A Privacy-preserving Edge Computing Framework for Group-level Classroom Learning-state Perception and Teacher Decision Support

Ziheng Zhang[1, *], Chenxi Jiang[2], Yihan Shen[3]
[1]College of Education, Zhejiang Normal University, Jinhua 321000, China
[2]College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua 321000, China
[3]School of Marxism, Zhejiang Normal University, Jinhua 321000, China
*Corresponding email: zzh061011@gmail.com

Abstract

Real-time classroom learning-state perception can help teachers notice changes in whole-class participation, but systems based on individual identification, facial analysis, or cloud transmission create privacy and governance risks. This paper presents a privacy-preserving edge computing framework for group-level classroom learning-state perception and teacher decision support. The study does not report completed classroom-effectiveness results. Instead, it develops a design framework that can be tested in later pilots. The framework was derived from learning analytics, classroom engagement, teacher decision-support, multimodal learning analytics, edge computing, and artificial intelligence (AI) governance literature. It specifies a five-layer architecture, a local data-flow boundary, window-level feature definitions, indicator computation formulas, a rule-based teacher-prompt mechanism, and a pilot evaluation protocol. The system processes millimeter-wave radar, speech-energy patterns, and interaction logs at a classroom edge node, then outputs aggregated indicators such as attention trend, activity level, participation balance, and rhythm fluctuation. Raw signals are deleted after feature extraction, and teacher prompts are advisory, dismissible, and traceable. The contribution of the paper is a more explicit and testable design specification for privacy-first classroom analytics, rather than an unverified claim of instructional effectiveness.

Keywords

Edge computing, Group-level perception, Privacy-preserving edge computing framework, Teacher decision support

Introduction

Artificial intelligence is increasingly used in educational technology, formative assessment, classroom observation, and teacher-support tools. The United States Department of Education Department of Education argued that AI systems in education should keep humans in the loop and should remain inspectable and overridable by educators [1].

United Nations Educational, Scientific and Cultural Organization (UNESCO) guidance on generative AI and its teacher competency framework also placed teacher agency, privacy, and institutional responsibility at the center of educational AI deployment [2,3]. These policy positions define a practical boundary for classroom analytics: The system may provide evidence, but it should not replace teachers' judgment.

In ordinary classrooms, teachers must attend to lesson pacing, student responses, classroom atmosphere, task completion, and unexpected events. Learning analytics provides concepts and tools for collecting and analyzing learning-process data [4]. Yet classroom analytics is useful only when indicators enter teaching decisions. A numerical dashboard that does not help a teacher decide whether to slow down, ask a question, or change an activity remains a display device.

A second difficulty is privacy. Many intelligent-classroom products depend on high-definition cameras, face recognition, individual behavior tracking, or cloud-based inference. These routes may increase recognition accuracy, but they also raise governance concerns in classrooms involving minors. The General Data Protection Regulation (GDPR) requires purpose limitation and data minimization, and the European Union (EU) AI Act defines and regulates several high-risk AI uses in education, including emotion-recognition systems [5,6].

A classroom learning-state system should therefore ask a narrower question: What group-level evidence is sufficient for teaching support without identifying individual students?

This paper proposes a privacy-preserving edge computing framework for group-level classroom learning-state perception and teacher decision support. The earlier term crowd-edge computing is avoided here because the present design does not yet implement cross-class model aggregation, collaborative inference, or teacher-crowd knowledge mining. The paper is positioned as a system architecture and design framework paper. It specifies what should be sensed, how signals should be processed locally, how group indicators should be computed, how prompts should be triggered, and how a later pilot should evaluate the design. Classroom teaching creates a demanding attention-allocation problem.

A teacher must explain content, monitor comprehension, manage transitions, distribute opportunities to participate, and decide whether an apparent change in classroom behavior requires intervention. Conventional observation can provide rich contextual judgment, but it is difficult to sustain every learner and every moment. Learning analytics may therefore be useful when it compresses weak and heterogeneous traces into a small number of interpretable signals rather than asking the teacher to inspect another continuous data stream. This paper treats analytics as a support for noticing and inquiry, not as a substitute for professional judgment.

The central design problem is not simply how to recognize classroom behavior more accurately. It is how to generate timely and actionable information without turning ordinary instruction into continuous individual surveillance. Camera-based identity, facial-expression classification, speech transcription, and cloud storage may produce detailed records, but much of that detail is unnecessary for decisions about class pace, participation breadth, or activity rhythm. Data minimization therefore changes the unit of analysis: The proposed framework estimates changes in a class-level time window and deliberately avoids named-student profiles. This choice accepts lower individual diagnostic resolution in exchange for a clearer educational purpose, a smaller privacy surface, and a more defensible governance boundary.

Three tensions motivate the framework. First, sensing resolution can conflict with proportionality: A system capable of identifying every learner may process more personal data than the instructional purpose requires. Second, predictive sophistication can conflict with interpretability: A high-dimensional model may be difficult for a teacher to question during a lesson. Third, centralized computation can conflict with both responsiveness and confidentiality because raw streams must travel across networks and remain available to remote services. Edge processing, group aggregation, persistent-trend rules, and teacher confirmation are introduced as a coordinated response to these tensions rather than as independent technical features.

Accordingly, the paper makes four contributions: Section 1, it specifies a privacy boundary for local and short-lived raw signals; defines interpretable group-level indicators; connects persistent changes to low-interruption teacher prompts; and proposes a staged evaluation protocol spanning technical, educational, privacy, ethical, and deployment outcomes. These contributions form a testable design artifact rather than evidence of instructional effectiveness. Section 2 synthesizes the literature and research questions. Section 3 derives design requirements. Sections 4-6 specify the architecture, indicator pipeline, and teacher interaction. Section 7 presents the pilot protocol. And Sections 8-9 discuss implications, limitations, and conclusions.

Literature review and research questions

Learning analytics and classroom engagement

Learning analytics concerns the measurement and interpretation of learner and context data for understanding and improving learning. Its relevance to classroom perception is not the availability of more traces but the capacity to connect evidence to pedagogically meaningful action. Gašević, Dawson, and Siemens warned that analytics can drift toward data collection detached from learning, while engagement research shows that behavioral, emotional, and cognitive participation cannot be reduced to a single observable signal [7,8]. The resulting requirement is to present cautious group-level proxies that teachers interpret in context, rather than automated labels that claim to reveal an internal state.

Teacher decision support

AI-in-education research has expanded from tutoring and assessment to teacher dashboards and classroom support, yet reviews continue to identify limited attention to educators, pedagogy, and real decision processes [9, 10]. Teacher-facing systems are more credible when they improve awareness while preserving professional judgment, as demonstrated by mixed-reality teacher-support work [11]. Scaffolding and situated-learning perspectives further emphasize timely support embedded in authentic practice, while deliberately designed analytics interventions connect evidence to purposeful action [12-14]. A teacher interface should therefore identify a possible concern, provide a comprehensible reason, and leave the timing and form of response to the educator.

Multimodal classroom perception

Multimodal learning analytics combines bodily, verbal, interactional, and environmental traces to represent complex learning activity [15,16]. Its advantage is complementary evidence; its deployment cost is synchronization error, missing modalities, calibration burden, and interpretive ambiguity. Lighting, occlusion, room acoustics, seating arrangement, lesson format, and teacher style may all change feature distributions. The design implication is therefore selective multimodality: Each signal must serve an explicit classroom decision, carry a quality value, and be removable without causing the system to manufacture certainty.

Edge computing and privacy-preserving analytics

Edge computing places computation close to data source, supporting latency-sensitive operation and limiting routine transmission of raw classroom streams [17,18]. Federated learning can further decentralize model updates, but it does not by itself establish purpose limitation, deletion, access control, or acceptable classroom use [19,20]. Dashboard research likewise shows that visualization alone does not guarantee educational value [21]. A privacy-preserving classroom framework must therefore combine local processing with explicit retention rules, interpretable outputs, teacher control, and institutional accountability.

Conceptual boundaries and research gaps

Four connected gaps remain. Classroom analytics often assume that individual recognition is necessary. Multimodal work frequently under-specifies the operational path from noisy signals to classroom use. Teacher-support systems may display information without defining low-interruption action, and privacy research may stop at technical decentralization without specifying deletion, consent, access, or secondary-use controls. These gaps require an artifact that aligns data minimization, interpretable group inference, teacher-controlled action, and verifiable governance.

The framework consequently treats observable participation, movement, response timing, and speech-energy variation as contextual proxies rather than direct measurements of attention, emotion, or understanding. Aggregation lowers the risk of durable individual labeling, but it does not make inference harmless: Small classes, stable seating zones, and repeated observation can still reveal patterns. The relevant research gap is therefore sociotechnical. A complete design must define what is sensed, what is discarded, what uncertainty prevents an output, how a prompt reaches a teacher, and how the institution limits reuse.

Engagement is commonly described as behavioral, emotional, and cognitive, but these dimensions cannot be read directly from a sensor. Observable participation, movement, response timing, and speech-energy variation are at most contextual proxies. A quiet class may be concentrating, confused, reading, or waiting for instruction; a noisy class may be disengaged or productively collaborating. The framework therefore avoids claims that a short signal window reveals an internal mental state. The term learning-state perception denotes a cautious estimate of observable group conditions that may warrant teacher attention. This distinction is consistent with the multidimensional engagement literature and with the argument that learning analytics must remain connected to learning theory and pedagogical interpretation.

The group-level unit also changes the meaning of error. In an individual classifier, an error may incorrectly label a student. In the proposed framework, an error is primarily a mistimed or unhelpful class-level prompt. The latter can still waste attention or contribute to surveillance concerns, but it is less likely to create a durable personal judgment when raw signals, identities, and individual scores are excluded. This does not make aggregation harmless. Small classes, sparse participation, stable seating zones, and repeated observations may still permit inference. Privacy must consequently be evaluated as a property of the complete sociotechnical system rather than inferred from the absence of names.

Table 1 maps these gaps to design requirements and corresponding framework responses. The mapping is intentionally one-to-one: Each technical component is justified by a pedagogical or governance problem.

The research questions that follow cover the privacy boundary, signal-to-indicator transformation, teacher-facing use, and multidimensional evaluation. Together they connect the literature review to a testable artifact while keeping effectiveness claims outside the scope of the present conceptual study.

Table 1. Research gaps mapped to design requirements and framework responses.

Existing research	Unresolved limitation	Design requirement	Framework response
Learning analytics	Indicators may optimize what is easy to count	Pedagogical interpretability	Group proxies with explicit semantic limits
Multimodal sensing	Noise, missingness, and calibration are under-specified	Quality-aware selective fusion	Per-modality quality gates and abstention
Teacher dashboards	Information does not automatically become action	Low-interruption decision support	Persistent, explainable, dismissible prompts
Privacy-preserving AI	Decentralization alone does not govern use	Operational data and governance boundaries	Local processing, deletion, audit, and purpose limits

Research questions

The study addresses four research questions (RQ).

RQ1 asks how classroom learning-state perception can be designed without identifying or profiling individual students.

RQ2 asks how heterogeneous local signals can be transformed into interpretable group-level indicators.

RQ3 asks how those indicators can support time-sensitive teacher decisions without creating excessive interruption or automation bias.

RQ4 asks how the framework should be evaluated across technical, educational, privacy, ethical, and deployment dimensions.

These questions connect the literature review to the artifact requirements and prevent the system architecture from being evaluated only by recognition accuracy.

The review therefore establishes both the practical necessity for developing the specific artifact and the clear boundaries of valid outcomes that this work may reasonably claim to deliver. The next section systematically translates these identified research gaps into explicit functional requirements.

Research design

This section thoroughly explains the core reasons why the proposed framework finally takes its current present form. It clearly defines the design-science object, research scope, functional requirements, basic assumptions, and sensor-selection logic in advance before the overall architecture further specifies which concrete components implement those deliberate design choices. Detailed information is shown in Figure 1.

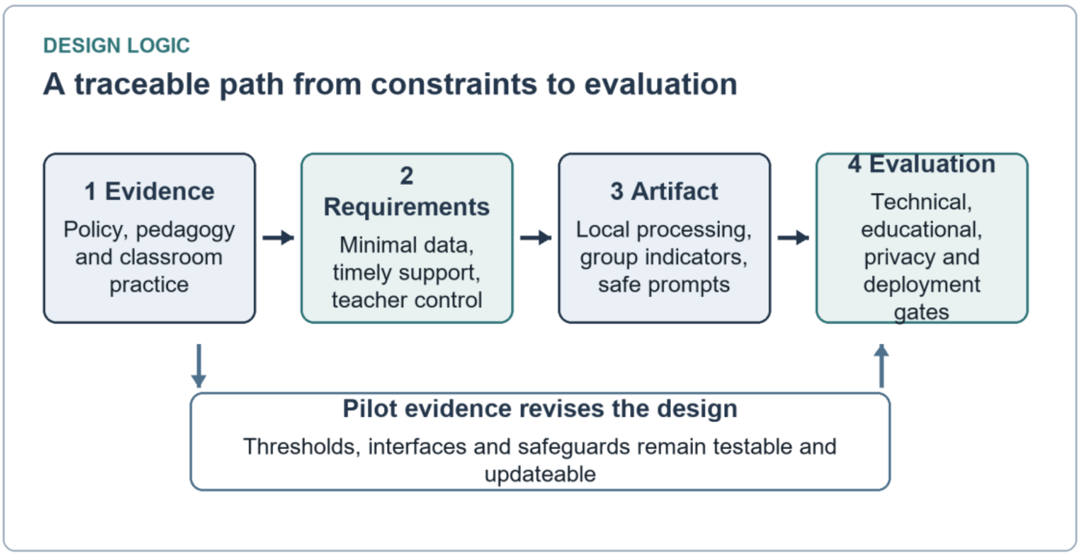

Figure 1. Design logic from evidence and policy constraints to a testable classroom artifact.

Research object

The research object is a privacy-preserving edge computing framework for group-level classroom learning-state perception and teacher decision support. The object is not an individual student, a single sensor, or a completed commercial product. It is a system design artifact that specifies how classroom signals can be locally processed, aggregated into group-level indicators, and converted into teacher-facing prompts.

This object was selected because it concentrates the central contradiction identified in the literature review: Classrooms need timely state information, but systems based on individual identification, cloud transmission, or emotion inference create privacy and governance risks. A framework-level object is appropriate at this stage because the study aims to define a testable design before claiming empirical instructional effects.

Research scope

The application scope is ordinary Kindergarten through 12th grade（K-12) classroom teaching, especially scenarios involving minors, teacher-led instruction, limited equipment budgets, and the need for low-interruption feedback. The framework focuses on whole-class indicators such as attention trend, activity level, participation balance, and rhythm fluctuation. It excludes individual identity recognition, individual emotion classification, student ranking, and punitive evaluation.

The material scope includes policy documents on educational AI governance, peer-reviewed studies on learning analytics, classroom engagement, teacher decision support, multimodal learning analytics, edge computing, and privacy-preserving analytics, together with the project scenario of a classroom edge node using millimeter-wave radar, speech-energy features, and interaction logs. The study does not include complete classroom trial data. Its conclusions are therefore limited to system design, technical specification, and pilot-evaluation planning.

Research method

This study adopts a design science-oriented conceptual system design method rather than a completed design-based research cycle. Design science research emphasizes the construction and evaluation of artifacts that address identified problems [22]. In this paper, the artifact is the proposed framework, including its architecture, feature definitions, prompt rules, privacy boundary, and pilot evaluation protocol.

The study uses literature and policy analysis to identify educational, technical, and governance requirements, then applies conceptual system design to translate those requirements into a layered architecture and operational pipeline. The method addresses four questions: privacy-preserving perception without individual profiling; transformation of heterogeneous local signals into interpretable group indicators; time-sensitive teacher support without excessive interruption or automation bias; and evaluation across technical, educational, privacy, ethical, and deployment dimensions.

The procedure included four steps. First, relevant literature and policy documents were reviewed to identify constraints on learning validity, teacher agency, privacy, and deployment. Second, the classroom project scenario was analyzed to define data sources and design boundaries. Third, the system architecture, group-level indicators, formulas, and prompt-triggering rules were specified. Fourth, a pilot evaluation protocol was designed so that future studies can test technical accuracy, teacher usability, student perception, and deployment cost.

Sensor selection rationale

Sensor selection follows a minimum-sufficient-evidence rule. The first prototype needs signals that can describe room-level movement, non-semantic acoustic activity, and observable interaction distribution without producing conventional images, speech transcripts, voiceprints, or continuously worn personal records. A 60 Gigahertz (GHz) millimeter-wave radar can represent motion through range-Doppler or point-cloud features and is less dependent on visible lighting than Red-Green-Blue (RGB) imaging, prior Institute of Electrical and Electronics Engineers (IEEE) research demonstrates its non-contact activity-sensing capability while also showing that radar outputs still require calibration and model validation [23].

Speech energy is selected only as a non-semantic channel: The edge node derives amplitude envelope, voice-activity ratio, and turn-density estimates, then discards the waveform. Interaction logs contribute event counts or classroom-zone distributions when an existing response system is available. The three sources are complementary but not individually diagnostic. Radar may be affected by occlusion and furniture; audio energy by ventilation and corridor noise; logs by incomplete coverage of classroom participation. Their use is conditional on per-modality quality checks.

Alternatives were not rejected categorically. RGB and depth cameras provide richer posture information but create a larger visual-data surface and more complex consent, access, and secondary-use risks. Wi-Fi Channel State Information (Wi-Fi CSI) can support device-free activity sensing with existing wireless infrastructure, yet surveys report sensitivity to environment, hardware, placement, and cross-domain generalization, as well as continuing security and privacy concerns [24].

Bluetooth Low Energy (BLE) often requires carried devices or coarse proximity inference; wearables can provide individual physiological data but introduce compliance, maintenance, equity, and profiling concerns. Table 2 therefore treats the selected combination as a defensible starting point for group-level feasibility, not as a universal optimum.

Table 2. Sensor-selection trade-offs for an initial privacy-first classroom prototype.

Modality	Privacy/exposure	Operational trade-off	Design decision
RGB/depth camera	Rich visual or body-shape record	Strong spatial detail; lighting, occlusion, and governance burden	Excluded from the initial operational design
Wi-Fi CSI/BLE	No conventional image; signals may still reveal occupancy or behavior	Infrastructure-dependent; domain shift; BLE may require devices	Candidate for a separately governed comparison
Wearables	Individual physiological or motion traces	Charging, compliance, equity, and maintenance burden	Excluded from routine group perception
60 GHz radar + speech energy + logs	Room-level motion, non-semantic audio features, aggregate events	Requires local calibration and quality gating	Selected as minimum-sufficient complementary evidence

Design-science process and requirement derivation

The study follows a design-science logic in which an artifact is constructed to address a defined problem and evaluated against explicit requirements. The process contains six linked activities: Problem identification, objective definition, requirement derivation, architecture design, scenario-based demonstration, and evaluation planning. The present paper completes the conceptual design and evaluation-planning activities. Prototype implementation, classroom calibration, and outcome evaluation remain future work.

This boundary is important because a well-specified artifact may be research-relevant before its effectiveness is established, but it must not be described as a validated intervention.

Requirements were derived by triangulating four sources. Educational literature requires indicators to be pedagogically interpretable and connected to possible action. Privacy and AI-governance documents require purpose limitation, data minimization, human oversight, security, and accountability [25].

Classroom practice requires low interruption, tolerance of noise, and compatibility with changing lesson formats. Edge-computing literature requires bounded resource use, predictable latency, and continued local operation when connectivity is weak. These requirements yield five design principles: Collecting the least revealing signal suitable for the purpose, computing locally by default, aggregating before persistence; trigger only on stable trends and keeping the teacher in control of interpretation and response.

Unit of analysis, assumptions, and exclusions

The analytical unit is a classroom-level state vector for a fixed time window, not a student. The initial scenario assumes a teacher-led K-12 lesson, one protected classroom edge node, sensors positioned to observe the room, and a teacher interface unavailable to students or external observers during the lesson. The core service must continue without an internet connection. Configuration updates may be downloaded outside lesson time, but raw classroom streams are not uploaded for routine inference. Each classroom requires local calibration because acoustics, layout, class size, subject, and teaching style affect feature distributions.

Several uses are explicitly excluded. The framework does not perform facial recognition, speaker identification, emotion recognition, automated discipline, student ranking, attendance enforcement, teacher-performance scoring, or high-stakes assessment.

It does not claim to identify confusion, motivation, boredom, or learning gain from movement and sound. If a school later proposes any excluded use, that proposal constitutes a new purpose requiring a separate necessity assessment, risk analysis, consent process, technical design, and governance decision.

Function creep cannot be justified by the fact that sensors are already installed.

These choices establish the rationale and boundary of the design. This section now specifies the five-layer artifact that implements the requirements without introducing algorithmic detail prematurely.

System architecture

This section describes what the system contains and how data moves between components. It focuses on layer responsibilities, interfaces, retention, and governance boundaries; feature equations and prompt algorithms are reserved for next section.

Overall architecture

The framework contains five layers, summarized in Figure 2. The edge-processing layer validates, aligns, and transforms short-window inputs before deleting raw buffers. The group-state layer produces bounded indicators and quality metadata.

The teacher-support layer applies eligibility rules and presents advisory prompts or abstains. The evaluation-and-governance layer stores only necessary aggregate outcomes, prompt events, teacher responses, deletion status, and audit information.

The architecture is deliberately modest. It does not require face recognition, individual identity tracking, or individual emotion classification. If future prototypes include visual posture features, they should be converted at the edge into window-level group statistics and should not retain restorable images.

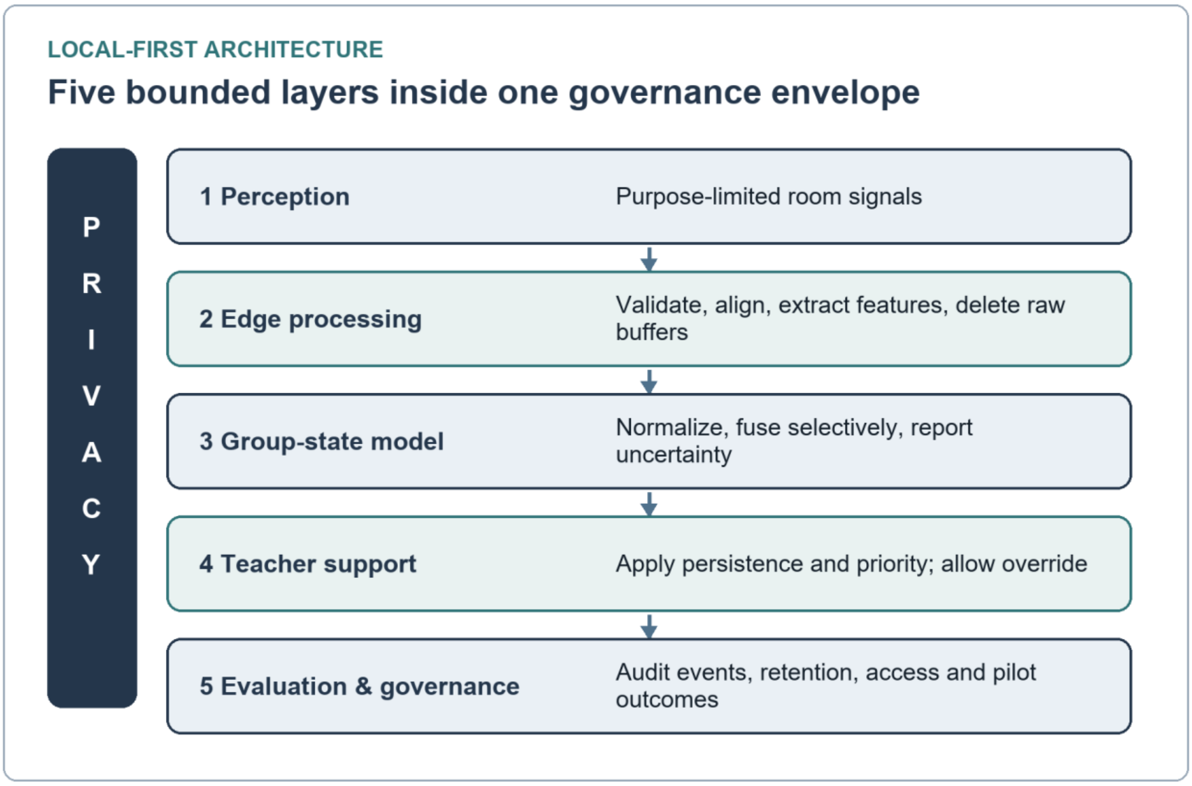

Figure 2. Five-layer edge architecture and privacy-governance boundary.

Policy and ethical constraints

Policy documents are used here as rigid design constraints rather than superficial and dispensable background decoration. The GDPR's data-minimization principle leads directly to necessary local feature extraction and thorough raw-data deletion to comply with legal mandates.

The EU AI Act's restrictions around high-risk educational AI and emotion recognition support the decision to avoid individual emotion labels. UNESCO's guidance on generative AI supports teacher agency and meaningful professional autonomy as well as necessary human oversight throughout educational practice.

The National Institute of Standards and Technology Artificial Intelligence Risk Management Framework (NIST AI Risk Management Framework) provides a useful, practical structure for comprehensive governance, systematic mapping, quantitative measurement, and holistic management of potential system risks.

Data flow and privacy boundary

The data flow has one rule: Raw signals stay local and short-lived. Radar streams are converted into group movement intensity. Speech streams are converted into energy and turn-density features rather than speech content. Interaction logs are aggregated into distributions, not individual rankings. Only aggregated indicators, prompt events, teacher responses, and audit logs are retained. This boundary responds to earlier ethical concerns in learning analytics, where researchers had warned that transparency, student control, institutional responsibility, and trusted data practice could not be treated as optional add-ons [26]. If summaries are exported across classes in later work, additional privacy techniques such as differential privacy noise or federated updating should be tested rather than assumed. Broader United Nations Educational, UNESCO and Organization for Economic Co-operation and Development (OECD) governance frameworks reinforce institutional responsibility, stakeholder protection, and accountable educational-AI deployment.

Layer responsibilities and interfaces

The perception layer acquires room-level traces. A 60 GHz millimeter-wave radar supplies range-Doppler or point-cloud summaries from which collective micro-movement can be estimated without producing an optical image. The audio channel supplies amplitude envelope, voice-activity ratio, and turn-density features; speech content, transcripts, and voiceprints are excluded. Interaction events may include anonymous counts of teacher questions, response-pad events, or classroom-zone participation when such sources are available. Environmental variables such as sound pressure or temperature are contextual controls rather than direct indicators of engagement.

The edge-processing layer validates timestamps, rejects corrupt frames, filters noise, segments windows, extracts features, and deletes raw buffers after a short processing interval. The state-estimation layer normalizes features against a classroom-specific baseline and produces indicators with quality metadata.

The decision-support layer applies persistence, cooldown, and priority rules before presenting a suggestion. The governance layer records configuration changes, prompt events, teacher responses, access attempts, deletion status, and software versions. Interfaces between layers carry only what the next layer requires. In particular, the decision layer receives indicators and quality flags, not reconstructable radar or audio data.

End-to-end data lifecycle

A lesson begins with a visible system-status check and a configuration profile selected for the planned activity, such as direct instruction or group discussion. Sensors write to volatile buffers on the edge node. After feature extraction, raw buffers are overwritten according to a fixed deletion policy.

Window features are retained only long enough to support smoothing and prompt logic. The longer-lived record contains time-bucketed indicators, quality flags, prompts, teacher responses, and audit events. Retention periods should be configured in days rather than assumed indefinitely, and schools should document who can export summaries and for what purpose.

The privacy boundary is enforced technically and procedurally. Services run under least-privilege accounts; storage is encrypted; administrator access is logged; model and rule packages are signed; exports exclude raw features when the instructional purpose can be met with aggregate indicators. A physical or software stop control allows the teacher to suspend sensing. Students and parents need an understandable notice describing signals, purposes, retention, access, and withdrawal options. These measures are implemented rather than merely cite the ethical principles in learning analytics.

Operational scenario and abstention

Consider a lecture segment in which interaction frequency and distributed participation decline across three windows while movement variability remains within the normal range. The system does not label the class inattentive. It presents a low-priority pace-check suggestion with the reason that response activity has declined relative to the local lesson baseline. The teacher may know that students are completing a silent reading task and dismiss the prompt as contextually inappropriate. That response becomes evidence for configuration review, not evidence about any learner.

In a second case, the audio channel is saturated during group work and radar quality falls because furniture has been moved. The quality gate marks the multimodal state estimate as unreliable. Rather than combine poor inputs into a confident score, the system abstains and displays only a neutral sensor-status indicator. Abstention is a designed outcome. It limits false authority, protects teacher attention, and makes missingness visible for technical evaluation.

The architecture therefore defines a bounded path from local sensing to teacher-controlled action and makes non-output an explicit system state to avoid ambiguous idle behaviours during the whole operational cycle. The next section specifies how windows, features, indicators, quality values, and decision rules operationalize that path with quantifiable metrics and executable logical constraints, so readers can fully grasp the concrete implementation logic of the overall framework step by step.

Technical design of group-level perception

This section explains how the framework operates in a complete end-to-end workflow. It defines the local feature vector, preprocessing and normalization, quality-aware fusion, indicator semantics, persistence rules, uncertainty handling, and safe degradation required for a testable prototype, and further elaborates on the logical correlations among these modules as well as the boundary conditions that guarantee stable and repeatable experimental verification in practical deployment.

Signal sources and window segmentation

Signals are segmented into fixed windows of 30 or 60 seconds. For each window t, the edge node extracts radar-based group micro-movement *M(t)*, speech-energy intensity *V(t)*, interaction frequency *I(t)*, participation distribution *D(t)*, and short-term change vector *C(t)*. The window length should be selected during pilot calibration. Shorter windows may react faster but create more false alarms; longer windows are more stable but may miss teaching moments.

Group-level feature definitions

The framework uses group-level features rather than individual labels. *M(t)* is the normalized intensity of group micro-movement. *V(t)* is the normalized speech-energy score after removing transient noise. *I(t)* is the number of classroom interaction events within the window. *D(t)* is the distribution of participation across interaction sources or classroom zones. *C(t)* is the difference between the current state vector and a smoothed previous state. These features are not claims about any named student.

Indicator computation

For feature x_j in window t, let median j and IQR_j denote the classroom-local median and interquartile range, ε (a small stabilizing constant), and c (the clipping bound). Robust normalization is defined in eq. (1):

$$z_j(t) = \text{clip}\left(\frac{x_j(t) - \text{median}_j}{\text{IQR}_j + \varepsilon}, -c, c\right) \tag{1}$$

Each normalized feature has a quality score $q_j(t)$ in [0, 1]. The eligible modality set is defined in eq. (2), where q_{min} is the calibrated minimum quality:

$$E(t)=\{j|q_j(t)\geq q_{min}\}. \tag{2}$$

The quality-aware fused state in eq. (3) uses non-negative weights that sum to one over the eligible modalities:

$$F(t) = \frac{\sum_{j\in E(t)} w_j q_j(t) z_j(t)}{\sum_{j\in E(t)} w_j q_j(t)} \tag{3}$$

The evidence-sufficiency score is given by eq. (4). It is not a probability of student attention. If $Q(t) < q_{min}$, the system abstains rather than producing a teacher prompt:

$$Q(t) = \sum_{j\in E(t)} w_j q_j(t) \tag{4}$$

Let E_A(t) be the eligible movement, speech-energy, and interaction subset. Activity level L(t) is computed by eq. (5):

$$L(t) = \frac{\sum_{j\in \text{E\_A}(t)} w_j q_j(t) z_j(t)}{\sum_{j\in \text{E\_A}(t)} w_j q_j(t)} \tag{5}$$

Participation balance *P(t)* and rhythm fluctuation *R(t)* are defined in eqs. (6) and (7), respectively. *D(t)* is the observable contribution distribution, and *S(t - 1)* is the preceding smoothed state:

$$P(t) = 1 - \text{Gini}(D(t)) \tag{6}$$

$$R(t) = \|F(t) - S(t-1)\|_2 \tag{7}$$

Attention trend *A(t)* is the robust Theil-Sen slope of the prespecified responsiveness score *r(τ)* across k eligible windows, as specified in eq. (8):

$$A(t) = \text{TheilSenSlope}(\{r(\tau)\}_{\tau=t-k+1}) \tag{8}$$

The smoothed state follows eq. (9), where $0 < \alpha \leq 1$ controls responsiveness to the current fused state.

$$S(t) = \alpha F(t) + (1-\alpha) S(t-1) \tag{9}$$

These expressions are design specifications rather than validated models.

Window length, clipping constant c, smoothing coefficient α, weights, thresholds, and q_{min} must be calibrated using complete lesson sequences and reported with uncertainty and missing-modality analyses. Table 3 defines the intended semantics and teacher-facing use of each indicator.

Table 3. Operational definitions and limits of group-level indicators.

Indicator	Operational definition	Primary evidence	Teacher-facing interpretation
Attention trend	Robust slope across k eligible responsiveness windows	Interaction timing with quality-gated contextual features	Possible need to check pace or task clarity; not an attention diagnosis
Activity level	Quality-weighted movement, speech-energy, and interaction score	Radar, non-semantic audio, event counts	Possible rhythm check relative to the selected lesson profile
Participation balance	1-Gini coefficient of observable contribution distribution	Anonymous events or classroom zones	Possible need to broaden the response format
Rhythm fluctuation	Distance between current fused state and smoothed baseline	All eligible group features	Possible transition, disruption, or context change requiring interpretation

Prompt triggering rule

Algorithm 1 specifies that a prompt is eligible only when an abnormal trend persists. A single noisy window is insufficient. The rule engine evaluates smoothed indicators across k consecutive eligible windows, checks calibrated thresholds and evidence quality, and then applies priority and cooldown constraints. The teacher may accept, dismiss, defer, or mark a false alarm; the response enters a configuration-review log rather than a student record.

Algorithm 1 defines window-level prompt triggering rule. Input: Indicators $A(t)$, $P(t)$, $R(t)$, and $L(t)$. Calibrated thresholds θ_A, θ_P, θ_H, and θ_L; persistence length k. Step 1: Smooth indicators across k eligible windows. Step 2: If $A(t)$ remains below θ_A for k windows, generate a pace-check prompt. Step 3: If $P(t)$ remains below θ_P for k windows, generate a participation-broadening prompt. Step 4: If $L(t)$ exceeds θ_H or drops below θ_L, generate a rhythm-check prompt. Step 5: Recording the teacher's response and updating the rule-preference log. If the evidence test in eq. (4) fails, return abstain.

Preprocessing and temporal alignment

Preprocessing must be modality specific. Radar processing removes static clutter, rejects implausible points, and aggregates motion energy by room zone without tracking trajectories. Audio processing applies band-limited noise reduction and voice-activity detection before discarding the waveform.

Interaction events are de-duplicated and assigned to the same monotonic edge clock. Features are aligned to non-overlapping or sliding windows. Sliding windows improve responsiveness but introduce dependence between observations; this dependence must be considered when thresholds and confidence intervals are evaluated.

Every feature is accompanied by a quality value. Radar quality may reflect frame completeness and clutter ratio; audio quality may reflect saturation and signal-to-noise ratio; log quality may reflect source availability. A modality below its quality threshold contributes no value to the fused indicator. Remaining weights are renormalized only when the minimum evidence rule is satisfied. Otherwise, the system abstains. This prevents missing audio, for example, from being silently interpreted as classroom silence.

Baseline, normalization, and context profiles

Raw magnitudes cannot be compared across rooms or lesson formats. The robust normalization in eq. (1) should use parameters estimated from an initial calibration period and bounded to reduce outlier influence. Min-max normalization may be easier to explain but is more sensitive to extreme sessions. The selected method, calibration duration, clipping range, and update frequency must be reported in a pilot. Baselines should never be updated from a single anomalous lesson.

Context profiles reduce systematic false alarms. Direct instruction, individual practice, whole-class questioning, and group work have different expected sound, movement, and participation patterns. The teacher can select the planned profile, or the system can remain in a conservative general profile. Automatic activity recognition is deliberately avoided in the first prototype because a mistaken lesson-mode classification could contaminate every downstream threshold. Profile changes are logged and remain reversible.

Feature semantics and limits

Group micro-movement *M(t)* summarizes short-range changes in radar energy after static clutter removal. It may indicate a transition, collective writing, restlessness, or a shift in seating, but it does not reveal attention by itself. Speech-energy intensity *V(t)* summarizes non-semantic acoustic activity and is sensitive to ventilation, corridor noise, and teaching style. Interaction frequency *I(t)* counts defined events and therefore depends on which interactions the classroom technology can observe. Participation distribution *D(t)* estimates how evenly observable contributions are spread across sources or zones, not whether every student has understood lesson.

The change vector *C(t)* captures the distance between the current normalized state and a smoothed recent baseline. A large change can be educationally meaningful even when no individual feature crosses a fixed threshold. Conversely, a stable but undesirable condition may not produce a large change. For this reason, the framework retains both level and trend indicators. Teacher interpretation supplies the contextual explanation that the sensing system lacks.

Fusion, weighting, and uncertainty

The first prototype should use transparent late fusion. Each modality produces a bounded feature and quality value; weighted sums are then computed with non-negative weights that total one.

Late fusion makes it possible to inspect contribution, remove an unavailable modality, and explain why a prompt occurred. End-to-end deep fusion may eventually improve prediction, but it would require representative data, stronger governance, and explanation methods. The conceptual paper does not assume that complexity is equivalent to educational value.

Weights can be initialized from expert elicitation and tuned in a pilot using nested validation that separates lessons, teachers, and classrooms. Randomly splitting adjacent windows would exaggerate performance because neighboring windows are correlated. Calibration should optimize a utility function that penalizes false alarms and interruptions, not accuracy alone. Results should include uncertainty intervals, sensitivity to window length, and performance under missing modalities. No universal weights are proposed because transfer across subjects and class sizes remains an empirical question.

Persistence, cooldown, and conflict resolution

A persistence rule requires an indicator to remain beyond a threshold for k eligible windows. Exponential smoothing should follow eq. (9); a higher α responds faster but passes more noise. The pilot should compare α, k, window length, and threshold combinations on complete lesson sequences. The main operational metric is not only event-level precision but false prompts per lesson, because repeated low-quality alerts can make an otherwise accurate system unusable.

After a prompt, a cooldown interval suppresses equivalent prompts unless the condition worsens materially. When multiple rules fire, the system applies a priority order: Safety or sensor-status messages, participation breadth, pace, and general rhythm. Only one instructional prompt should be visible at a time. A prompt expires if it is no longer relevant. These constraints prevent a dashboard from becoming a second classroom requiring management. Detailed information is shown in Figure 3.

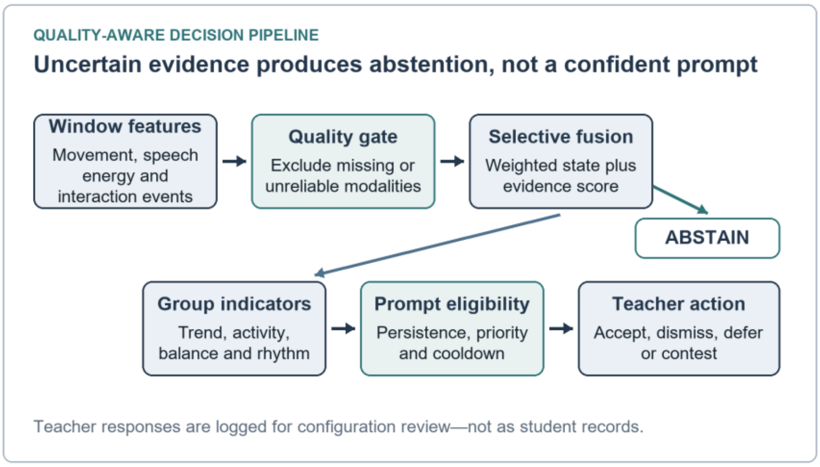

Figure 3. Quality-aware indicator generation, prompt eligibility, abstention, and teacher feedback.

Failure modes and safe degradation

Expected failures include sensor obstruction, moved furniture, clock drift, audio saturation, edge-node overload, software update failure, power loss, and a lesson format outside the calibrated profile.

Each failure maps to a safe response: drop the modality, reduce confidence, abstain from prompts, preserve a minimal audit event, or stop processing. The interface must distinguish "no abnormal trend detected" from "insufficient evidence". Conflating these states would create false reassurance.

Security failures require equally explicit behavior. Repeated unauthorized access attempts lock export functions without preventing the teacher from stopping sensing. Unsigned model packages are rejected. A deletion failure raises an administrative alert and blocks the next recording session until resolved. These controls embody the NIST risk-management emphasis on governance, measurement, and management across the AI lifecycle.

The technical design thus makes uncertainty, persistence, and safe degradation observable rather than hiding them behind a single score. The section turns these outputs into a teacher interaction model that preserves agency and contestability.

Teacher-in-the-loop decision support

This section explains how the system's bounded outputs enter classroom work. The interface is designed for rapid noticing, optional action, feedback, and challenge; it does not convert aggregate trends into automated judgments about students or teachers.

Interaction design and explanation

The teacher interface should communicate the minimum information necessary for action: prompt type, persistence duration, direction of change, data-quality status, and a short reason. It should avoid student-level visualizations, emotive labels, red warning colors for ordinary instructional variation, and pseudo-precise probabilities. A suggested message is "participation breadth has remained below this lesson profile for three windows; consider widening the response format". The wording identifies evidence and offers an option without claiming a diagnosis.

Prompts may appear on a small teacher-only display, wearable, or existing classroom console. They should be glanceable, silent by default, and dismissible with one action. Teachers need controls to snooze a prompt, mute a category for the lesson, switch context profile, and pause sensing. The system should cap the number of instructional prompts per lesson. Interface logs record only the interaction needed for evaluation and tuning; they must not be repurposed to score teacher compliance.

Feedback and controlled adaptation

A response taxonomy improves learning from feedback. "Accepted" means the suggestion was considered useful. "Dismissed context" means the signal may be correct but the prompt was inappropriate for the activity. "False alarm" means the sensed trend was not credible, and "later" means the timing was poor. Free-text notes should be optional because they increase burden and may introduce identifiable information. Aggregate response patterns can guide periodic threshold reviews.

Adaptation must be bound. The system should not immediately lower a threshold whenever a teacher accepts a prompt, because feedback is sparse and affected by workload. A safer process accumulates evidence across lessons, proposes a configuration change, simulates its effect on historical aggregate features, and requires authorized approval. The previous configuration remains recoverable. This separation between feedback collection and model update supports auditability.

Managing automation bias and contestability

Human oversight is meaningful only when the human has information, time, authority, and an easy way to disagree. Automation bias may lead users to follow a suggestion despite contradictory context, whereas algorithm aversion may lead them to ignore a useful tool after visible errors.

The interface should therefore expose limitations, use calibrated language, and periodically show that no conclusion about individual students is being made. Teacher training should include examples in which the correct action is to dismiss the prompt.

Contestability extends beyond the lesson. Teachers, students, and parents need channels to question what is collected and how summaries are used. Administrators need a procedure to inspect prompt logic, retention compliance, access logs, and configuration history. A complaint or withdrawal request should not require technical expertise. The system remains decision support only if institutional practice respects the same boundary as the interface.

Illustrative teaching scenarios

During whole-class questioning, participation may concentrate in one classroom zone. If the pattern persists and data quality is adequate, the system can suggest broadening the response method. During silent individual practice, low speech energy and low movement are expected, so the selected profile suppresses rhythm prompts. During group work, high sound energy is not automatically abnormal; the rule emphasizes distribution and abrupt change rather than absolute volume. These scenarios show why context profiles and teacher control are not optional interface refinements but components of validity.

Teacher-in-the-loop support is therefore an operational allocation of authority: The system notices bounded patterns, while the teacher supplies pedagogical context and decides whether to act. The following evaluation protocol tests whether that allocation remains useful, understandable, and proportionate in practice.

Pilot evaluation protocol

This section systematically specifies rigorous metrics for how a future functional prototype should be properly evaluated without incorrectly implying that full empirical validation has already practically occurred in the current research. It prioritizes a measurable primary feasibility outcome, clearly distinguishes secondary quantified outcomes from tentative exploratory measures, and strictly requires staged progressive progression through well-defined explicit privacy and reliability assessment gates.

Outcome hierarchy

(1) Primary outcome

The initial teacher-visible pilot should prioritize useful-prompt rate at an acceptable interruption cost. A prompt is useful when the teacher understands its reason, judges it contextually appropriate, and reports that it supported or confirmed a feasible instructional action. The primary analysis should jointly report useful prompts and false or disruptive prompts per lesson.

(2) Secondary outcome

Prespecified secondary outcomes include end-to-end latency, dropped-window rate, raw-buffer deletion success, offline availability, prompt comprehension, teacher workload, trust calibration, time to response, student perceived surveillance, and hardware or maintenance burden. These measures test whether technical operation, educational usability, and governance constraints are simultaneously acceptable.

(3) Exploratory measures

Exploratory analyses may compare subjects, activity profiles, class sizes, rooms, teachers, window lengths, smoothing settings, fusion rules, and missing modalities. Student engagement questionnaires, classroom observations, and learning measures may be collected to estimate future study requirements. Table 4 separates these levels.

Table 4. Prioritized outcome hierarchy for a future pilot.

Level	Construct	Illustrative measure	Decision use
Primary	Prompt utility at acceptable interruption cost	Useful-prompt rate together with false/disruptive prompts per lesson	Decide whether teacher-visible testing should continue
Secondary	Technical, human-factor, privacy, and deployment feasibility	Latency; deletion success; comprehension; workload; surveillance perception; cost	Identify blocking risks and redesign requirements
Exploratory	Transportability and candidate educational effects	Variation by subject, class size, profile, modality, and observation measure	Generate hypotheses and power a later controlled study

Pilot setting, phases, and comparison

A future pilot should proceed in phases rather than exposing classrooms immediately to active prompts.

Phase 1 is a bench test with scripted signals and network faults.

Phase 2 is a silent classroom deployment in which indicators are computed but not shown, allowing baseline and quality assessment.

Phase 3 is a limited teacher-visible feasibility study.

Phase 4, justified only if earlier privacy and reliability gates are met, compares instructional conditions.

A feasible initial sample could span several teachers, subjects, grade levels, class sizes, and lesson formats; the final sample size should follow a power analysis tied to a prespecified primary outcome rather than a convenient number.

Possible comparison conditions include ordinary teaching without the system, a passive dashboard that displays trends, and persistent-trend prompts. A crossover design can reduce between-teacher variation, but order and carryover effects must be addressed. Randomization should occur at a hierarchical level that effectively limits cross-group contamination, and adjacent windows from the same lesson must not be treated as statistically independent participants. The pilot protocol, thresholds, exclusions, and analysis plan should be formally preregistered when practically feasible. See Figure 4 for details.

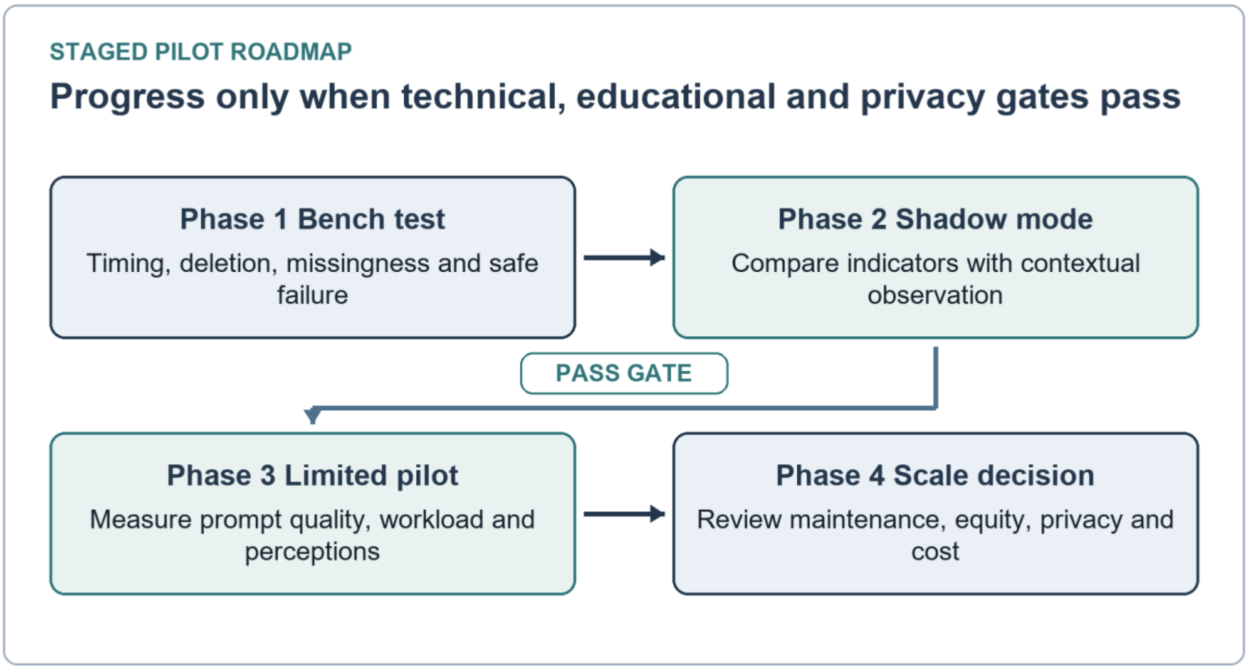

Figure 4. Staged pilot roadmap with progression gates.

Ground truth and technical analysis

Ground truth for group conditions cannot be assumed to be a single observer's judgment. Two or more trained observers can code lesson segments using a predefined scheme for activity format, response opportunities, transitions, and visible participation breadth. Inter-rater agreement should be reported, disagreements resolved transparently, and constructs kept at the same group level as the system output. Teacher post-lesson annotations may add context but should not be treated as infallible labels.

Technical evaluation should report end-to-end latency distributions, dropped-window rate, Central Processing Unit（CPU) or accelerator load, memory use, energy consumption, storage growth, offline availability, deletion success, modality quality, and recovery time after failure. Prompt evaluation should include precision, recall, false prompts per lesson, missed-event analysis, calibration, and subgrouped performance by subject, class size, room, activity profile, and teacher. Confidence intervals and complete confusion matrices are more informative than a single accuracy value.

Educational and human-factors analysis

Educational evaluation asks whether a prompt supports a productive act at an acceptable attentional cost. Outcomes may include prompt comprehension, acceptance by category, time to response, observed instructional adjustment, perceived usefulness, trust calibration, workload, interruption, and post-lesson recall.

Student measures may include perceived surveillance, classroom climate, and engagement questionnaires, but the study should avoid claiming learning improvement without suitable assessment design and duration.

Qualitative interviews can thoroughly explain why identical prompts prove useful in one specific context and disruptive in another distinct setting. Analysis should systematically compare accepted, ignored, contextually dismissed, and false-alarm episodes.

Video may be useful for research coding only if separately justified and properly governed; it is not a core component of the operational framework. Mixed-method triangulation can connect system events, classroom observation, and teacher firsthand accounts while preserving reasonable uncertainty when evidence clearly conflicts.

Privacy, ethics, and deployment gates

Privacy evaluation must test actual behavior rather than rely on architecture diagrams. Auditors should verify raw-buffer deletion, network destinations, account privileges, encryption, export contents, log retention, stop controls, withdrawal handling, and small-class inference risk. Threat modeling should consider a curious administrator, stolen edge device, compromised update package, and secondary use of prompt logs. Student and parent perceptions should be collected before and after exposure because technically minimized data can still create a feeling of monitoring.

Deployment evaluation covers installation time, sensor calibration, hardware and maintenance cost, update burden, teacher training, accessibility, and compatibility with existing school networks. Prespecified gates should determine progression: For example, no unresolved raw-data retention failure, acceptable false prompts per lesson, a minimum proportion of teachers who correctly interpret prompt limitations and demonstrated offline operation. The precise thresholds require stakeholder agreement; listing them in advance prevents a technically functioning prototype from being declared successful despite unacceptable educational or privacy costs.

The protocol therefore treats educational usefulness, privacy protection, and operational reliability as co-equal conditions of progression. The discussion now positions this framework against alternative classroom-analytics designs and draws out its theoretical and practical implications.

Discussion

This section interprets the framework rather than repeating its components. It compares the design with common alternatives, identifies theoretical and practical contributions, and states the trade-offs, limitations, and research sequence that delimit the paper's claims.

Design guidelines

The framework yields four design guidelines. First, use time-windowed group indicators rather than individual-level labels. Second, process raw signals locally and retain only aggregated indicators and audit logs. Third, trigger prompts only after persistent trends, not after single noisy events. Fourth, make teacher response part of the rule-adjustment process.

Comparative positioning and design trade-offs

(1) Camera-based systems

RGB and depth imaging can provide fine spatial detail and support posture or interaction analysis, but they also create a richer and more reusable record of identifiable classroom activity. The proposed framework sacrifices some diagnostic resolution to reduce visual exposure and keep its output at the group-window level. This is a proportionality choice, not a claim that radar is intrinsically anonymous or universally more accurate.

(2) Cloud-based analytics

Central services can provide elastic computation, fleet-wide updates, and cross-site model development. Their trade-off is routine transmission, dependency on connectivity, a larger breach surface, and more complex control over retention and secondary use. Edge processing narrows that surface and supports timely offline operation, while shifting responsibility to schools for device security, software updates, and local maintenance.

(3) Emotion-recognition systems

Emotion labels promise concise interpretations but risk converting ambiguous expressions, motion, or voice into high-authority psychological claims, especially in high-stakes educational contexts. The present design reports observable trends and uncertainty, avoids named-student inference, and treats teacher interpretation as necessary. It consequently cannot supply individual diagnoses, which is a deliberate limitation.

(4) Teacher dashboards

Persistent dashboards can support reflection and comparison, yet they may require continuous visual attention and leave the path from information to action unspecified. The prompt model reduces display density through persistence, priority, cooldown, explanation, and expiry. The trade-off is that summarization may omit detail a teacher would value after the lesson; optional aggregate review should therefore remain separate from real-time prompting.

(5) Individual recognition

Individual profiles can identify which learner appears to need support, but they also increase misclassification consequences, equity concerns, and the possibility of ranking or disciplinary reuse. Group aggregation cannot detect every individual difficulty and must supplement rather than replace teacher-student interaction. Its contribution is to support low-risk classroom-level adjustments without normalizing continuous personal profiling.

Theoretical contributions

The first theoretical contribution is granularity alignment: The granularity of collected data, inferred state, and supported action should be mutually consistent. When the intended action concerns lesson pace, response format, or activity rhythm, collecting an identifiable learner profile may be disproportionate. The second contribution is minimum-sufficient classroom evidence: System quality depends not on maximizing observable detail but on identifying the least revealing combination that can support a bounded pedagogical question. The third contribution is abstention as a legitimate educational-AI output. Insufficient quality, conflicting modalities, or an out-of-profile activity should prevent a prompt rather than be hidden inside an apparently precise score.

Human-AI collaboration and governance implications

The framework shifts the primary analytical object from an individual learner profile to a time-bounded classroom condition. This is not only a privacy technique. It changes what the system can legitimately claim and what kind of intervention it supports. Rather than diagnosing a person, the artifact helps a teacher ask whether pace, participation opportunity, or activity rhythm merits a contextual check. The design therefore aligns granularity of data, granularity of inference, and granularity of action. This position extends the principle that learning analytics should lead to pedagogically meaningful action.

A group-level signal is valuable only when its uncertainty and limits remain visible and when a teacher can translate it into a low-risk response. The system's most important output may sometimes be abstention. Designing for abstention, cooldown, and contestability resists the common tendency to evaluate educational AI only by how often it produces an answer.

Practical implications

For schools, deployment should begin with purpose definition, stakeholder notice, a local risk assessment, a retention schedule, and a stop procedure before sensors are installed. Procurement should include edge-node security, deletion verification, maintenance ownership, and an exit plan, not only hardware accuracy claims.

For teachers, the framework offers a glanceable prompt that can be accepted, dismissed, snoozed, or used as a cue to gather better evidence. Training should include examples in which the correct response is to ignore the system, thereby calibrating trust rather than maximizing compliance.

For developers, the main requirement is traceable transformation: each prompt should be explainable through eligible features, quality values, context profile, persistence duration, and rule version. Missingness and abstention need first-class interface states, and feedback should trigger periodic review rather than immediate opaque adaptation.

For education leaders and regulators, group aggregation should not be treated as automatic anonymization. Governance must address purpose change, access to prompt logs, small-class inference, cross-class comparison, model updates, complaints, and deletion evidence. Oversight is credible only when technical controls and institutional procedures enforce the same boundary.

Relationship to emerging multimodal classroom systems

Recent authentic-classroom research shows that real-time multimodal analytics can affect teacher awareness and support, while also revealing practical problems in orchestration and interpretation. Privacy-preserving cloud-edge classroom monitoring further demonstrates the relevance of processing location and data exposure to engagement analytics.

Teacher-dashboard studies report that combining modalities can generate useful insight but increases the need for careful visualization, co-design, and contextual explanation. Long-term in-the-wild deployment work similarly shows that stakeholder involvement, maintenance, and adaptation are central rather than peripheral to system performance.

The present framework is more conservative than systems that attempt individual engagement or emotion classification. That conservatism is intentional. Radar and non-semantic audio features reduce visual and linguistic exposure, but no sensor is intrinsically privacy-preserving. Aggregated features can still reveal routines, enable comparison, or be repurposed. The contribution is therefore the combination of signal minimization, local computation, group aggregation, explicit retention, teacher control, and governance, not a claim that one modality solves privacy.

Teacher agency, trust, and automation risk

Research on teacher trust in AI-powered educational technology indicates that trust is shaped by perceived usefulness, reliability, understanding, and professional learning rather than by accuracy alone. Human-factors research on automation bias shows that decision aids can create omission and commission errors when users over-rely on automated cues.

These findings support a design in which prompts provide reasons, avoid authoritative labels, expire, and can be contested. Training must address both uncritical reliance and reflexive rejection.

A newer K-12 study of multimodal learning-analytics dashboards emphasizes orchestration and teacher decision making, reinforcing the need to evaluate how analytics fits the temporal organization of classroom work. The framework contributes a privacy-first complement: Before asking which dashboard visualization is most informative, designers should ask which underlying data are necessary, whether raw signals must leave the classroom, and whether the proposed decision can be supported at group level.

Trade-offs and generalizability

Group aggregation reduces individual exposure but can hide a learner who needs help. It should therefore supplement, not replace, teacher-student interaction and established support processes. Edge computation reduces transmission and latency but creates device-management, update, and physical-security responsibilities. Interpretable rules simplify audit but may be less flexible than learned models. Persistent thresholds reduce noise but can delay a useful prompt. These are design trade-offs to be measured, not defects that can be eliminated by rhetoric.

Generalization is likely to be limited. A participation distribution meaningful in a seminar may be inappropriate in direct instruction; movement patterns in a science laboratory differ from those in reading; small classes increase inference risk; and cultural norms influence visible participation. Multi-site evaluation should therefore examine transportability and recalibration rather than pool all classrooms into one global threshold. Schools should be able to operate the core service without contributing data to a central model.

Limitations and staged future research

The paper remains a conceptual design and has not established construct validity, optimal weights, acceptable prompt frequency, causal teaching effects, or student acceptance. Signal distributions may change across cultures, subjects, grades, room layouts, class sizes, languages, and pedagogical formats. Small classes and stable seating zones can weaken the protection offered by aggregation. The assumed radar, audio-energy, and interaction pipeline may also be limited by hardware cost, calibration expertise, maintenance, accessibility, and teacher attention. These constraints prevent universal thresholds or claim of general effectiveness.

Future research should proceed through the staged route shown in Figure 4. A minimal prototype should first be tested with scripted or synthetic signals and explicit deletion and fault checks. Silent classroom deployment should then establish signal quality, calibration stability, and stakeholder acceptability without presenting prompts.

A limited teacher-visible pilot should evaluate the primary utility/interruption outcome and redesign the interface where necessary.

Only after technical, privacy, and usability gates are satisfied should a preregistered controlled study examine instructional processes and learning outcomes. Multi-site work should test recalibration and transportability rather than assume a universal model, and large-scale deployment should remain conditional on independent governance review.

Conclusion

This paper presents a privacy-preserving edge computing framework for group-level classroom learning-state perception and teacher decision support. Its research contribution is a traceable design artifact that connects policy and literature requirements to sensor choice, a five-layer architecture, quality-aware group indicators, bounded prompt logic, teacher feedback, and a staged evaluation protocol. The framework is deliberately testable but not yet validated as an instructional intervention.

The theoretical contribution is to align data, inference, and action at a classroom-group granularity. This alignment reframes privacy as a design constraint on what the system may legitimately claim, not merely a post-processing technique. Minimum-sufficient evidence, visible uncertainty, abstention, cooldown, and contestability further position human-AI collaboration as an allocation of authority in which the system supports noticing while the teacher retains contextual interpretation and decision responsibility.

The practical contribution is a deployment logic for schools, teachers, developers, and education leaders. Raw radar and audio-derived streams remain local and short-lived; longer-lived records are limited to necessary indicators, prompt events, responses, and audit evidence; and teacher controls include dismissal, snoozing, profile selection, and suspension. These safeguards reduce exposure but do not eliminate small-class inference, maintenance, function-creep, or institutional-governance risks.

Future research should implement the artifact and move from bench validation to silent deployment, a teacher-visible feasibility pilot, a controlled study, and only then multi-site scaling. Progression should depend on deletion verification, acceptable false-prompt burden, calibrated teacher understanding, student and parent acceptability, secure offline operation, and a prespecified educational rationale. Until such evidence exists, the framework's defensible claim is limited: privacy-first, group-level, teacher-controlled classroom analytics can be specified in a form that is sufficiently explicit for responsible empirical testing.

Funding

This work was supported by Provincial Undergraduate Training Program on Innovation and Entrepreneurship (Grant No: S202610345058).

Acknowledgements

The authors would like to show sincere thanks to those technicians who have contributed to this research.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Wang, P., Viberg, O., Law, E. L. C., Greiff, S. (2026) When can ai models explain learning? Validity criteria for ai as cognitive models in education. *Educational Psychology Review*, 38(1), 106.

[2] Radwan, A., Mcginty, J. (2024) Toward a conceptual generative AI ethical framework in teacher education. *Exploring New Horizons: Generative Artificial Intelligence and Teacher Education*, 88.

[3] Guberina, T., Procházka, F. (2026) Human-AI collaborative teaching: generative artificial intelligence (Gen-AI) as co-teacher. *Social Science Chronicle*, 6, 1-19.

[4] Romero, C., Ventura, S. (2020) Educational data mining and learning analytics: an updated survey. *Wiley Interdisciplinary Reviews: Data Mining And Knowledge Discovery*, 10(3), e1355.

[5] Finck, M., Biega, A. J. (2021) Reviving purpose limitation and data minimisation in data-driven systems. *Technology and Regulation*, 2021, 44-61.

[6] Sunyaev, A., Benlian, A., Pfeiffer, J., Jussupow, E., Thiebes, S., Maedche, A., Gawlitza, J. (2025) High-Risk Artificial Intelligence: Ali Sunyaev et al. *Business & Information Systems Engineering*, 67(6), 981-994.

[7] Joksimović, S., Kovanović, V., Dawson, S. (2019) The journey of learning analytics. *HERDSA Review of Higher Education*, 6, 27-63.

[8] Isohätälä, J., Näykki, P., Järvelä, S. (2020) Cognitive and socio-emotional interaction in collaborative learning: Exploring fluctuations in students' participation. *Scandinavian Journal of Educational Research*, 64(6), 831-851.

[9] Kim, J. (2024) Leading teachers' perspective on teacher-AI collaboration in education. *Education and Information Technologies*, 29(7), 8693-8724.

[10] Cabral, L., Pinto, R., Gonçalves, G. (2025) AI-powered learning analytics dashboards: a systematic review of applications, techniques, and research gaps. *Discover Education*, 4(1), 525.

[11] Fry, M., Wang, C. X., Vasquez, S., Shaffer, J., Clark, A., Wolfe, R., Reynolds, E. (2026) Faculty Lived Experiences with Mixed-Reality Simulation in Teacher Preparation: An Interpretive Phenomenology. *Journal of Technology and Teacher Education*, 34(1), 5-46.

[12] Vargas, E. G., Chiappe, A., Durand, J. (2024) Reshaping education in the era of artificial intelligence: insights from Situated Learning related literature. *Journal of Social Studies Education Research*, 15(2), 1-28.

[13] Wright, E., Taffe, S., Luxton, S., Wragg, N. (2026) Embedding authentic learning: A case study in curriculum transformation. *Education Sciences*, 16(3), 392.

[14] Van Zyl, L. E. (2026) The AI-IARA framework: How to cultivate human agency before artificial intelligence optimizes it a (ny) way. *The Journal*

of Positive Psychology, 1-19.

[15] Guerrero-Sosa, J. D., Romero, F. P., Menéndez-Domínguez, V. H., Serrano-Guerrero, J., Montoro-Montarroso, A., Olivas, J. A. (2025) A comprehensive review of multimodal analysis in education. *Applied Sciences*, 15(11), 5896.

[16] Becerra, A., Cobos, R., Lang, C. (2026) Enhancing online learning by integrating biosensors and multimodal learning analytics for detecting and predicting student behaviour: a review. *Behaviour & Information Technology*, 45(9), 1885-1910.

[17] Chiang, Y., Zhang, Y., Luo, H., Chen, T. Y., Chen, G. H., Chen, H. T., Chou, C. T. (2023) Management and orchestration of edge computing for IoT: a comprehensive survey. *IEEE Internet of Things Journal*, 10(16), 14307-14331.

[18] Tarapder, S. A. (2026) EDGE artificial intelligence based automation for ultra-low-latency control in industrial robotic systems. *Review of Applied Science and Technology*, 5(01), 01-37.

[19] Chen, J., Yan, H., Liu, Z., Zhang, M., Xiong, H., Yu, S. (2024) When federated learning meets privacy-preserving computation. *ACM Computing Surveys*, 56(12), 1-36.

[20] Bellavista, P., Foschini, L., Mora, A. (2021) Decentralised learning in federated deployment environments: a system-level survey. *ACM Computing Surveys (CSUR)*, 54(1), 1-38.

[21] Matcha, W., Uzir, N. A. A., Gašević, D., Pardo, A. (2019) A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, 13(2), 226-245.

[22] Hevner, A. R., March, S. T., Park, J., Ram, S. (2004) Design science in information systems research1. *MIS Quarterly*, 28(1), 75-106.

[23] Hossain, F., Anagnostou, E. N., Dinku, T., Borga, M. (2004) Hydrological model sensitivity to parameter and radar rainfall estimation uncertainty. *Hydrological Processes*, 18(17), 3277-3291.

[24] Zhao, C., Shen, W. (2023) Federated domain generalization: a secure and robust framework for intelligent fault diagnosis. *IEEE Transactions on Industrial Informatics*, 20(2), 2662-2670.

[25] Li, C. (2025) AI-driven governance: Enhancing transparency and accountability in public administration. *Digital Society & Virtual Governance*, 1(1), 1-16.

[26] McCrary, Q. D. (2026) Are we ghosts in the machine? AI, agency, and the future of libraries. *The Journal of Academic Librarianship*, 52(1), 103181.