Journal of Artificial Intelligence and Technology, 2026, 6, 686-695
https://doi.org/10.37965/jait.2026.1252

ISTP
RESEARCH ARTICLE

Artificial Intelligence Chatbots for Inclusive Pedagogical Practices: A PRISMA Systematic Review and Bibliometric Mapping (2021–2025)

Paula Andrea Estrada Palencia, Karen Patricia Agudelo Arteaga, and Tobías Alfonso Parodi Camaño
Universidad de Córdoba, Montería, Colombia

(Received 10 March 2026; Revised 16 May 2026; Accepted 05 June 2026; Published online 08 July 2026)

***Abstract**:* The integration of artificial intelligence (AI) in education has accelerated, yet evidence on how AI chatbots support inclusive pedagogical practices remains fragmented. This study synthesizes recent literature on the contributions of AI chatbots to inclusive education, emphasizing accessibility, equity, and pedagogical adaptation. We conducted a systematic review guided by PRISMA 2020. A sensitive search in November 2025 retrieved 338 records from Scopus ($n = 59$) and Dimensions ($n = 279$). After removing duplicates ($n = 13$) and applying eligibility criteria (2021–2025; open access; relevance to inclusive education and AI chatbots), 11 studies were included for in-depth analysis. In parallel, bibliometric mapping was conducted using VOSviewer to describe publication trends, keyword co-occurrence, collaboration networks, and citation/co-citation structures. Findings suggest three contributions of AI chatbots to inclusive pedagogical practices: (i) personalization and instructional adaptation, where chatbots operate as intelligent tutors and scaffold learning pathways for diverse learners; (ii) functional accessibility, including speech-to-text and text-to-speech affordances that reduce participation barriers for students with sensory or communication needs; and (iii) equity-oriented support, in low-resource contexts where chatbots may partially compensate for limited instructional capacity. However, the evidence base is constrained by small samples, short-term implementations, and heterogeneous outcomes, limiting generalizability. A persistent gap concerns support for students with reduced mobility, alongside ethical risks related to privacy, academic integrity, dependency, and potentially perpetuating medicalized views of disability. We conclude that AI chatbots can strengthen inclusive education when implemented under principles of equity, transparency, and pedagogical ethics and when aligned with Universal Design for Learning and reasonable accommodation (reasonable adjustments).

***Keywords**:* artificial intelligence; bibliometric mapping; chatbots; inclusive education; PRISMA 2020

I. INTRODUCTION

Artificial intelligence (AI) chatbots are being adopted in higher education, but their real contribution to inclusive teaching is still scattered across disciplines and study designs. We systematically reviewed studies published from 2021 to 2025, following PRISMA 2020 guidance. From an initial corpus of 338 records, 11 studies met the inclusion criteria for qualitative synthesis. Overall, the evidence suggests that AI chatbots can support inclusion mainly by personalizing explanations and feedback, providing accessible interaction channels, and assisting teachers with scaffolding and monitoring learning.

At the same time, the literature remains limited in scope, with uneven coverage of disability profiles, learning contexts, and evaluation outcomes. We highlight methodological gaps and propose priorities for more rigorous, equity-focused research and implementation. Methodologically, the studies included in this review comprise systematic reviews, scoping reviews, meta-analyses, adoption studies, implementation-focused research, and contextual qualitative investigations. Most studies adopted qualitative descriptive, analytical, or mixed-method approaches, while others emphasized systematic evidence synthesis.

The methodologies frequently relied on narrative thematic analysis, perception-based evaluations, contextual interpretation, and exploratory implementations of AI-supported educational interventions. Overall, findings consistently indicate that AI chatbots may strengthen inclusive educational practices by facilitating adaptive learning pathways, accessible communication channels, and scaffolded instructional support. However, researchers also report concerns related to privacy, bias, academic integrity, technological dependency, and unequal access to digital infrastructure.

Despite the rapid growth of research in this field, significant gaps remain regarding long-term effectiveness, participation-oriented outcomes, and the implementation of AI chatbots for underrepresented learner populations, particularly students with mobility-related needs. Therefore, this systematic review addresses the following research question: What contributions of AI chatbots can be evidenced in inclusive pedagogical practices? The objectives are to (1) identify and synthesize the main contributions of AI chatbots to inclusion, accessibility, and equity, (2) describe research trends through bibliometric mapping, and (3) analyze the theoretical, methodological, pedagogical, and ethical gaps reported in the literature.

The remainder of this article is organized as follows. Section II presents the Materials and Methods, including the PRISMA 2020 protocol, search strategy, eligibility criteria, and bibliometric procedures. Section III presents the Results, including descriptive findings, bibliometric analyses, and thematic synthesis of the 11 included studies. Section IV discusses the pedagogical implications, scaffolding contributions, accessibility challenges, ethical considerations, and research gaps associated with AI chatbot integration in inclusive education. Finally, Section V presents the conclusions, highlighting the potential of AI chatbots to strengthen inclusive

Corresponding author: Paula Andrea Estrada Palencia (e-mail: pestrada@correo.unicordoba.edu.co).

pedagogical practices when implemented through equity-oriented, accessible, and ethically grounded educational frameworks.

II. LITERATURE REVIEW

Inclusive education is widely recognized as a priority for educational quality and equity, as reflected in Sustainable Development Goal 4 (SDG 4) and the Education 2030 Framework for Action, which advocate for inclusive and equitable quality education for all learners [1,2]. In this study, inclusive pedagogical practices are understood as educational approaches aimed at removing barriers to learning and participation, promoting accessibility, and ensuring meaningful engagement for diverse learners within mainstream educational settings [3]. These practices are closely associated with Universal Design for Learning (UDL), which recommends multiple means of engagement, representation, and action/expression in order to address learner variability from the outset [4]. In cases where barriers persist, reasonable accommodations are also necessary to guarantee equitable participation and educational access [5]. Contemporary perspectives on disability further emphasize participation, contextual adaptation, and rights-based approaches instead of deficit-oriented interpretations [6,7].

Parallel to these developments, AI has emerged as a transformative force in education. Among AI applications, chatbots have gained increasing attention due to their conversational capabilities, scalability, and potential to provide personalized learning support. Educational institutions and researchers have increasingly explored how AI chatbots can enhance teaching, accessibility, and learner engagement while supporting inclusive educational environments. UNESCO's ICT Competency Framework for Teachers also highlights the importance of digital competencies and ethical technology integration in inclusive educational practices [8].

Recent literature demonstrates growing interest in the relationship between AI chatbots and inclusive education. For instance, Ali D *et al.* [9] conducted a systematic review showing that ChatGPT can support tutoring, feedback, and adaptive explanations for diverse learners when pedagogically mediated. Alyoussef I *et al.* [10] examined factors influencing inclusive AI adoption in higher education and found that institutional support, collaboration, and perceived usefulness strongly shape implementation processes. Coughlan T *et al.* [11] analyzed disability-related student narratives and suggested that AI systems may help identify and organize support strategies, although ethical concerns regarding oversimplification remain. Beyond empirical studies, policy frameworks also shape how these tools are adopted: Colombia's Decree 1421 (2017), for instance, establishes regulatory expectations for inclusive education and disability support [12].

Similarly, Dizon J and Prudente M [13] demonstrated through a meta-analysis that AI-powered chatbots can improve achievement when integrated as scaffolding, feedback, and guided-practice tools. Complementing these classroom-focused findings, emerging studies propose teacher professional-development approaches specifically oriented toward inclusive AI integration [14]. Hamzah R *et al.* [15] proposed an autism-focused individualized education chatbot model designed to support communication and developmental profiles. In a related line of work, culturally tailored health-promotion chatbots have also been examined, illustrating how accessibility considerations can be embedded for diverse and underserved populations [16]. In low-resource educational contexts, Kouam A and Muchowe R [17] argued that AI technologies may help mitigate educational inequities by extending learning support where institutional resources are limited.

Other studies further reinforce the potential of AI for accessibility and inclusion. Jaime-Vargas J [18] conducted a systematic review indicating that generative AI can enhance accessibility in higher education through personalized explanations and alternative content formats. Lin T and Riccomini P [19] explored technology-enabled inclusive learning in rural mathematics education, highlighting improved participation opportunities for students with disabilities. In addition, competency frameworks addressing educators' digital skills and ICT use offer a further lens for guiding responsible implementation [20]. Pagliara S *et al.* [21] synthesized evidence showing that AI integration can support adaptive learning, accessibility tools, and learner-centered assistance in inclusive environments. Likewise, Pergantis P *et al.* [23] found that AI chatbots may contribute to executive-function scaffolding and cognitive regulation processes, while Potapiuk L and Sasiuk A [24] emphasized the importance of ethical safeguards, teacher preparation, and accessibility adaptation when implementing AI in inclusive educational systems.

Collectively, these studies suggest that AI chatbots can contribute to inclusive pedagogical practices through three major dimensions: personalization and instructional adaptation, functional accessibility, and equity-oriented educational support. In particular, scaffolding emerges as a recurring pedagogical mechanism, since chatbots can provide guided explanations, adaptive feedback, executive-function support, and differentiated learning assistance for students with diverse educational needs. Nevertheless, the evidence base remains fragmented due to methodological heterogeneity, limited longitudinal evaluation, and uneven representation of disability profiles and educational contexts.

III. MATERIALS AND METHODS

A. REPORTING STANDARDS AND PROTOCOL

This systematic review was conducted and reported in accordance with PRISMA 2020 [22]. The PRISMA flow diagram is provided (Fig. 1), and the PRISMA 2020 [22] checklist is provided as Supplementary Material S3 (PRISMA 2020 Checklist). The annual distribution of publications in the initial corpus is shown in Fig. 2.

Protocol registration: The review protocol was not prospectively registered.

B. STUDY DESIGN

We conducted a documentary systematic review guided by PRISMA 2020, complemented by bibliometric analyses in VOSviewer to characterize the research landscape.

C. DATA SOURCES AND SEARCH STRATEGY

A sensitive search was performed in November 2025 using Scopus and Dimensions. The initial retrieval yielded 338 records (Scopus, $n = 59$; Dimensions, $n = 279$). Eligibility screening then applied the review's temporal criterion (2021–2025) and an open-access filter (as defined by each database), and records outside the time window were removed prior to screening (see Fig. 1). To preserve reproducibility, both databases were queried using the same conceptual core ("inclusive education" AND "artificial intelligence chatbots"), after which each platform's native filters were used to delimit publication years, document types, and subject categories. In Dimensions, the search incorporated Education-related categories together with Artificial Intelligence, Machine Learning, Psychology, and adjacent fields, and restricted results to open-access articles, chapters, and edited-book records. In Scopus, the search was executed in TITLE-

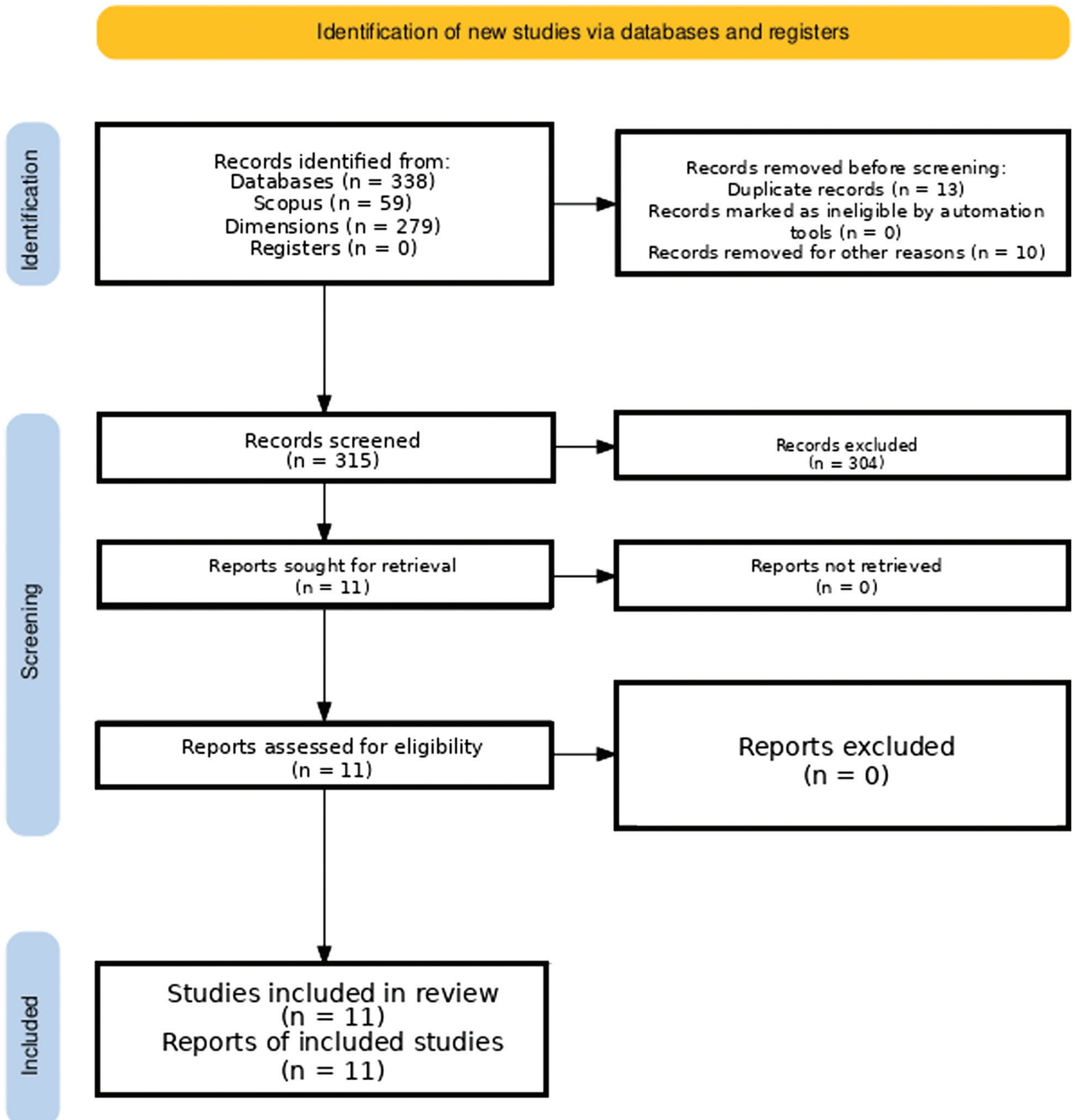

Fig 1. PRISMA 2020 flow diagram for the identification, screening, eligibility assessment, and inclusion of studies in the systematic review ($n = 11$).

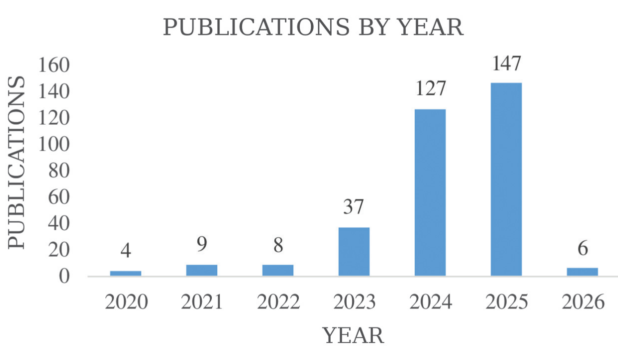

Fig 2. Annual distribution of publications in the initial corpus ($n = 338$); records outside 2021–2025 were removed prior to screening (see Fig. 1).

ABS-KEY and limited to 2021–2025 open-access records from relevant subject areas (e.g., social sciences, computer science, engineering, psychology, chemistry, and chemical engineering) and selected document types. This broader "landscape" search served two complementary purposes: first, to identify studies for PRISMA screening; and second, to assemble a corpus suitable for bibliometric mapping of research growth, keywords, and collaboration patterns. The exact search strings, filters, export decisions, and database-specific limits are reproduced in Supplementary Material S1 so that readers can distinguish the broad bibliometric corpus from the smaller set retained for qualitative synthesis.

D. ELIGIBILITY CRITERIA

Inclusion criteria covered (i) empirical or review studies addressing AI chatbots in education, (ii) explicit relevance to inclusive education, accessibility, or equity, (iii) publication types including articles, books, or book chapters (as returned by the databases), and (iv) open-access availability. Exclusion criteria removed non-educational AI development papers, opinion pieces without evidence, and inclusive education studies without reference to AI chatbots or closely related technologies.

E. QUALITY APPRAISAL AND RISK-OF-BIAS CONSIDERATIONS

To strengthen methodological transparency, the included studies ($n = 11$) were appraised using a structured methodological coherence checklist aligned with the review's research question and inclusion focus. The checklist was derived from the screening questions already defined by the review team and covered the following domains: (1) clarity and alignment of the title/aim with the topic; (2) study context and author affiliation; (3) publication type and relevance; (4) adequacy and credibility of sources used; (5) explicitness of conceptual/theoretical framework; (6) sample characteristics and appropriateness; (7) data collection techniques and instruments; (8) internal methodological coherence among design, sampling, instruments, and analysis; (9) claims of generalizability and supporting evidence; and (10) relevance and sufficiency of conclusions in relation to the review objective and inclusive education implications.

Each domain was coded qualitatively (Yes/Partially/No) to identify potential threats to validity and interpretive limitations rather than to exclude studies a priori. Given the heterogeneity of designs and outcomes across the evidence base, no single quantitative risk-of-bias score was calculated. Instead, appraisal outcomes informed the narrative synthesis by indicating where conclusions should be treated with caution (e.g., small samples, short-term implementations, and limited outcome measures).

F. SCREENING AND SELECTION PROCESS (PRISMA FLOW)

From 338 records, 13 duplicates were removed in Zotero, leaving 325 unique records. Records were then restricted to the 2021–2025 window, excluding 10 records ($n = 315$). Title and abstract screening focused on whether each record explicitly connected AI chatbots or closely related conversational AI systems with education, inclusion, accessibility, disability support, or equity. Full-text eligibility checks then verified the educational context, the relevance of the chatbot component, and the presence of evidence or analytical substance beyond opinion-based commentary. After this staged process, 11 studies met the inclusion criteria for rigorous analysis. This distinction is important: the initial corpus ($n = 338$) was used to map the broader research landscape, whereas the final PRISMA-included set ($n = 11$) was used to synthesize evidence on inclusive pedagogical practices.

G. DATA EXTRACTION AND SYNTHESIS

For each included study, we extracted educational level/context, target populations, inclusion/accessibility focus, chatbot function, study design, sample characteristics, instruments, outcomes, and reported limitations. Extraction was organized around a common coding sheet to reduce interpretive dispersion across heterogeneous study designs (Table I). The synthesis followed a narrative thematic approach centered on inclusion-related contributions (personalization, accessibility, and equity), alongside reported gaps and ethical challenges. In addition, we coded whether the chatbot was described as a general-purpose conversational tool or as a more specialized assistive or advisory system, because this distinction shaped both the type of pedagogical contribution claimed and the nature of the limitations reported by authors. We also noted whether the study foregrounded learner-facing uses, teacher-facing uses, or institutional/administrative support functions, which helped clarify the level at which "inclusion" was being operationalized.

H. BIBLIOMETRIC ANALYSIS

All RIS files were imported into Zotero and then processed in VOSviewer to generate (i) keyword frequency and co-occurrence networks, (ii) co-authorship by country, (iii) author citation networks, and (iv) co-citation maps. Before visualization, terms were standardized where necessary to reduce duplication caused by singular/plural variants, translation differences, and near-synonymous labels. The VOSviewer maps were interpreted descriptively: node size indicates relative occurrence or citation strength, link thickness indicates the strength of connections, and clusters suggest topical or intellectual proximity. These bibliometric outputs were not treated as evidence of effectiveness; rather, they contextualized the growth and structure of the field. Accordingly, the maps are presented as a landscape complement to the PRISMA synthesis, helping distinguish between a broad publication surge and the much smaller set of studies that explicitly address inclusion, accessibility, or equity in a way that met the review criteria.

The full search strings and database filters for Scopus and Dimensions are provided in Supplementary Material S1.

Table I. Technical coding scheme used to characterize AI chatbot approaches and implementation attributes

Category	Coding and operationalization
(1) Study metadata & context	Year; venue/source; country/region; educational level/discipline; study type (design/implementation, empirical evaluation, review).
(2) Inclusion & accessibility scope	Target learners/needs (e.g., disability/NEE); barriers addressed; accessibility features (TTS/STT, captions, simplified language, multimodal supports, assistive compatibility).
(3) Use case & interaction	Primary purpose (tutoring, practice/feedback, Q&A, guidance/support); modality (text/voice/multimodal); platform (web/mobile/LMS/messaging); personalization/adaptation (if reported).
(4) AI approach & architecture	Model family (rule/retrieval/classical ML/neural/transformer/LLM/hybrid); grounding (KB/ontology, RAG); tool use/agents; prompt or context strategies (if reported).
(5) Data & resources	Data sources (public/proprietary; curriculum/KB); dataset size/labeling (if reported); privacy handling (anonymization/consent) when applicable.
(6) Evaluation design & metrics	Design (experimental/quasi/observational/qualitative/mixed); baseline/comparator; sample (n), duration, setting (lab/field); metrics (learning outcomes, engagement logs, usability/satisfaction, accessibility-related measures, qualitative themes).
(7) Dependability, safety & governance	Reported/assessed risks (hallucinations, bias, privacy leakage, academic integrity, robustness/jailbreak, transparency); mitigations (human oversight, guardrails/filtering, data minimization, secure storage, auditing/logging, risk-based evaluation); key findings, limitations, and future directions mapped to gaps.

IV. RESULTS

A. DESCRIPTIVE RESULTS AND RESEARCH GROWTH

Across the initial corpus ($n = 338$), publication activity shows a steep recent increase, with peaks in 2024 and 2025 (as reported in the dataset trend analysis) (Fig. 2). This pattern suggests that AI chatbots and related conversational AI tools have become a rapidly expanding topic in educational research, although the final set of 11 included studies shows that only a small subset of this growth explicitly examines accessibility, equity, or inclusive pedagogical practice. Therefore, Fig. 2 should be read as evidence of field expansion, not as evidence that robust inclusion-focused interventions are already consolidated.

The distribution also helps explain the small number of studies retained for qualitative synthesis: much of the recent literature discusses AI in education broadly, while fewer publications meet the narrower criteria of chatbot relevance plus explicit inclusion, accessibility, or equity focus.

B. BIBLIOMETRIC HIGHLIGHTS (KEYWORDS AND NETWORKS)

Within the PRISMA-included studies ($n = 11$), the most frequent keywords include AI, chatbots, and inclusive education, suggesting a shared conceptual core around AI-mediated inclusion. Additional terms such as disability, accessibility, students' perceptions, higher education, ChatGPT, and adaptive technologies show how the corpus connects technical tools with participation barriers, learner diversity, and pedagogical adaptation. Keyword frequency is summarized in Fig. 3.

This visualization clarifies that the review is not centered on chatbots as generic automation tools but on their relationship with inclusive education, disability support, accessibility, and adaptive learning. The presence of "students' perceptions" also indicates that part of the evidence base remains perception-oriented, which reinforces the need for stronger outcome-based studies.

Network-based bibliometric visualizations were generated from the initial corpus ($n = 338$): author citation (Fig. 4), author co-citation (Fig. 5), country co-authorship (Supplementary Fig. S1), and keyword co-occurrence (Supplementary Fig. S2). In these maps, larger nodes indicate authors or terms with stronger presence in the corpus, while links indicate citation, co-citation, collaboration, or keyword association depending on the map. The networks are therefore useful for locating visible authors, intellectual proximity, and emerging conceptual clusters, but they should not be interpreted as direct evidence that a given chatbot intervention improves inclusion.

Taken together, Figs. 4 and 5 show that the broader field is still organizing around dispersed author groups rather than a single consolidated research tradition. This supports the need for the qualitative synthesis in Table II, which narrows the interpretation to the studies that directly address inclusive pedagogical practices.

C. CHARACTERISTICS OF INCLUDED STUDIES

The 11 studies comprise a balanced mix of secondary evidence syntheses (systematic/scoping reviews and a meta-analysis) and primary empirical work, plus one implementation-focused contextual study. This composition is important because it reveals a field that is conceptually active but still empirically consolidating: several papers synthesize emerging evidence or propose interpretive frameworks, while fewer studies evaluate sustained intervention effects under clearly defined inclusive conditions. Geographically, the evidence spans 10countries (two studies from the USA and one each from the Philippines, Saudi Arabia, Italy, Malaysia, Zimbabwe, Greece, Mexico, the UK, and Ukraine), suggesting broad international interest but also a fragmented knowledge base shaped by very different educational systems and resource conditions. Most studies are situated in higher education or inclusive schooling contexts and examine chatbot-mediated personalization, communication support, or equity-oriented assistance in resource-constrained settings. Across the sample, chatbots are described either as general-purpose LLM-based tools (e.g., ChatGPT) used for learning support and feedback, or as task-specific systems tailored to particular needs (e.g., autism support and executive-function coaching). Methodologically, the sample also reflects uneven maturity: some studies are broad reviews synthesizing patterns across settings [9,18,21,23], some are adoption

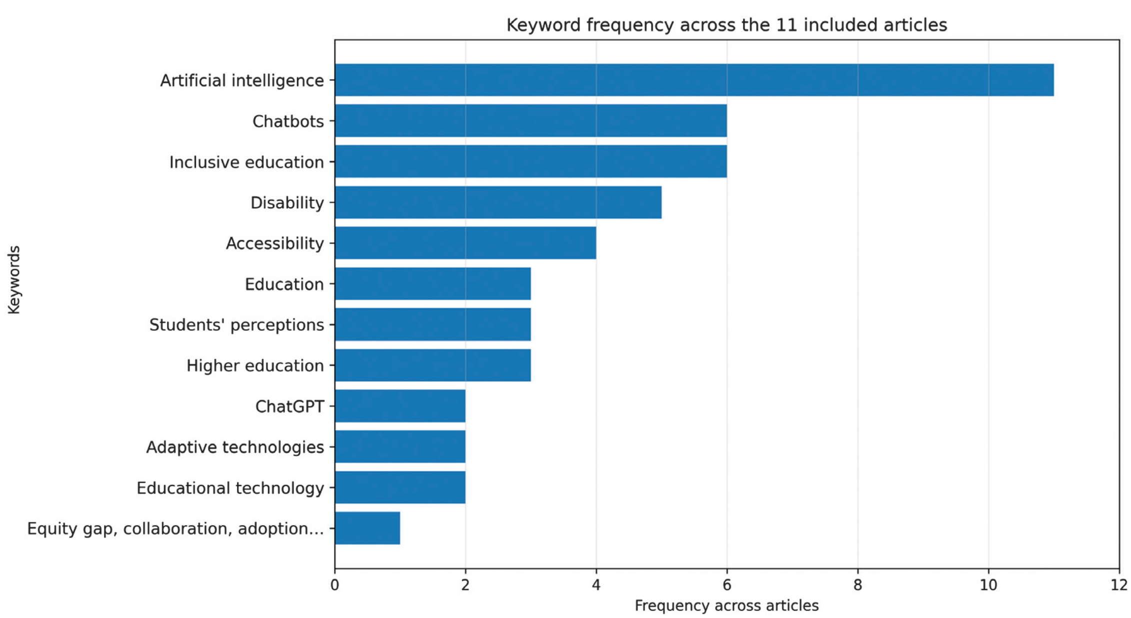

Fig 3. Keyword frequency across the 11 included studies ($n = 11$). Author keywords were extracted from each included paper and standardized (translation where needed, plus singular/plural harmonization) for consistency.

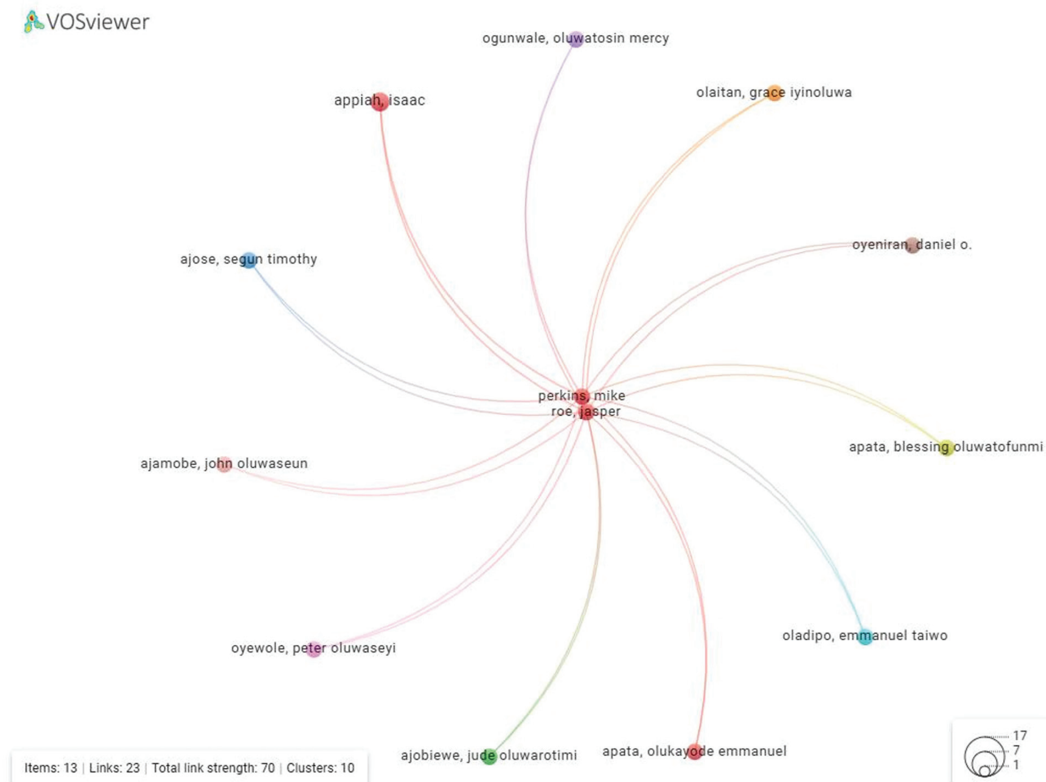

Fig 4. Author citation network (VOSviewer) derived from the initial corpus ($n = 338$).

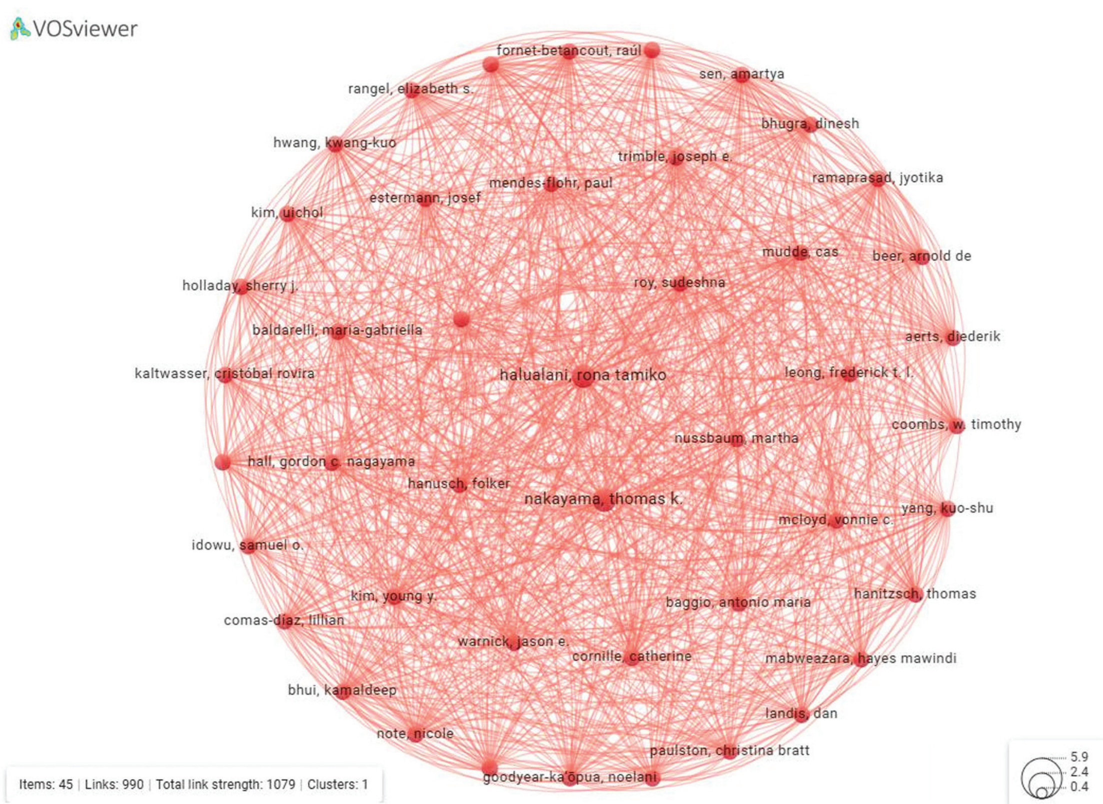

Fig 5. Author co-citation network (VOSviewer) derived from the initial corpus ($n = 338$).

Table II. Summary of key findings from the 11 studies included in the qualitative synthesis

Study	Type/context	Main inclusion-related finding	Caution or gap
Ali *et al.* [9]	Systematic review of ChatGPT in teaching and learning	ChatGPT can support explanations, feedback, tutoring, and flexible learning interactions that may assist diverse learners when pedagogically mediated.	Evidence is heterogeneous; inclusive value depends on instructional design, teacher guidance, and safeguards.
Alyoussef *et al.* [10]	Higher education adoption study	Inclusive learning adoption is shaped by collaboration, perceived usefulness, institutional support, and technology acceptance factors.	Adoption readiness does not by itself demonstrate improved participation or accessibility outcomes.
Coughlan *et al.* [11]	Analysis of disability descriptions and student suggestions	Learner-generated descriptions help identify barriers and inform support strategies that AI systems could help organize or route.	AI-based interpretation of disability narratives may oversimplify needs or reproduce deficit-based assumptions if not governed carefully.
Dizon and Prudente [13]	Meta-analysis of AI-powered chatbots in science education	Chatbots can improve achievement when they are pedagogically integrated as scaffolding, practice, or feedback tools.	Most effects are not specific to disability or inclusion; transfer to inclusive education requires targeted evaluation.
Hamzah *et al.* [15]	Autism-focused individualized education chatbot model	Specialized chatbot designs can be tailored to communication and developmental profiles, supporting individualized educational interaction.	Prototype and modeling emphasis; classroom validity, usability, and long-term impact require further testing.
Kouam and Muchowe [17]	Equity-focused contextual study in Zimbabwe	AI technologies may help reduce access gaps by extending learning support where institutional resources are limited.	Infrastructure, connectivity, cost, language, and digital competence can reproduce inequity if not addressed.
Jaime-Vargas [18]	Systematic review on generative AI, accessibility, and inclusion in higher education	Generative AI can support accessible content, alternative explanations, and personalized learning assistance in higher education.	Risks include privacy, bias, academic integrity, over-reliance, and uneven access to premium tools.
Lin and Riccomini [19]	Technology-enabled inclusive learning in rural math education	Technology can expand access and support improved math learning outcomes for students with disabilities in rural contexts.	Findings depend on local infrastructure, teacher mediation, and contextual scalability.
Pagliara *et al.* [21]	Scoping review of AI integration in inclusive education	AI can contribute to inclusive education through adaptive supports, accessibility tools, and learner-centered assistance.	The evidence remains fragmented across tools, populations, and outcome definitions.
Pergantis *et al.* [23]	Systematic review on AI chatbots and cognitive control	Chatbot interaction may support executive-function scaffolding, cognitive regulation, and guided learning processes.	Cognitive benefits require rigorous measurement and careful attention to dependency and autonomy.
Potapiuk and Sasiuk [24]	Implementation-focused study in Ukrainian inclusive education	AI is positioned as an advanced learning tool for inclusive educational environments and differentiated support.	Local readiness, teacher preparation, ethical safeguards, and accessibility adaptation remain decisive conditions.

or implementation studies focused on attitudes and enabling conditions [10,17,24], and others are practice-oriented or prototype-oriented contributions exploring how chatbot functions might be adapted for specific learners or instructional problems [11,15,19]. Table II summarizes the key inclusion-related findings and cautions from each included study, while full study-level characteristics are reported in Supplementary Material S2.

The table shows that the most consistent evidence concerns personalization, accessibility affordances, and equity-oriented support, while the most persistent weaknesses concern short-term designs, limited accessibility metrics, and uneven coverage of disability profiles. These patterns inform the thematic synthesis that follows.

D. THEMATIC SYNTHESIS: THREE INCLUSION-CENTERED CONTRIBUTIONS

1). PERSONALIZATION AND PEDAGOGICAL ADAPTATION. AI chatbots are consistently described as supporting individualized instruction and adaptive learning pathways, including intelligent tutoring roles and cognitive scaffolding for students facing learning/participation barriers. Across the included studies, personalization operates in at least three ways. First, chatbots can reformulate explanations, generate examples, and provide immediate feedback at different levels of complexity, which may help learners who require repeated clarification or alternative representations of content [9,13,18]. Second, they can support teachers in differentiating instruction by suggesting materials, structuring practice opportunities, or drafting accessible prompts aligned with learner needs [19]. Third, in more specialized designs, conversational agents can be configured to respond to particular communication or developmental profiles, as illustrated by autism-focused individualized education proposals [15]. Taken together, the literature suggests that the main pedagogical value of chatbots is not merely speed or automation, but their potential to make instructional adaptation more continuous, responsive, and feasible within ordinary teaching workflows.

2). FUNCTIONAL ACCESSIBILITY (OPERATIONAL AND COMMUNICATIONAL). Evidence highlights accessibility gains through assistive affordances such as text-to-speech (TTS) and speech-to-text (STT) for learners with sensory, speech, or communication limitations, contributing to participation and autonomy. Several studies frame accessibility not only as the presence of technical features but also as the availability of alternative modes of interaction that reduce friction in academic tasks and communication [18,21,24]. In that sense, chatbots can serve as mediators between learners and instructional content by offering conversational prompts, language simplification, reminders, or supportive dialog that makes participation less dependent on a single format. Teacher- and institution-facing studies also suggest that chatbot systems may reduce administrative burden and guide students toward more appropriate forms of support, although the validity and fairness of such guidance still require careful evaluation [11]. The strongest pattern here is that conversational AI may widen access when it expands communicative options, but only if those options are designed and evaluated as accessibility supports rather than as generic convenience features.

3). EQUITY-ORIENTED SUPPORT IN LOW-RESOURCE CONTEXTS. Chatbots are positioned as compensatory agents where institutional resources are constrained, offering scalable support that may improve engagement and access to learning guidance. This contribution appears most clearly in studies that discuss disadvantaged or under-resourced settings, where access to specialist support, timely feedback, or individualized tutoring is limited [13,17,19]. In such contexts, conversational AI is described as a potentially lower-cost layer of assistance that can partially extend teacher capacity and provide learners with additional opportunities to ask questions, rehearse ideas, or obtain clarification. However, the equity promise is explicitly conditional. The same literature warns that premium access models, infrastructure gaps, language barriers, and uneven digital competence can reproduce exclusion rather than alleviate it [17]. Thus, equity-oriented value does not derive from the chatbot itself, but from whether institutions can integrate it in ways that are affordable, accessible, and pedagogically meaningful for populations facing structural disadvantage.

E. REPORTED LIMITATIONS AND GAPS

Across the 11 included studies, the evidence base remains methodologically constrained. Most contributions rely on small or convenience samples, brief implementation windows, and self-reported outcomes, with limited use of validated accessibility measures or participation-oriented indicators. These constraints reduce robustness, limit causal inference, and make it difficult to determine whether reported benefits translate into sustained learning gains and durable participation. Coverage of inclusion needs is also uneven: studies most frequently address sensory and communication-related barriers, while mobility-related needs and intersectional barriers (e.g., language and socioeconomic constraints) are rarely examined. Heterogeneity is visible at several levels: the reviewed papers differ in educational stage, chatbot architecture, target population, outcome definitions, and even in what counts as "inclusive" use. Some focus on adoption perceptions, others on learning outcomes, and others on conceptual framing or prototype development, which complicates cross-study comparison and weakens the basis for strong generalizations. Future work would benefit from aligning outcome definitions and reporting with participation-oriented disability frameworks and rights-based inclusion constructs, rather than treating disability primarily as an individual deficit [6,7]. It would also benefit from clearer reporting on implementation conditions—such as teacher mediation, access infrastructure, language settings, and prior digital competence—because these contextual variables likely influence whether a chatbot functions as an inclusive support or as an additional layer of exclusion.

For this reason, the synthesis avoids treating the presence of a chatbot as equivalent to inclusive impact. Instead, the interpretation emphasizes the pedagogical mechanism, the accessibility barrier addressed, the learner population considered, and the implementation conditions reported by each study.

V. DISCUSSION

This review indicates that AI chatbots can contribute to inclusive pedagogical practices primarily by enabling (i) personalization, (ii) accessibility through assistive modalities, and (iii) equity-oriented support under resource constraints. Importantly, these contributions should not be interpreted as "automatic inclusion." The same technologies can create new barriers if implementation ignores UDL principles, contextual constraints, or the lived realities of disability and exclusion.

A. PEDAGOGICAL IMPLICATIONS (INCLUSION-FIRST INTEGRATION)

From an inclusion-first perspective, chatbots should be framed as supports for participation and learning rather than as substitutes for instruction. Aligning chatbot use with UDL principles (multiple means of engagement, representation, and expression) helps ensure that design choices accommodate learner variability and reduce the need for reactive, individualized fixes [4]. However, the review also reinforces that technology does not deliver inclusion automatically: inclusion depends on pedagogical intent, barrier removal, and ongoing adaptation in context—consistent with inclusive pedagogy frameworks that prioritize participation and avoid separating "most" from "some" learners [3,6]. Accordingly, implementation should specify inclusion objectives (e.g., participation, autonomy, and accessibility), define who benefits and under which conditions, and combine chatbot use with teacher mediation and accessible learning resources. A practical implication is that educators and institutions should begin with barrier analysis rather than tool enthusiasm: before adopting a chatbot, they should clarify which participation barrier is being addressed, what alternative modality or scaffold is required, how human oversight will be maintained, and how benefits will be evaluated for different learner groups. Under this logic, chatbots become one component within a broader inclusive design strategy, not the centerpiece of it.

B. ETHICAL AND PRACTICAL RISKS

The reviewed literature flags risks that are directly relevant to the trustworthy and responsible deployment of AI-driven chatbots: academic integrity concerns (including plagiarism), privacy and data protection challenges, potential over-reliance that may undermine learner autonomy, and bias or inaccuracies in chatbot outputs. These issues become especially sensitive in inclusive education settings, where learners may be more exposed to harm from misinformation, intrusive data practices, or stigmatizing recommendations. Related reviews also highlight cognitive and executive-function considerations when chatbots are used as learning supports [23], and context-specific work in inclusive environments underscores the need to align AI tools with local accessibility realities [24]. Framing these risks within established AI ethics and governance frameworks can strengthen both interpretation and recommendations, linking concerns about bias, privacy, transparency, and integrity to actionable principles such as human oversight, accountability, data minimization, and risk-based evaluation [25–27]. An additional concern emerging from the included literature is representational harm: when systems infer needs from disability descriptions, user prompts, or interaction histories, they may oversimplify complex support requirements or encode narrow assumptions about what counts as appropriate participation [11,21]. For inclusive education, this means that technical performance alone is insufficient; governance must also address fairness of interpretation, contestability of recommendations, and the possibility of unintended stigmatization.

C. RESEARCH AGENDA

Future work should move beyond describing "use" of chatbots and instead test how, when, and for whom chatbots contribute to inclusion. This requires treating inclusion as barrier removal and participation (not merely usability) and explicitly linking chatbot functions to inclusive pedagogical mechanisms (e.g., UDL-aligned multiple means of engagement/representation/action). The evidence reviewed here suggests that methodological progress will depend not only on larger samples but also on sharper conceptualization: studies need to identify what kind of chatbot is under examination, which barrier it is intended to address, which population is expected to benefit, and what inclusion-sensitive outcomes will indicate success. To strengthen the evidence base, future studies should prioritize:

- Longer-term and larger-scale evaluations (multi-site when possible) with clearly defined inclusion outcomes and accessibility metrics, including participation, autonomy, and equity indicators rather than satisfaction-only measures. In particular, studies should distinguish short-term usability gains from sustained pedagogical effects and should report implementation conditions in enough detail to support replication.
- Targeted research for underexplored populations, especially learners facing mobility-related barriers, and broader disability profiles beyond sensory/communication needs, with attention to intersectional constraints (e.g., disability plus socioeconomic or language barriers). Comparative work across rural/urban, low-resource/high-resource, and different linguistic settings would be especially valuable for testing the equity claims often made about conversational AI.
- More explicit theoretical integration (UDL plus participation-oriented disability and inclusive pedagogy frameworks) to avoid defaulting to deficit-based or medicalized interpretations of disability, and to operationalize inclusion as barrier removal and participation [3–7]. Studies should state the adopted framework, specify hypothesized mechanisms (e.g., scaffolding, reduced cognitive load, alternative modality access), and align measures accordingly. This would also help prevent the common slippage between "accessible," "personalized," and "inclusive" terms that are related but not interchangeable.
- Governance models for safe deployment (transparency, human oversight, data minimization, and integrity safeguards) grounded in established AI ethics and risk management frameworks [25–27]. At minimum, future studies should report data handling, bias mitigation steps, and human-in-the-loop oversight practices to enable reproducibility and comparability. Research should also consider institutional readiness, teacher professional development, and procurement/access models, because implementation quality may be as consequential as technical capability for inclusive outcomes.

D. LIMITATIONS OF THIS REVIEW

This review is limited by database scope (Scopus and Dimensions), open-access filtering, and the small number of eligible studies ($n = 11$), which restricts the ability to generalize across educational systems and disability groups. The open-access restriction may have introduced selection bias by underrepresenting relevant evidence not indexed as OA. In addition, heterogeneity in study designs, populations, and chatbot architectures limits direct comparability across studies and constrains the feasibility of quantitative synthesis. Finally, bibliometric maps reflect the initial corpus ($n = 338$) and should be interpreted as a landscape view rather than as evidence of effectiveness, which is based on the PRISMA-included set ($n = 11$). A further limitation is that some included studies examined broader AI-enabled inclusion ecosystems rather than chatbot interventions in isolation; although these studies remained relevant to the review question, their broader scope may blur boundaries between chatbot-specific effects and more general digital-support dynamics. This was handled through narrative interpretation, but it should be considered when reading the synthesis.

V. CONCLUSION

AI chatbots show promise for inclusive education by supporting personalized instruction, functional accessibility (e.g., TTS/STT), and equity-oriented assistance in low-resource contexts. However, the current evidence base is still small and methodologically heterogeneous, and benefits should not be assumed to generalize across learners, contexts, or time. A key implication of this review is that "inclusion" should be treated as a theory-driven construct: chatbot adoption is most defensible when embedded within inclusive pedagogical designs—anchored in UDL and reasonable accommodation—and when outcomes are evaluated in terms of participation, autonomy, and accessibility rather than usability alone [4–7]. Finally, responsible scaling requires governance that addresses bias, privacy, transparency, and academic integrity through risk-based approaches and explicit human oversight [25–27]. In practical terms, the literature reviewed here suggests that the question is no longer whether conversational AI can enter inclusive education, but under what pedagogical, ethical, and institutional conditions it can do so without reproducing the very barriers it claims to reduce.

ACKNOWLEDGMENTS

The authors have no acknowledgments to declare.

FUNDING

The authors received no specific funding for this work.

DATA AVAILABILITY STATEMENT

All data underlying the findings are available within the manuscript and in the Supplementary Materials (S1–S3) and Supplementary Figures (S1–S2), including the full search strategies and the extracted evidence table.

CONFLICT OF INTEREST STATEMENT

The author(s) declare that they have no conflicts of interest to report regarding the present study.

ETHICAL STATEMENT

Ethical approval was not required because this study is a systematic review of published literature and did not involve human participants, animals, or identifiable personal data.

SUPPLEMENTARY MATERIAL CAPTIONS

Supplementary Material S1. Full database search strings and filters (Scopus and Dimensions).

Supplementary Material S2. Characteristics of the 11 included studies (study type, context/population, AI/chatbot function, inclusion-related contribution, and reported gaps/limitations).

Supplementary Material S3. PRISMA 2020 checklist.

Supplementary Fig. S1. Country co-authorship network (VOSviewer) derived from the initial corpus (n = 338).

Supplementary Fig. S2. Keyword co-occurrence network (VOSviewer) derived from the initial corpus (n = 338). Terms were standardized for consistency (translation and singular/plural harmonization).

REFERENCES

[1] United Nations, "*Transforming our world: The 2030 Agenda for Sustainable Development*," (A/RES/70/1), [Internet]. New York: United Nations, 2015 [cited 2026 Jan 9], Available: https://undocs.org/en/A/RES/70/1

[2] UNESCO, "*Education 2030: Incheon Declaration and Framework for Action for the implementation of Sustainable Development Goal 4*," [Internet]. Paris: UNESCO, 2016 [cited 2026 Jan 9]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000245656

[3] L. Florian and J. Spratt, "Enacting inclusion: A framework for interrogating inclusive practice," *Eur. J. Spec. Needs Educ.*, vol. 28, no. 2, pp. 119–135, 2013, DOI: 10.1080/08856257.2013.778111

[4] CAST, "*Universal Design for Learning Guidelines version 3.0*," [Internet]. Wakefield (MA): CAST, 2024 [cited 2026 Jan 9]. Available: https://udlguidelines.cast.org/

[5] United Nations, "*Convention on the Rights of Persons with Disabilities*," [Internet]. New York: United Nations, 2006, [cited 2026 Jan 9]. Available: https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-persons-disabilities

[6] World Health Organization., "*International classification of functioning, disability and health: ICF*". Geneva: World Health Organization, 2001. Available: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health

[7] M. Oliver, *The Politics of Disablement*, London: Macmillan Education, 1990.

[8] UNESCO, "*Marco de competencias de los docentes en materia de TIC (versión 3)*" [Internet], 2019, [cited 2026 Jan 6]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000371024/PDF/371024spa.pdf.multi

[9] D. Ali *et al.*, "ChatGPT in teaching and learning: A systematic review," *Educ. Sci.*, vol. 14, no. 6, p. 643, 2024, DOI: 10.3390/educsci14060643.

[10] I. Alyoussef *et al.*, "AI adoption for collaboration: Factors influencing inclusive learning adoption in higher education," *IEEE Access*, vol. 13, pp. 81690–81713, 2025, DOI: 10.1109/ACCESS.2025.3567656.

[11] T. Coughlan, F. Iniesto, and J. Carr, "Analysing disability descriptions and student suggestions as a foundation to overcome barriers to learning," *J. Interact. Media Educ.*, vol. 2024, no. 1, p. 4, 2024, DOI: 10.5334/jime.836.

[12] Departamento Administrativo de la Función Pública (CO). "*Decreto 1421 de 2017: Por el cual se reglamenta, en el marco de la educación inclusiva, la atención educativa a la población con discapacidad*," Aug. 29, 2017, [Internet]. [cited 2026 Jan 6]. Available: https://www.funcionpublica.gov.co/eva/gestornormativo/norma_pdf.php?i=87040

[13] J. Dizon and M. Prudente, "AI-powered chatbots in science education: Meta-analysis of pedagogical integration and student achievement," *Int. J. Chem. Educ. Res (IJCER)*, vol. 9, no. 2, pp. 134–150, 2025, DOI: 10.20885/ijcer.vol9.iss2.art4.

[14] P. Estrada, K. Agudelo, and E. Flórez, "Artificial intelligence and inclusive education: A professional development model for chemistry teachers in secondary education," *Educ. Process: Int. J.*, vol. 17, p. e2025380, 2025, DOI: 10.22521/edupij.2025.17.380.

[15] R. Hamzah *et al.*, "Convolutional neural network modelling for autistic individualized education chatbot. IAES," *Int. J. Artif. Intell. (IJ-AI).*, vol. 14, no. 1, pp. 109–118, 2025, DOI: 10.11591/ijai.v14.i1.pp109-118.

[16] M. Kim *et al.*, "A culturally tailored artificial intelligence chatbot (K-Bot) to promote human papillomavirus vaccination among Korean Americans: Development and usability study," *Asian Pac. Island Nurs. J.*, vol. 9, p. e71865, 2025, DOI: 10.2196/71865.

[17] A. Kouam and R. Muchowe, "Assessing the role of AI technology in mitigating the equity gap in educational access in Zimbabwe: Barriers and implications," *J. Appl. Learn. Teach.*, vol. 8, no. 1, pp. 154–163, 2025, DOI: 10.37074/jalt.2025.8.1.15.

[18] J. Jaime-Vargas, "Generative AI for enhancing accessibility and inclusion in higher education: A systematic review," *HighTech Innov. J.*, vol. 6, no. 2, pp. 723–734, 2025, DOI: 10.28991/HIJ-2025-06-02-022.

[19] T. Lin and P. Riccomini, "Inclusive education in rural settings: Leveraging technology for improved math learning outcomes among students with disabilities," *J. Spec. Educ. Technol.*, vol. 40, no. 2, pp. 219–226, 2024, DOI: 10.1177/01626434241263045.

[20] Ministerio de Educación Nacional de Colombia, "*Competencias TIC para el desarrollo profesional docente*" [Internet]. [cited 2026 Jan 6]. 2013 Available: https://www.mineducacion.gov.co/1621/articles-339097_archivo_pdf_competencias_tic.pdf

[21] S. Pagliara *et al.*, "The integration of artificial intelligence in inclusive education: A scoping review," *Information*, vol. 15, no. 12, p. 774, 2024, DOI: 10.3390/info15120774.

[22] M. J. Page *et al.*, "Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas," Rev. Esp. Cardiol., vol. 74, no. 9, pp. 790–799, 2021, DOI: 10.1016/j.recesp.2021.06.016.

[23] P. Pergantis *et al.*, "AI chatbots and cognitive control: Enhancing executive functions through chatbot interactions: A systematic review," *Brain Sci.*, vol. 15, no. 1, p. 47, 2025, DOI: 10.3390/brainsci15010047.

[24] L. Potapiuk and A. Sasiuk, "Integrating artificial intelligence as an advanced learning tool into Ukrainian inclusive educational environment," *Educol. Disc.*, vol. 47, no. 4, pp. 35–43, 2024, DOI: 10.28925/2312-5829.2024.4.4.

[25] UNESCO, "*Recommendation on the Ethics of Artificial Intelligence*," [Internet], Paris: UNESCO; 2021 [cited 2026 Jan 9], Available: https://unesdoc.unesco.org/ark:/48223/pf0000380455

[26] OECD, "*Recommendation of the Council on Artificial Intelligence*," [Internet], Paris: OECD, 2019 [cited 2026 Jan 9], Available: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[27] National Institute of Standards and Technology (NIST), "*Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1*," [Internet], Gaithersburg (MD): NIST, (2023), [cited 2026 Jan 9]. Available: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf