

Beyond AI Tutors: A Human-Centred Framework for Goal-Driven Learning Agents in Education

Ghodakhe Shrirang Keshav¹, Prof. Y.G.Suryawanshi², Dr. H.B.Jadhav³

Student, Department of Computer Engineering¹

Guide, Department of Computer Engineering²

HOD, Department of Computer Engineering³

Adsul's Technical Campus, Chas, Ahilyanagar, India

Abstract: Artificial intelligence is changing education by offering new ways to explain concepts, support teachers, and personalise student learning. However, many conventional educational chatbots focus on answering individual questions rather than coordinating the complete learning process. Students often need more than an answer: they need a learning plan, suitable practice activities, feedback on mistakes, and guidance about what to study next. Agentic artificial intelligence offers a possible way to coordinate these activities around a defined learning objective.

This paper reviews selected research published between 2024 and 2026 on pedagogical AI agents, large language model agents, and agentic systems in education. The reviewed literature highlights opportunities for personalized tutoring, adaptive feedback, assignment support, and teacher assistance. It also identifies concerns involving inaccurate outputs, student dependency, privacy, fairness, accountability, and limited real-world evaluation. To address these challenges, this paper proposes a human-centred framework that connects learning-goal identification, task planning, trusted educational resources, instructional support, progress evaluation, and adaptive follow-up. The framework centres student independence, teacher judgement, and responsible data use. It is a conceptual proposal and has not yet been experimentally evaluated. Future research should assess its educational effectiveness, reliability, accessibility, and suitability for different learning environments..

Keywords: Agentic Artificial Intelligence, Goal-Driven Learning, Educational AI Agents, Personalised Learning, Human-AI Collaboration, Intelligent Tutoring Systems

I. INTRODUCTION

Artificial intelligence (AI) has become an increasingly visible part of modern education. Students can use AI-powered applications to understand difficult concepts, practise questions, prepare assignments, and receive explanations outside classroom hours. Teachers can also use these tools to prepare learning materials, organise activities, and generate preliminary feedback. These developments create opportunities to make education more flexible and responsive to individual learning needs.

Despite these advantages, conventional educational chatbots have limitations. They commonly depend on user prompts and may provide an answer without organising the learner's next steps, checking whether the student understands the explanation, or adapting later activities according to performance. For example, a student may receive a definition of database normalization but still be unable to apply the concept to a practical problem.

Goal-driven or agentic artificial intelligence offers a possible way to address this limitation. An AI agent can be designed to work toward a defined objective by dividing it into smaller tasks, selecting relevant resources, carrying out activities, checking intermediate results, and revising its plan when necessary. In education, these capabilities could support a more continuous learning experience.

Consider a student preparing for a programming examination. Instead of explaining only one programming concept, a goal-driven agent could organize the topic into smaller lessons, recommend relevant course materials, generate practice



problems, review the student's answers, and suggest further revision. This approach could help the learner understand what to study next instead of depending entirely on individual prompts.

However, greater autonomy does not automatically lead to better education. An agent may generate inaccurate explanations, misunderstand a student's difficulties, recommend unsuitable activities, or encourage excessive dependence on automated assistance. Educational systems must therefore be designed around learning objectives, student needs, and appropriate human supervision.

Recent research has explored pedagogical AI agent design, agentic AI in education, and language-model agents for learning [1]–[3]. Research published in 2026 has also examined the expanding landscape of agentic educational systems and the need for stronger connections between technical development and educational practice [4].

Motivated by these developments, this paper examines how goal-driven educational agents can support learning without replacing teachers or weakening student independence.

The main contributions of this paper are:

- A comparative review of selected research published between 2024 and 2026.
- An analysis of the opportunities and limitations of goal-driven educational AI agents.
- A conceptual human-centred framework connecting learning goals, planning, educational resources, feedback, and adaptation.
- A discussion of future research priorities for reliable, inclusive, and responsible educational AI.

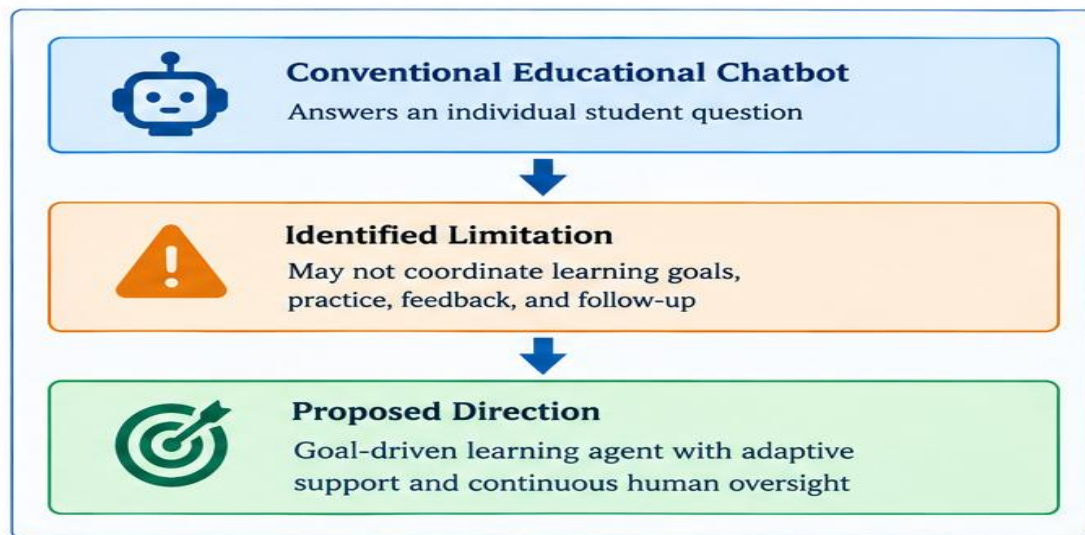


Fig. 1. Overview of the research problem and proposed direction.

II. REVIEW METHODOLOGY

This paper uses a focused narrative literature review to examine recent developments in agentic AI for education. The review concentrates on publications from 2024, 2025, and 2026 that discuss pedagogical AI agents, large language model agents, educational applications, and the relationship between agent capabilities and learning design.

Relevant research was identified through scholarly publication sources, including IEEE Xplore, the ACL Anthology, academic publisher websites, and research repositories. The selected publications were examined in terms of their research objectives, approaches, contributions, educational applications, and reported limitations.

The selected works represent complementary perspectives: pedagogical agent design, broad reviews of agentic AI in education, technical surveys of educational language-model agents, and studies of teaching-capability evaluation [1]–



[7]. Additional research on educational workflows and programming education was considered to understand practical implementation and human oversight [8]–[10].

The findings were compared thematically to identify common research directions, differences in emphasis, and unresolved challenges. The resulting analysis informed the conceptual framework proposed in Section IV.

This review is intended to provide a focused understanding of the field rather than a complete systematic review. It does not claim to include every relevant publication, and it does not introduce new experimental data. The proposed framework is therefore a literature-informed design proposal that requires empirical validation.

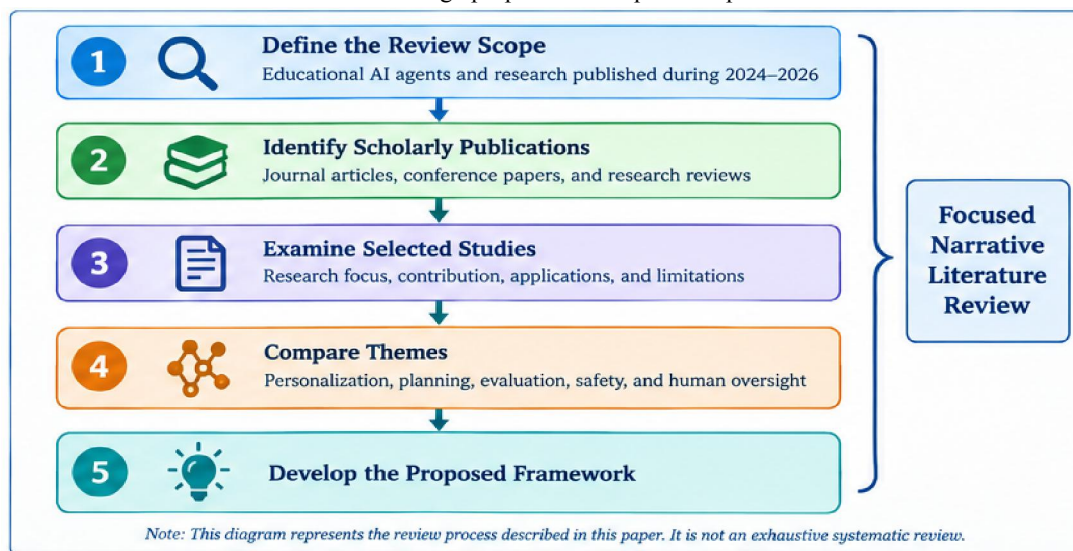


Fig. 2. Workflow of the focused narrative literature review.

III. LITERATURE REVIEW AND EXISTING SYSTEM

A. Pedagogical AI Agent Design in 2024

Lan and Chen [1] examined the design of pedagogical AI agents in the era of large language models and generative AI. Their work highlights the importance of considering teachers' roles and instructional needs when developing educational agents.

The central lesson is that an educational agent should not be designed only around its ability to generate text. Its activities should be connected to learning objectives, appropriate teaching strategies, and the needs of the intended learners.

For example, a language-learning agent may introduce a concept, provide examples, ask students to construct sentences, identify errors, and offer corrective feedback. The sequence of activities matters because effective learning requires more than simply receiving information.

Zhang et al. [5] reviewed pedagogical agent design for K–12 education. Their work provides another perspective on the design of educational agents and the importance of considering learners and their educational context.

These studies offer useful design foundations. However, a conceptual design or literature review does not by itself establish that an agent will improve learning outcomes. Different subjects, age groups, and classroom environments require further evaluation.



B. Agentic AI in Education in 2025

Kostopoulos et al. [2] reviewed the state of agentic AI in education, examining applications and concerns associated with its use. Their work connects the technical capabilities of agentic systems with educational opportunities and implementation challenges.

Chu et al. [3] surveyed large language model agents for education, examining technical approaches and educational tasks, including learning support and feedback generation.

Together, these works demonstrate growing interest in AI systems that can coordinate multiple educational activities. However, completing a task is not the same as helping a learner understand a concept.

For instance, an agent may generate a study timetable automatically. That timetable becomes educationally useful only when it reflects the course requirements, the student's objectives, available study time, and actual learning needs.

Yusuf et al. [6] proposed a conceptual framework and surveyed pedagogical AI conversational agents in higher education. Their work further highlights the need to connect conversational technology with educational context.

Yang et al. [9] provided a comprehensive survey of large-language-model-based educational agents, contributing to the broader understanding of this developing research area.

C. Research on Agent Evaluation and Teaching Capabilities

Shi et al. [7] introduced EducationQ, a framework for evaluating language models' teaching capabilities through multi-agent dialogue. This direction is important because an AI system may produce persuasive explanations without necessarily teaching accurately or responding appropriately to a learner's misconception.

Evaluation should therefore consider more than whether an answer was generated. It should examine explanation quality, the suitability of feedback, the handling of errors, and the relationship between system performance and learning outcomes.

Zhang [8] discussed trends and a practical workflow for AI agents in education. This work provides a practical perspective on how educational agents might be organized and used.

Research on language-model applications in programming education has also emphasized the need to balance automation with human oversight [10]. This is especially relevant in education, where students need opportunities to practise and develop independent problem-solving skills.

D. Emerging Research Directions in 2026

Wang et al. [4] conducted a scoping review of generative-AI-powered agentic systems in education. The study mapped research characteristics, agent roles, architectures, and capabilities.

The review highlights the expanding research landscape and the need for stronger connections between technical development, educational theory, and real-world validation.

These findings suggest that educational AI research should move beyond demonstrating that agents can perform sequences of actions. Researchers must also investigate whether those actions are educationally appropriate, whether feedback is reliable, and whether the systems contribute to meaningful learning outcomes.

The work is particularly relevant to this paper because it supports the need for frameworks that combine agentic capabilities with instructional design and human oversight.

E. Existing System Architecture

A common simplified approach to educational chatbots follows a prompt-and-response workflow. The student enters a question, the language model processes the prompt, and the system returns a generated answer.

This can be useful for quick explanations and basic learning support. However, a basic prompt-and-response workflow does not necessarily include long-term goal tracking, structured task planning, verified course-resource retrieval, or progress-based adaptation.



It is important to note that not every existing educational chatbot has these limitations. Some current systems already provide tutoring workflows, memory, retrieval, or adaptive features. The architecture below represents a simplified conventional chatbot, used as a comparison point for the proposed framework.

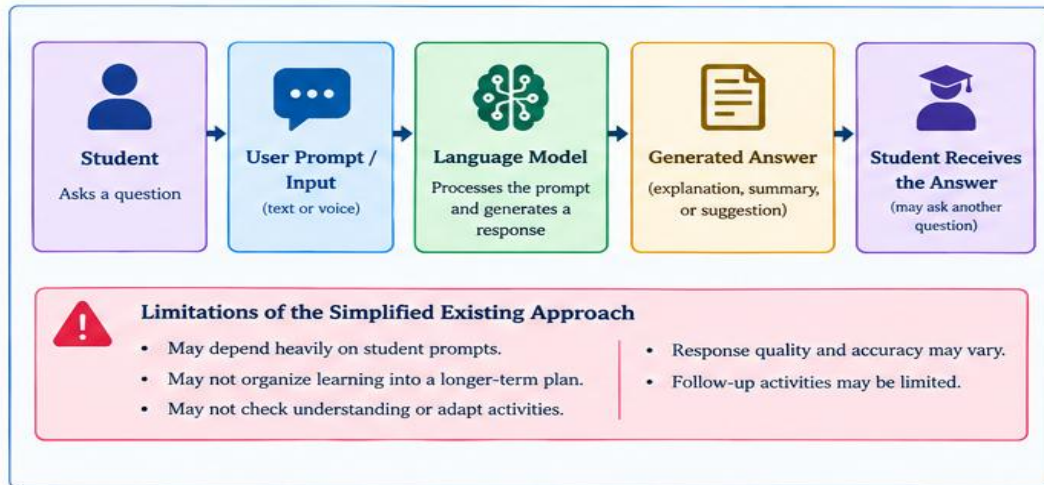


Fig. 3. Simplified architecture of an existing prompt-based educational chatbot.

Limitations of the simplified existing approach:

- The learning process may depend heavily on the student's prompts.
- Learning activities may not be organized into a longer-term plan.
- The system may not check whether the student can apply the concept independently.
- The quality and accuracy of the response may vary.
- Follow-up activities may not be adapted to demonstrated learning progress.
- These limitations motivate the design of the proposed framework.

IV. PROPOSED HUMAN-CENTRED FRAMEWORK

A. Framework Overview

This paper proposes a human-centred framework for goal-driven learning agents in education. The framework is intended to help an AI agent coordinate learning activities while keeping students and teachers actively involved. The central idea is that an educational agent should not focus only on completing tasks. It should help the student progress toward a learning objective, provide understandable feedback, and adjust its recommendations according to evidence of learning.

The proposed framework contains six connected stages:

- Learner and learning-goal identification.
- Planning and task decomposition.
- Trusted educational resource retrieval.
- Instructional support and practice.
- Learning progress evaluation.
- Adaptive planning and follow-up activities.

Human oversight operates throughout the process. Students retain control over learning choices, teachers can review important outputs, and privacy and safety safeguards apply at every stage.



B. Proposed System Architecture

The proposed architecture differs from the simplified existing system by organizing learning support as a continuing process.



Fig. 4. Proposed human-centred architecture for goal-driven learning agents in education.

C. Framework Components

1. Learner and Learning-Goal Identification

The process begins by identifying what the student wants to learn. Relevant inputs may include the topic, learning objective, existing knowledge, previous responses, and available study time.

For example, a student may want to understand database normalization before an examination. The agent should use this goal to organize relevant activities instead of producing unrelated explanations.

The system should collect only information necessary for the learning task and avoid making unsupported assumptions about the student's ability.

2. Planning and Task Decomposition

The agent divides the learning objective into smaller, manageable activities. A complex topic can be organized into introductory explanations, worked examples, practice questions, and revision.

Task decomposition helps establish a logical learning sequence. The plan should remain flexible so that the agent can modify it when the student demonstrates difficulty or completes a task earlier than expected.

3. Trusted Educational Resource Retrieval

The agent retrieves relevant course notes, approved textbooks, teacher-provided materials, exercises, and other suitable resources.

Using trusted materials can help reduce unsupported explanations and keep learning activities aligned with the curriculum. Where appropriate, the system should identify the sources supporting its explanations so that students and teachers can check them.

Resource retrieval does not guarantee correctness. Important information still requires verification.

4. Instructional Support and Practice

The agent provides explanations, examples, hints, practice questions, and corrective feedback. The amount of assistance should depend on the learning activity and the student's demonstrated needs.

Instead of immediately giving the complete solution to every problem, the system may first offer a hint, ask a guiding question, or present a similar example. This approach is intended to encourage reasoning and independent problem-solving.



5. Learning Progress Evaluation

After an instructional activity, the agent examines the student's response and identifies possible misunderstandings. It may compare the response with an answer key, assessment rubric, or teacher-approved criteria.

The evaluation should be treated as an estimate rather than a final judgement of the student's ability. Ambiguous responses and consequential assessments require suitable human review.

6. Adaptive Planning and Follow-Up

The agent uses available evidence to recommend the next activity. If the student demonstrates understanding, it may introduce a more challenging problem. If difficulties remain, it may offer another explanation, a simpler example, or additional practice.

This feedback cycle allows the learning sequence to change as the student progresses. The goal is to provide appropriate support while gradually encouraging greater independence.

D. Human Oversight and Responsible Operation

Human oversight is a central feature of the proposed framework rather than a separate final step.

Student agency allows learners to review recommendations, ask questions, challenge feedback, and choose suitable learning activities. Teacher oversight supports curriculum alignment, correction of inaccurate information, and review of important educational decisions.

Privacy, security, fairness, and accountability safeguards must apply throughout the system. High-impact decisions, including final grading and decisions affecting student progression, should not be delegated to an AI agent without appropriate human review.

The level of autonomy should depend on the risk of the activity. Generating a practice question is generally lower risk than making a decision that affects a student's academic record.

E. Working Example

Consider a student preparing for a database management examination.

The student identifies normalization as a topic requiring revision. The agent divides the topic into functional dependencies, first normal form, second normal form, and third normal form. It retrieves relevant course notes and begins with a short explanation of functional dependencies.

The agent then provides an example and asks the student to identify a possible normalization problem. If the response is incorrect, it offers a hint and a simpler practice question. If the student answers correctly, the agent moves to the next concept.

After several activities, the agent summarizes the student's progress and recommends further revision where necessary. The student can question the explanation, and the teacher can review the learning material and assessment criteria.

This example demonstrates how an agent could coordinate a sequence of learning activities. It does not establish that the proposed framework improves examination performance; that claim would require an empirical study.

V. EDUCATIONAL APPLICATIONS**A. Personalized Learning**

Goal-driven agents can organize learning activities according to a student's objectives, progress, and demonstrated difficulties. Students who need foundational explanations may receive additional examples, while learners who demonstrate understanding may proceed to more challenging activities.

Personalization should be based on relevant evidence and should allow students and teachers to correct inappropriate recommendations.



B. Intelligent Tutoring

Educational agents can support learners by explaining concepts, demonstrating problem-solving methods, providing hints, and asking follow-up questions. Their ability to coordinate multiple steps may be useful in programming, mathematics, science, and language learning.

An effective tutoring agent should encourage understanding rather than simply supply answers. It should also acknowledge uncertainty and avoid presenting incorrect information as established fact.

C. Assignment and Assessment Support

Agents can help students interpret assignment requirements, organize tasks, practise related concepts, and understand feedback. They may also assist teachers with draft rubrics and preliminary feedback.

However, AI-generated assessments may contain errors or reflect inappropriate criteria. Final grading and consequential academic decisions should follow institutional policies and appropriate human review.

D. Teacher Assistance

Teachers may use goal-driven agents to prepare lesson outlines, develop exercises, organize resources, and generate alternative explanations for difficult topics.

These applications may reduce repetitive preparation work and provide additional instructional options. The teacher should retain responsibility for selecting appropriate materials, understanding students' needs, and making professional decisions.

E. Academic Planning and Revision

An agent can help students divide a broad examination syllabus into smaller study tasks, schedule revision, and identify topics that need additional practice.

The recommendations should remain realistic and adjustable. A study plan should support the student's own goals rather than impose a fixed schedule without considering individual circumstances.

F. Accessibility and Continuous Learning

Educational agents may provide repeated explanations, alternative examples, and learning support outside classroom hours. They may also help students navigate digital learning materials and access supplementary resources.

The actual benefits depend on language support, accessible interfaces, device availability, and the quality of educational content.

VI. CHALLENGES AND LIMITATIONS**A. Inaccurate Information**

Language-model-based agents can produce incorrect explanations, unreliable references, and unsuitable recommendations. Errors can also propagate when an agent uses an incorrect intermediate result to determine later actions.

Possible safeguards include retrieving trusted resources, checking important claims, maintaining clear records of actions, and requiring human review for high-impact outputs.

B. Student Overreliance

Excessive dependence on AI may reduce opportunities for students to develop independent reasoning, problem-solving, and self-assessment skills.

Agents should therefore provide hints, encourage reflection, and gradually reduce assistance when appropriate. Students should remain active participants rather than passive recipients of generated answers.



C. Privacy and Security

Educational agents may process student responses, learning histories, and assessment records. Inadequate protection could expose sensitive information or allow unauthorized access.

Institutions should establish appropriate data-minimization, access-control, retention, and security procedures. Students should receive clear information about how their data are used.

D. Bias and Fairness

Agent performance may vary across languages, cultural settings, disability-related needs, and levels of digital access. A system that works well for one group may provide less reliable support to another.

Evaluation should include diverse learners and examine differences in accuracy, feedback quality, accessibility, and educational outcomes.

E. Accountability and Human Control

It may be difficult to determine responsibility when an agent produces harmful advice, selects unsuitable materials, or makes an incorrect assessment.

The system should make its actions understandable, provide opportunities to challenge recommendations, and allow teachers to intervene. Clear accountability arrangements should be established before educational deployment.

F. Implementation and Cost

Institutions may face difficulties related to computing resources, software integration, staff training, system maintenance, and the availability of reliable digital materials.

Successful implementation requires institutional planning and continuing evaluation. A technically capable agent may still be unsuitable if it is too costly, difficult to use, or incompatible with existing educational practices.

G. Limitations of the Present Paper

This paper presents a focused narrative review rather than an exhaustive systematic review. The selected publications provide useful perspectives, but they do not represent every development in agentic AI for education.

The proposed framework is conceptual and has not been implemented or tested with students. Therefore, no claims are made about measured improvements in academic performance, student satisfaction, or teaching efficiency.

Future empirical research is necessary to evaluate the framework's effectiveness, reliability, fairness, and practical feasibility.

VII. FUTURE RESEARCH DIRECTIONS**A. Long-Term Educational Evaluation**

Future studies should examine whether goal-driven agents improve knowledge retention, conceptual understanding, problem-solving, and independent learning over extended periods. Evaluations should compare AI-supported learning with appropriate alternatives and consider differences between learner groups.

B. Integration of Educational Theory

Agent design should be guided by established principles of instructional design, formative assessment, scaffolding, and self-regulated learning.

Research should investigate how planning and adaptation can be aligned with learning objectives rather than relying solely on a model's ability to generate plausible responses.



C. Explainable and Verifiable Agent Actions

Students and teachers should be able to understand why an agent recommends a particular activity and which educational resources support its explanation.

Future systems should be evaluated for their ability to identify uncertainty, detect errors, and correct unsuitable recommendations.

D. Collaboration Between Specialized Agents

Multiple specialized agents may collaborate to provide explanations, generate practice activities, evaluate responses, and check the quality of feedback.

However, additional agents may introduce inconsistent recommendations, duplicated work, and increased computational cost. Research should investigate reliable coordination and clear responsibility among components.

E. Responsible Governance

Future educational agents should incorporate privacy protection, secure access, transparent decision-making, and appropriate human approval for consequential actions.

Institutions should also establish processes for reporting errors, correcting outputs, and reviewing system performance.

F. Multilingual and Inclusive Education

Further research should explore educational agents that support regional languages, accessible learning materials, and students with different learning requirements.

Evaluation should examine whether the quality of assistance remains consistent across groups and whether the technology is practical for institutions with limited digital resources.

G. Standardized Evaluation Methods

The field needs evaluation methods that examine more than task completion. Useful measures include factual accuracy, quality of feedback, learning outcomes, student independence, fairness, accessibility, computational cost, and safety.

Shared evaluation practices would help researchers compare systems and determine which agent designs provide meaningful educational benefits.

H. Classroom-Based Validation of the Proposed Framework

The proposed framework should be implemented in a limited educational setting and evaluated through a carefully designed pilot study. Researchers could examine one subject, such as introductory programming or database management, and compare student progress under different forms of learning support.

The study should define its sample, evaluation criteria, data-protection procedures, and comparison method before collecting data. Results should be reported transparently, including unsuccessful outcomes and limitations.

VIII. CONCLUSION

Goal-driven learning agents offer a promising direction for educational AI because they can coordinate multiple activities toward a learning objective rather than simply respond to individual questions. Research published between 2024 and 2026 demonstrates growing interest in pedagogical agent design, educational language-model agents, personalized support, and the broader development of agentic systems. However, technical progress alone cannot guarantee meaningful educational improvement. Inaccurate information, student overreliance, privacy concerns, bias, and limited real-world validation remain important challenges. To address these issues, this paper proposes a human-centred framework that combines learner-goal identification, task planning, trusted resource retrieval, instructional support, progress evaluation, and adaptive learning under continuous human oversight. The framework is a conceptual contribution and requires further implementation and empirical testing. Future educational AI should be evaluated not only by its ability to complete tasks but also by its capacity to strengthen student understanding, preserve learner



independence, support teachers, and operate responsibly. The long-term goal should be to develop AI agents that complement human teaching and help students become more confident, capable, and independent learners.

ACKNOWLEDGMENT

The author acknowledges the researchers whose published work informed this literature review. The author also recognizes the importance of responsible AI use, academic integrity, and human supervision in the development of educational technologies.

REFERENCES

- [1] Y.-J. Lan and N.-S. Chen, "Teachers' agency in the era of LLM and generative AI: Designing pedagogical AI agents," *Educational Technology & Society*, vol. 27, no. 1, 2024, doi: 10.30191/ETS.202401_27(1).PP01.
- [2] G. Kostopoulos, V. Gkamas, M. Rigou, and S. Kotsiantis, "Agentic AI in Education: State of the Art and Future Directions," *IEEE Access*, vol. 13, pp. 177467–177491, 2025, doi: 10.1109/ACCESS.2025.3620473.
- [3] Z. Chu et al., "LLM Agents for Education: Advances and Applications," in *Findings of the Association for Computational Linguistics: EMNLP 2025*, pp. 13782–13810, 2025, doi: 10.18653/v1/2025.findings-emnlp.743.
- [4] N. Wang, D. Zou, H. Xie, and S. J. Qin, "A scoping review of generative AI-powered agentic AI in education: Research landscape, agentic capabilities, and insights from the frontier agent paradigm, exemplified by OpenClaw," *Computers and Education: Artificial Intelligence*, vol. 11, Art. no. 100653, 2026, doi: 10.1016/j.caeai.2026.100653.
- [5] S. Zhang, C. D. Jaldi, N. L. Schroeder, A. A. López, J. R. Gladstone, and S. Heidig, "Pedagogical agent design for K-12 education: A systematic review," *Computers & Education*, vol. 223, Art. no. 105165, 2024, doi: 10.1016/j.compedu.2024.105165.
- [6] H. Yusuf, A. Money, and D. Daylamani-Zad, "Pedagogical AI conversational agents in higher education: a conceptual framework and survey of the state of the art," *Educational Technology Research and Development*, vol. 73, pp. 815–874, 2025, doi: 10.1007/s11423-025-10447-4.
- [7] Y. Shi, R. Liang, and Y. Xu, "EducationQ: Evaluating LLMs' Teaching Capabilities Through Multi-Agent Dialogue Framework," in *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics, Volume 1: Long Papers*, pp. 32799–32828, 2025, doi: 10.18653/v1/2025.acl-long.1576.
- [8] J. Zhang, "AI Agents in Education: Four Trends and a Practical Workflow," *Proceedings of the AAAI Symposium Series*, vol. 5, no. 1, pp. 131–132, 2025, doi: 10.1609/aaaiss.v5i1.35578.
- [9] J. Yang, M. Wang, X. Du, and N. Ra, "A Comprehensive Survey on Large-Language-Model-Based Agents for Education," *IEEE Transactions on Learning Technologies*, vol. 18, pp. 898–913, 2025, doi: 10.1109/TLT.2025.3617909.
- [10] G. Pitts, A. P. Hridi, and A. B. L. Narayanan, "A Survey of LLM-Based Applications in Programming Education: Balancing Automation and Human Oversight," in *Proceedings of the Fourth Workshop on Bridging Human–Computer*

