AI Hub for Education

In Brief: A Forward-Looking Research Agenda for AI and Learning

Building the Evidence Base for Genuine K-12 Transformation

AI adoption is moving faster than the evidence. A recent review identified more than 1,600 studies relevant to AI in K–12 education, but only a small fraction provide strong causal evidence. System leaders are choosing tools and setting procurement, governance, staffing, and instructional policies that may shape students' experiences for years, despite research that has already documented learning losses from some uses of AI. This agenda sets out the evidence needed to guide those decisions and expand what students can become and what educators and schools can help them achieve.

AI's place in education should be earned, not assumed

Research can identify where AI improves learning and development, where it adds little or causes harm, and where human connection and judgment are irreplaceable. A finding that an AI approach does not work, or works no better than the alternative, is as valuable as a finding that it does. Evidence helps most when it arrives while decisions are still open, which requires research cycles that move closer to the pace of adoption.

Four guiding principles: What counts as progress and how to build the evidence

Most research on AI in education asks a narrow question: does this tool improve or hurt outcomes? This agenda asks a broader one: whether, and how, AI can strengthen students' long-term development, and how education systems can evolve to support it. That shift requires rethinking both what counts as success and how schools are designed to pursue it. Four guiding principles set that standard. These principles are the agenda's foundation, and they hold even as tools change, questions are answered, and hypotheses are revised.

Table 1. Four guiding principles and what they require

	Guiding principle	How it shapes the agenda	What it requires
1	**Improvement means building a broad range of students' capacities**	Sets broad, durable student development, not short-term performance alone, as the standard for improvement.	Designing research that detects effects on the full range of student capacities, along with harms such as overreliance or the displacement of human connection and judgment.
2	**Data, measurement, and evaluation are essential infrastructure**	Makes high-quality data, valid measures, and credible evaluation prerequisites for progress.	Building data on strong instruction and how students learn, shared and validated measures, and credible, independent evaluation so evidence can accumulate across settings.
3	**Progress involves scaling what works and inventing what's next**	Pursues evidence-backed approaches and emerging models in parallel.	Testing whether proven approaches retain their benefits as they reach more students, and evaluating new models over timelines long enough for them to develop.
4	**What works should reach students least well served first**	Makes equitable benefit a measure of transformation from the outset.	Examining who gains access to effective approaches and who benefits, along with the resource, role, and organizational changes needed to reach under-resourced students and schools.

Three research bets: Hypotheses about where research on AI could make a difference for students

This agenda applies the four principles to three promising research bets about how AI could address persistent challenges in education. The research bets span students' learning experiences, educators' work, and the organization of schooling. Each bet includes questions about tool design, use in schools, and the conditions that shape outcomes. Research will examine whether benefits emerge,

for whom, and at what cost, alongside potential harms. The bets depend on one another and can advance together.

Research Bet 1: How can AI help create learning experiences that strengthen the full range of student capacities?

Most personalized learning has adjusted the pace of standardized content and is optimized for what is easy to measure. Meanwhile, high-quality individualized support such as tutoring has remained scarce and unequally distributed. AI could make responsive support possible at scale without a one-to-one supply of expert adults. It could also make visible capacities that traditional assessments miss, such as persistence, curiosity, and sensemaking. Some of the priority questions are whether AI-supported learning produces durable gains across the full range of capacities, what mix of human and AI support works best for which students, and whether it builds independence rather than overreliance.

Research Bet 2: How can AI help enable fundamentally different educator roles that improve instruction and student support?

Research has produced strong evidence about effective instruction, yet it rarely reaches daily practice. One reason is that educators spend much of their time on routine feedback and administrative work. AI could take on routine tasks and surface evidence-based guidance at the moment instructional decisions are made, freeing educators for relationships, targeted intervention, and instructional design. It could also make new staffing models feasible, such as an experienced teacher leading a team with real-time AI progress monitoring. Some of the priority questions are whether time saved is reinvested in high-value work, and whether AI strengthens or erodes professional judgment.

Research Bet 3: How can AI help enable new models of schooling that improve student development and expand educational opportunity?

Age-graded classrooms, fixed schedules, and one-teacher-per-classroom staffing leave many students under-challenged or under-supported, and decades of reform have done little to change that model. AI could change how schools use their scarcest resources: adult expertise, student time, information, and money. That could make long-proposed alternatives feasible at scale, including competency-based progression, flexible learning time, team-based staffing, and integrated tiered support. Early evidence is thin and mixed. Research can compare redesigned models with incremental AI use within existing school structures, examining student development, educational opportunity, and cost. Following models over time can reveal how design, implementation, and governance shape outcomes as models develop and expand.

Moving the agenda forward

Advancing each research bet involves work on two horizons. In the near term, research can focus on approaches that have strong evidence behind them, asking whether and how AI can extend them to

more students, educators, and schools. At the same time, research can test new models of learning, teaching, and schooling that AI may make possible, where the answers may take years.

Table 2. Examples of near-term and transformational work across core domains

	Scale what works (near-term)	Invent what's next (transformational)
Learning experiences	Can AI extend high-impact tutoring and personalized support to students who have never had access?	How might AI enable new forms of instruction that build the full range of student capacities?
Educators' work	Can AI offload routine tasks and help educators use known-effective practices more consistently?	How might AI make fundamentally redesigned educator roles and staffing models workable?
School systems	Can AI optimize efficiency within existing school structures?	How might schooling itself need to evolve, and can AI make new models feasible?

No single study can answer these questions alone. Progress depends on shared measures, independent evaluation, studies that span many sites and years, and partnerships between researchers, developers, and education systems. Every study is a chance to learn what works, for whom, and under what conditions.

For the full principles, priority research questions, and illustrative study designs, see A Forward-Looking Research Agenda for AI and Learning.