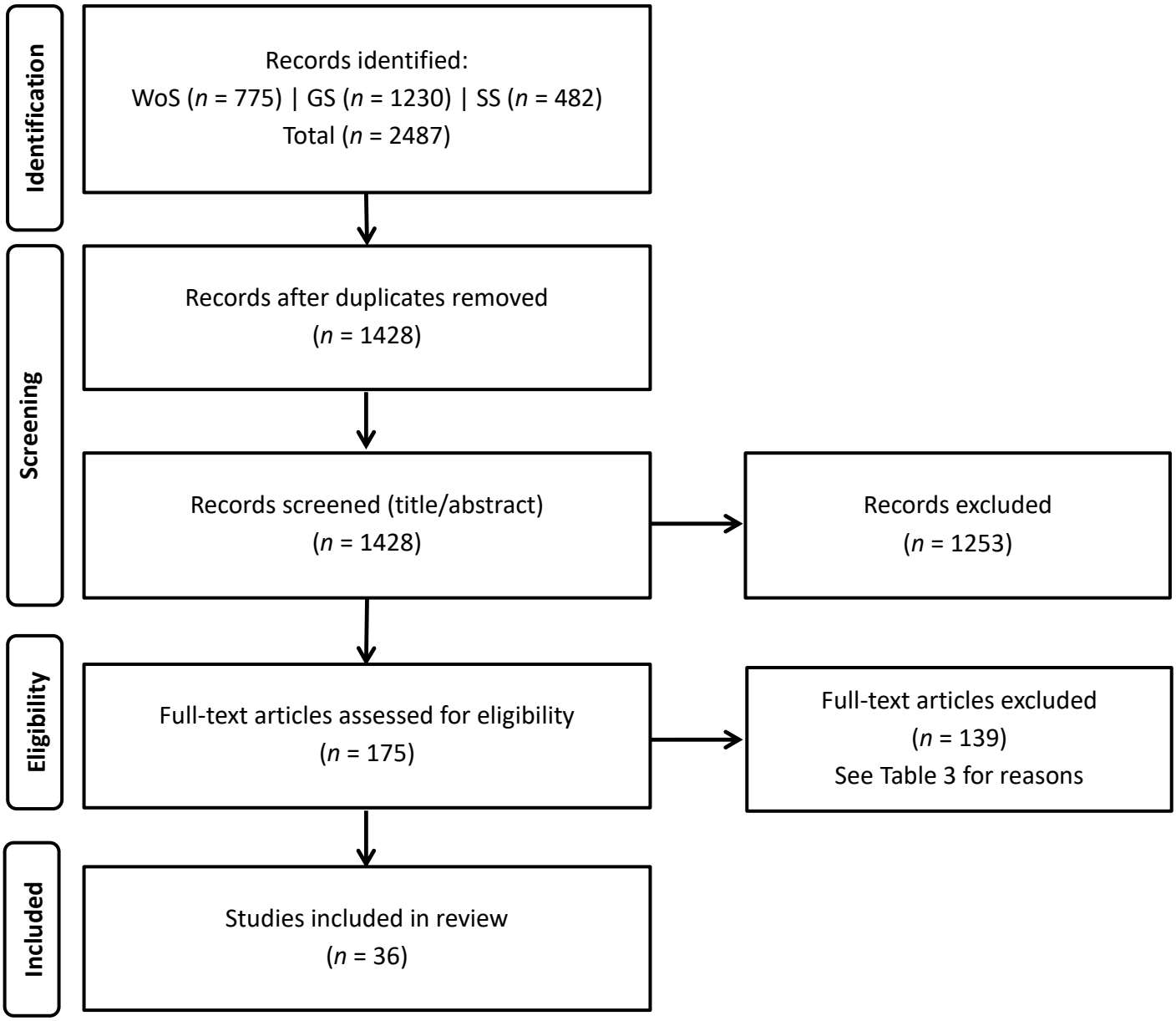
Appendix

Figure 1

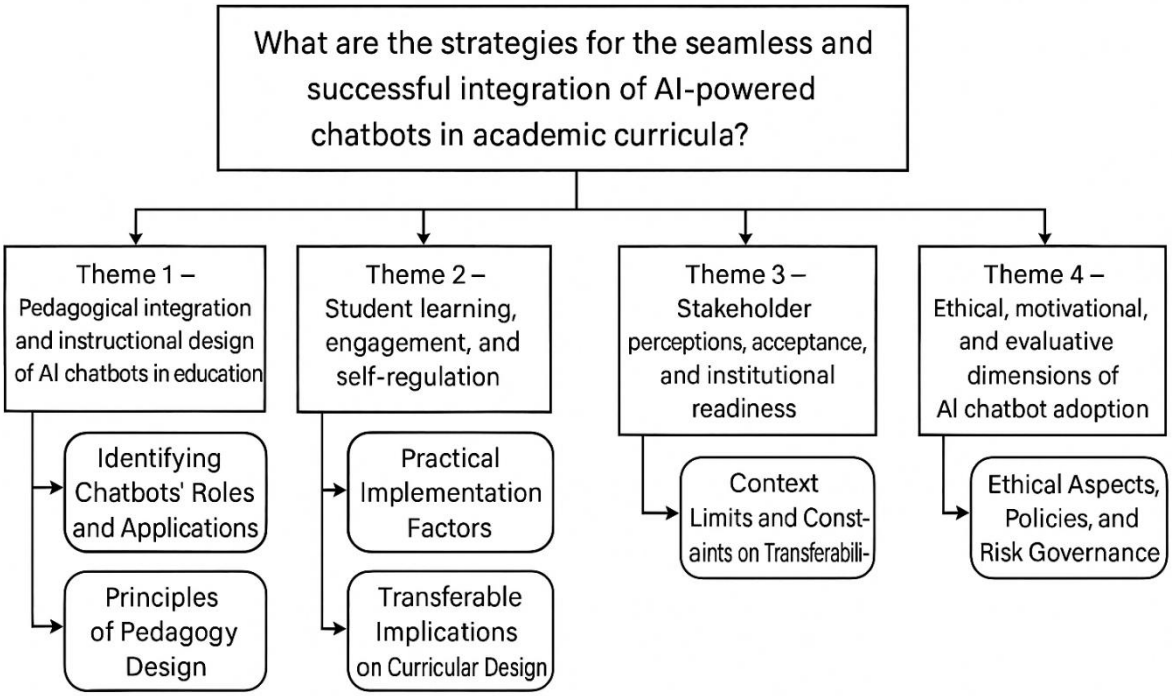
Simplified Search-Flow PRISMA-Style Diagram



Source: authors' own work.

Figure 2

Conceptual Framework for the Integration of AI Chatbots in Education



Source: authors' own work.

Table 1

Boolean Search Queries

Database	Boolean query
Web of Science (Core Collection)	("AI chatbots in education" OR chatbot* OR "AI chatbot*") AND (education OR "higher education" OR educat*) AND (teaching OR learning OR student OR learner OR teacher OR instructor) AND (acceptance OR motivation OR assessment OR evaluation OR perception)
Google Scholar	("AI chatbots in education" AND chatbots AND "AI chatbots") AND (education OR "higher education") OR (teaching OR learning) OR (student OR learner) OR (teacher OR instructor) OR (acceptance OR motivation OR assessment OR evaluation OR perception)
Semantic Scholar	("AI chatbots in education") AND (chatbots OR "AI chatbots") AND (education OR "higher education" OR teaching OR learning)

Source: authors' own work.

Table 2

Inclusion and Exclusion Criteria for the Systematic–Narrative Hybrid Review

Population / Educational Context	Studies conducted in education.	Studies not conducted in non–educational setting.
Intervention / Technology	Studies examining AI chatbots, conversational agent**s** and assistants used for teaching, learning, assessment, feedback, tutoring, advising, or student support in education.	Studies not involving AI chatbots conversational agent**s** and assistants used for teaching, learning, assessment, feedback, tutoring, advising, or student support in education.
Purpose / Focus of Study	Research addressing pedagogy, learning outcomes, engagement, assessment, student support, teaching practices, acceptance, motivation, or user experience related to AI chatbots.	Technical-only papers focusing on algorithms, NLP model development, ML architectures, or chatbot engineering without an educational application.
Study Type	Peer-reviewed journal articles or full conference papers with empirical, conceptual, or theoretical contributions.	Editorials, commentaries, opinion pieces, news items, workshop abstracts, non-peer-reviewed content, or studies lacking scholarly rigor.
Research Contribution	Studies providing empirical data (qualitative, quantitative, mixed methods), conceptual frameworks, or theoretical analyses related to chatbot use in higher education.	Articles lacking sufficient methodological detail or not providing extractable data relevant to the review objectives.
Publication Years	Published between 2021 and 2025.	Published outside this timeframe (if any appear despite filters).
Language	Full text available in English.	Non-English publications.
Accessibility	Full text accessible and complete for review and data extraction.	Full text inaccessible, corrupted, incomplete, or behind paywalls with no metadata available.
Duplicates	Unique studies contributing new data or insights.	Duplicate conference/journal versions or substantially overlapping publications.

Source: authors' own work.

Table 3

Aggregated Reasons for Full-Text Exclusion

Exclusion Reason Category	Description	Count
Not focused on higher education	Chatbots used in K–12, corporate, healthcare, marketing, etc.	42
Not about AI chatbots or conversational agents	Articles focused on other AI tools, general LLMs, or unrelated technology.	26
Technical-only study with no educational application	NLP/ML/algorithm papers without pedagogical context.	23
Irrelevant topic despite keyword match	Chatbots for administration, mental health, customer service, marketing, etc.	18
Insufficient methodological detail / missing data	Full text lacked extractable data or methodological clarity.	12
Duplicate at full-text level	Conference + journal duplicates or near-identical publications.	8
Full text inaccessible or incomplete	Unavailable, corrupted, or incomplete documents.	10
Total excluded after full-text screening		139

Source: authors' own work.

Table 4

Articles Included in Narrative Synthesis Supported by a Structured Coding Procedure

Article title	Reference	Research design	Context	Objectives	Main findings
Online chat and chatbots to enhance mature student engagement in higher education	Abbas et al, 2022	Mixed-methods case study (Exploratory pilot using an online survey, focus groups, a semi-structured interview, and platform usage data)	Lifelong Learning Centre (LLC) at the University of Leeds, UK; focusing on mature students (aged 21+ at undergraduate start) during the 2020/2021 academic year, which was impacted by the COVID-19 pandemic.	To explore the impact of using the online chat platform 'Differ' (including its chatbot 'Bo') and student-led online communities on enhancing mature students' engagement and sense of belonging.	The platform was effective for making initial connections, particularly for mature and part-time students, and increased their sense of belonging. Key to success was the role of trained Digital Student Mentors. Students valued a separate, safe space for university communication, though many later migrated to more familiar apps like WhatsApp for ongoing communication.
Modeling Students' Perceptions of Chatbots in Learning: Integrating Technology Acceptance with the	Al-Abdullatif, 2023	Quantitative study using a cross-sectional online survey. Data analysis performed with Structural Equation Modeling (SEM-PLS)	Saudi higher education; 432 undergraduate and postgraduate students from three universities across various academic	To investigate the factors driving students' acceptance of chatbots for learning by integrating the Technology	Perceived usefulness, attitude, and perceived value were the strongest direct predictors of chatbot acceptance. Perceived enjoyment and perceived ease of use significantly influenced perceived value. Perceived risk

Value-Based Adoption Model			disciplines (e.g., humanities, social sciences, computer science).	Acceptance Model (TAM) with the Value-Based Adoption Model (VAM).	was not a significant predictor of attitude or perceived value in this context. The integrated TAM-VAM model effectively explains students' acceptance.
Implementing the Bashayer chatbot in Saudi higher education: measuring the influence on students' motivation and learning strategies	Al-Abdullatif et al. 2023	Quasi-experimental	60 postgraduate students in a Master of Educational Technologies program at a Saudi university.	To investigate the effect of the Bashayer chatbot (on WhatsApp) on postgraduate students' learning motivation and learning strategies.	Postgraduate students using the Bashayer chatbot showed significantly higher levels of: 1. Motivation (task value and self-efficacy). 2. Learning strategies (cognitive and metacognitive self-regulation).
Implications of AI Chatbots in Education: Challenges and Solution	Awad & Moosa, 2024	Quantitative descriptive (Questionnaire-based survey)	100 respondents (students, faculty, industry) in Bahrain; focuses on ethical implications.	1. Identify the implications of AI Chatbots in education. 2. Investigate societal awareness of AI Chatbot implications. 3. Propose a framework for	- Identifies key ethical challenges (bias, data security, academic misconduct, etc.). - Survey shows divided opinions on ethics, especially for writing reports/dissertations. - Proposes a 4-pillar framework (Regulations, Course Assessments, Technology/Tools, Awareness) to ensure ethical use.

				enhancing AI Chatbots utilization in education.	
Exploring the Possible Use of AI Chatbots in Public Health Education: Feasibility Study	Baglivo et al, 2023	Quantitative feasibility study comparing AI chatbots and medical students on medical multiple-choice questions	University of Pisa, Italy; fifth-year medical students in a "Hygiene and Public Health" course.	To evaluate whether AI chatbots can serve as educational support tools in public health training and assess their performance on complex medical questions.	AI chatbots significantly outperformed medical students on vaccination-related questions and were rated highly by students as effective and engaging educational aids, demonstrating their feasibility as supportive learning tools.
Enough of the chit-chat: A comparative analysis of four AI chatbots for calculus and statistics	Calonge et al., 2023	Comparative analysis evaluating four AI chatbots (ChatGPT, GPT-4, Bard, LLAMA) on 12 calculus and statistics prompts	The study focuses on the fields of mathematics education and statistics, evaluating the chatbots' performance on specific problem-solving tasks.	To evaluate and compare the features, functionalities, and potential applications of these AI platforms to enhance student learning in calculus and statistics.	GPT-4 outperformed the other chatbots in accuracy for calculus and statistics prompts, though all models had limitations. Chatbots have the potential to transform higher education by offering personalized support, but they can be verbose and provide incorrect answers.

Educational Design Principles of Using AI Chatbot That Supports Self-Regulated Learning in Education: Goal Setting, Feedback, and Personalization	Chang et al., 2023	Conceptual paper proposing a theoretical framework	Higher education, focusing on integrating AI chatbots into teaching and learning.	To propose educational design principles for using AI chatbots to support and foster student Self-Regulated Learning (SRL).	The paper proposes three key principles: teaching effective prompting for goal setting, using "reverse prompting" for feedback and self-assessment, and leveraging learning analytics for personalization to develop students' self-regulation.
Does AI have utility in medical student surgical education? A comparative analysis of chatbots in answering standardized surgical multiple-choice questions	DaFonte et al., 2025	Comparative quantitative study	Medical student surgical education; assessment using 20 NBME Surgery Sample Items.	To assess and compare the accuracy of four AI chatbots (ChatGPT-4o mini, ChatGPT o3-mini, Gemini 2.0 Flash, Claude 3.5 Sonnet) on standardized surgery questions.	ChatGPT o3-mini and Claude 3.5 Sonnet achieved 100% accuracy, demonstrating high potential as supplementary study tools. ChatGPT-4o mini and Gemini 2.0 Flash showed lower, inconsistent performance and critical reasoning errors.
AI Chatbots in Education: Challenges and Opportunities	Davar et al., 2025	Narrative literature review	General higher education and online learning settings, focusing on the integration of AI	To explore the benefits and challenges of AI chatbots in education and how they can	AI chatbots offer benefits like personalized learning, 24/7 support, and administrative automation. Key challenges include academic integrity

			chatbots like ChatGPT.	address learning barriers.	issues, data hallucinations, data security risks, and over-reliance hindering critical thinking.
The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education	Essel et al., 2022	Quasi-experimental pretest-posttest design combined with qualitative focus groups. Sixty-eight undergraduate students were randomly allocated into experimental (chatbot) and control (instructor) cohorts	Final-year undergraduate students in a multimedia programming course at Kwame Nkrumah University of Science and Technology (KNUST), Ghana. The context featured a high student-instructor ratio.	To investigate the impact of a zero-coding virtual teaching assistant (KNUSTbot) on students' academic performance and to explore their perceptions of using the chatbot for learning support.	Students in the experimental cohort who interacted with the chatbot performed significantly better academically than those in the control cohort. Students reported high satisfaction, appreciating the immediate, 24/7 feedback, though some noted a lack of in-depth explanations and the absence of a human element.
An Overview of the Use of Chatbots in Medical and Healthcare Education	Frangoudes et al., 2021	Systematic literature review following the PRISMA methodology. Analysed 19 studies from databases including ACM, IEEE, PubMed, and	Focuses on the application of chatbots in medical and healthcare education, including virtual patient simulations, patient	To identify the primary uses of chatbots in medical education, understand their development frameworks, and	Chatbots are primarily used as virtual patients for clinical skills training, for patient education and counselling, and as course assistants. The review found their efficacy has not been thoroughly tested and there are

		Springer, published between 2015 and 2020	education, and course assistance in Higher Education Institutions (HEIs).	examine the metrics used for their evaluation (e.g., usability, accuracy, user skills).	limited examples of their integration into European healthcare curricula, calling for further research.
AI chatbots in programming education: Students' use in a scientific computing course and consequences for learning	Groothuijsen et al., 2024	Mixed-methods case study	A Master's course in Scientific Computing for Mechanical Engineering at a Dutch university; 29 students.	To investigate how students use AI chatbots for a programming assignment and how this use influences their learning.	Students used ChatGPT for six tasks (e.g., code explanation, debugging). Use negatively impacted pair programming and collaboration, and the teacher perceived a decline in code quality and learning of pragmatic skills.
Potentials of ChatGPT in computer programming: insights from programming instructors	Husain, 2024	Qualitative study (in-depth interviews and thematic analysis)	12 programming instructors from the IT faculty at AL al-Bayt University, Jordan.	To explore instructors' perceptions of ChatGPT's benefits, drawbacks, and its role in enhancing programming instruction.	Instructors identified key benefits (e.g., debugging, personalized learning) and drawbacks (e.g., inaccurate code, over-reliance). They recommend a complementary role for ChatGPT alongside traditional teaching methods to balance its potential.

A review of opportunities and challenges of chatbots in education	Hwang & Chang, 2023	Systematic literature review of 29 articles from SSCI journals	Analysis of chatbot studies in education from the Web of Science database (2005-2020).	To identify research trends, opportunities, and challenges of using chatbots in education.	Most studies used quantitative methods and focused on language learning. Research is still in an early stage, with guided learning being the primary strategy and a lack of empirical studies on advanced learning designs.
Effects of Generative Chatbots in Higher Education	Ilieva et al., 2023	Mixed-methods study involving a student survey and an experimental evaluation of chatbot performance	Higher education at Plovdiv University, Bulgaria; survey of 131 students and testing of seven AI chatbots on academic tasks.	To propose a framework for AI chatbot integration in university teaching and to assess student perceptions and chatbot effectiveness.	Most students (89.3%) had used chatbots and showed high intention for future use. Chatbots performed variably on academic tasks, with ChatGPT Plus achieving the highest scores, but human oversight is still essential.
Effects of AI chatbots on EFL students' communication skills	Kim et al. 2021	Quantitative (Pre-test/post-test design)	49 South Korean university students in a general English course, divided into low and intermediate proficiency levels.	To examine the effects of AI chatbots on EFL students' speaking skills and their perceptions of the technology.	Both groups significantly improved pronunciation, intonation, and stress. Only the intermediate group improved in fluency. Students reported reduced anxiety and a more comfortable learning environment.

Chatbots and academic writing for doctoral students	Krumsvik, 2025	Self-study approach	Doctoral supervision	To examine how interactions with a Generative AI chatbot align with, or diverge from, authentic feedback practices of supervisor.	The chatbot complements supervision but is not a substitute.
Educational chatbots for project-based learning: investigating learning outcomes for a team-based design course	Kumar, 2021	Mixed-method quasi-experimental study	60 second-year undergraduate students in a Bachelor of Education program in Malaysia, taking an Instructional Design course.	To investigate the impact of a suite of educational chatbots on learning outcomes, including performance, teamwork, and affective-motivational factors, in a team-based project.	The chatbot group showed significantly improved learning performance and teamwork compared to the control group. However, affective-motivational outcomes (perception of learning, need for cognition, motivation, creative self-efficacy) showed no significant difference.
Role of AI chatbots in education: systematic literature review	Labadze et al., 2023	Systematic literature review	Analysis of 67 relevant studies on AI chatbots in education, with a	To identify the key advantages for students and educators, and the	For students, benefits include homework assistance, personalized learning, and skills development. Key concerns are

			focus on newer models like ChatGPT.	main concerns regarding AI chatbot integration in education.	reliability, fair assessment, and ethical issues like data privacy.
Integration of AI-Powered Chatbots in Nursing Education: A Scoping Review of Their Utilization, Outcomes, and Challenges	Labrague et al., 2025	Scoping Review	A review of 14 studies on the use of AI-powered chatbots in nursing education across multiple countries.	To explore the use, implementation, and impact of AI-powered chatbots on nursing students' learning, engagement, and skill development.	Chatbots improved knowledge retention, engagement, and skill development. Key applications included simulations, educational support, tutoring, and assessment.
Adapting Self-Regulated Learning in an Age of Generative Artificial Intelligence Chatbots	Lai, 2024	Conceptual / Methodological Paper	Proposes a framework for analyzing how learners use chatbots for Self-Regulated Learning (SRL).	To adapt SRL models for GenAI chatbots and develop a method to classify and analyse learner prompts to understand SRL processes.	A process-action framework was developed to tag chatbot prompts with SRL processes. Process mining of these tags can provide insights into a learner's SRL strategies and inform personalized support.
CHAT-ACTS: A pedagogical framework for	Lin & Chang,, 2023	Conceptual paper proposing a theoretical framework.	General higher education, focusing on integrating	To propose a pedagogical framework (CHAT-	The framework positions the learner at the center of three integrated modes: Personalized

personalized chatbot to enhance active learning and self-regulated learning			personalized chatbots into teaching and learning.	ACTS) that guides the integration of personalized chatbots to enhance active learning and self-regulated learning (SRL).	Chatbot (evaluation, feedback, plan), SRL (goal setting, study tactics), and Active Learning (activity-based strategies). It provides principles for designing chatbots to improve engagement, motivation, and learning outcomes.
Do AI chatbots impact motivation? Insights from a preliminary longitudinal study	Liu & Reinders, 2025	Preliminary longitudinal study using piecewise mixed effects models	24 first-year English majors in a compulsory academic listening course at a university in China.	To examine the motivational trajectories of students interacting with a pre-scripted chatbot versus a generative AI (ChatGPT) chatbot designed to support self-regulated learning (SRL).	A significant upward trend in motivation was observed after switching to the generative AI chatbot. The study also revealed substantial inter-individual differences, with the generative AI phase amplifying these differences more than the pre-scripted phase.
Educational AI Chatbots for Content and Language Integrated Learning	Mageira et al., 2022	Mixed-methods study (field experiments with pre-test/post-test and questionnaires)	61 high school students in Greece learning English or French as a foreign	To design, implement, and evaluate an AI chatbot (AssaaraBot)	The chatbot was found to be engaging and user-friendly, effectively supporting cultural content learning. However, it

			language, focusing on cultural content about the Minoan Civilization.	for teaching cultural content and a foreign language simultaneously using the CLIL approach.	was less effective for language learning compared to other ICT tools, and no statistically significant difference in overall knowledge gain was found between the chatbot and control groups.
Navigating the Moral Maze: Ethical Challenges and Opportunities of Generative Chatbots in Global Higher Education.	Mariyono & Hidayatullah, 2025	Systematic literature review using a Hybrid Thematic SWOT (HT-SWOT) analysis	Global higher education, focusing on the ethical integration of generative AI chatbots like ChatGPT.	To explore the ethical challenges and opportunities of generative chatbots and provide strategic insights for their responsible integration.	Generative AI offers strengths like personalized learning and opportunities for inclusivity, but also poses weaknesses such as algorithmic bias and threats to academic integrity. The study emphasizes the need for ethical policies, AI literacy, and equitable access.
Exploring the Role of Chatbots and Messaging Applications in Higher Education: A Teacher's Perspective	Merelo et al., 2022	Quantitative study using online surveys	Higher education instructors, primarily based in Spain and Spanish-speaking countries.	To investigate teachers' use of messaging apps, identify useful chatbot features, and understand their	Over half of teachers use institution-supported messaging apps. The most valued chatbot features were facilitating agenda formation and answering FAQs. Teachers preferred small chat

				preferred interaction styles with students.	groups with only their students from a specific course.
Analysis of artificial intelligence chatbots and satisfaction for learning in mathematics education	Moral-Sánchez et al., 2023	Descriptive cross-sectional study using a quantitative questionnaire	A Didactics of Geometry course for 120 primary education trainee teachers at the University of Málaga, Spain.	To analyse chatbots created by students and evaluate their satisfaction with the tool for learning.	Students reported high satisfaction with the chatbots they created using Snatchbot and its integration with Telegram. The activity improved their digital competence, and the vast majority would recommend using chatbots as teaching assistants in other subjects.
Exploring the Effectiveness of Advanced Chatbots in Educational Settings: A Mixed-Methods Study in Statistics	Navas et al., 2024	Mixed-methods study (Quantitative error analysis & Qualitative Grounded Theory interviews)	Undergraduate engineering statistics education; evaluation of ChatGPT 3.5, ChatGPT 4.0, and Bing Chat.	To investigate the effectiveness of advanced chatbots in solving statistical problems and understand the causes of their errors.	Chatbots had high error rates (65% calculation errors), especially in probability and inference. Qualitative analysis revealed patterns like chatbot inflexibility and provided recommendations for educational use.
Unlocking educational potential: exploring students' satisfaction and sustainable	Ngo et al., 2024	Quantitative research approach employing a cross-sectional survey research design	The study enlisted 435 valid responses from students across eight Vietnamese	To investigate the factors affecting students' satisfaction and continuous usage	The study highlighted the pivotal role of ChatGPT in amplifying students' intrinsic motivation

engagement with chatgpt using the ecm model			universities with a minimum of three months of prior experience utilizing ChatGPT for educational purposes.	of ChatGPT in an educational context.	during their English as a second language learning.
Chatbots applications in education: A systematic review.	Okonkwo & Ade-Ibijola, 2021	Systematic Literature Review	Analysis of 53 research articles on educational chatbots from six digital databases.	To review the status, benefits, challenges, and future research areas of chatbot applications in education.	Chatbots are primarily used for teaching/learning (66%). Key benefits include engagement and quick access; major challenges are ethical issues and user attitude.
Chatbots and AI in Education (AIEd) tools: The good, the bad, and the ugly	Ifelebuegu et al., 2023	Exploratory qualitative study using a systematic literature review and thematic analysis	General higher education and research settings, examining the integration of various AI chatbots and tools.	To explore the beneficial applications (the good), negative impacts (the bad), and ethical challenges (the ugly) of using AI chatbots in education and research.	The "good" includes personalized learning, administrative automation, and research support. The "bad" and "ugly" encompass overreliance, bias reinforcement, misinformation, plagiarism, data privacy issues, job displacement threats, and the erosion of human connection.

ChatGPT in Higher Education: Supporting Academic Literacy Through ChatGPT-Based Activities	Raptopoulou, 2025	Qualitative case study	A philosophy of education course for first-year Education students at a Swedish university.	To explore how ChatGPT can be integrated to support academic literacy and to identify associated challenges and opportunities.	When used with clear pedagogical goals, ChatGPT can support academic literacy by summarizing concepts and providing writing structures. Collaborative, critical evaluation of AI outputs with teacher guidance is essential to mitigate inaccuracies and deepen learning.
ChatGPT Review: A Sophisticated Chatbot Models in Medical & Health-related Teaching and Learning	Razak et al., 2023	Systematic review	Review about utilization of chatbots, specifically ChatGPT, in medical and health sciences education.	Deeper understanding of the current and potential future uses of chatbots.	The findings confirm the exponential growth and potential of ChatGPT and chatbots in the areas of healthcare and medical education.
The impact of a chatbot working as an assistant in a course for supporting student learning and engagement	Roca et al., 2024	Quantitative study using the Technology Acceptance Model (TAM) and structural equation modeling (SEM)	A first-year Computer Engineering course at a university in Guatemala, focusing on a Java	To illustrate chatbot integration for assisting students and to analyse factors influencing its perceived	Perceived usefulness significantly influenced students' intention to use the chatbot. Students valued its 24/7 availability and user-friendly interface for error detection and learning support.

			programming activity.	usefulness and adoption.	
Battle of AI chatbots: Graduate students' perceptions of ChatGPT versus Gemini for learning purposes in Egyptian higher education	Sobaih & Abuelnasr, 2025	Sequential mixed-methods (Survey based on UTAUT model, followed by in-depth interviews)	410 graduate students from tourism and hospitality programs across six public Egyptian universities.	To compare graduate students' perceptions, acceptance, and use of ChatGPT versus Gemini for learning purposes using the UTAUT framework.	Students had positive perceptions of both chatbots, but showed significantly higher acceptance and usage intention for Gemini. This preference was due to its user-friendly interface, free features, and perceived accuracy. A lack of institutional support and guidelines for AI use was also noted.
The ethical implications of using generative chatbots in higher education	Williams, 2023	Perspective/Conceptual paper	General higher education, focusing on the integration of generative AI chatbots like ChatGPT.	To outline and discuss the key ethical concerns of using generative chatbots in education, including data privacy, bias, and academic integrity.	Identifies major ethical risks: data privacy issues, algorithmic bias, plagiarism, and AI hallucinations. Argues for clear policies and guidelines to ensure responsible use.

Source: authors' own work.

Text Analysis and Coding Procedure

Thematic analysis of narrative analytic synthesis, analysing the content and context of text, identifies recurring patterns in the literature and understands shared themes. It involves reading, coding segments for key ideas, grouping codes into potential themes, refining these themes, and then constructing a compelling narrative report that explains these patterns across the dataset, offering insights beyond individual accounts (Braun & Clarke, 2006; Nowell et al., 2017). The procedure first involves data extraction to identify patterns across studies, followed by coding, in which each included study serves as a unit of analysis. Coding involved two stages. Initially, key findings, conceptual statements, and interpretations from each study were manually coded and labelled to reflect their educational focus (e.g., instructional use, learner experience, ethical perspective). With the subsequent axial coding, the initial codes are organised into conceptual categories, including pedagogical integration, motivation and engagement, evaluation and feedback, and teacher perception. To capture the field's main conceptual trends, higher-order themes were inductively derived.

The coding protocol combined inductive (data-driven) and deductive (theory-informed) reasoning to ensure both the recent finding and conceptual alignment with prior chatbot-in-education literature. Coding was done manually and independently by two researchers. All disagreements between them were discussed and resolved by agreement. Due to resource limitations, inter-rater reliability coefficients for potential bias were not determined. Agreement was achieved through iterative review and discussion, enhancing interpretive validity.

The synthesis followed a transparent flow from codes → categories → themes → implications, allowing replication and auditability of the analytic reasoning process, Table 5. The process concludes with writing the narrative, explaining how the thematic analysis aligns with the research question.

Table 5

Codebook for Thematic Coding of Narratives Analytic Synthesis of Included Studies (appendix)

Codes (Initial Open Coding)	**Axial Categories (Conceptual Clusters)**	**Higher-order themes**
Engineering, medical, programming, and language learning contexts; higher education focus	Educational domains and instructional contexts	Theme 1 – Pedagogical integration and instructional design of AI chatbots in education
Chatbots as tutors, virtual assistants, feedback tools, and question generators	Functional and pedagogical roles of chatbots	
Project-based learning, CLIL, SRL scaffolds, and blended settings	Learning design and instructional strategies	
Learning performance, knowledge gains, problem-solving accuracy	Cognitive learning outcomes	Theme 2 – Student learning, engagement, and self-regulation
Motivation, enjoyment, engagement, and confidence	Affective outcomes (motivation and engagement)	
Self-regulation, study habits, and metacognitive reflection	Behavioral and self-regulatory outcomes	
Student perceptions, attitudes, acceptance, and trust	Learner perceptions and acceptance	Theme 3 – Stakeholder perceptions, acceptance, and institutional readiness
Teacher perspectives, readiness, concerns, and training	Educator attitudes and professional readiness	
Institutional guidelines, support structures, and policy gaps	Institutional readiness and governance	
Academic integrity, plagiarism, data privacy, bias, and fairness	Ethical and risk dimensions	Theme 4 – Ethical, motivational, and evaluative

Accuracy, hallucination, and reliability of outputs	Technical and quality concerns	dimensions of AI chatbot adoption
Over-reliance, critical-thinking decline, and misuse of generative AI	Pedagogical risks and learner dependence	
Frameworks, reviews, and conceptual syntheses (CHAT-ACTS, SRL, ethics)	Meta-level theoretical and evaluative perspectives	

Note. This thematic codebook presents the analytical hierarchy from open codes to axial categories and final themes. Themes 1–4 summarize pedagogical, cognitive-behavioral, stakeholder, and ethical dimensions of AI chatbot adoption in education, reflecting the selective integration step of thematic analysis.

Source: authors' own work based on "Using thematic analysis in psychology", V. Braun, & V. Clarke, 2006, *Qualitative Research in Psychology*, *3*(2), 77–101 (https://doi.org/10.1191/1478088706qp063oa).