Innovative Educational Technologies and Pedagogies
Vol. 1 No. 1 (2026) pp. 72-96
Online ISSN: XXXX-XXXX

Original Research

Beyond Adoption: Evidence, Values, and Governance in AI and Emerging-Technology Education

Arthit Srisawat[1]✉, Pimchanok Charoensuk[1], Nattaya Boonmee[1]
[1] Department of Educational Technology and Communications, Faculty of Education, Chulalongkorn University, Bangkok 10330, Thailand
✉ arthitssrisawat@gmail.com

10.65840/ietp.vxix.xx

Article Info
Submitted:
11/25/2025
Revised:
02/20/2026
Accepted:
04/20/2026
Available Online:
04/24/2026

Abstract*. Artificial intelligence and emerging technologies are expanding through education faster than the field can establish which uses are pedagogically effective, normatively acceptable, and institutionally sustainable. This study conducts a prospective thematic synthesis of a Scopus export containing ten records indexed for 2027 and available on 25 July 2026. One non-educational record was excluded, leaving nine publications. Metadata, available abstracts, and accessible full texts were coded by technology family, educational level, research design, evidence maturity, institutional reach, teacher mediation, governance orientation, and normative purpose. The corpus divided into three equally sized clusters: AI and large language model integration; broad digital transformation, Industry 5.0, and sustainability; and domain-specific enactments involving technology diffusion, immersive simulation, computational thinking, and health education. Five studies were empirical or practice-oriented, two were reviews, and two were conceptual or speculative. Higher and professional education accounted for five studies, whereas primary and secondary education appeared only once each. The most consistent cross-study pattern was a movement beyond adoption: every included publication connected technology to a non-technical purpose such as agency, safety, academic integrity, cultural values, motivation, public health, or the Sustainable Development Goals. However, the frontier remained methodologically immature, with no longitudinal, comparative, or large-scale implementation evaluation. The article develops the Responsible Educational Technology Orchestration framework, which requires four jointly adequate conditions before scale: credible evidence, pedagogical mediation, value alignment, and accountable governance. This framework converts a fragmented horizon scan into a stage-gated research and implementation agenda for educational institutions and policymakers.*

Keywords*: artificial intelligence in education; emerging technologies; responsible innovation; educational technology governance; thematic synthesis; Industry 5.0*

1. Introduction

Artificial intelligence, large language models, immersive environments, learning analytics, and automated assessment have moved from peripheral experiments to visible components of educational strategy. Their spread is occurring across institutions whose pedagogical capacity, data governance, infrastructure, and regulatory safeguards vary substantially. Global guidance therefore increasingly distinguishes technological availability from educationally justified use. UNESCO's human-centred guidance calls for institutional validation, age-appropriate protection, and the preservation of human agency, while the OECD frames effective digital transformation as an ecosystem problem involving infrastructure, professional competence, interoperability, evidence, and trust (Miao & Holmes, 2023; OECD, 2023, 2026). Research reviews confirm both the rapid expansion and the fragmentation of the field. Earlier artificial-intelligence-in-education scholarship was concentrated in technical and higher-education settings, often with limited participation from educators and weak attention to ethics or organisational implementation (Zawacki-Richter et al., 2019). More recent reviews identify broader applications in teaching, assessment, feedback, prediction, and learner support, but also persistent uncertainty about causal learning effects, transfer, transparency, and equity (Chiu et al., 2023; Crompton & Burke, 2023). A meta-review of higher-education research similarly found substantial duplication among reviews, uneven methodological rigour, and an urgent need for stronger ethics and collaboration (Bond et al., 2024). Rapid reviews of generative AI add a further warning: plausible affordances are frequently inferred from tool capability or short-term perception data rather than demonstrated through durable learning outcomes (Lo, 2023). The central challenge is thus no longer whether educational technology will be adopted, but whether institutions can distinguish compelling novelty from robust educational value.

Conventional adoption theories explain only part of this transition. The Technology Acceptance Model links uptake to perceived usefulness and ease of use, diffusion theory explains how innovation spreads through social systems, and the Unified Theory of Acceptance and Use of Technology incorporates performance expectancy, effort expectancy, social influence, and facilitating conditions (Davis, 1989; Rogers, 2003; Venkatesh et al., 2003). These frameworks remain useful for predicting intention and early use, but they do not determine whether a technology improves learning, redistributes professional authority, or produces acceptable institutional consequences. Educational implementation also depends on teachers' capacity to align technological, pedagogical, and disciplinary knowledge, as formalised in the TPACK framework (Mishra & Koehler, 2006). Decades of technology-integration research show that access alone rarely changes practice because first-order barriers such as time, equipment, and support interact with second-order barriers involving beliefs, confidence, curriculum, and professional identity (Ertmer, 1999; Hew & Brush, 2007). These problems intensify with generative AI because students may adopt tools before institutions establish shared rules, assessment redesign, or reliable forms of teacher oversight. Historical analyses caution that AI in education repeatedly cycles through strong promises of personalisation and efficiency while underestimating institutional complexity and the political character of automation (Williamson & Eynon, 2020). Critical scholarship likewise argues that educational technology should be examined through questions of power, labour, dependency, and public purpose rather than treated as an autonomous driver of improvement (Selwyn, 2019). An adoption-centred account is therefore analytically incomplete: it measures entry into use but not the quality of educational orchestration after entry.

A stronger analytical approach must connect evidence, pedagogy, values, and governance. Human-centred AI requires systems that remain reliable, safe, understandable, and subject to meaningful human control (Shneiderman, 2020).

Responsible-innovation theory adds anticipation, reflexivity, inclusion, and responsiveness, while value-sensitive design requires human values to be examined throughout design and deployment rather than appended after technical development (Friedman et al., 2008; Stilgoe et al., 2013). In education, these principles imply that learning objectives, teacher judgement, student agency, cultural context, accessibility, data rights, and institutional accountability must be treated as design requirements. The field has begun to move in this direction: community-wide ethical frameworks identify obligations across researchers, developers, institutions, and users; policy models connect pedagogy with governance and operations; and explainable-AI research emphasises the need for interpretations that are meaningful to educational stakeholders (Chan, 2023; Holmes et al., 2022; Khosravi et al., 2022). The European Industry 5.0 agenda similarly reframes innovation around human-centricity, sustainability, and resilience, while ethical guidelines for educators prioritise transparency, fairness, privacy, and informed professional judgement (European Commission, 2021, 2022). UNESCO's global technology report and AI competency frameworks reinforce the same direction by asking whose terms govern technology, who benefits, and which human capacities are strengthened or displaced (UNESCO, 2023, 2024). Against this background, the small set of publications indexed ahead of their 2027 issue dates offers a useful horizon signal. This article asks: (1) which thematic and methodological configurations characterise this emerging frontier; (2) how mature and institutionally extensive is its evidence; and (3) what integrated framework can guide responsible research and implementation? It contributes a prospective thematic synthesis and develops the Responsible Educational Technology Orchestration (RETO) framework, positioning adoption as the beginning rather than the endpoint of educational innovation.

2. Conceptual Framework

2.1 From technology acceptance to educational orchestration

Technology acceptance concerns whether an individual or organisation is willing and able to use an innovation. Educational orchestration concerns how multiple components are deliberately coordinated so that use produces defensible educational and public value. The distinction matters because an accepted tool can remain pedagogically superficial, ethically problematic, or operationally fragile. Direct access may reduce time and cost, for example, yet still require rules for academic integrity, user support, and appeals. An immersive simulation may increase engagement but fail to improve transfer if the activity is disconnected from debriefing and assessment. A language model may support ideation while simultaneously obscuring authorship, amplifying bias, or displacing productive struggle. Orchestration therefore expands the unit of analysis from the user-tool dyad to the interaction among tasks, teachers, learners, curricula, data, institutional rules, and social values. It is not synonymous with central control. Rather, it refers to the transparent allocation of responsibility across human and technical actors, with mechanisms for adaptation when evidence or conditions change.

This expanded unit of analysis is consistent with research showing that educational technology becomes effective through pedagogical design rather than technical presence. TPACK identifies teacher knowledge as a situated integration of content, pedagogy, and technology, while active-learning research demonstrates that learning improves when students engage in structured cognitive work rather than merely receive information (Freeman et al., 2014; Mishra & Koehler, 2006). Self-determination theory adds that motivation depends on autonomy, competence, and relatedness, which means that convenience or novelty cannot substitute for meaningful learner agency (Deci & Ryan, 2000; Ryan & Deci, 2020). Digital tools can support these needs when they provide useful feedback, multiple representations, accessible participation, and opportunities for collaboration; they can undermine them when automation narrows choice, creates opaque surveillance, or reduces teachers to compliance monitors. Research on students' use of digital technologies similarly finds that perceived usefulness depends on the relationship between tools, study practices, and institutional context rather than on technological sophistication alone (Henderson et al., 2017). Educational orchestration therefore treats pedagogy as a mediating mechanism and specifies that technology must be judged against the quality of the activity system it enables.

2.2 Evidence maturity, institutional reach, and normative purpose

Three analytical dimensions are required to compare a heterogeneous emerging literature. Evidence maturity concerns the distance between an idea and a robust claim. Conceptual proposals and reviews can identify possibilities and risks, but they cannot establish effectiveness in a specific setting. Pilot studies can demonstrate feasibility, yet small samples, short exposure, and local novelty effects limit transfer. Comparative, longitudinal, and multi-site designs provide stronger evidence about durability, mechanisms, and variation. Institutional reach concerns the level at which an innovation operates: learner or course, programme, institution, or system. A tool may be effective at course level but require new procurement, data, staffing, and accountability arrangements when scaled. Normative purpose concerns the values used to justify or constrain implementation. These may include learning, safety, academic integrity, inclusion, cultural fit, motivation, sustainability, or public health. The three dimensions should not be collapsed into a single quality score. A conceptually sophisticated system-level framework can have broad reach but limited empirical maturity, while a carefully tested course intervention can have strong local evidence but narrow institutional relevance.

The distinction between evidence and reach prevents two common errors. The first is premature scaling: an institution generalises from a successful local pilot without testing implementation variability, opportunity costs, or unintended effects. The second is evidentiary dismissal: policymakers ignore conceptual and qualitative work because it does not estimate an effect size, even when such work reveals values, responsibilities, and failure modes that experiments are not designed to capture. Responsible research needs a portfolio in which speculative inquiry anticipates possible futures, stakeholder studies identify acceptable conditions, pilots test feasibility, and longitudinal evaluations assess sustained effects. Design fiction can surface tacit assumptions before systems are built; teacher interviews can identify constraints hidden from technical designers; reviews can map recurrent concerns; and trials can estimate learning or performance outcomes. The problem is not methodological diversity but disconnection among methods. A mature field should show cumulative movement across stages, with later studies testing propositions generated by earlier work.

2.3 The RETO framework

The Responsible Educational Technology Orchestration framework integrates four conditions. Evidence asks whether claims are valid, educationally meaningful, robust across contexts, and transferable beyond novelty. Pedagogy asks whether the tool supports a clear learning task, preserves productive cognitive activity, enables teacher mediation, and aligns assessment with intended outcomes. Values ask whether agency, fairness, accessibility, cultural legitimacy, safety, and integrity are embedded in design and use. Governance asks who controls data, how systems are procured and audited, which responsibilities remain human, how grievances are addressed, and under what conditions scaling can be reversed. These conditions are sequential for decision-making but recursive in practice. Weak evidence should halt scale; pedagogical failure should trigger redesign; value conflicts should change objectives or safeguards; governance failures should restrict or discontinue deployment. Monitoring then returns implementation evidence to the next design cycle. RETO consequently treats technology as a socio-technical intervention whose educational legitimacy depends on joint adequacy rather than on isolated technical performance.

2.4 Analytical propositions and failure modes

RETO yields three analytical propositions for emerging-technology research. First, evidence maturity and institutional reach should be treated as independent dimensions. A system-level proposal can be theoretically important while remaining empirically untested, whereas a well-designed course pilot can generate credible local evidence without supporting institutional scale. Second, normative breadth should increase rather than decrease as reach expands. A classroom experiment may legitimately concentrate on learning and usability, but institution-wide implementation must also address access, workload, data protection, procurement, appeals, and public accountability. Third, teacher mediation should be examined as a causal and organisational mechanism rather than reported as background context. Teachers select tasks, frame acceptable use, interpret outputs, provide feedback, protect productive struggle, and identify when automation has displaced rather than supported learning. These propositions establish a disciplined progression from exploratory possibility to bounded efficacy and then to responsible scale. They also prevent the common inference that a positive local outcome automatically validates the surrounding platform, policy, or business model.

The framework distinguishes four corresponding failure modes. Evidentiary failure occurs when claims exceed the design, such as inferring durable learning from satisfaction or immediate task completion. Pedagogical failure occurs when a technically capable system weakens the cognitive, dialogic, or disciplinary processes through which learning develops. Value failure occurs when benefits are produced through unacceptable trade-offs in agency, integrity, safety, cultural legitimacy, or inclusion. Governance failure occurs when an institution cannot explain decisions, control data, allocate responsibility, manage incidents, or exit a technological dependency. These failures can coexist and can be concealed by successful adoption. High usage may reflect convenience, compulsory policy, or lack of alternatives rather than educational value; favourable attitudes may coexist with declining authorship clarity or increasing teacher verification work. RETO therefore directs evaluation toward the complete arrangement of people, tasks, technologies, rules, and resources. A technology becomes educationally responsible not when every risk disappears, but when relevant risks are visible, proportionate safeguards exist, decisions can be contested, and implementation can be revised or stopped in response to evidence.

3. Methodology

3.1 Design and corpus construction

The study used a prospective thematic synthesis, designed for a small, current, and conceptually heterogeneous corpus. It was not presented as a conventional bibliometric study because the number of records was insufficient for stable co-citation, co-authorship, or keyword-network inference. The data source was a Scopus CSV export supplied for this project and dated 25 July 2026. It contained ten records indexed with a 2027 publication year, including articles and conference papers available online ahead of issue assignment. The analysis treated these records as an emerging-frontier snapshot: a bounded signal of topics entering indexed publication rather than a representative sample of all educational-technology research. Exact titles, authorship, source information, document type, page range, DOI, and indexing status were retained from the export. Source pages, DOI landing pages, available abstracts, and accessible full texts were consulted to clarify study purpose, educational setting, method, and stated contribution.

Screening followed relevance rather than journal ranking. Records were included when education was the substantive domain and technology, digital innovation, artificial intelligence, computational thinking, or technology-enabled curriculum was central to the research problem. One article on United States soft power in India was excluded because its presence resulted from broad search retrieval rather than conceptual relevance. Nine studies remained. No record was excluded because of method, country, or document type, since methodological and geographic diversity were themselves analytical interests. The final corpus comprised five conference papers and four journal articles. This inclusion strategy maximised sensitivity to an emerging frontier while preserving a clear boundary around educational innovation.

Table 1. Included studies and primary analytical characteristics

Study	Educational setting	Technology or innovation	Design	Primary contribution
Pees et al. (2027)	Cross-sector	AI-enhanced education and design fiction	Value-oriented design evaluation	Makes values explicit before implementation
Kumar et al. (2027)	Doctoral / higher education	Direct diffusion of plagiarism-detection tools	Empirical survey and	Links access structure to integrity, time, and cost

Study	Educational setting	Technology or innovation	Design	Primary contribution
			comparative process analysis	
Narala et al. (2027)	Nursing education	360° immersive virtual reality	Application redesign and pilot testing	Tests feasibility for patient-safety training
Reichert et al. (2027)	Secondary education	LLM integration	Teacher-informed framework	Specifies conditions for classroom integration
Pehlić & Tufekčić (2027)	Cross-sector	Intelligent systems and Industry 4.0/5.0	Conceptual pedagogical analysis	Connects innovation to human-centred development
Angraini et al. (2027)	Pre-service teacher education	Computational thinking in geometry	Pretest-posttest intervention with qualitative support	Integrates cognitive and religious values
Farrokhnia et al. (2027)	Higher education	Student use of AI	Scoping review	Synthesises perceived benefits and concerns
Thinwiangthong et al. (2027)	Primary education	Health curriculum supported by evidence mapping	Bibliometric and mixed-methods needs analysis	Aligns curriculum design with public health and SDGs
Petrukha et al. (2027)	Higher education	Digital and innovative pedagogies	Literature review	Connects motivation, active pedagogy, inclusion, and SDGs

Note. The corpus represents records indexed for 2027 and available by the data cut-off. The classifications are analytical summaries based on metadata and available source content.

The corpus is sufficiently diverse to reveal directional patterns but too small to support claims about prevalence in the broader field. Its value lies in the juxtaposition of studies that are usually discussed separately: generative AI, direct technology diffusion, immersive simulation, computational thinking, values, public-health curriculum, and sustainable digital pedagogy. Analysing them together makes it possible to identify whether a shared implementation logic is emerging beneath technological heterogeneity.

3.2 Coding and analytical procedure

A structured codebook was developed deductively from the conceptual framework and refined through a second reading of the corpus. Each study was coded for technology family, educational level, stakeholder focus, research design, evidence maturity, institutional reach, teacher mediation, governance orientation, and normative purpose. Technology families were non-exclusive because a study could combine, for example, artificial intelligence with design methods or digital pedagogy with sustainability. Educational level used the primary setting reported by the study. Research design distinguished conceptual or speculative work, evidence synthesis, qualitative or framework development, survey or process analysis, pilot or intervention, and mixed-methods needs analysis. Teacher mediation was coded when teachers, teacher education, instructional design, or professional judgement formed an explicit part of the proposed mechanism. Governance was coded when the study addressed policy, access structures, institutional rules, data, accountability, or scaling.

Table 2. Analytical codebook

Dimension	Operational question	Categories	Interpretive function
Technology family	What form of innovation is central?	AI/LLM; immersive; broad digital; direct diffusion; computational or disciplinary integration	Identifies thematic concentration and diversity
Evidence maturity	How far has the claim moved from proposition to robust evaluation?	1 = speculative; 2 = conceptual/review; 3 = framework/qualitative; 4 = pilot/intervention; 5 = comparative/longitudinal	Locates evidentiary development without ranking overall quality
Institutional reach	At what organisational level would implementation operate?	1 = learner; 2 = course; 3 = programme; 4 = institution; 5 = system/policy	Separates local feasibility from scaling demands
Mediation and governance	Are teachers and institutional rules part of the mechanism?	Teacher mediation; access and support; data and accountability; policy or scaling	Tests whether technology is treated as a socio-technical intervention
Normative purpose	Which non-technical value justifies or constrains use?	Agency; learning; safety; integrity; inclusion; cultural fit; motivation; sustainability; public health	Reveals the shift from adoption to responsible use

Note. Evidence-maturity and institutional-reach scores are structured interpretive codes. They are not bibliometric impact measures or journal-quality ratings.

Coding was conducted in two passes. The first pass assigned descriptive categories from the study title, abstract, method, and contribution. The second pass checked consistency against the operational definitions and reconciled ambiguous classifications by privileging the study's actual evidence over its aspirational language. For example, a literature review proposing system change was coded as broad in reach but conceptual in evidence maturity; a local intervention reporting post-test improvement was coded as stronger in local evidence but narrow in reach. This rule prevented policy ambition from being mistaken for implementation evidence.

Three forms of analysis were performed. First, descriptive counts summarised document type, educational level, research design, and thematic cluster. Second, a study-by-dimension matrix identified recurring combinations among technology, empirical evaluation, teacher mediation, governance, and normative purpose. Third, studies were positioned on a two-dimensional maturity-reach map, with bubble size representing normative breadth. The scoring scale was intentionally coarse to avoid false precision. The synthesis then compared the locations of studies and identified empty areas, particularly high-reach research with mature comparative evidence. Finally, the recurring dimensions were integrated into the RETO framework and converted into a stage-gated research agenda.

The analysis is interpretive and non-causal. It cannot estimate technology effects across education systems, and the small corpus prevents statistical generalisation. Some records were analysed from abstracts and metadata where full text was not openly available, although source claims were not extended beyond accessible information. Future-dated indexing may also reflect publication workflows rather than a complete view of 2027 scholarship. These limitations are integral to the design: the article provides a transparent horizon scan and conceptual synthesis, not a systematic review of the entire literature. Ethical approval was not required because only publicly indexed bibliographic information and published research were analysed; no personal or confidential data were used.

3.3 Validity, transparency, and reproducibility

Several procedures were used to strengthen interpretive validity. Bibliographic fields were normalised before coding, and DOI or publisher records were used to verify titles, publication status, and source identity. Where a full text was available, coding was checked against the method and results rather than inferred from the title. Where access was limited to metadata and an abstract, the code was restricted to claims explicitly supported by those materials. The evidence-maturity score described the design actually reported, not the sophistication of the technology or the prestige of the source. Institutional reach was likewise based on the level implicated by the proposed intervention, while the analysis explicitly separated intended reach from demonstrated implementation. A study could therefore receive broad reach and modest maturity without contradiction. This separation reduced halo effects and made the map analytically useful rather than evaluative of author or journal quality.

Reproducibility was supported through an auditable study table, operational definitions, explicit exclusion reasoning, and transparent derived counts. The synthesis did not calculate a pooled effect because outcomes, populations, technologies, and designs were not commensurable. It also avoided treating citation counts as quality indicators because future-dated records had unequal exposure time. Sensitivity was examined by asking whether the principal conclusion changed when conceptual and review papers were considered separately from empirical studies. It did not: the empirical subset still showed bounded settings and limited follow-up, while the conceptual subset still showed high institutional ambition without system evaluation. The small corpus remains a limitation, but its boundedness allows close comparison of assumptions that would be obscured in a large automated mapping exercise. The resulting claims are therefore analytical propositions about an emerging configuration, not prevalence estimates for global educational-technology scholarship.

4. Results

4.1 A heterogeneous frontier with three converging clusters

The nine studies formed three clusters of equal size. The first centred on AI and LLMs: value-oriented design fictions, a teacher-informed framework for secondary-school integration, and a scoping review of higher-education students' perceptions and concerns (Farrokhnia et al., 2027; Pees et al., 2027; Reichert et al., 2027). The second addressed broad transformation and sustainability through Industry 4.0/5.0, digitally supported motivation in higher education, and an SDG-oriented health curriculum for primary schools (Pehlić & Tufekčić, 2027; Petrukha et al., 2027; Thinwiangthong et al., 2027). The third examined domain-specific enactment: direct diffusion of plagiarism-detection software, 360° virtual reality for patient-safety training, and computational thinking integrated with Islamic values in geometry learning (Angraini et al., 2027; Kumar et al., 2027; Narala et al., 2027). The equal distribution is analytically significant. It shows that the emerging frontier is not reducible to generative AI; it combines general-purpose systems with disciplinary and organisational innovations.

Educational settings were unevenly distributed. Five studies focused on higher or professional education, two were cross-sector conceptual contributions, one addressed secondary education, and one addressed primary education. Early childhood, technical and vocational education, adult basic education, special education, and non-formal learning were absent. This concentration reproduces a long-standing pattern in which digitally intensive research appears first in institutions with stronger infrastructure, research capacity, and discretionary autonomy (Crompton & Burke, 2023; Zawacki-Richter et al., 2019). The corpus nevertheless showed movement beyond generic university surveys. Nursing, doctoral research integrity, pre-service teacher education, secondary computing education, and endemic-disease prevention each located innovation within a concrete practice problem. Domain specificity therefore increased even as system coverage remained narrow.

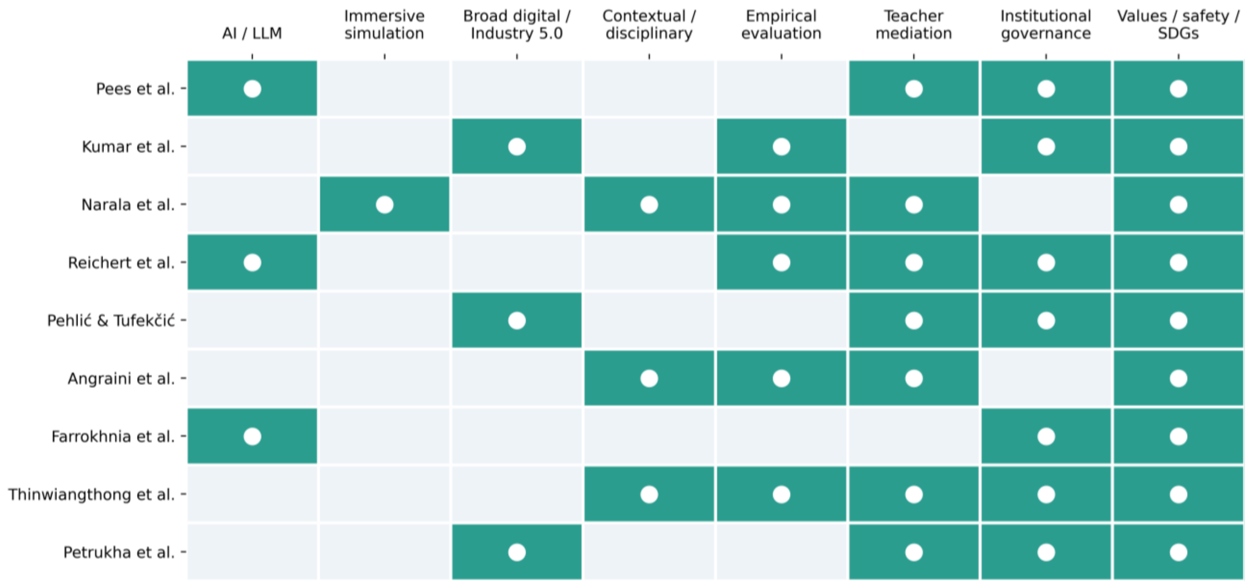

Figure 1. thematic and implementation architecture of the emerging research frontier

Figure 1 shows that normative purpose was the only dimension present across the entire corpus. Empirical evaluation appeared in five studies, teacher mediation in six, and institutional governance in six. AI/LLM studies were consistently connected to values or governance, whereas immersive and contextual studies were more strongly connected to direct empirical testing. The matrix therefore suggests a division of labour: general-purpose AI research is currently clarifying acceptable conditions, while domain-specific technologies are more often being piloted within bounded tasks. A mature field would need these trajectories to converge, producing AI studies with robust classroom evidence and domain pilots with explicit governance and equity analysis.

Publication form reinforced this division. Conference papers carried much of the anticipatory and design-oriented work, enabling rapid circulation of frameworks, values, and early prototypes, whereas journal articles more often provided extended empirical or review-based accounts. This is not a quality hierarchy; it reflects different functions in a fast-moving research ecology. Rapid venues are valuable when models and policy conditions change quickly, but their propositions require later replication and fuller methodological reporting. Journal publication can provide that depth, yet longer production cycles may leave technology descriptions obsolete unless studies report mechanisms and implementation conditions that remain interpretable across versions. Across both forms, the strongest contributions were those that specified an educational problem before describing the technology. Patient safety, academic integrity, geometry understanding, public-health education, and student motivation generated clearer outcome logics than generic claims about transformation. The finding suggests that future research should organise comparison around educational functions and mechanisms, not merely around branded tools.

4.2 Evidence maturity lags behind institutional ambition

The methodological profile contained five empirical or practice-oriented studies, two reviews, and two conceptual or speculative contributions. This composition may appear balanced, but the empirical studies were largely local and short-term. The 360° virtual-reality study piloted a redesigned safety-training application; the computational-thinking study tested a bounded instructional intervention; the health-curriculum study combined bibliometric mapping with stakeholder needs analysis; the direct-diffusion study examined a specific academic-integrity process; and the LLM framework was grounded in computing-teacher insights. These designs are useful for feasibility, mechanism identification, and contextual validity. None, however, tested multi-site implementation, long-term retention, comparative cost-effectiveness across institutions, or sustained organisational change. The reviews widened thematic coverage but depended on heterogeneous primary evidence, while the conceptual papers articulated values and system directions without implementation tests.

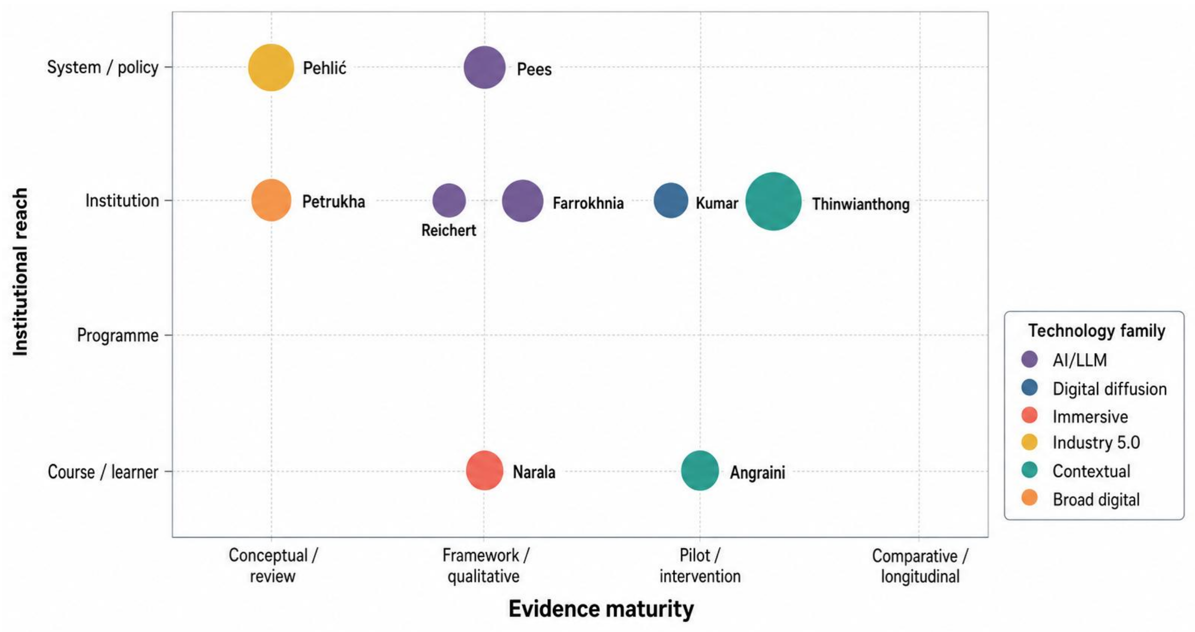

Figure 2. Evidence maturity, institutional reach, and technological orientation of the included studies.

The maturity-reach map makes the imbalance visible. Several studies occupy the institution or system level in ambition, but no study reaches the comparative or longitudinal evidence category. The empty upper-right area represents the field's central deficit: research capable of supporting system-scale decisions with robust evidence. This is not a reason to discount the existing work. Speculative and review studies are necessary for anticipating risks, while local pilots identify design constraints. The problem is the lack of an explicit cumulative pathway connecting these forms. Without replication and follow-up, promising results remain isolated demonstrations; without prospective governance research, system frameworks remain principled but untested.

Outcome measurement was also narrower than the purposes invoked. The empirical studies offered useful indicators of feasibility, conceptual understanding, perceptions, process efficiency, or curriculum need, but the corpus rarely combined immediate performance with retention, transfer, equity, workload, and cost. This creates a measurement asymmetry: institutional claims are multidimensional, while evaluation remains concentrated on one or two proximal outcomes. For AI-supported work, a credible design would distinguish assistance from learning by assessing unaided performance, delayed recall, transfer to novel tasks, and the learner's ability to explain or critique generated output. For immersive simulation, evaluation should connect presence and usability to authentic performance and safety behaviour. For broad digital transformation, indicators should include teacher workload, participation gaps, service reliability, and opportunity costs. The missing measures do not invalidate the studies; they identify the next evidentiary step. A shared minimum outcome architecture would make replication cumulative while still permitting disciplinary specificity.

4.3 The normative turn: from tool capability to educational purpose

Every included study justified innovation through a non-technical purpose. Pees et al. (2027) evaluated the values embedded in imagined AI-enhanced educational futures. Reichert et al. (2027) positioned computing teachers as interpreters of classroom feasibility rather than passive recipients of LLM policy. Farrokhnia et al. (2027) synthesised student benefits and concerns, making perception and legitimacy part of implementation. Kumar et al. (2027) evaluated direct diffusion through academic integrity, time, cost, and user empowerment. Narala et al. (2027) anchored immersive technology in patient safety, where simulation quality has consequences beyond engagement. Angraini et al. (2027) treated computational thinking as compatible with moral and religious formation rather than as a culturally neutral skill. The remaining studies linked innovation to human-centred Industry 5.0, student motivation, sustainable development, and public-health prevention. Technology was thus consistently framed as a means to an educational or public end.

This normative turn has two interpretations. Optimistically, the field is responding to criticism that educational technology has been driven by capability and market availability rather than by public purpose. Values, safety, integrity, and sustainability are entering research questions earlier. More cautiously, normative language can become ceremonial when it is not connected to measures, trade-offs, or decision rules. A study may invoke inclusion without disaggregating access, or cite sustainability without estimating energy, cost, maintenance, or institutional durability. The included studies rarely operationalised conflicts among values. Direct access can improve efficiency but weaken institutional oversight; automation can support feedback but reduce authorship clarity; local cultural integration can increase relevance while complicating transfer; immersive simulation can improve safety preparation but impose equipment and accessibility costs. The next research stage must therefore move from naming values to specifying how they are prioritised, measured, and governed.

Teacher agency emerged as the principal bridge between normative purpose and implementation. Six studies explicitly located teachers, teacher educators, or instructional designers within the mechanism, and the remaining studies still implied professional decisions about task selection, feedback, integrity, or curriculum alignment. This pattern challenges two symmetrical errors: treating teachers as barriers to innovation and treating them as final-stage implementers of decisions made elsewhere. In a responsible model, teachers contribute problem definition, contextual judgement, learner support, and detection of unintended effects. Their participation is particularly important where AI outputs are probabilistic, where multilingual or cultural interpretation matters, and where assessment decisions affect progression. Yet teacher mediation also has costs. Verification, prompt design, redesign of assessment, incident management, and student guidance can create substantial hidden labour. Research that celebrates efficiency without measuring this redistribution may overstate institutional benefit. Professional development should therefore be evaluated not only by confidence or acceptance but by whether teachers can diagnose appropriate and inappropriate use, preserve disciplinary standards, and exercise meaningful authority over implementation.

4.4 The RETO framework and the missing implementation chain

The synthesis generated a four-condition implementation chain. Evidence requires more than positive perception or immediate performance; it includes validity, learning gain, robustness, retention, transfer, and distributional effects. Pedagogy requires a defensible task structure, teacher mediation, learner activity, feedback, and assessment alignment. Values require explicit attention to agency, integrity, safety, inclusion, culture, motivation, and sustainability. Governance requires data rules, procurement standards, professional development, human accountability, appeal mechanisms, and reversible scaling. The sequence is stage-gated because failure in any condition can invalidate expansion. A technically accurate system without pedagogical purpose should not scale; an effective tool that produces unacceptable surveillance or exclusion should be redesigned; a value-aligned pilot without institutional capacity should remain bounded until governance is adequate.

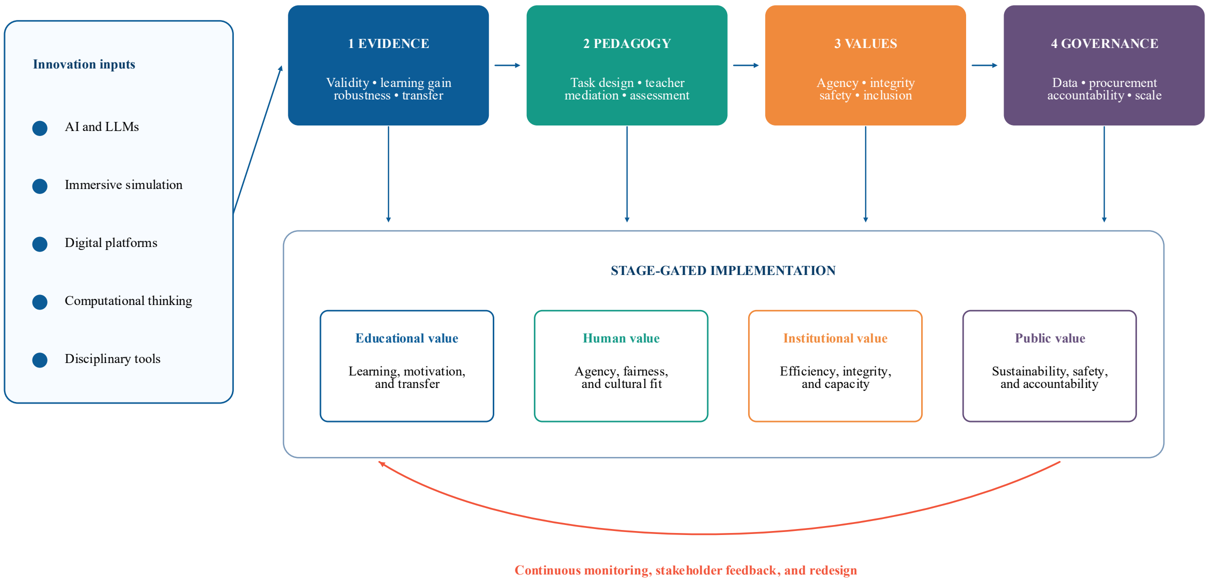

Figure 3. The Responsible Educational Technology Orchestration (RETO) Framework

The framework also clarifies why the included studies are complementary. Design fiction contributes anticipation to the values stage; teacher-informed LLM frameworks connect pedagogy with governance; direct diffusion studies reveal how access structures alter outcomes; immersive and computational-thinking pilots provide evidence about bounded pedagogies; student-perception reviews identify legitimacy concerns; and SDG-oriented curriculum studies connect innovation to public value. No single study covers the full chain. RETO is therefore not a summary score for individual articles but a programme architecture that links different types of research into a cumulative implementation process.

Table 3. Corpus profile and principal gaps

Analytical feature	Observed pattern	Interpretation	Priority gap
Educational level	5 higher/professional; 2 cross-sector; 1 secondary; 1 primary	Innovation remains institutionally concentrated	Early childhood, TVET, inclusive, adult, and non-formal settings
Research design	5 empirical/practice-oriented; 2 reviews; 2 conceptual	Diverse methods but weak cumulative sequencing	Replication, comparison, retention, and longitudinal implementation
Technology cluster	3 AI/LLM; 3 transformation/SDG; 3 domain-specific enactment	General-purpose and disciplinary technologies are co-evolving	Cross-technology comparisons using common educational outcomes
Normative orientation	All 9 studies include a non-technical purpose	The field is moving beyond adoption	Operational measures for values, trade-offs, and accountability
Institutional reach	Several institution/system proposals; no mature high-reach evaluations	Ambition exceeds evidence maturity	Multi-site governance and cost-effectiveness studies

Note. Counts refer to the nine-study horizon corpus and should not be generalised as prevalence estimates for the wider literature.

Table 3 converts the descriptive findings into a research agenda. The recurring deficit is not a lack of innovation but a missing chain from concept to sustained institutional value. Stronger research should follow technologies across stages, preserve common outcome definitions, document implementation cost and professional workload, and report which groups benefit or are burdened. This would allow the field to compare not only whether tools work, but also where, for whom, under which governance arrangements, and at what opportunity cost.

Table 4. Stage-gated research and implementation agenda

Gate	Minimum question	Indicative evidence	Decision	Failure response
1. Evidence	Does use improve a defined educational outcome?	Validated measures; comparison; retention; subgroup analysis	Proceed to bounded implementation	Revise claim, task, or tool
2. Pedagogy	Is the mechanism aligned with learning and assessment?	Task analysis; teacher observation; learner process data	Integrate into course or programme design	Redesign activity and professional support
3. Values	Are agency, fairness, safety,	Stakeholder deliberation;	Authorise with explicit safeguards	Restrict use or change objectives

Gate	Minimum question	Indicative evidence	Decision	Failure response
	integrity, and culture protected?	accessibility and bias audit; scenario testing		
4. Governance	Can the institution operate, monitor, and reverse the system responsibly?	Data and procurement review; cost model; accountability map; incident protocol	Scale gradually with monitoring	Pause, contain, or discontinue

Note. Gates are cumulative. Passing one gate does not compensate for failure at another, and monitoring continues after implementation.

The agenda is deliberately conservative about scale but not about experimentation. It supports iterative pilots, disciplinary adaptation, and participatory design while requiring claims to remain proportional to evidence. By making pause and discontinuation legitimate outcomes, it also counters the sunk-cost logic that often turns pilot adoption into permanent infrastructure before educational value has been established.

4.5 An integrated programme for cumulative research

The corpus supports a sequenced programme rather than a single preferred methodology. Anticipatory studies should first identify plausible benefits, affected groups, value conflicts, and failure scenarios. Co-design and qualitative inquiry should then specify tasks, professional roles, and acceptable conditions with teachers and learners. Bounded pilots should test usability, process, immediate learning, and implementation burden against a meaningful comparison. Replication should examine whether the mechanism survives changes in institution, discipline, language, infrastructure, and learner profile. Only after these stages should research evaluate programme- or system-level effects, including retention, transfer, equity, workload, cost, procurement, data incidents, and organisational adaptation. This sequence does not require every project to conduct all stages. It requires projects to state their position in the chain and generate outputs usable by the next stage. Shared protocols, preregistered outcome definitions, open instruments, and structured reporting of null or adverse findings would reduce repeated pilotism and improve the field's capacity to learn across rapidly changing tools.

A common measurement core could provide continuity across technologies while preserving disciplinary outcomes. At minimum, studies should report the educational target, baseline condition, nature and intensity of use, human mediation, comparison condition, immediate outcome, delayed or transfer outcome where feasible, subgroup participation, professional workload, material cost, and governance events. AI studies should additionally document model version, access mode, prompting or scaffolding conditions, and whether assessment was completed with or without assistance. Immersive studies should report accessibility, adverse physical effects, equipment reliability, and transfer to authentic performance. Institutional evaluations should record training, support demand, procurement constraints, data flows, complaints, and exit feasibility. This architecture would allow policymakers to compare complete interventions rather than isolated effect sizes. It would also make responsible innovation empirically testable: value alignment would be observed through access, agency, integrity, and safety indicators, while governance quality would be observed through traceability, response capacity, and reversibility.

The programme has particular relevance for education systems that are expanding digital services while operating with uneven infrastructure, multilingual classrooms, decentralised administration, and strong dependence on external platforms. In such settings, a uniform technology mandate can amplify rather than reduce institutional variation. Well-resourced institutions may convert the same platform into richer learning through specialist staff and reliable connectivity, whereas others experience interrupted access, reduced teacher autonomy, or a transfer of costs to families. Comparative research should therefore treat context not as a control variable to be removed but as part of the intervention mechanism. Multi-site studies need to document infrastructure reliability, language of interface and content, disability access, staff-to-learner ratios, professional support, procurement conditions, and local data regulation. They should also examine whether implementation strengthens public capacity or creates dependency on vendors and short-term projects. UNESCO and OECD guidance increasingly emphasises human capacity, ecosystem coordination, and local validation, while the Industry 5.0 perspective adds resilience and sustainability to human-centred innovation (European Commission, 2021; Miao & Holmes, 2023; OECD, 2023, 2026). Translating these principles into evidence requires regional replication networks in which institutions share protocols but retain authority to adapt pedagogy and safeguards. Such networks could identify which mechanisms travel across contexts, which require linguistic or cultural redesign, and which should remain local. Their evaluations should include not only average learning gains but also service continuity, affordability, accessibility, institutional workload, and the capacity to maintain or discontinue a system after external funding ends. Public reporting of these dimensions would enable governments and institutions to learn from implementation failures instead of repeating them behind proprietary boundaries. It would also support procurement based on demonstrated educational functions rather than vendor claims. Cross-country work should preserve disaggregated results so that apparent system improvement is not produced by widening gaps between institutions, languages, disability groups, or socioeconomic strata. It should additionally distinguish temporary project capacity from capabilities embedded in public institutions, teacher education, and routine budgets. The result would be a more credible alternative to both technological universalism and contextual exceptionalism: common standards for evidence and accountability combined with explicit analysis of the conditions under which educational value is produced.

5. Discussion

The prospective corpus indicates a substantive shift from technology adoption to responsible orchestration. Earlier acceptance and diffusion models remain relevant to willingness, access, and spread, but the included studies increasingly ask what educational purpose a system serves and which human arrangements make that purpose credible. AI design fictions expose embedded values before deployment, teacher-informed LLM research positions professional judgement inside the implementation mechanism, and student-perception synthesis treats legitimacy as an educational condition rather than a communication problem. Domain studies reach the same conclusion from a different direction: plagiarism detection is shaped by diffusion structure, virtual reality requires safety-centred task design, computational thinking can be culturally situated, and digital motivation depends on active pedagogy rather than on devices alone. This pattern supports the critical claim that technology is never introduced into an institutional vacuum (Selwyn, 2019; Williamson & Eynon, 2020). It also extends human-centred and responsible-innovation frameworks by showing how broad principles can be operationalised in educational stages. RETO's contribution is to make evidence, pedagogy, values, and governance jointly necessary. Existing frameworks often privilege one layer: TPACK focuses on teacher knowledge, acceptance models on user intention, ethical frameworks on principles, and digital-ecosystem models on infrastructure and governance. Each is indispensable, but none alone can decide whether a particular technology should scale. Joint adequacy matters because failure can migrate across layers. Weak pedagogy can be misread as technical failure; opaque governance can destroy trust in an effective tool; strong motivation can mask shallow learning; and short-term achievement can coexist with inequitable access or unsustainable workload. The emerging literature is therefore conceptually ahead of a simple adoption narrative, even though its empirical designs have not yet caught up with its institutional ambitions.

The main weakness is the absence of cumulative evidence at scale. Five practice-oriented studies create a useful base, but most are bounded by one domain, one institution, a short intervention, or a specific user group. The field's upper-right quadrant—high institutional reach combined with comparative or longitudinal evidence—remains empty. This gap is especially important for generative AI because tools change rapidly, use often occurs outside institutional platforms, and apparent productivity gains may redistribute rather than reduce work. Robust evaluation should therefore separate immediate output quality from learning, retention, transfer, metacognition, authorship, and teacher workload. It should also compare guided and unguided use, document model and prompt conditions, and report adverse events or non-use. Similar principles apply to immersive and digital innovations. VR research needs transfer to authentic performance, accessibility, simulator sickness, maintenance, and cost; computational-thinking interventions need delayed assessment and comparison with alternative pedagogies; direct-diffusion studies need governance analysis across technologies and countries; SDG-oriented curricula need evidence that motivation and knowledge translate into behaviour or community outcomes (Makransky & Petersen, 2021; Radianti et al., 2020; Wing, 2006). Reviews should become living evidence infrastructures rather than repeated inventories, using common taxonomies and explicit links to primary studies. The corpus also reveals sectoral blind spots. Higher education dominates, while early childhood, TVET, disability-inclusive education, adult learning, and low-connectivity settings are absent. Unless research capacity is deliberately distributed, responsible-AI principles may be defined by settings with the greatest technical resources and then exported to institutions with different languages, infrastructures, legal protections, and professional conditions.

For institutions and policymakers, particularly in middle-income and ASEAN contexts, the practical implication is to replace generic AI strategies with portfolio governance. Different innovations require different evidence and safeguards: a generative writing assistant, a clinical simulation, a plagiarism-detection service, and a public-health curriculum should not pass through the same approval template. Yet all should answer a common sequence of questions. First, what educational problem is being addressed, and what outcome would count as improvement? Second, which pedagogy and human expertise mediate the technology? Third, which values could be advanced or harmed, including linguistic and cultural fit, disability access, integrity, privacy, and student agency? Fourth, who is accountable for procurement, data, incidents, appeals, professional development, and discontinuation? OECD and UNESCO guidance increasingly emphasises this ecosystem and human-capacity approach, but implementation will depend on local evidence and transparent decision rights (Miao & Holmes, 2023; OECD, 2023, 2026; UNESCO, 2024). Institutions should create stage-gated sandboxes rather than institution-wide mandates, require baseline and follow-up measures, include teachers and students in design review, publish model and data documentation, and budget for training, maintenance, accessibility, and evaluation. Procurement should require interoperability, audit access, data minimisation, and exit provisions so that dependency on a vendor does not become irreversible policy. At system level, funding should support replication networks and shared measurement rather than isolated demonstrations. The objective is not to slow educational innovation but to increase its reliability: technologies should move faster through learning cycles and more cautiously through irreversible scaling decisions.

6. Conclusion

The 2027 horizon corpus shows an emerging research frontier that is technologically diverse but conceptually convergent. AI and LLMs, immersive simulation, digital platforms, computational thinking, technology diffusion, and SDG-oriented curricula are increasingly evaluated through purposes that exceed adoption: agency, safety, academic integrity, cultural values, motivation, sustainability, and public health. This is a significant maturation in problem framing. At the same time, the evidence base remains uneven. Higher and professional education dominate, local and short-term studies carry much of the empirical load, and no included publication combines system-level reach with comparative or longitudinal evaluation. The field therefore has strong propositions, useful pilots, and increasingly explicit values, but an incomplete pathway from innovation to sustained educational and public value.

The Responsible Educational Technology Orchestration framework addresses this gap by requiring four jointly adequate conditions before scale: credible evidence, pedagogical mediation, value alignment, and accountable governance. It converts diverse research designs into a cumulative sequence rather than forcing them into a single hierarchy. Future studies should replicate promising interventions across institutions, measure retention and transfer, report subgroup and workload effects, operationalise value conflicts, and test governance arrangements under real procurement and data conditions. Educational institutions need not wait for perfect evidence, but they should keep claims proportional, deployment reversible, and monitoring continuous. The decisive question is not whether a technology is innovative or widely accepted. It is whether the complete socio-technical arrangement enables better learning and human development without transferring hidden costs, risks, or exclusions to students, teachers, institutions, or the public.

Declarations

Author Contributions

Arthit Srisawat: Conceptualization, methodology, formal analysis, theoretical synthesis, and writing—original draft. Pimchanok Charoensuk: Data curation, validation, visualization, and writing—review and editing. Nattaya Boonmee: Literature review, coding verification, project administration, and writing—review and editing. All authors critically reviewed the manuscript, approved the final version, and agreed to be accountable for the integrity of the work.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare no conflict of interest.

Ethics Approval

Ethical approval was not required because the study analysed publicly available bibliographic metadata and published research and did not involve human participants, personal data, or confidential records.

Data Availability

The Scopus metadata file analysed in this study and the derived coding matrix can be made available as supplementary material or from the corresponding author, subject to database licensing conditions.

Use of Generative Artificial Intelligence

Generative artificial intelligence was used to support language refinement and visualization scripting. The authors retained responsibility for corpus screening, analytical coding, source verification, interpretation, and final approval of the manuscript.

References

Ahmed, H., Daim, T., & Basoglu, N. (2007). Information technology diffusion in higher education. Technology in Society, 29(4), 469–482. https://doi.org/10.1016/j.techsoc.2007.08.011

Angraini, L. M., Kusuma, D. A., Wahyuni, R., Noto, M. S., Wahyuni, A., & Wahyuni, P. (2027). Integrating computational thinking and Islamic values in geometry learning: Effects on pre-service primary teachers' conceptual understanding. ASEAN Journal of Educational Research and Technology, 6(1), 67–78.

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

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20, 38. https://doi.org/10.1186/s41239-023-00408-3

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. Computers and Education: Artificial Intelligence, 4, 100118. https://doi.org/10.1016/j.caeai.2022.100118

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. International Journal of Educational Technology in Higher Education, 20, 22. https://doi.org/10.1186/s41239-023-00392-8

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: Strategies for technology integration. Educational Technology Research and Development, 47, 47–61. https://doi.org/10.1007/BF02299597

European Commission. (2021). Industry 5.0: Towards a sustainable, human-centric and resilient European industry. Publications Office of the European Union. https://doi.org/10.2777/308407

European Commission. (2022). Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators. Publications Office of the European Union.

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. Innovations in Education and Teaching International, 61(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

Farrokhnia, M., Salimi, G., & Ahmadshahi, A. (2027). Students, AI, and the future of higher education: A scoping review of perceptions and concerns. In E. G. Blanchard, G. Chen, M. Chi, & S. Isotani (Eds.), Artificial intelligence in education (pp. 363–377). Springer. https://doi.org/10.1007/978-3-032-29773-0_28

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Friedman, B., Kahn, P. H., Jr., & Borning, A. (2008). Value sensitive design and information systems. In K. E. Himma & H. T. Tavani (Eds.), The handbook of information and computer ethics (pp. 69–101). Wiley. https://doi.org/10.1002/9780470281819.ch4

Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. The Internet and Higher Education, 7(2), 95–105. https://doi.org/10.1016/j.iheduc.2004.02.001

Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. Educational Researcher, 42(1), 38–43. https://doi.org/10.3102/0013189X12463051

Henderson, M., Selwyn, N., & Aston, R. (2017). What works and why? Student perceptions of 'useful' digital technology in university teaching and learning. Studies in Higher Education, 42(8), 1567–1579. https://doi.org/10.1080/03075079.2015.1007946

Hew, K. F., & Brush, T. (2007). Integrating technology into K–12 teaching and learning: Current knowledge gaps and recommendations for future research. Educational Technology Research and Development, 55, 223–252. https://doi.org/10.1007/s11423-006-9022-5

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, 32, 504–526. https://doi.org/10.1007/s40593-021-00239-1

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Stadler, M. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martínez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. Computers and Education: Artificial Intelligence, 3, 100074. https://doi.org/10.1016/j.caeai.2022.100074

Kumar, A., Paliwal, J., Singh, M., Pendse, V., Saini, A., Gade, R., Dhaigude, R., Bora, A., Nimbalkar, S., & Raibagkar, S. (2027). Direct technology diffusion in higher education: Evidence from doctoral students. ASEAN Journal of Educational Research and Technology, 6(1), 79–92.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. Education Sciences, 13(4), 410. https://doi.org/10.3390/educsci13040410

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1–16). ACM. https://doi.org/10.1145/3313831.3376727

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. Educational Technology Research and Development, 69, 937–958. https://doi.org/10.1007/s11423-020-09852-2

Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Narala, S. V. R., Gayam, S. L. R., Kalyanam, J., Miranda, J., & Zhong, J. (2027). Improving and piloting the application of 360° virtual reality technology in nursing education: Challenges and recommendations for patient safety training. In Communications in Computer and Information Science (Vol. 2907, pp. 95–105). Springer. https://doi.org/10.1007/978-3-032-29003-8_7

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence, 2, 100041. https://doi.org/10.1016/j.caeai.2021.100041

OECD. (2023). OECD digital education outlook 2023: Towards an effective digital education ecosystem. OECD Publishing. https://doi.org/10.1787/c74f03de-en

OECD. (2026). OECD digital education outlook 2026: Exploring effective uses of generative AI in education. OECD Publishing. https://doi.org/10.1787/062a7394-en

Pees, A., Oelscher, L., Specht, M., Rietz, C., & Lehnert, F. (2027). From speculation to evaluation: Exploring values in design fictions about AI-enhanced education. In E. G. Blanchard, G. Chen, M. Chi, & S. Isotani (Eds.), Artificial intelligence in education (pp. 396–405). Springer. https://doi.org/10.1007/978-3-032-29773-0_31

Pehlić, I., & Tufekčić, N. (2027). Educational innovations for Industry 4.0/5.0: Pedagogical perspectives on intelligent systems and emerging technologies. In New technologies, development and application IX (pp. 644–654). Springer. https://doi.org/10.1007/978-3-032-30107-9_60

Petrukha, N., Volskyi, D., Pomaza-Ponomarenko, A., Poroka, S., & Kharchenko, N. (2027). Innovative and digital approaches in higher education for student motivation and Sustainable Development Goals (SDGs): A literature review. ASEAN Journal of Educational Research and Technology, 6(1), 171–182.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. Computers & Education, 147, 103778. https://doi.org/10.1016/j.compedu.2019.103778

Redecker, C. (2017). European framework for the digital competence of educators: DigCompEdu. Publications Office of the European Union. https://doi.org/10.2760/159770

Reichert, H., Yasir, T., Satyavolu, M., & Barnes, T. (2027). A framework for LLM integration in secondary education: Insights from computing teachers. In E. G. Blanchard, G. Chen, M. Chi, & S. Isotani (Eds.), Artificial intelligence in education (pp. 107–121). Springer. https://doi.org/10.1007/978-3-032-29763-1_8

Rogers, E. M. (2003). Diffusion of innovations (5th ed.). Free Press.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. Contemporary Educational Psychology, 61, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. Polity.

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe and trustworthy. International Journal of Human–Computer Interaction, 36(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist, 57(10), 1510–1529. https://doi.org/10.1177/0002764213479366

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. Research Policy, 42(9), 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

Thinwiangthong, S., Silanoi, L., Srisuruk, P., Tungkasamit, A., Boonmee, S., Thangsathirasima, S., Supadol, P., & Jaratmetakul, P. (2027). Biology and health education curriculum on liver fluke disease and cholangiocarcinoma prevention among primary school students in endemic areas to support Sustainable Development Goals (SDGs): Bibliometric and mixed-methods needs analysis. ASEAN Journal of Educational Research and Technology, 6(1), 127–152.

UNESCO. (2021). AI and education: Guidance for policy-makers. UNESCO.

UNESCO. (2023). Global education monitoring report 2023: Technology in education—A tool on whose terms? UNESCO.

UNESCO. (2024). AI competency framework for teachers. UNESCO.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478. https://doi.org/10.2307/30036540

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, 28(2), 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. Learning, Media and Technology, 45(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995

Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33–35. https://doi.org/10.1145/1118178.1118215

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