FROM GENERATIVE AI TO AGENTIC LEARNING SYSTEMS: A HUMAN-AI COLLABORATION FRAMEWORK FOR THE FUTURE OF DISTANCE HIGHER EDUCATION

Yakhshiboev Rustam Erkinboy ugli
PhD, Associate Professor, Tashkent State University of Economics, Tashkent, Uzbekistan
r.yaxshiboyev@tsue.uz

Abstract: *Generative AI entered universities as a tool that students consult on demand. The next generation of systems, AI agents that plan, remember and act over several steps, will take over parts of teaching, and distance education will feel this first. This article asks how responsibility should be divided between people and such agents in distance higher education. Following design science methodology, it combines an integrative review of experimental and survey evidence published in 2019-2026, regulatory documents and enrolment data for Uzbekistan. The evidence points in one direction: the same models raise or lower learning depending on how they are constrained. An AI tutor built around expert-written scaffolding produced learning gains 0.73 to 1.3 standard deviations larger than active learning in class, while unrestricted GPT-4 access raised practice scores by 48% but lowered later exam scores by 17%. Student use of AI is close to universal (92-95% in UK surveys, 60-72% in Uzbek samples), yet only about a third of students report receiving institutional training. In Uzbekistan, correspondence and distance students account for 32% of 1.85 million students, with 36 students per teacher. The proposed framework defines five levels of AI agency, from tool to orchestrating agent, and assigns each distance learning task a maximum level using three criteria: the stakes of the decision, the reversibility of errors and the strength of evidence. Tutoring and formative feedback may reach the tutor agent level under teacher oversight, whereas grading, progression and admission remain human decisions. Recommendations for universities and national AI policy follow.*

Keywords: *generative AI, agentic AI, AI tutoring, human-AI collaboration, distance education, higher education, human oversight, levels of automation, AI literacy, Uzbekistan.*

INTRODUCTION

Within three years of the public release of ChatGPT, generative AI became an ordinary part of student life. In the United Kingdom, 92% of full-time undergraduates reported using AI tools at the end of 2024, up from 66% a year earlier, and 88% had used generative AI in assessed work (Freeman, 2025). By 2026 the figures had reached 95% and 94% (Stephenson & Armstrong, 2026). A survey of more than 3,800 students in 16 countries found that 86% used AI in their studies, while 58% felt they lacked the knowledge and skills to use it well (Digital Education Council, 2024).

Most of this use still follows a simple pattern: the student asks and the model answers. The systems now being built work differently. Agents based on large language models break a goal into steps, call external tools, keep memory between sessions and act with limited supervision (Wang et al., 2024; Yao et al., 2023). Sapkota et al. (2026) separate single-purpose AI agents from agentic AI, in which several agents coordinate,

share memory and divide tasks among themselves as the work unfolds. In a university course such a system could plan a student’s week, generate practice items, comment on drafts, contact students who fall behind and report to the teacher, without a person approving each step.

Distance education is where these capabilities matter most. Moore (1993) described the central problem of distance learning as transactional distance, the gap in communication and understanding that grows when dialogue between learner and teacher is scarce. An agent that is available at any hour could narrow that gap. It could also widen it, if it replaces the teacher instead of supporting them, or if students hand their thinking over to it. The experimental record already contains both outcomes (Bastani et al., 2025; Kestin et al., 2025).

For Uzbekistan the question is practical. The Law "On Education" recognises distance learning as a separate form of study (Republic of Uzbekistan, 2020), and in 2026 correspondence and distance students made up about a third of higher education enrolment (Ta’lim xabarlari, 2026). The national AI strategy to 2030 sets targets for training AI specialists and opening AI laboratories at universities (President of the Republic of Uzbekistan, 2024), but it says little about how AI should be used in teaching and assessment.

The scientific problem is the absence of a principled way to decide how much agency an AI system should have over a given educational task and who remains accountable for the result. Existing frameworks either describe levels of automation in general terms (Molenaar, 2022; Parasuraman et al., 2000) or offer broad ethical guidance (Miao & Holmes, 2023). Neither tells the head of a distance learning programme which tasks can be delegated to an agent next semester and which cannot.

The aim of this study is to develop and apply a human-AI collaboration framework that specifies, for each core task of distance higher education, the maximum level of AI agency and the human role that must accompany it. The objectives are (1) to synthesise experimental and survey evidence on generative AI in learning, (2) to define levels of AI agency relevant to education, (3) to formulate decision rules for allocating tasks to these levels and (4) to apply the framework to distance higher education in Uzbekistan. The object of the study is distance higher education; the subject is the division of responsibilities between teachers, students and AI systems. The study rests on the proposition that the learning value of AI depends less on the capability of the model than on the pedagogical constraints and human oversight built around it. The novelty lies in linking levels of AI agency to concrete distance learning tasks through explicit criteria and in grounding this allocation in recent controlled experiments and in regulatory risk classes.

LITERATURE REVIEW

Research on AI in higher education predates generative models. Zawacki-Richter et al. (2019) reviewed 146 studies and grouped applications into four areas: profiling and prediction, assessment and evaluation, adaptive systems and personalisation, and intelligent tutoring systems. They also found weak links to pedagogical theory and almost no critical reflection on risks, which explains the

question in their title about where the educators were. Five years later Bond et al. (2024), in a meta systematic review, still called for more ethical reflection, more collaboration across disciplines and more methodological rigour. Generative AI arrived in a field that had not yet settled these questions.

Controlled studies of generative AI in learning are few, but the best of them are informative. Kestin et al. (2025) ran a randomised crossover trial with 194 students of an introductory physics course at Harvard. A GPT-4 tutor that followed expert-written, step-by-step solutions and guided students through each part of a problem in sequence produced learning gains 0.73 to 1.3 standard deviations larger than an in-class active learning session, in less time. The authors stress that the result depended on carefully prepared prompts and materials. Bastani et al. (2025) gave close to 1,000 high school students in Türkiye access to one of two GPT-4 interfaces. A standard chat interface raised practice scores by 48%, but once access was removed these students scored 17% lower on exams than students who never had it. A tutor version with prompts designed to protect learning raised practice scores by 127% and largely avoided the harm. Read together, the two studies show the same model producing gains or losses depending on how its behaviour is constrained.

Other work points to a quieter risk. In a randomised experiment with 117 university students, Fan et al. (2025) found that ChatGPT support improved essay scores more than support from a human expert or from writing analytics tools, but did not improve knowledge gain or transfer. The authors warn of "metacognitive laziness", a dependence that lets students skip the regulation of their own learning. Lee et al. (2025) surveyed 319 knowledge workers who described 936 cases of generative AI use and found that greater confidence in the AI went with less critical thinking, while greater self-confidence went with more. These are not studies of distance students, but distance students work alone more often and have fewer occasions for a teacher to notice the problem.

The evidence base is young, and some of it has not held up. A meta-analysis of 51 studies by Wang and Fan (2025) reported a large positive effect of ChatGPT on learning performance ($g = 0.867$) and was widely cited. In May 2026 the journal retracted it, citing discrepancies that undermined confidence in the analysis; critics had pointed to inconsistent study counts, improper weighting and the inclusion of a retracted study, and estimated the effect closer to 0.5 (Plagiarism Today, 2026). The episode supports the call by Bond et al. (2024) for more rigour and suggests that policy should rest on individual well-designed trials rather than on early pooled estimates.

The question of how to divide work between people and machines has a longer history in human factors research. Parasuraman et al. (2000) distinguished four functions that can be automated to different degrees: information acquisition, information analysis, decision selection and action implementation. Shneiderman (2020) rejected the idea of a single scale running from human control to machine control and argued that systems can, and should, combine high automation with high human control. Molenaar (2022) carried a six-level model from the automotive industry into education, from full teacher control at level 0 to full automation at level 5, and concluded that the most promising designs lie at the intermediate levels, where

teacher and technology share control. Mollick and Mollick (2023) took a more practical route and described seven roles that AI can play for students: mentor, tutor, coach, teammate, student, simulator and tool. None of these frameworks was written with autonomous agents in mind.

Agents differ from chat tools because they take actions as well as give answers. Yao et al. (2023) showed that language models can interleave reasoning with actions such as searching or calling tools. Wang et al. (2024) describe the typical architecture of such agents, with modules for profile, memory, planning and action. Sapkota et al. (2026) draw a line between AI agents, which automate defined tasks, and agentic AI, in which multiple agents decompose goals, keep persistent memory and coordinate with each other. For education this means that an AI system may soon carry a decision from analysis to action without a teacher seeing the intermediate steps, which is exactly the situation that Parasuraman et al. (2000) and Shneiderman (2020) treat as needing the most deliberate design.

Distance education theory adds a further consideration. In Moore’s (1993) account, transactional distance depends on dialogue, on the structure of the course and on learner autonomy. The Community of Inquiry model of Garrison et al. (2000) describes successful online learning through three kinds of presence: cognitive, social and teaching. An AI tutor can supply dialogue and structure at any hour. Whether it can supply social and teaching presence is far less clear, and a framework that ignores this would trade one kind of distance for another.

Policy documents have moved quickly. UNESCO’s guidance on generative AI calls for a human-centred approach, validation of tools before use and a minimum age of 13 for independent use (Miao & Holmes, 2023). Its competency frameworks define twelve AI competencies for students and fifteen for teachers (Miao & Cukurova, 2024; Miao et al., 2024). The EU AI Act classifies AI systems used for admission, for evaluating learning outcomes, for assessing the level of education a person can access and for monitoring students during tests as high-risk, and it prohibits emotion recognition in educational institutions (European Union, 2024). Under the Digital Omnibus on AI, which entered into force on 27 July 2026, the obligations for these high-risk systems apply from 2 December 2027 instead of August 2026 (Lewis Silkin, 2026).

Research from Uzbekistan is still small in scale. Qosimov et al. (2025) surveyed 257 students at a Tashkent university: 72% used AI for academic purposes, interviews suggested that the time-saving argument often masked procrastination, and students’ knowledge of AI was mostly operational rather than critical. Pirniyazov (2025) found that more than 60% of 200 students at three universities used AI tools frequently and that many accepted outputs without checking them. Neither study concerns distance learners specifically, and neither proposes how teaching and assessment should be organised once AI is present.

Three gaps follow from this review. The automation frameworks are general and predate agents. The empirical studies show when AI helps or harms, but they do not translate into rules for institutions. Research on Uzbekistan describes student use but not its governance. This article addresses all three by building a task-level framework

for distance higher education and applying it to Uzbek data.

METHODOLOGY

The study followed the design science research methodology of Peffers et al. (2007), which moves through problem identification, definition of objectives, design and development, demonstration, evaluation and communication. This approach was chosen because the intended result was an artefact, a framework for allocating tasks, rather than a test of a causal hypothesis, and because the methodology requires the artefact to be justified by existing knowledge and demonstrated in a concrete setting.

The knowledge base was assembled through an integrative review. Sources were searched in Google Scholar, publisher databases and arXiv, and on the websites of HEPI, the Digital Education Council, UNESCO, the European Union and Uzbek government portals (lex.uz, gov.uz), using combinations of the terms "generative AI", "ChatGPT", "AI tutor", "randomized", "AI agents", "agentic AI", "levels of automation", "distance education" and "Uzbekistan". The period covered was 2019 to October 2026; foundational works on distance education and automation were added regardless of date. Empirical studies were included if they used a control group or surveyed at least 100 respondents. The retraction status of every empirical source was checked; one retracted meta-analysis was retained only as evidence of the fragility of early estimates. The selection was purposive rather than exhaustive, so the review is integrative, not systematic. It produced four controlled or large-sample studies of AI effects, five surveys of student AI use and eight policy and legal documents.

Data on Uzbekistan came from two sources. Enrolment by form of study, the number of teachers and the number of higher education institutions were taken from HEMIS data published in October 2026 (Ta’lim xabarlari, 2026). Education-related targets were taken from the AI strategy to 2030 (President of the Republic of Uzbekistan, 2024). Ratios such as students per teacher were calculated from these figures.

The framework was developed in four steps. First, levels of AI agency were defined by aligning the six levels of Molenaar (2022) with the four automated functions of Parasuraman et al. (2000) and with the distinction between AI agents and agentic AI proposed by Sapkota et al. (2026). Second, core tasks of distance higher education were listed, starting from the four application areas of Zawacki-Richter et al. (2019) and adding administrative, proctoring and admission functions. Third, each task was rated on three criteria: stakes (high if the AI Act places the related decision among high-risk uses, medium if it affects learning but not the student record, low otherwise), reversibility (whether an error can be corrected before it affects the student) and evidence (positive under constraints, mixed, or absent). Fourth, four decision rules were applied:

R1. If a task produces or directly determines a decision about grades, progression, admission or misconduct, AI agency is capped at level 2, and a named person takes and signs the decision.

R2. If errors are reversible and controlled evidence shows learning benefits only under pedagogical constraints, agency may reach level 3, provided the constraints are

implemented and unassisted performance is measured.

R3. If a task is low-stakes and administrative, agency may reach level 4 with periodic human audit.

R4. No task is assigned to level 5, and practices prohibited by law are excluded altogether.

In the demonstration step the framework was applied to the structure of distance higher education in Uzbekistan. Evaluation was analytical: each allocation was checked for consistency with the reviewed evidence and with the regulatory classification. The framework has not yet been tested in a field experiment. Calculations were done in Python 3.

ANALYSIS AND RESULTS

Table 1 summarises the controlled and large-sample evidence on the effects of generative AI on learning and thinking.

Table 1

Experimental and survey evidence on the effects of generative AI

Study	Setting and sample	AI condition	Main result
Kestin et al. (2025)	Introductory physics, Harvard; 194 students; randomised crossover trial	GPT-4 tutor with expert-written solutions and sequential scaffolding	Learning gains 0.73-1.3 SD larger than in-class active learning; median time 49 min against 60 min of instruction
Bastani et al. (2025)	High school mathematics, Türkiye; about 1,000 students; field experiment	GPT Base (standard chat) and GPT Tutor (learning-protective prompts)	Practice scores +48% (Base) and +127% (Tutor); without access, Base group scored 17% lower than control; harm largely avoided with Tutor
Fan et al. (2025)	University writing task; 117 students; randomised experiment	ChatGPT, human expert, writing analytics or no support	ChatGPT improved essay scores but not knowledge gain or transfer
Lee et al. (2025)	319 knowledge workers; 936 cases; survey	Everyday use of generative AI	Higher confidence in AI associated with less critical thinking
Wang & Fan (2025)	Meta-analysis of 51 studies	ChatGPT interventions	Reported g = 0.867; retracted in May 2026

In both experiments that included a pedagogically constrained AI condition, that condition produced the largest gains. The only condition that reduced later unassisted performance was unrestricted chat access. In the writing experiment, a better product was not accompanied by more learning. The AI tutor in the physics trial needed 18% less time than classroom instruction.

Table 2 shows how widely students already use AI.

Table 2

Student use of AI in higher education

Source	Sample	Use of AI	Other indicators
Digital Education Council (2024)	More than 3,800 students, 16 countries	86% in studies; 54% weekly	58% lack AI knowledge and skills; 60% worry about fairness of AI evaluation
Freeman (2025), HEPI	1,041 UK undergraduates, Dec	92% (66% in 2024)	88% use generative AI for assessments; 36% received AI

	2024		skills training; 42% say staff are well-equipped
Stephenson & Armstrong (2026), HEPI	1,054 UK undergraduates	95%	94% use AI for assessed work
Qosimov et al. (2025)	257 students, Tashkent	72% for academic purposes	Knowledge of AI mainly operational, not critical
Pirniyazov (2025)	200 students, three Uzbek universities	More than 60% use frequently	Many use outputs without verification

Use rose by 26 percentage points in one year in the UK and use for assessments by 35 points. In every survey that reported it, the share of students with training or adequate skills was well below the share using AI.

Table 3 presents the structure of higher education enrolment in Uzbekistan.

Table 3

Higher education in Uzbekistan by form of study, 2026

Indicator	Value	Share of students by form, %
Full-time students	1,126,685	62.0
Correspondence students	492,154	27.1
Evening students	108,166	6.0
Distance students	89,991	5.0
Total by form of study	1,816,996	100.0
Total reported	1,845,426	x
Correspondence and distance students	582,145	32.0
Teaching staff	51,239	x
Students per teacher	36.0	x
Correspondence and distance students per teacher	11.4	x
Higher education institutions	205	x

If each correspondence and distance student received one hour of individual consultation per month, the system would need 582,145 teacher-hours monthly, or 11.4 hours for every teacher in the country, including those who teach only full-time students. The AI strategy to 2030 sets targets for the number of AI specialists trained (from 572 to 1,000), AI laboratories in higher education and research institutions (10 by 2030) and AI professors with doctoral degrees at IT universities (40 by 2030). It contains no indicators on the use of AI in teaching, learning or assessment (President of the Republic of Uzbekistan, 2024).

Table 4 defines the levels of AI agency that form the core of the framework.

Table 4

Levels of AI agency in distance higher education

Level	What the AI does	What people do	Oversight mode	Closest level in the six-level model of Molenaar (2022)
L0. Human only	Nothing	Teacher performs all functions	Not applicable	0 (full teacher control)
L1. AI as tool	Answers or generates content on request;	Formulate every request, judge every	Human in command	1

	keeps no goals	output		
L2. AI as assistant	Prepares drafts, feedback or recommendations	Approve, edit or reject each output before use	Human in the loop	2
L3. AI as tutor agent	Conducts multi-step dialogues within a set scope, uses course materials, escalates uncertain cases	Design scope and constraints, answer escalations, review samples	Human on the loop	3
L4. AI as orchestrating agent	Plans and carries out sequences of tasks across tools and agents; keeps memory	Set goals and limits, audit logs periodically	Human audit	4
L5. Autonomous system	Sets and pursues goals without review	None	None	5 (full automation)

Generative AI as most students use it today corresponds to L1. Levels L3 and L4 describe agent-based systems; L4 corresponds to agentic AI in the sense of Sapkota et al. (2026). Moving up the scale changes what people do more than what the AI does: at L1 and L2 a person judges each output, at L3 a person designs the limits and handles exceptions, and at L4 a person sets goals and inspects records after the fact. Table 5 shows the result of applying the decision rules to nine core tasks.

Table 5

Allocation of distance learning tasks to levels of AI agency

Task	Stakes / reversible / evidence	Maximum level	Human responsibility	Main safeguards
Explaining content, answering course questions	Low / yes / positive with scaffolding	L3	Teacher curates materials, answers escalated questions	Answers grounded in course materials; hints before solutions
Practice and formative feedback	Medium / yes / positive only with constraints	L3	Teacher sets rubrics, checks a weekly sample	Unassisted checkpoints; no full solutions to graded items
Writing support	Medium / yes / mixed	L2	Student remains the author and declares AI use	Version history; short reflection on AI use
Course design and content generation	Medium / yes / limited	L2	Teacher validates all content	Fact checking; alignment with learning outcomes
Learning analytics and early warning	Medium / partly / limited	L4 for alerts; L2 for interventions	Tutor decides whether and how to contact the student	Explainable flags; no automatic sanctions
Administrative support (schedules, reminders,	Low / yes / not required	L4	Office audits logs monthly	Clear hand-off to staff

enquiries)				
Summative assessment and grading	High / no / contested	L2	Teacher assigns and signs the grade	Double marking of samples; right of appeal
Exam proctoring	High / no / not applicable	L2	A person reviews every flag	No emotion recognition
Admission and progression decisions	High / no / not applicable	L1	Committee decides	AI supplies information only

Source: developed by the author by applying rules R1-R4.

Of the nine tasks, three are high-stakes, four are medium-stakes and two are low-stakes. Controlled evidence of benefit exists for two tasks, and in both cases only under pedagogical constraints. Two of the nine tasks reach L3, one reaches L4 fully and one only for alerts, four are capped at L2, one stays at L1 and none is assigned to L5. All three task groups that correspond to high-risk uses under the AI Act are capped at L2 or below.

DISCUSSION

The study set out to determine how responsibility should be divided between people and AI agents in distance higher education. The reviewed evidence supports the proposition behind the framework. In the experiment by Bastani et al. (2025), the same underlying model improved or damaged later performance depending on the prompts around it, and the gains in Kestin et al. (2025) came from scaffolding that experts had written for each problem. For a university this shifts the main investment from access to models towards the design of the constraints, the materials and the checks that surround them. Buying licences is the easy part.

The framework also fits distance education theory. In Moore's (1993) terms, a tutor agent at L3 adds dialogue at any hour while keeping the structure of the course intact, which is why tutoring is placed there. Students in the AI condition of Kestin et al. (2025) reported higher engagement than those in class, which is consistent with this reading. The Community of Inquiry model (Garrison et al., 2000) explains why tutoring is not placed at L4: an agent that runs the course without the teacher in view might deliver content efficiently, but social and teaching presence would thin out. Molenaar (2022) argued that the potential of AI in education lies at the intermediate levels of automation. The present framework agrees for learning tasks and extends the argument to agents by allowing L4 only for low-stakes administration and for alerts that a person then acts upon.

What changes with agents is the visibility of the chain from analysis to action. A generative tool at L1 shows its answer and stops. An agent at L4 may analyse a student's activity, decide that the student is at risk and send messages, all in one pass, so an error at the start can travel unnoticed to the end. Parasuraman et al. (2000) recommended keeping decision selection and action implementation at lower levels of automation when the consequences are serious, and Shneiderman (2020) showed that high automation need not mean low control. Rule R1 and the logs, escalation paths and audits listed in Table 5 apply these ideas to education: an agent may prepare a decision, but a person takes it.

The evidence on cognitive offloading adds a requirement that is easy to overlook. If students in the writing experiment of Fan et al. (2025) produced better essays without learning more, then course outcomes measured only through AI-assisted work will overstate learning. The same mechanism appears in the drop of unassisted exam scores reported by Bastani et al. (2025) and in the link between confidence in AI and reduced critical thinking found by Lee et al. (2025). This is why rule R2 makes unassisted checkpoints a condition for L3. For Uzbek students, the operational rather than critical AI literacy reported by Qosimov et al. (2025) and the unverified use of outputs reported by Pirniyazov (2025) make this condition more pressing. The training gap in Table 2 suggests that the UNESCO competency frameworks (Miao & Cukurova, 2024; Miao et al., 2024) should be treated as a precondition for moving students and teachers beyond L1, not as an optional addition.

The retraction of the meta-analysis by Wang and Fan (2025) carries a lesson of its own. Large pooled effects travelled quickly into presentations and policy notes, while the trials that showed harm under poor design received less attention. Bond et al. (2024) had already asked for more rigour in this field. Universities deciding on AI tutors should rely on well-designed trials, treat vendor claims with caution and, where possible, evaluate pilots locally with unassisted post-tests.

For Uzbekistan the framework has three implications. First, the scale of the problem is real. With 582,145 correspondence and distance students and 36 students per teacher overall, one hour of individual consultation per student per month would cost more than eleven hours of every teacher’s time. L3 tutor agents grounded in course materials are the most realistic way to provide regular dialogue to these students, provided teachers design their scope and handle escalations. Second, national policy measures the supply of AI specialists but not the quality of AI use in teaching. Indicators such as the share of distance programmes with validated tutor agents, the share of teachers trained under the UNESCO framework and unassisted learning outcomes would close this gap. Third, although the EU AI Act does not apply in Uzbekistan, its risk classes give a ready reference for national rules, and the postponement of high-risk obligations to December 2027 leaves time to align with it. Tutor agents will have to work in Uzbek and Russian, and their quality in Uzbek should be tested locally before they are used at scale.

The framework also changes the work of teachers. At L3 a teacher spends less time answering routine questions and more time writing scaffolds, checking samples of AI feedback and handling the cases that the agent escalates. This is skilled work, and the HEPI survey suggests that most teaching staff are not yet ready for it: in 2025 only 42% of students said staff were well-equipped to work with AI (Freeman, 2025). Workload models in distance programmes will need to count this design and review time, otherwise the agent will be configured once and left unattended.

Persistent memory, which distinguishes agents from simple chat tools (Wang et al., 2024; Sapkota et al., 2026), also raises a question of data. A tutor agent that remembers a student’s errors across a semester holds a detailed record of that student’s learning. In Uzbekistan such records fall under the Law "On Personal Data" (Republic of Uzbekistan, 2019), and universities will need clear rules on what agents may store,

for how long, and who may see it. Rule R1 already excludes the use of these records for automatic sanctions; a retention limit and student access to their own records would be natural additions.

Several limitations should be noted. The review was integrative and purposive, not systematic, and the number of controlled studies is small. Two of the three experiments were conducted outside distance education, one of them in a secondary school. The framework was evaluated analytically and has not yet been tested in a field experiment. The enrolment data come from a media report of HEMIS figures with an unexplained difference of 28,430 students between the total and the sum by form of study. Finally, the decision rules simplify judgements that will sometimes require context, and the technology is changing faster than the evidence about it.

CONCLUSION

Generative AI is already part of how students learn, and agentic systems will soon be able to run sequences of teaching tasks with little supervision. The evidence reviewed here shows that the effect of these systems on learning depends mainly on how they are constrained and supervised. The proposed framework turns this finding into five levels of AI agency and four decision rules that assign each distance learning task a maximum level and a responsible person. Tutoring and formative feedback can be delegated to tutor agents under teacher oversight; grading, progression and admission remain human decisions.

The contribution of the study is a framework that connects levels of AI agency to concrete tasks of distance higher education through explicit criteria of stakes, reversibility and evidence, and that can be applied by an institution without specialised technical knowledge.

For universities in Uzbekistan the recommendations are to start distance programmes at L2 and L3 with tutor agents grounded in course materials, to keep summative assessment and progression decisions with teachers, to measure learning through unassisted checkpoints, and to train teachers and students under the UNESCO AI competency frameworks before extending AI agency. At the national level, adding indicators on AI use in teaching to the AI strategy and adopting rules for high-risk educational uses would give universities a clear basis for these decisions.

The next step is empirical: a randomised trial in Uzbek distance programmes comparing L2 and L3 tutoring, with unassisted post-tests and measurement of teacher workload, would show whether the framework delivers what the reviewed evidence suggests.

REFERENCES

1. Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences, 122*(26), e2422633122. https://doi.org/10.1073/pnas.2422633122
2. Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence

in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education, 21*, 4. https://doi.org/10.1186/s41239-023-00436-z
3. Digital Education Council. (2024). *Digital Education Council Global AI Student Survey 2024*. https://www.digitaleducationcouncil.com/post/what-students-want-key-results-from-dec-global-ai-student-survey-2024
4. European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*, L, 2024/1689.
5. Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology, 56*(2), 489-530. https://doi.org/10.1111/bjet.13544
6. Freeman, J. (2025). *Student generative AI survey 2025* (HEPI Policy Note 61). Higher Education Policy Institute.
7. Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2-3), 87-105. https://doi.org/10.1016/S1096-7516(00)00016-6
8. Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports, 15*, 17458. https://doi.org/10.1038/s41598-025-97652-6
9. Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Article 1121). ACM. https://doi.org/10.1145/3706598.3713778
10. Lewis Silkin. (2026, July 27). *The Digital Omnibus on AI enters into force today*. https://www.lewissilkin.com/insights/2026/07/27/the-digital-omnibus-on-ai-enters-into-force-today-102nedo
11. Miao, F., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO.
12. Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.
13. Miao, F., Shiohira, K., & Lao, N. (2024). *AI competency framework for students*. UNESCO.
14. Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education, 57*(4), 632-645. https://doi.org/10.1111/ejed.12527
15. Mollick, E. R., & Mollick, L. (2023). *Assigning AI: Seven approaches for students, with prompts* (arXiv:2306.10052). https://doi.org/10.48550/arXiv.2306.10052
16. Moore, M. G. (1993). Theory of transactional distance. In D. Keegan (Ed.), *Theoretical principles of distance education* (pp. 22-38). Routledge.
17. Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and

levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 30*(3), 286-297. https://doi.org/10.1109/3468.844354
18. Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45-77. https://doi.org/10.2753/MIS0742-1222240302
19. Pirniyazov, A. A. (2025). The use of artificial intelligence by students in the learning process in the context of Uzbekistan: Opportunity or risk? *International Journal of Inventive Engineering and Sciences, 12*(8), 15-20. https://doi.org/10.35940/ijies.H1119.12080825
20. Plagiarism Today. (2026, July 15). *Study claiming AI helps students learn retracted*. https://www.plagiarismtoday.com/2026/07/15/study-claiming-ai-helps-students-learn-retracted/
21. President of the Republic of Uzbekistan. (2024). *Sun’iy intellekt texnologiyalarini rivojlantirishning 2030-yilgacha bo‘lgan davrga mo‘ljallangan strategiyasini tasdiqlash to‘g‘risida* [On the approval of the Strategy for the development of artificial intelligence technologies until 2030] (Resolution No. PQ-358, 14 October 2024). https://lex.uz/docs/7159258
22. Qosimov, A., Xolmatov, A., & Rustamova, N. (2025). Navigating AI in higher education: Student perceptions, motivations, and concerns. *Philology Matters, 53*(2), 185-201. https://doi.org/10.36078/987655531
23. Republic of Uzbekistan. (2019). *Shaxsga doir ma’lumotlar to‘g‘risida* [On personal data] (Law No. ZRU-547, 2 July 2019). https://lex.uz
24. Republic of Uzbekistan. (2020). *Ta’lim to‘g‘risida* [On education] (Law No. ZRU-637, 23 September 2020). https://lex.uz
25. Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2026). AI agents vs. agentic AI: A conceptual taxonomy, applications and challenges. *Information Fusion, 126*, 103599. https://doi.org/10.1016/j.inffus.2025.103599
26. Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction, 36*(6), 495-504. https://doi.org/10.1080/10447318.2020.1741118
27. Stephenson, R., & Armstrong, C. (2026). *Student generative artificial intelligence survey 2026* (HEPI Report 199). Higher Education Policy Institute.
28. Ta’lim xabarlari. (2026, October 8). *O‘zbekistonda talabalar soni 2 millionga yaqinlashdi* [The number of students in Uzbekistan is approaching 2 million]. https://talimxabarlari.uz/80847/
29. Wang, J., & Fan, W. (2025). The effect of ChatGPT on students’ learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. *Humanities and Social Sciences Communications*. https://doi.org/10.1057/s41599-025-04787-y (Retracted May 2026)
30. Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science, 18*(6), 186345.

https://doi.org/10.1007/s11704-024-40231-1
31. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *Proceedings of the International Conference on Learning Representations (ICLR 2023)*. https://doi.org/10.48550/arXiv.2210.03629
32. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, 39. https://doi.org/10.1186/s41239-019-0171-0