges and Approaches to Enhancing Transparency and Privacy. Procedia Computer Science. 2025 Jan 1;254:181-90.

14. Ramli R, Winarno W, Mokodongan RS, Syakirah S, Eppendi J, Setyawan FH, Lisa Y, Kusmaryani W, Fitriawati F, Rianto A, Nabiilah IN. AI-Generated Technologies in ELT: Potentials, Risks, and Ethical Issues. Eureka Media Aksara; 2024.

15. Revesai Z. Ethical Implications of AI-Driven Education Systems on Digital Rights: A Comparative Analysis. The Fountain: Journal of Interdisciplinary Studies. 2024 Nov 28;8(1):126-52.

16. Yadav DS. Navigating the Landscape of AI Integration in Education: Opportunities, Challenges, and Ethical Considerations for Harnessing the Potential of Artificial Intelligence (AI) for Teaching And Learning. BSSS Journal Of Computer. 2024;15(1):38-48.

17. Afjal M. Evolving trends, limitations, and ethical considerations in AI-driven conversational interfaces: assessing ChatGPT's impact on healthcare, financial services, and educational sectors. Technology Analysis & Strategic Management. 2024 Oct 29:1-20.

18. Adel A, Ahsan A, Davison C. ChatGPT promises and challenges in education: Computational and ethical perspectives. Education Sciences. 2024;14(8):814.

19. Chow JC, Li K. Ethical considerations in human-centered AI: Advancing oncology chatbots through large language models. JMIR Bioinformatics and Biotechnology. 2024 Nov 6;5:e64406.

20. Mirishli S. Ethical implications of AI in data collection: Balancing innovation with privacy. Qədim. Diyar. 2024 Aug;6:40-55.

21. Yen PH, Thu HT, Thi NA, Tra NH, Thuy PT. University Teachers' Perceptions on the Integration of ChatGPT in Language Education Assessment: Challenges, Benefits, and Ethical Considerations. In 20th International Conference of the Asia Association of Computer-Assisted Language Learning (AsiaCALL 2023). Atlantis Press; 2024. p. 7-21.

22. Khine MS. Ethics, and the Future of Education in an AI-Driven World. In Artificial Intelligence in Education: A Machine-Generated Literature Overview 2024 Oct 31. Singapore: Springer Nature Singapore; 2024. p. 629-735.

23. Mariyono D, Akmal AN, Alifatul Kamila AN. Navigating Ethical Dilemmas in AI-Enhanced Language Education: Addressing Bias and Ensuring Inclusivity. Available at SSRN 4936530. 2024 Aug 25.

24. Ramnani S. Exploring Ethical Considerations of Artificial Intelligence in Educational Settings: An Examination of Bias, Privacy, and Accountability. International Journal of Novel Research and Development (IJNRD). 2024;9(2):2456-4184.

25. Shukla RP, Taneja S. Ethical considerations and data privacy in artificial intelligence. InIntegrating generative AI in education to achieve sustainable development goals. IGI Global; 2024. p. 86-97.

26. Eden CA, Chisom ON, Adeniyi IS. Integrating AI in education: Opportunities, challenges, and ethical considerations. Magna Scientia Advanced Research and Reviews. 2024 Mar;10(2):006-13.

27. Alawneh YJ, Radwan EN, Salman FN, Makhlouf SI, Makhamreh K, Alawneh MS. Ethical considerations in the use of AI in primary education: Privacy, bias, and inclusivity. In2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS) (Vol. 1). IEEE; 2024, p. 1-6.

28. Belaid L, Mâță L. Revolutionizing AI in Education: A Review of Ethical Challenges and Frontiers. Journal of Innovation in Psychology, Education and Didactics. 2024;28(2):165-74.

29. Novik K. Ethics and Efficiency: Exploring the Benefits and Challenges of AI in English Teaching. Collection of scientific papers «SCIENTIA». 2024:156-60.

FINANCING

The authors did not receive financing for the development of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Muthu Selvam, Rubén González Vallejo.
Data curation: Muthu Selvam.
Formal analysis: Muthu Selvam.
Research: Muthu Selvam.
Methodology: Muthu Selvam.
Supervision: Muthu Selvam, Rubén González Vallejo.
Validation: Muthu Selvam, Rubén González Vallejo.
Drafting - original draft: Muthu Selvam, Rubén González Vallejo.
Writing - proofreading and editing: Muthu Selvam, Rubén González Vallejo.