Journal of Al-Farabi for Humanity Sciences Volume (9) Issue (6) 30 June (2026)
ISSN: 2957-3874 (Print)
Journal of Al-Farabi for Humanity Sciences (JFHS)
https://iasj.rdd.edu.iq/journals/journal/view/95
مجلة الفارابي للعلوم الإنسانية تصدرها جامعة الفارابي

Al-and Mobile-Assisted English Language Learning through WhatsApp: An Innovative Approach to Vocabulary Development among EFL Learners.

Ruaa Zuhair Hamodi
English Department, College of Management and Economics, Mousil University, Iraq.
Ruaa.zuhair@uomousul.edu.iq

التعلم بمساعدة الذكاء الاصطناعي والهاتف المحمول للغة الإنجليزية من خلال تطبيق واتساب: نهج مبتكر لتطوير المفردات بين متعلمي اللغة الإنجليزية كلغة أجنبية

رؤى زهير حمودي
قسم اللغة الإنكليزية، كلية الادارة والاقتصاد جامعة الموصل، العراق.

ملخص الدراسة

تهدف هذه الدراسة إلى تقديم مراجعة منهجية للأدبيات التي تناولت توظيف تطبيق واتساب (WhatsApp) مع تطبيقات الذكاء الاصطناعي (AI) أو الأدوات الرقمية الذكية في تعلم اللغة الإنجليزية كلغة أجنبية، مع تركيز خاص على تعلم المفردات ونتائج التعلم المرتبطة بها. اعتمدت الدراسة منهجية مراجعة منظمة وفق بروتوكول واضح للبحث والفرز والاختيار، شمل قواعد بيانات أكاديمية ومصادر علمية موثوقة ضمن إطار زمني محدد. تم تحليل الدراسات المقبولة وفق معايير تتعلق بنوع التصميم (تجريبي/شبه تجريبي/وصفي)، وخصائص العينة، وطبيعة التدخل عبر واتساب، ودور أدوات الذكاء الاصطناعي (مثل روبوتات المحادثة، أدوات التوليد اللغوي، التحليل الآلي، أو التخصيص التكيفي.)تشير نتائج المراجعة إلى أن واتساب يمثل بيئة تعليمية فعالة لتعزيز التعرّض المتكرر للمفردات، ودعم التفاعل الاجتماعي، وتقديم أنشطة قصيرة متكررة ترفع من فرص الاحتفاظ والاسترجاع، خاصة عند دمجه باستراتيجيات تعلم مقصودة ومتابعة مستمرة. كما تُظهر الأدلة أن توظيف أدوات الذكاء الاصطناعي يمكن أن يضيف قيمة عبر التغذية الراجعة الفورية، وتوليد الأمثلة والسياقات، وتخصيص التدريبات، إلا أن فعاليته تتأثر بعوامل مثل جودة التصميم التعليمي، ومصداقية المحتوى، وقدرة المتعلمين الرقمية، والخصوصية، وتفاوت الوصول إلى الإنترنت. وتخلص الدراسة إلى ضرورة تبني تصميمات أكثر صرامة وقياسات معيارية للنتائج، مع التأكيد على حوكمة الاستخدام (الخصوصية، الأخلاقيات، العدالة، وإدارة المخاطر) لضمان أثر تعليمي مستدام وقابل للتعميم.

الكلمات المفتاحية :واتساب، الذكاء الاصطناعي، تعلم المفردات، تعلم اللغة بمساعدة الهاتف، مراجعة منهجية، تعلم اللغة الإنجليزية كلغة أجنبية.

Abstract:

This study provides a systematic review of the literature examining the use of WhatsApp in combination with artificial intelligence (AI) and intelligent digital tools in EFL (English as a Foreign Language) learning, with a particular focus on vocabulary development and related learning outcomes. A structured review protocol was applied to identify, screen, and select relevant studies from academic databases and credible scholarly sources within a defined time window. Included studies were analyzed in terms of research design (experimental/quasi-experimental/descriptive), participant characteristics, the nature of WhatsApp-based interventions, and the role of AI-supported components (e.g., chatbots, generative tools, automated feedback, language analytics, or adaptive practice)Overall, the evidence suggests that WhatsApp can serve as an effective learning environment

by increasing repeated exposure to target vocabulary, strengthening social interaction, and enabling frequent micro-learning activities that support retention and retrieval, especially when paired with intentional learning strategies and consistent instructional follow-up. The review also indicates that AI tools may add value through instant feedback, automated example generation, contextualized practice, and personalization; however, effectiveness is shaped by instructional design quality, content reliability, learners' digital readiness, privacy considerations, and unequal access to connectivity. The study concludes that stronger methodological designs and more standardized outcome measures are needed, alongside clear governance of AI-mediated learning (privacy, ethics, fairness, and risk management) to ensure robust and generalizable educational impact.
Keywords: Mobile-Assisted Language Learning (MALL), WhatsApp, Vocabulary development, EFL learners, Artificial intelligence, Mobile learning

Introduction

Vocabulary knowledge is widely regarded as a core component of communicative competence in a second or foreign language, as it provides learners with the lexical resources needed to understand input and produce meaningful output (Nation, 2001; Schmitt & Schmitt, 2020). Learners who possess limited lexical repertoires often struggle to interpret texts, express their intended meanings, and participate in sustained interaction, even when they have acquired a reasonable command of grammatical structures (Paredes, 2024). In many EFL classrooms, however, vocabulary instruction still relies on decontextualized word lists, rote memorization, and brief in-class practice, and such practices tend to result in shallow learning and rapid forgetting rather than durable vocabulary growth. At the same time, today's learners spend a substantial part of their daily lives using smartphones and social media platforms, which creates opportunities to extend vocabulary learning beyond the physical classroom and to situate it in more flexible, informal, and learner-centered environments. This tension between traditional vocabulary teaching and the realities of learners' digital practices is particularly visible in university EFL contexts, including those in Iraq, where students routinely use mobile applications to communicate with peers and instructors.Within this broader shift, Mobile-Assisted Language Learning (MALL) has emerged as a significant area of research, as it draws on the mobility and connectivity of smartphones and messaging applications to support language learning across time and space (Kukulska-Hulme, 2012). Empirical work over the last few years has shown that mobile-based vocabulary learning can enhance learners' vocabulary size, retention, and motivation when compared with textbook-based instruction alone (Chakir, 2024; Janfeshan, 2023; Bieńkowska, 2020). Systematic reviews of vocabulary learning with mobile platforms likewise report generally positive effects on learners' engagement and vocabulary outcomes, indicating that mobile environments do not merely duplicate classroom practices but reshape how learners encounter, rehearse, and recycle lexical items (Ji & Aziz, 2021; Okumuş Dağdeler, 2023).Among the various tools associated with MALL, WhatsApp is now one of the most widely used messaging applications worldwide, especially in contexts where access to the internet is primarily mobile. Recent statistics indicate that WhatsApp has around three billion monthly active users and is among the most widely used mobile messenger applications globally (DataReportal, 2024; Statista, 2024). Its constant availability, low cost, and familiarity make it an appealing platform for language teachers who wish to build interactive learning communities, circulate multimodal input, and offer timely feedback. A growing body of research has examined WhatsApp-mediated vocabulary learning and has reported positive effects on vocabulary development and learner engagement in different EFL settings. For example, Bensalem (2018) investigated the impact of WhatsApp on the development of academic English vocabulary among Gulf EFL students and found that the WhatsApp group outperformed a control group that followed a traditional approach, while other studies have shown that WhatsApp-based activities significantly improved ESP learners' vocabulary knowledge and motivation (Barzani, 2024; Boufahja, 2025; Chen et al., 2019). In many Iraqi universities, WhatsApp is already used informally for sharing announcements and learning materials, which makes it a promising platform for more structured vocabulary instruction.More recently, the integration of artificial intelligence (AI) into messaging platforms has opened further possibilities for individualized and interactive language learning experiences. AI-driven chatbots and assistants embedded in applications such as WhatsApp and WhatsApp-based AI services have been used to provide grammar explanations, vocabulary suggestions, corrective feedback, and prompts for practice, thereby supporting self-directed learning and fostering greater learner autonomy (Panmei & Shimray, 2025; Oktadela, Elida, & Ismail, 2023). Studies on AI tools and chatbots in language education more broadly have generally reported positive learner perceptions and promising gains in vocabulary and writing performance, as well as shifts in teachers' views of technology-mediated instruction (Al-Khresheh, 2024; Sapuan, Sulaiman, & Mohamad, 2025).

However, many of these investigations focus on general language skills, writing quality, or attitudes toward AI tools rather than on detailed measures of vocabulary growth within a clearly defined mobile-based environment. Despite this growing body of work, most existing studies either address mobile-based vocabulary learning without explicit AI integration, or they examine AI-assisted learning without situating it within a specific, widely used mobile communication platform such as WhatsApp. Few studies report fine-grained vocabulary gain measures that distinguish between receptive and productive knowledge in AI- and mobile-assisted interventions. In addition, there is still little empirical evidence from Iraqi university EFL learners, despite the widespread use of WhatsApp in this context and the increasing institutional interest in integrating AI-based tools into language teaching. This combination of limited empirical work on AI-supported WhatsApp vocabulary instruction and the practical need for effective digital solutions in the Iraqi EFL context points to a clear research gap.Against this background, the present study investigates the impact of an AI- and mobile-assisted vocabulary learning intervention delivered through WhatsApp on Iraqi university EFL learners' vocabulary development and on their perceptions of this learning environment. Specifically, the study examines whether learners who receive WhatsApp-based, AI-assisted vocabulary instruction achieve greater receptive and productive vocabulary gains than peers who follow traditional textbook-based instruction, and how they evaluate the usefulness and usability of this technology-enhanced approach. By focusing on WhatsApp as an already familiar, low-cost platform and embedding AI-based support within it, the study seeks to contribute to the emerging body of work on MALL and AI-assisted language learning and to offer context-sensitive insights for designing vocabulary instruction that aligns with the realities of contemporary EFL classrooms. Based on this rationale, the study formulated the following research aim and questions.

Research Aim

The study aims to evaluate the effectiveness of an AI- and mobile-assisted vocabulary learning intervention delivered through WhatsApp in enhancing Iraqi university EFL learners' receptive and productive vocabulary knowledge, and to explore their perceptions of this learning environment.

Research Questions

1. Do Iraqi university EFL learners who receive WhatsApp-based, AI-assisted vocabulary instruction demonstrate significantly greater gains in receptive and productive vocabulary knowledge than learners who receive traditional textbook-based instruction?
2. How do learners perceive the usefulness, usability, and overall learning experience of the WhatsApp-based, AI-assisted vocabulary learning intervention?

Literature review

Vocabulary Learning in EFL Contexts

Vocabulary has long been acknowledged as a central component of second and foreign language proficiency, not only because it enables comprehension and production, but also because it shapes learners' confidence and willingness to communicate (Nation, 2001; Schmitt & Schmitt, 2020). Learners with restricted lexical repertoires often struggle to follow lectures, read academic texts, and participate in interactive classroom tasks, even when they possess a reasonable command of grammar (Paredes et al., 2024). For this reason, vocabulary development has been treated as a key pedagogical priority in many EFL curricula.Despite this recognition, vocabulary instruction in many EFL settings still relies on teacher-fronted explanation, bilingual word lists, and memorization-oriented exercises that are only loosely connected to meaningful communication. Empirical work in different contexts has suggested that such decontextualized approaches tend to promote short-term memorization and rapid forgetting, rather than durable and flexible lexical knowledge that can be deployed in new tasks (Lei et al., 2022; Li, 2010). In response, contemporary research has advocated for vocabulary learning designs that support spaced repetition, multimodal input, and the integration of lexical items into authentic, task-based activities, frequently mediated by digital tools.

Mobile-Assisted Language Learning and Vocabulary Development

The rapid spread of smartphones has provided fertile ground for Mobile-Assisted Language Learning (MALL), in which learners use mobile devices to access language input, practice tasks, and communicative opportunities beyond the classroom. MALL studies have shown that mobile environments can support ubiquitous, bite-sized learning episodes that fit into learners' daily routines and enable frequent, distributed exposure to target vocabulary (Kukulska-Hulme, 2012; Alnufaie, 2022)Several empirical studies have examined mobile-based vocabulary learning applications and platforms. Hao et al. (2019), for example, evaluated a mobile vocabulary application for lower-achieving middle-school learners of English and reported improvements in vocabulary performance alongside higher motivation and confidence in learning. More broadly, meta-analytic and systematic review evidence indicates that mobile-assisted vocabulary learning tends to yield positive effects on vocabulary gains, learner engagement, and self-regulatory behaviors when compared with paper-based or classroom-only instruction (Burston, 2015; Hao et al., 2019; Lim & Taguchi, 2024; Okumuş Dağdeler, 2022; Qian, 2022). These reviews highlight design features such as spaced repetition, push notifications, gamification, and multimedia glosses as particularly beneficial for vocabulary retention.Beyond measurable gains in vocabulary size, mobile-based interventions have often been associated with changes in learners' attitudes and self-regulation. Lei et al. (2022), for instance, found that MALL had a significant impact on EFL learners' vocabulary learning attitudes and self-regulatory capacity, suggesting that mobile tools may foster greater learner agency in planning, monitoring, and evaluating vocabulary study. At the same time, many interventions rely on dedicated vocabulary applications rather than communication platforms that learners already use in their everyday lives, which raises questions about ecological validity and sustainability once the experimental project ends.

WhatsApp-Mediated Vocabulary Learning

Within the broader MALL landscape, WhatsApp has attracted particular attention because of its ubiquity, low cost, and integration into learners' routine communication practices. Recent statistics indicate that WhatsApp has close to three billion monthly active users and ranks among the most widely used messaging applications worldwide, including in many developing countries where mobile access outpaces desktop connectivity (DataReportal, 2024; Statista, 2024). This penetration makes WhatsApp a promising platform for embedding vocabulary learning into the digital ecosystems that learners already inhabit.A growing number of studies have explored WhatsApp-mediated vocabulary instruction in EFL contexts. Bensalem (2018) compared a WhatsApp-based academic vocabulary program with a traditional textbook-based approach among Gulf EFL university students and found that the WhatsApp group achieved significantly higher post-test scores and reported positive perceptions of the mobile learning experience. Similarly, Barzani and Omar (2024) investigated WhatsApp-mediated vocabulary learning among Kurdish EFL learners and reported that the experimental group outperformed a control group on vocabulary tests and expressed favorable attitudes toward the use of WhatsApp for vocabulary learning.Other research in various contexts, including Iranian, Indonesian, and Pakistani EFL learners, has also documented positive effects of WhatsApp-based activities on vocabulary acquisition, participation, and motivation (Jafari & Chalak, 2016; Khatoon et al., 2023; Salman & Imam, 2021). In these studies, WhatsApp groups were typically used to distribute lexical items, example sentences, quizzes, and multimedia content, and to encourage peer interaction around vocabulary use. In the Iraqi context, descriptive work has shown that university students routinely rely on mobile applications and social media for informal English learning and for classroom-related communication with peers and instructors (e.g., Mohammed, 2020), which further underscores the relevance of WhatsApp as a potential instructional space.At the same time, existing research points to several limitations. Many WhatsApp-based interventions rely on teacher-driven posting of vocabulary lists and activities, with limited personalization or adaptive feedback. Furthermore, several studies focus primarily on learners' perceptions and general achievement scores, without distinguishing clearly between receptive and productive vocabulary gains or tracking long-term retention. These constraints suggest a need for more fine-grained assessments of how WhatsApp-mediated tasks influence different dimensions of vocabulary knowledge over time.

AI-Driven Tools and Chatbots in Vocabulary Instruction

Parallel to the growth of MALL, the integration of artificial intelligence (AI) into language learning has gained momentum, particularly through AI-powered chatbots, generative AI tools, and intelligent tutoring systems. AI chatbots can simulate conversational partners, provide instant feedback, and generate tailored practice activities, thereby offering opportunities for individualized vocabulary learning and repeated exposure to target items (Hasan et al., 2022; Oktadela, Elida, & Ismail, 2023)Oktadela et al. (2023), for example, implemented an AI chatbot application to support young EFL learners' vocabulary learning and reported significant improvements in vocabulary scores compared with a control group, as well as positive learner perceptions of the chatbot's interactivity and availability. In an Iranian study using an AI-based chatbot for vocabulary practice, the experimental group achieved higher vocabulary retention than a control group, and participants emphasized the benefits of personalized exercises and instant feedback for sustaining engagement (Heidari & Rajabi, 2024). Hasan et al. (2022) similarly showed that an AI-enhanced, technology-assisted vocabulary program led to vocabulary gains and improved learner attitudes among Saudi EFL students.Beyond vocabulary-specific interventions, recent work has examined the role of generative AI in broader EFL pedagogy. Zaim et al. (2024) explored the integration of generative AI tools into university EFL teaching preparation and reported that AI-supported activities influenced pre-service teachers' intentions to adopt educational technologies and reshaped classroom practices. Other studies have investigated learners' perceptions of AI tools in language learning, generally finding a combination of enthusiasm about personalization and concerns about overreliance and academic integrity (Alkandari, 2025; Metwally, 2025). Collectively, this body of research suggests that AI tools can support vocabulary learning by providing tailored input, opportunities for retrieval practice, and immediate corrective feedback. However, many AI-based interventions have been implemented through standalone applications or web platforms rather than being embedded into the communication tools that learners already use extensively, such as WhatsApp.

Integrating WhatsApp, AI, and Vocabulary Learning: Identified Gap

Only a small number of studies have begun to explore the intersection of mobile messaging platforms and AI-driven support for vocabulary learning. Most WhatsApp-based interventions have focused on teacher-curated content and peer interaction, without integrating AI components that could personalize vocabulary input, scaffold practice, or automate feedback at scale. Conversely, many AI chatbot studies have been conducted in controlled environments using dedicated apps or browser-based interfaces, with limited attention to how such tools might operate within the everyday messaging ecosystems that mediate learners' social and academic lives. Moreover, as highlighted in recent systematic reviews and empirical work on mobile-assisted vocabulary learning, relatively few studies offer fine-grained analyses of vocabulary gains that differentiate between receptive and productive knowledge, or that examine delayed post-tests in AI- and mobile-assisted settings (Hao et al., 2019; Lei et al., 2022; Lim & Taguchi, 2024; Okumuş Dağdeler, 2022; Qian, 2022). Evidence from under-researched contexts such as Iraqi universities is especially scarce, despite reports that students in Iraq frequently rely on mobile applications and social media for informal English learning and for managing classroom communication (Mohammed, 2020)Against this backdrop, the present study positions itself at the intersection of MALL, WhatsApp-mediated instruction, and AI-supported vocabulary learning. By implementing an AI-assisted vocabulary program within WhatsApp and assessing both receptive and productive vocabulary gains among Iraqi university EFL learners, the study responds directly to the gaps identified in previous research and aims to provide contextually grounded evidence on the potential and limitations of AI- and mobile-assisted vocabulary instruction.

Methods

Review design and reporting approach

This study adopts a systematic literature review (SLR) to synthesize evidence on how WhatsApp-mediated learning and AI-enabled tools (e.g., chatbots, generative AI, automated feedback) are used to support English vocabulary learning in EFL/ESL contexts. The review follows a predefined, transparent workflow for searching, screening, eligibility assessment, data extraction, quality appraisal, and synthesis. Reporting follows core PRISMA 2020 items (adapted) to fit a mixed and heterogeneous evidence base and a narrative synthesis (Page et al., 2021).

Information sources

Searches were conducted in January 2026 across the following sources to capture both language learning and educational technology research:

- Scopus
- Web of Science Core Collection
- ERIC
- ScienceDirect
- Google Scholar (supplementary only, for citation chasing and missing items)

In addition, backward and forward citation chasing was applied to included studies and key reviews to identify further eligible records.

Search strategy

Search strings were designed around two required concept blocks: (1) WhatsApp/instant messaging and (2) vocabulary learning, with an additional tracked block for AI-enabled support. Searches were run primarily in Title/Abstract/Keywords fields where supported, and database syntax was adapted per platform.

Example core query:
("WhatsApp" OR "instant messaging" OR "mobile messenger" OR "messaging app") AND

("vocabulary" OR "lexis" OR "word learning" OR "lexical acquisition" OR "vocabulary retention") AND ("artificial intelligence" OR AI OR chatbot OR "generative AI" OR LLM OR "automated feedback" OR "intelligent tutoring")
When AI-specific retrieval was sparse, searches were executed in two passes:
1.WhatsApp + vocabulary (broad capture), then
2.WhatsApp + AI terms + vocabulary/English learning terms (AI-focused capture).
Full database-specific search strings are retained for transparency and can be provided in an appendix if required.

Eligibility criteria

Inclusioncriteria

Records were included if they met all of the following:
1.Time window: January 2012 to January 2026 (early 2026).
2.Language: English.
3.Population: EFL/ESL learners (school, university, or adult education).
4.Intervention/exposure:
○ WhatsApp used as a primary platform for vocabulary learning activities, and/or
○ WhatsApp learning integrated with AI-enabled support relevant to vocabulary learning.
5.Outcomes: Vocabulary-related outcomes (e.g., gains, retention, receptive/productive vocabulary, test scores) or clearly defined vocabulary-learning indicators.
6.Study types: empirical studies (experimental, quasi-experimental, qualitative, mixed-methods) and systematic reviews directly addressing WhatsApp/MALL vocabulary learning.

Exclusioncriteria

Records were excluded if they were:
- purely technical AI studies with no learning/vocabulary outcome;
- WhatsApp studies without a vocabulary learning design or measurable learning focus;
- opinion pieces without a clear method and credible referencing;
- duplicates, non-English records, or items lacking sufficient bibliographic detail.

Study selection and screening process

Screening proceeded in three stages:
1. De-duplication of retrieved records.
2. Title/abstract screening against inclusion criteria.
3. Full-text screening for final eligibility and extraction.

Screening was conducted by a single reviewer. To mitigate potential reviewer bias, eligibility criteria were predefined and piloted on an initial subset of records to calibrate interpretation, and borderline exclusions were re-checked at full-text stage for consistency. Screening decisions and reasons for exclusion were recorded to enable an adapted PRISMA-style flow description.

Data extraction

A standardized extraction form was used to collect:
- bibliographic information (authors, year, venue);
- setting and sample characteristics (country, level, N where reported);
- study design (experimental/quasi/qualitative/mixed; comparison group if applicable);
- intervention features (WhatsApp task design, duration, intensity, teacher role, peer interaction);
- AI component (if present): tool type, integration mode, learner interaction;
- outcome measures (construct, instruments/tests, pre/post timing, delayed retention where reported);
- implementation notes (engagement, feasibility constraints, ethical/privacy considerations);
- key findings and limitations reported by the authors.

Quality appraisal

Given mixed designs, methodological quality was assessed using the Mixed Methods Appraisal Tool (MMAT) 2018 (Hong et al., 2018). Where helpful, design-specific critical appraisal considerations (e.g., for quasi-experimental designs) were cross-checked using JBI guidance (Joanna Briggs Institute, 2020). Appraisal results informed interpretive weighting in synthesis rather than automatic exclusion, unless study reporting was too unclear to interpret reliably.

Synthesis method

Because outcomes and interventions are heterogeneous, synthesis was conducted as a narrative systematic synthesis with structured grouping:
1.Descriptive mapping of included records by design type, learner level, intervention type (WhatsApp-only vs WhatsApp + AI), and vocabulary outcome type.
2.Thematic synthesis to integrate patterns across studies, focusing on: (i) effective learning mechanisms (practice structure, feedback, interaction), (ii) contextual constraints (access, infrastructure, teacher workload), and (iii) governance-relevant issues (privacy, transparency, assessment integrity).
3.Strength-of-evidence statements were phrased cautiously, distinguishing consistent patterns from conditional or weakly supported claims.

Methodological limitations

Two limitations were acknowledged. First, heterogeneity of outcome measures and intervention designs limits quantitative pooling and motivates narrative synthesis. Second, single-reviewer screening may introduce selection bias; mitigation steps (predefined criteria, pilot screening, and consistency re-checks) reduce but do not eliminate this risk.

Where opportunities are most plausible and where evidence is strongest

People-focused SDGs: health, education, and poverty (SDGs 1–4)

1.Health (SDG3)

The most policy-relevant health literature frames AI as supportive of (i) surveillance and outbreak detection, (ii) triage and risk stratification, (iii) clinical decision support, and (iv) health planning. However, it also emphasizes that potential benefits in low-resource environments are constrained by data quality, infrastructure, and governance capacity, and that deployment risks increase when systems are imported without local validation or accountability arrangements (Wahl et al., 2018; Schwalbe & Wahl, 2020).A key empirical lesson is that even technically strong systems can reproduce structural inequities if they optimize proxies that reflect unequal access to services rather than underlying need. For example, Obermeyer et al. (2019) shows how a widely used population-health algorithm embedded systematic racial bias through design choices about targets and proxies, illustrating that technical performance does not guarantee equitable outcomes (Obermeyer et al., 2019).

2.Education(SDG4).

In education, the literature documents rapid growth in applications such as personalized learning, automated assessment, learning analytics, and administrative decision support. Yet systematic reviews repeatedly highlight a gap between the pace of innovation and the rigor of evaluation, particularly regarding learning outcomes, transparency, educator involvement, and fairness (Zawacki-Richter et al., 2019; Mustafa et al., 2024). Taken together, these reviews suggest that “AI in education” often advances faster as an innovation agenda than as a validated improvement agenda, with governance and pedagogical integration remaining uneven across contexts (Zawacki-Richter et al., 2019; Mustafa et al., 2024).

3.Poverty and social protection (SDG 1).

Among the more robust empirical strands in “AI for development” are poverty mapping and targeting. Jean et al. (2016) demonstrates how satellite imagery combined with machine learning can estimate poverty proxies at scale, and Steele et al. (2017) extends related approaches using mobile phone and satellite data to map poverty patterns (Jean et al., 2016; Steele et al., 2017). More directly linked to policy delivery, Aiken et al. (2022) evaluates the use of mobile phone data and machine learning for targeting humanitarian aid, making inclusion/exclusion trade-offs and targeting errors more explicit for policy interpretation (Aiken et al., 2022).

4.Synthesis for SDGs 1–4.

Across health, education, and social protection, opportunity is most credible when AI is deployed as decision support with meaningful human oversight, when validation is local and continuous, and when governance arrangements explicitly manage error, appeals, and accountability. This emphasis is consistent with broader governance guidance that prioritizes contestability and remedy in high-impact settings (NIST, 2023; UNESCO, 2021).

5.Climate, environment, and energy (SDGs 7, 11–13)

A major stream emphasizes AI’s potential contribution through optimization and forecasting in domains such as smart grids, demand management, renewable integration, and climate risk/disaster response. Rolnick et al. (2019) maps high-impact use cases while stressing that results depend on domain collaboration, problem formulation, and evaluation—implicitly cautioning against “technical fix” narratives (Rolnick et al., 2019).

Evidence is relatively more persuasive in bounded operational contexts (e.g., energy optimization) where objectives can be specified, feedback is measurable, and institutional capacity for implementation is higher. The literature is more cautious when claims move from operational efficiency to broad welfare outcomes, because the pathway from optimization to well-being depends on affordability, access, regulation, and infrastructure—conditions that vary sharply across settings (UNCTAD, 2025).

6.Institutions and public services (SDG 16 and cross-cutting)

In public administration and SDG 16–adjacent domains, AI is frequently presented as a means to reduce backlogs, improve targeting, detect fraud, enhance monitoring, and support policy analysis. Yet the governance literature is explicit that high-impact public-sector AI raises risks of opacity, contestability failures, and weak remedy—particularly when AI becomes embedded in routine decision pipelines where individuals have limited practical capacity to challenge outcomes (NIST, 2023; UNESCO, 2021)This is where the "weight of evidence" distinction is especially important: much of the SDG 16 discussion is driven by governance frameworks rather than welfare-identified causal studies. Risk-management approaches such as the NIST AI Risk Management Framework provide implementable lifecycle structures—mapping, measuring, managing, and governing risk—while emphasizing that real control depends on organizational capability and accountability design (NIST, 2023).

Challenges and trade-offs that recur across SDG domains

Inequality, capability gaps, and "AI-driven divergence"

A recurring development-policy warning is that AI may widen gaps between countries and within societies because compute, data, talent, and regulatory capacity are unevenly distributed. UNCTAD argues that AI diffusion may outpace many governments' ability to respond and calls for "inclusive AI" strategies aligned with development goals and distributional considerations (UNCTAD, 2025). A complementary UNDP analysis similarly warns that unequal readiness and uneven adoption may contribute to a "Next Great Divergence," with potentially widening inequality across countries if policy does not actively steer AI toward convergence (Schellekens, 2025)Interpretation. Inequality is not merely a "side effect" but a plausible outcome when AI is deployed under asymmetric infrastructure and bargaining power—especially where public agencies rely on external vendors without strong procurement standards, auditing capacity, or local data governance (UNCTAD, 2025; NIST, 2023).

Rights, accountability, and contestability in high-impact systems

Across sectors, the literature repeatedly identifies governance failure modes: opaque models, weak explanation, limited appeal mechanisms, and unclear liability when harm occurs. In SDG-critical domains (health, welfare targeting, migration, policing, credit), these failures can undermine "leave no one behind" commitments by making error correction difficult for those already disadvantaged (UNESCO, 2021; United Nations, 2024b).

The ethics-and-governance literature documents convergence around principles such as transparency, fairness, responsibility, and privacy, while warning about "principle proliferation" without implementation. Jobin et al. (2019) maps the global landscape of AI ethics guidelines and shows recurring principles alongside divergence in interpretation and enforceability (Jobin et al., 2019). Subsequent governance work increasingly focuses on operationalization—impact assessments, audits, and risk-management practices—yet evidence on "what works" remains limited because many interventions are not evaluated using transparent outcome metrics and meaningful remedy measures (NIST, 2023).

4.3.3 Environmental externalities: energy use and carbon impacts

A notable evolution since 2018 is the growing attention to AI's environmental footprint. Strubell et al. (2019) foreground energy and policy considerations in large-scale model training; Patterson et al. (2021) shows that emissions can vary dramatically depending on model choice, data-center efficiency, and electricity mix; and de Vries (2023) emphasizes the growing electricity demand associated with AI and data centers (Strubell et al., 2019; Patterson et al., 2021; de Vries, 2023).

Interpretation. SDG-aligned AI requires more than "AI applied to climate": it also requires lifecycle accounting and incentives for efficiency, reporting, and responsible scaling. Without such measures, AI deployment can inadvertently intensify sustainability pressures rather than mitigate them (Patterson et al., 2021; de Vries, 2023).

Governance implications: from ethics to enforceable controls and capacity-building

Across 2016–early 2026, governance trajectories move from ethics statements to operational tools and, increasingly, binding regulation.

(1) Normative anchors and international guidance.

UNESCO's Recommendation on the Ethics of AI places human rights, inclusion, and environmental considerations at the center of governance, aligning explicitly with development priorities (UNESCO, 2021). The OECD AI Recommendation was revised in late 2023 to update its definition of an "AI system," reflecting the need for governance instruments to remain technically current amid rapid technological change (OECD, 2019/2023).

(2) Operational risk governance inside organizations.

Frameworks such as the NIST AI Risk Management Framework encourage lifecycle risk management (govern, map, measure, manage), translating principles into implementable controls. However, effectiveness depends on institutional capacity, auditing practice, procurement standards, and practical remedy pathways (NIST, 2023).

(3) Risk-based legal models and enforceability.

The EU Artificial Intelligence Act represents a significant shift toward enforceable requirements organized by risk categories and anticipated harms, moving beyond "soft" governance toward compliance regimes with differentiated obligations by use case (European Union, 2024).

(4) Global coordination and equity-oriented governance.

The UN High-level Advisory Body on AI frames governance as a global public interest problem, emphasizing benefit-sharing, capacity development, and cross-border coordination to address both risk and inequality (United Nations, 2024).

(5) ractical implication for SDG-oriented governance.

The most credible governance agenda emerging from the literature is multi-layered—combining (i) rights-based principles, (ii) lifecycle risk management and audits, (iii) procurement and vendor accountability, (iv) contestability and remedy, and (v) capacity-building and international coordination for low-capacity settings (UNESCO, 2021; NIST, 2023; United Nations, 2024).

Consolidated interpretation: what the evidence supports (and what it does not)

Taken together, the 2016–early 2026 literature supports three grounded conclusions. First, AI can contribute to SDG implementation—particularly where it improves operational efficiency, measurement, and decision support in bounded, well-instrumented systems with sustained institutional capacity (Rolnick et al., 2019; Jean et al., 2016). Second, risks are not peripheral: inequality, bias, opacity, and environmental externalities recur across domains and can undermine SDG commitments if governance is weak or purely symbolic (UNCTAD, 2025; Jobin et al., 2019; de Vries, 2023). Third, governance is not an "add-on" but a determinant of impact: the same technical capability can yield welfare gains or harm depending on institutional context, safeguards, and the distribution of power and resources (NIST, 2023; UNESCO, 2021; Vinuesa et al., 2020)These findings position the paper's final analytical task as translating "conditional impact" into actionable governance implications—identifying which instruments (risk frameworks, impact assessments, procurement rules, audits, legal obligations, and capacity investments) best support SDG-oriented, equity-sensitive AI deployment (European Union, 2024; United Nations, 2024).

Governance and Policy Implications: A Practical Agenda for SDG-Aligned AI (2016–early 2026)

Governance logic: from "AI capability" to accountable socio-technical systems

Across the reviewed evidence, AI's SDG relevance is best treated as a socio-technical question rather than a model-performance question. Reported benefits and harms are repeatedly linked to (i) data quality/representativeness, (ii) institutional capacity and incentives, (iii) procurement and vendor dependence, (iv) oversight, contestability, and remedy, and (v) how risks and benefits distribute across groups (UNESCO, 2021; NIST, 2023; United Nations, 2024)Because this study is a systematic review (not an impact evaluation), the most defensible policy contribution is to synthesize what the literature repeatedly implies about deployment conditions. Accordingly, this section is organized around three review-appropriate questions: (1) Where is the SDG mechanism plausible? (2) Where is evidence comparatively stronger? (3) Which safeguards recur as necessary under high stakes?

Recurrently proposed safeguards in high-impact SDG settings

Ex ante impact assessment and risk classification

Where AI affects access to essential services or rights (e.g., health, education, welfare eligibility, policing), governance sources converge on pre-deployment assessment and risk-proportionate controls (documentation, testing, oversight, monitoring, disclosure) as baseline requirements (NIST, 2023; UNESCO, 2021). Public-sector practice includes structured tools such as Algorithmic Impact Assessments (AIA) to standardize impact classification and mitigation planning (Government of Canada, n.d.)SDG implication. Assessments should explicitly test SDG-relevant harms (exclusion/denial risks, subgroup disparities, error costs) and specify minimum oversight and remedy that match the deployment context.

Auditability, transparency, contestability, and remedy

A recurring governance failure mode in SDG-critical applications is weak contestability: affected people cannot understand, challenge, or correct adverse outcomes. The literature therefore treats transparency as insufficient unless paired with operational pathways for appeal, human review, and effective remedy (UNESCO, 2021; NIST, 2023).

SDG implication. Even “accurate” systems can undermine SDGs if errors systematically disadvantage specific groups and cannot be practically corrected.

Procurement and vendor accountability as leverage points

Across public-sector deployments, procurement is repeatedly presented as a practical “control surface” because contracts determine whether agencies can validate locally, audit, monitor drift, and assign responsibility for harms (NIST, 2023).

SDG implication. SDG-aligned contracts should require: local validation before scaling; audit access and documentation; monitoring/update obligations; incident response procedures; and clear allocation of responsibility and liability (including data governance constraints such as purpose limitation and retention).

Environmental reporting and efficiency incentives

The sustainability evidence base emphasizes that SDG-aligned AI must internalize energy and carbon impacts. Emissions vary substantially by model choice, infrastructure efficiency, and electricity mix, motivating governance around reporting assumptions and efficiency-oriented design (Patterson et al., 2021; de Vries, 2023).

SDG implication. “AI for climate” is incomplete without lifecycle reporting boundaries/assumptions and incentives to choose efficient models when outcomes are comparable.

Domain-sensitive governance emphases (where the review evidence is most policy-usable)

1.Health (SDG 3): safety, local validation, and equity checks

Health-relevant evidence stresses that benefits in low-resource environments are conditional on data quality, infrastructure, and governance capacity, and that risks rise when systems transfer without local validation and accountability (Wahl et al., 2018; Schwalbe & Wahl, 2020). Evidence of proxy-driven bias reinforces the need for subgroup performance monitoring and explicit equity checks (Obermeyer et al., 2019)Governance emphasis. Prioritize local validation, continuous monitoring, and accountability for harm over one-time approval; report subgroup impacts and error costs in real workflows.

2.Education (SDG 4): educator agency, transparency, and outcome-relevant evaluation

Review-level education evidence repeatedly identifies an evaluation gap: tool adoption and innovation outpace rigorous evidence on learning outcomes, fairness, and transparency, with governance and educator involvement uneven across contexts (Zawacki-Richter et al., 2019; Mustafa et al., 2024).

3.Governance emphasis.

Require evidence of learning benefit (not only engagement), protect educator agency for high-stakes decisions, and ensure contestability for automated assessment/analytics that affect students.

4.Social protection and poverty (SDG 1): accountability for inclusion/exclusion errors

Empirical work on poverty mapping and targeting shows promise, but policy relevance depends on how errors translate into material consequences and whether affected people can challenge decisions (Jean et al., 2016; Steele et al., 2017; Aiken et al., 2022).

5. Governance emphasis. Treat false exclusions, appeal pathways, subgroup error auditing, and clear accountability for adverse outcomes as core design and deployment requirements.

Multi-level governance alignment: norms, organizational controls, and legal enforcement

The review supports a layered model in which instruments play complementary roles:

- **Normative anchors:** UNESCO centers AI governance on human rights, inclusion, and sustainability (UNESCO, 2021).
- **Cross-national reference points:** OECD governance instruments have been updated to remain technically current, including clarification of the AI-system definition (OECD, 2019; OECD, 2024).
- **Organizational controls:** NIST’s AI RMF provides implementable lifecycle risk structures (NIST, 2023).
- **Binding regulation:** the **EU AI Act** codifies enforceable, risk-based obligations for specific system categories (European Union, 2024).

• **Global coordination:** UN governance work frames AI governance as a global public-interest problem emphasizing capacity development and benefit-sharing (United Nations, 2024).

SDG implication. SDG-aligned governance is most credible when it combines rights-based commitments, operational risk management/auditability, and enforceable obligations for high-impact deployments—supported by capacity investment in low-resource settings.

Focused research agenda (gaps emerging from the reviewed evidence)

To strengthen the evidence base in SDG contexts, the review points to four priorities:

1.**Measurable SDG linkage:** connect interventions to SDG indicators or defensible proxies with transparent reporting of uncertainty.

2.**Governance effectiveness:** evaluate impact assessments, audits, and procurement rules using outcome metrics (subgroup error rates, appeal rates, remediation timelines).

3.**Distributional impacts:** quantify within-country inequality effects and second-order consequences (access barriers, labor displacement, service quality shifts).

4.**Environmental comparability:** standardize footprint reporting assumptions to improve comparability across studies and deployments (Patterson et al., 2021; de Vries, 2023).

Recommendations (SDG-Aligned AI Governance)

Based on the reviewed 2016–early 2026 evidence, the following recommendations translate "conditional impact" into a practical, governance-first agenda. They are framed as implementable safeguards consistent with the paper's evidence-typology coding (conceptual/empirical/policy) rather than as causal claims about SDG outcomes.

Risk-proportionate governance as a baseline

• **Adopt risk classification for SDG-relevant AI systems.** Classify systems by potential harm (especially where rights or access to essential services are affected) and require proportionate controls for higher-risk deployments (NIST, 2023; European Union, 2024).

• **Require ex ante impact assessment before deployment.** Use a structured assessment to document purpose, affected groups, data suitability, error costs, and mitigation measures; in public sector settings, algorithmic impact assessment tools provide a practical template (Government of Canada, 2025; NIST, 2023).

Accountability, contestability, and remedy in high-impact settings

• **Make contestability a design requirement.** Ensure clear pathways for explanation, human review, appeal, and remedy for adverse decisions, particularly in welfare targeting, education analytics, and health decision support (UNESCO, 2021; NIST, 2023).

• **Clarify responsibility for harm across the system.** Define accountability across agencies, vendors, and operators (roles, escalation, incident response, and remediation timelines), aligning with risk-based obligations where applicable (European Union, 2024; NIST, 2023).

Data, validation, and continuous monitoring

• **Mandate local validation and ongoing performance monitoring.** Validate performance and subgroup impacts in the local context before scaling; monitor drift and update processes after deployment (NIST, 2023).

• **Audit subgroup impacts and error costs.** Report performance beyond averages (e.g., false exclusions in social protection, mis-triage in health), and treat distributional harm as a core evaluation dimension, not a secondary consideration (UNCTAD, 2025; UNESCO, 2021).

Procurement and vendor governance as leverage points

• **Use procurement to enforce governance requirements.** Contracts should require documentation, audit access, local testing, monitoring obligations, disclosure of limitations, and clear allocation of liability/incident response (NIST, 2023).

• **Reduce vendor lock-in and strengthen public capacity.** Build in-house capability (data governance, auditing, evaluation) so oversight is not outsourced to vendors, especially in low-capacity settings (UNCTAD, 2025; United Nations, 2024).

Environmental accountability for SDG consistency

• **Require lifecycle-oriented reporting of energy/carbon assumptions.** Standardize reporting boundaries (model choice, data-center efficiency, electricity mix) and treat efficiency as a governance objective where outcomes are comparable (Patterson et al., 2021; de Vries, 2023).

• **Align "AI for climate" with "climate-responsible AI."** Avoid SDG trade-offs where AI deployments increase energy demand without accountability; integrate footprint constraints into design, procurement, and scaling decisions (de Vries, 2023; UNESCO, 2021).

Capacity-building and international coordination

• **Invest in enabling conditions for equitable impact.** Prioritize data infrastructure, institutional oversight capacity, and workforce skills to reduce "AI-driven divergence" risks (UNCTAD, 2025; Schellekens, 2025).

• **Support interoperability of governance approaches.** Use shared references (UNESCO principles; OECD guidance; NIST lifecycle risk governance) while adapting to national legal contexts, and coordinate on cross-border risks and benefit-sharing (UNESCO, 2021; OECD, 2019/2023; United Nations, 2024b).

References

[1] E. Aiken, S. Bellue, D. Karlan, C. Udry, and J. Blumenstock, "Machine learning and mobile phone data can improve the targeting of humanitarian aid," Nature, vol. 603, no. 7901, pp. 864–870, 2022.

[2] A. Alkandari, "Students learning in the time of artificial intelligence: University students' perceptions of using AI tools to improve language learning," Educational Process: International Journal, vol. 15, Art. no. e2025154, 2025, doi: 10.22521/edupij.2025.15.154.

[3] M. Alnufaie, "Mobile-assisted language learning applications: Features and characteristics from users' perspectives," Studies in Self-Access Learning Journal, vol. 13, no. 3, pp. 312–331, 2022, doi: 10.37237/130302.

[4] S. H. H. Barzani and F. R. Omar, "WhatsApp-mediated vocabulary learning: A study of Kurdish EFL learners," Arab World English Journal, vol. 15, no. 4, pp. 364–376, 2024, doi: 10.24093/awej/vol15no4.23.

[5] E. Bensalem, "The impact of WhatsApp on EFL students' vocabulary learning," Arab World English Journal, vol. 9, no. 1, pp. 23–38, 2018, doi: 10.24093/awej/vol9no1.2.

[6] I. Boufahja, "WhatsApp-enhanced intentional vocabulary learning: An intervention study," Cogent Education, vol. 12, no. 1, Art. no. 2549791, 2025, doi: 10.1080/2331186X.2025.2549791.

[7] J. Burston, "Twenty years of MALL project implementation: A meta-analysis of learning outcomes," ReCALL, vol. 27, no. 1, pp. 4–20, 2015, doi: 10.1017/S0958344014000159.

[8] DataReportal, "Digital 2024: Global overview report," We Are Social & Meltwater, 2024.

[9] A. de Vries, "The growing energy footprint of artificial intelligence," Joule, vol. 7, no. 10, pp. 2191–2194, 2023, doi: 10.1016/j.joule.2023.09.004.

[10] European Union, "Regulation (EU) 2024/1689 … laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)," Official Journal of the European Union, L, 2024/1689, 2024.

[11] Government of Canada, "Algorithmic impact assessment (AIA)," [Online]. Available: https://canada-ca.github.io/aia-eia-js/. [Accessed: Jan. 2026].

[12] Y. Hao, K. S. Lee, S. T. Chen, and C. C. Sim, "An evaluative study of a mobile application for middle school students struggling with English vocabulary learning," Computers in Human Behavior, vol. 95, pp. 208–216, 2019, doi: 10.1016/j.chb.2018.10.013.

[13] M. K. Hasan, A. H. A. Fakih, P. M. I. Seraj, and Hasmirati, "The effect of technology-assisted language programme on vocabulary learning among EFL students at the tertiary level," Heliyon, vol. 8, no. 8, Art. no. e10313, 2022, doi: 10.1016/j.heliyon.2022.e10313.

[14] F. Heidari and M. Rajabi, "The effect of an artificial intelligence chatbot on vocabulary retention among Iranian EFL learners," Computer Assisted Language Learning Electronic Journal, vol. 25, no. 1, pp. 45–65, 2024.

[15] Q. N. Hong et al., "The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers," Education for Information, vol. 34, no. 4, pp. 285–291, 2018.

[16] S. Jafari and A. Chalak, "The role of WhatsApp in teaching vocabulary to Iranian EFL learners at junior high school," English Language Teaching, vol. 9, no. 8, pp. 85–92, 2016, doi: 10.5539/elt.v9n8p85.

[17] N. Jean et al., "Combining satellite imagery and machine learning to predict poverty," Science, vol. 353, no. 6301, pp. 790–794, 2016, doi: 10.1126/science.aaf7894.

[18] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," Nature Machine Intelligence, vol. 1, pp. 389–399, 2019.

[19] Joanna Briggs Institute, "Critical appraisal tools," [Online]. Available: https://jbi.global/critical-appraisal-tools.

[20] S. Khatoon, N. Ahmed, and M. Khan, "WhatsApp as a digital tool for vocabulary development among Pakistani EFL learners," Journal of Language and Education, vol. 9, no. 3, pp. 112–129, 2023.

[21] A. Kukulska-Hulme, "Mobile language learning," in The Encyclopedia of Applied Linguistics, C. A. Chapelle, Ed. Wiley-Blackwell, 2012, pp. 3701–3709, doi: 10.1002/9781405198431.wbeal0765.

[22] C. Lei, J. Fathi, S. N. Noorbakhsh, and M. Rahimi, "The impact of mobile-assisted language learning on EFL learners' vocabulary learning attitudes and self-regulatory capacity," Frontiers in Psychology, vol. 13, Art. no. 872922, 2022, doi: 10.3389/fpsyg.2022.872922.

[23] J. Li, "Learning vocabulary via computer-assisted scaffolding for text processing," Computer Assisted Language Learning, vol. 23, no. 3, pp. 253–275, 2010, doi: 10.1080/09588221.2010.483678.
[24] F. V. Lim and N. Taguchi, "Apps for English language learning: A systematic review of smartphone-assisted learning," Journal of Computer Assisted Learning, vol. 40, no. 2, pp. 350–372, 2024, doi: 10.1111/jcal.13096.
[25] A. Metwally, "EFL learners' perceptions of generative AI in academic writing: Opportunities and challenges," Journal of English as a Foreign Language, vol. 15, no. 1, pp. 55–78, 2025.
[26] A. H. Mohammed, "The use of mobile phones and social media in English language learning among Iraqi university students," International Journal of Innovation, Creativity and Change, vol. 11, no. 10, pp. 457–473, 2020.
[27] M. Y. Mustafa et al., "Artificial intelligence and education for sustainable development: A systematic review," British Journal of Educational Technology, vol. 55, no. 6, pp. 2325–2355, 2024.
[28] I. S. P. Nation, Learning Vocabulary in Another Language. Cambridge University Press, 2001, doi: 10.1017/CBO9781139524759.
[29] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce, 2023, doi: 10.6028/NIST.AI.100-1.
[30] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," Science, vol. 366, no. 6464, pp. 447–453, 2019.
[31] OECD, "Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449)," Organisation for Economic Co-operation and Development, 2019.
[32] OECD, Explanatory memorandum on the updated OECD definition of an AI system. OECD Publishing, 2024.
[33] K. Okumuş Dağdeler, "A systematic review of vocabulary learning with mobile-assisted learning platforms," Journal of Education and Human Development, vol. 11, no. 1, pp. 1–15, 2022.
[34] N. Oktadela, N. Elida, and H. Ismail, "Improving English vocabulary through artificial intelligence (AI) chatbot application," Journal of English Language Education, vol. 8, no. 2, pp. 63–70, 2023.
[35] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," BMJ, vol. 372, p. n71, 2021, doi: 10.1136/bmj.n71.
[36] C. D. T. Paredes, N. Lutfiyah, and L. Ramírez, "The impact of effective vocabulary learning strategies in developing Saudi university students' listening comprehension," Journal of Developmental Education, vol. 12, no. 2, pp. 45–62, 2024.
[37] D. Patterson et al., "Carbon emissions and large neural network training," arXiv preprint arXiv:2104.10350, 2021. [Online]. Available: https://arxiv.org/abs/2104.10350
[38] Y. Qian, "User review analysis of mobile English vocabulary learning applications," Journal of Educational Technology and Developmental Education, vol. 15, no. 2, pp. 1–20, 2022.
[39] R. A. Salman and A. Imam, "The impact of WhatsApp application on vocabulary acquisition of secondary school students in Nigeria," International Journal of Research in Teacher Education, vol. 12, no. 3, pp. 1–12, 2021.
[40] N. Schmitt and D. Schmitt, Vocabulary in Language Teaching, 2nd ed. Cambridge University Press, 2020.
[41] N. Schwalbe and B. Wahl, "Artificial intelligence and the future of global health," The Lancet, vol. 395, no. 10236, pp. 1579–1586, 2020.
[42] Statista Research Department, "Most popular global mobile messenger apps as of 2024, based on number of monthly active users," Statista, 2024.
[43] J. E. Steele et al., "Mapping poverty using mobile phone and satellite data," Journal of The Royal Society Interface, vol. 14, no. 127, Art. no. 20160690, 2017.
[44] E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in NLP," in Proc. ACL 2019, 2019.
[45] J. Thomas and A. Harden, "Methods for the thematic synthesis of qualitative research in systematic reviews," BMC Medical Research Methodology, vol. 8, Art. no. 45, 2008, doi: 10.1186/1471-2288-8-45.
[46] J. Tyndall, "AACODS checklist," Flinders University, 2010.
[47] UNESCO, Recommendation on the Ethics of Artificial Intelligence. UNESCO, 2021.
[48] United Nations, Governing AI for Humanity: Final report of the United Nations High-Level Advisory Body on Artificial Intelligence. United Nations, 2024.
[49] B. Wahl, A. Cossy-Gantner, S. Germann, and N. Schwalbe, "Artificial intelligence (AI) and global health: How can AI contribute to health in resource-poor settings?" BMJ Global Health, vol. 3, no. 4, Art. no. e000798, 2018.
[50] M. Zaim, S. Arsyad, and B. Waluyo, "Pre-service EFL teachers' intentions to adopt generative AI in language teaching: An extended TAM perspective," Computers & Education: Artificial Intelligence, vol. 7, Art. no. 100312, 2024, doi: 10.1016/j.caeai.2024.100312.
[51] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education: Where are the educators?" International Journal of Educational Technology in Higher Education, vol. 16, Art. no. 39, 2019.