

GOVERNING EDUCATION WHEN AI BEGINS TO ACT

WHEN SOFTWARE STOPS WAITING

For most of the digital era, educational technology has occupied a largely obedient position. A teacher opens the platform. A student submits a prompt. An administrator requests a report. The system responds. Even the recent explosion of generative AI has largely preserved this sequence: a human initiates, the technology produces, and a human decides what happens next.

That relationship is beginning to change.

The emerging generation of AI systems can do more than answer, summarize or generate. Given an objective, an agentic system may be able to interpret its environment, formulate a sequence of actions, use external tools, retain relevant memory, adapt when circumstances change and continue working toward an outcome without requiring a fresh human instruction at every stage. The technological shift may appear incremental—another improvement in artificial intelligence—but institutionally it is far more consequential. Software is moving from **responding within educational processes to participating in their execution**.

Imagine a school in which an intelligent system notices a pattern of student absence, examines previous communications, identifies available support appointments, proposes an intervention, schedules a meeting, drafts communication for the family and records the next review point. Elsewhere, another agent reorganizes timetable conflicts after a teacher becomes unavailable. An instructional agent monitors curriculum progression and recommends resources for classes falling behind. A professional-learning agent identifies recurring instructional needs from observation data and coordinates suitable development opportunities. None of these actions necessarily requires science-fiction levels of intelligence. They require something more immediately significant: permission to act across connected systems.

That is where the conversation about artificial intelligence in schools changes.

The familiar debate has concentrated on whether students should use AI, whether teachers can trust generated material, whether assessment remains credible and whether educators need better AI literacy. Those questions remain important, but they belong largely to a world in which humans still initiate most consequential interactions. Agentic systems introduce a different governance problem because they make **delegation** technologically plausible.

> ***The central question is no longer merely whether AI can perform an educational task competently. It is whether an institution should authorize a system to determine the next action, initiate it and continue pursuing an objective on the institution's behalf.***

This distinction matters because schools are not collections of independent administrative tasks waiting to be optimized. They are moral, relational and public institutions in which decisions distribute opportunity, attention, resources, recognition and sometimes life-changing consequences. A timetable adjustment and a disciplinary judgment are both decisions, but they do not carry equivalent authority. Sending a routine reminder and determining that a student requires a safeguarding intervention cannot be governed by the same threshold of autonomy. Efficiency alone cannot tell us where delegation should stop.

The future-ready school will therefore need more than AI adoption. It will need an explicit architecture of **decision rights**: a way of determining which forms of agency may be transferred to intelligent systems, under what constraints, with which mechanisms for review, and which forms of authority must remain recognizably human.

An agentic school should not be understood as a school in which AI acts everywhere. It is a school sophisticated enough to decide **where action may be delegated and where institutional authorship must remain human**.

1. From Generative AI to Agentic Systems

The language surrounding artificial intelligence has become crowded enough to obscure an important distinction. Generative AI, AI assistants, automation and agentic AI are often discussed as though they represent different degrees of the same phenomenon. They do not. Their boundaries are evolving, and researchers do not yet use every term identically, but the central distinction concerns the relationship between **generation and action**.

A conventional generative system primarily produces an output in response to an instruction: an explanation, lesson plan, summary, image, assessment item or piece of feedback. An automated system may go further by executing predefined rules—sending a notification when attendance falls below a specified threshold, for example. Agentic systems introduce another layer. They can be given an objective and possess varying capacities for planning, memory, tool use, environmental observation, feedback processing and autonomous or semi-autonomous action. Contemporary research therefore cautions against calling every generative-AI application an "agent." A 2026 systematic review of GenAI agents in education explicitly distinguishes ordinary generative systems from systems incorporating delegation, goal pursuit, planning, memory, tool use and action.

This distinction is essential because **generating a recommendation and possessing authority to execute it are institutionally different acts**. An AI assistant might tell a school leader that three classes have scheduling conflicts. An agentic system might identify the conflicts, evaluate alternatives, consult calendars, reassign rooms, notify relevant staff and continue resolving secondary clashes produced by the original change. The first expands information available to human judgment. The second begins to exercise bounded operational agency.

Recent technical scholarship similarly conceptualizes agentic AI around systems capable of perceiving an environment, reasoning about changing conditions, planning goal-directed actions and executing those actions through available tools. Multi-agent architectures extend this further by allowing specialized agents to interact and coordinate around shared or competing objectives. Commercial definitions increasingly make the same distinction between assistants that remain principally reactive and agents that can proactively pursue multistep goals, although educational institutions should be especially cautious about importing corporate metaphors of autonomous "digital workers" without examining the very different responsibilities schools carry.

Education is already moving toward this territory, although the empirical field remains immature. A major 2026 review examining 33 empirical studies found educational GenAI agents operating across cognitive and epistemic support, self-regulated learning, affective support and human–AI interaction. Importantly, the authors also identified trust, governance, contextual adaptation and multi-agent ecosystems as unresolved priorities. A broader 2026 scoping review mapped 474 publications on generative-AI-powered agentic systems in education and concluded that the relationship between technical agentic capability and pedagogical design remains insufficiently understood.

That evidentiary caution matters. The technological trajectory is moving faster than educational institutions' ability to develop norms for it. Schools should therefore resist two symmetrical errors. The first is technological inflation: describing every chatbot, adaptive platform or automated workflow as agentic simply because the term has become fashionable. The second is institutional complacency: assuming that an increasingly autonomous system is merely another productivity tool because its interface resembles familiar educational software.

What changes with agency is not simply capability. It is **the locus of initiative**.

Once a system can decide how to pursue an assigned goal, select among possible actions and interact with other systems on the institution's behalf, leaders must begin asking questions that conventional technology procurement rarely foregrounds:

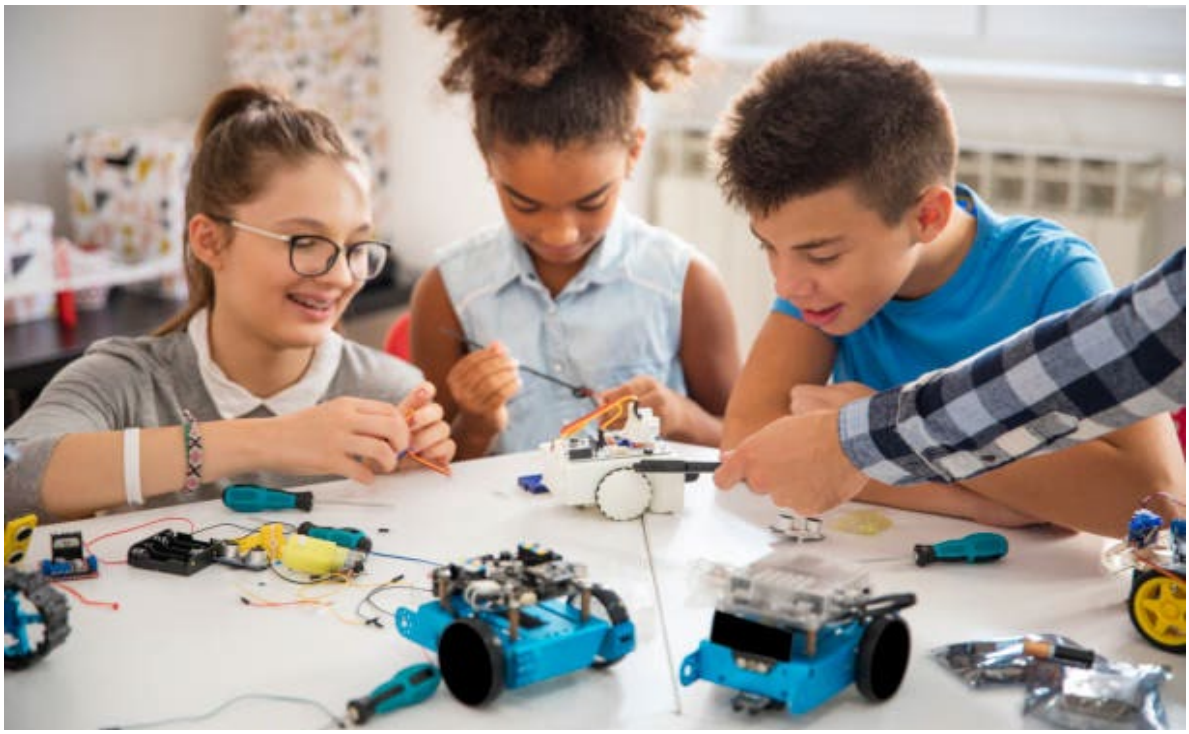
> ***What objective has the system actually been instructed to optimize?***
>
> ***Which information can it access?***
>
> ***Which actions can it initiate without approval?***
>
> ***What happens when its objectives collide with those of another system?***
>
> ***At what point must uncertainty trigger escalation rather than further autonomous action?***
>
> ***Most importantly, who remains accountable when an authorized system acts exactly as designed and the educational consequence is nevertheless unacceptable?***

These are questions of governance rather than functionality. They cannot be answered by model accuracy alone.

> ***The arrival of agentic systems therefore requires schools to abandon a simplistic binary between "human decision-making" and "AI automation."***

Future institutions are likely to contain multiple degrees of delegated agency, ranging from systems that merely surface information to systems authorized to coordinate and execute particular actions within carefully prescribed boundaries. The leadership challenge will be determining how much authority should accompany each capability.

That distinction will become foundational to the agentic school: **automation concerns what a system can perform; delegated agency concerns what the institution permits it to decide and do.**



2. THE SCHOOL AS AN ACTION ENVIRONMENT

Schools are unusually complex environments for agentic systems because they contain hundreds of recurring decisions distributed across teaching, administration, communication, student support and institutional governance. Much of this work already depends on digital platforms: learning-management systems, student-information systems, assessment databases, attendance records, scheduling software, communication tools and curriculum repositories. Agentic AI changes the significance of this infrastructure because these systems can potentially become connected surfaces upon which software does not merely retrieve information but coordinates activity across them.

The OECD's *Digital Education Outlook 2026* already describes generative AI as extending beyond classroom instruction into institutional operations, including curriculum alignment, assessment development, educational-resource classification, guidance and administrative workflows. The emerging agentic layer pushes this further. Instead of asking a system to draft a timetable, leaders could eventually authorize an agent to identify conflicts, examine constraints, negotiate alternatives across calendars, allocate available spaces and initiate changes within predefined parameters. Instead of merely summarizing attendance records, an agent might identify an emerging pattern, retrieve the school's intervention protocol, assemble relevant contextual information and recommend—or, where authorized, initiate—the appropriate next step.

The potential applications can be understood across several domains. In **instruction**, agents may help teachers locate curriculum-aligned resources, adapt materials, monitor learning progress or coordinate differentiated support. Current research already describes educational agents performing cognitive, self-regulatory, affective and human–AI interaction roles, while newer pedagogical-agent frameworks explore how AI might participate in knowledge construction and collaborative learning rather than functioning merely as a question-answering interface. In **administration**, agents may coordinate scheduling, documentation, routine communication, resource allocation and workflow follow-up. In **student support**, they could bring together evidence dispersed across multiple systems to help professionals recognize patterns that would otherwise remain fragmented. In **professional learning**, they might identify recurring instructional needs, recommend appropriate development opportunities and coordinate follow-up evidence.

These possibilities should not be confused with evidence that schools are already operating mature autonomous infrastructures. The research base remains emergent, and much of the empirical literature continues to concentrate on higher education rather than K–12 systems. Nevertheless, current work is moving beyond isolated conversational agents toward architectures in which specialized agents perform distinct functions, share information and coordinate activity. A recent scoping review of educational technology research describes multi-agent designs involving roles such as student, teacher and critic agents, while October 2026 research has proposed autonomous personalized-learning environments in which learner, teacher and institutional agents pursue differentiated objectives within a common system.

This movement matters because a school is not simply a collection of tasks. It is an **action environment** in which one decision can alter the conditions surrounding many others. A timetable change can affect teacher workload, student access, transport arrangements and support services. An attendance intervention can intersect with safeguarding, family circumstances and disability provision. An instructional recommendation can influence assessment, grouping and expectations. The more systems become interconnected, the less meaningful it becomes to evaluate an agent only according to whether its immediate action appears efficient.

The institutional question concerns scope of action. An agent that recommends three possible meeting times carries limited authority. An agent that sends invitations automatically carries more. An agent authorized to reschedule staff, alter student provision

> ***or trigger a formal intervention operates at a qualitatively different level. The technical architecture may be similar, but the educational consequences are not.***

This is where agentic schools will need to become more deliberate than many organizations have been during earlier waves of educational technology adoption. The United States Institute of Education Sciences is already encouraging district leaders to clarify decision rights, assess privacy and tool risk, document who should participate in AI decisions, and establish monitoring processes rather than treating adoption as a one-time procurement decision. That orientation becomes even more important once AI systems possess permission to act.

A future-ready school should therefore begin by seeing itself not merely as a place where AI tools are used, but as an institutional environment in which **actions carry different kinds of authority**. Some actions are routine and reversible. Others redistribute opportunity, create records, alter expectations or trigger formal processes. Before schools decide where agents belong, they must understand the architecture of their own decisions.

That leads to the article's next distinction: **delegation is not the same as automation**. The critical issue is no longer simply whether a task can be performed by software, but whether the authority embedded in that task should travel with it.

3. DELEGATION IS NOT THE SAME AS AUTOMATION

Automation is not new to schools. Attendance notifications, grade calculations, timetable generation, payroll processing and routine communications have long been partially automated. In these systems, the logic is usually predetermined: when a specified condition occurs, the system performs a defined action. The human institution designs the rule; the software executes it.

Delegated agency is different.

An agentic system may be authorized not merely to execute a prewritten instruction but to determine **how an objective should be pursued**. It may select among alternative actions, sequence those actions, revise its approach when conditions change, and interact with other tools or systems in order to complete the task. The difference is subtle in language but substantial in governance. Automation transfers labour. Delegated agency transfers a degree of **discretion**.

Consider a simple example. A conventional attendance system may automatically send a message after three unexplained absences. That is automation: the rule has already been decided. An agentic system, by contrast, might review attendance history, examine previous family communications, identify whether absences coincide with particular days, compare the pattern with existing support records, determine which intervention pathway appears appropriate, draft an individualized message and schedule follow-up. If the system is allowed to execute some or all of these actions independently, it is no longer merely following a workflow. It is exercising bounded judgment within that workflow.

This is why delegation cannot be evaluated solely through technical performance. A system may be highly accurate while still possessing more authority than the institution should reasonably transfer to it. Conversely, a low-risk administrative process may justify considerable autonomy because its consequences are easily reversible and its decision space tightly constrained.

> ***The governing question must therefore shift from "Can the system do this?" to "What kind of authority is embedded in doing this?"***

That distinction becomes especially important in education because many apparently ordinary processes contain hidden judgment. Course recommendations influence opportunity.

Intervention thresholds shape which students receive scrutiny. Grouping decisions affect access to challenge. Behavioural classifications influence how subsequent conduct is interpreted. Even resource-allocation decisions can privilege one educational priority over another. What appears operational from a systems perspective may be normative from an educational one.

Delegation therefore requires leaders to identify the **decision rights** embedded inside a task:

Who has traditionally had authority to make this judgment?

What knowledge is considered legitimate?

Which consequences follow from the decision?

Can the decision be challenged?

Can its effects be reversed?

Does it alter a learner's opportunity, status or institutional record?

These questions reveal that the same technology can be appropriate at one level of authority and unacceptable at another.

A useful distinction is between **execution authority** and **judgment authority**. Execution authority allows a system to carry out a decision whose parameters have already been established. Judgment authority allows the system to select among consequential alternatives. Schools may reasonably delegate considerable execution authority in low-risk environments while retaining human control over judgments involving rights, welfare, opportunity and high-stakes educational consequences.

This does not mean humans must manually approve every action. Such a model would eliminate much of the operational value of agentic systems. Instead, schools need what might be called **bounded delegation**: autonomy within clearly specified domains, constraints and escalation rules. An agent may act independently while conditions remain ordinary, but uncertainty, conflict, exceptional circumstances or predefined thresholds trigger human review.

That distinction also exposes the limitations of the familiar phrase **"human in the loop."** A person can remain nominally involved while having little meaningful control. If the system produces a recommendation that is rarely questioned, executes actions too quickly for review, or operates through reasoning that staff cannot meaningfully inspect, human presence may become ceremonial rather than authoritative. What matters is not whether a human appears somewhere in the process, but whether a human retains genuine capacity to **understand, intervene, override and assume responsibility**.

> ***For agentic schools, the central governance issue is not the elimination of autonomy but its distribution. Different actions warrant different levels of delegated authority.***

A system may reasonably be permitted to coordinate room bookings, reconcile calendars or send routine reminders. It may require explicit approval before altering instructional provision. And there will be decisions—particularly those involving safeguarding, discipline, high-stakes assessment, exclusion, psychological classification or substantial restriction of opportunity—for which institutional responsibility should remain fundamentally human.

The task for leadership is to make those boundaries explicit before autonomous capability makes them invisible.

This is where the article's central framework becomes necessary. If schools are to delegate agency responsibly, they need a systematic way to distinguish between systems that inform, systems that recommend, systems that coordinate, systems that execute within boundaries, and decisions whose authority should not be delegated at all.

4. WHO HOLDS THE DECISION RIGHT?

If agentic systems can operate at different degrees of autonomy, schools need a language more precise than "AI-assisted" or "AI-driven." These terms reveal little about the actual distribution of authority. A system that summarizes evidence for a principal and a system authorized to initiate an intervention may both be described as AI-assisted, yet the institutional implications are entirely different. The critical variable is **decision right**: who is permitted to determine what happens next.

This question is becoming increasingly important as education systems move toward formal AI governance. UNESCO's 2026 global consultation on education in the age of AI places institutional agency, public accountability and deliberative governance at the centre of the policy problem, arguing that education systems must retain the capacity to shape technological adoption according to their own purposes rather than simply accommodate technological diffusion. OECD analysis similarly shows that emerging education policies increasingly emphasize human oversight and professional judgment rather than unrestricted automation.

For schools, however, "human oversight" remains too imprecise unless leaders can specify how much authority a system possesses before human intervention becomes necessary. The Educational Agency Ladder offers a practical way to make that distinction.

The Educational Agency Ladder

Level	AI Role	Decision Relationship
Inform	Detects, retrieves, summarizes or organizes information.	Humans interpret the evidence and decide what happens.
Recommend	Proposes an action, response or set of alternatives.	Humans explicitly accept, reject or modify the recommendation.
Coordinate	Organizes low-risk activity across systems according to established institutional rules.	Humans supervise the operating environment and intervene when necessary.
Execute Within Bounds	Initiates and completes actions autonomously inside a narrowly defined mandate.	Humans establish permissions, escalation thresholds, review mechanisms and override authority.
Reserved Human Authority	Decisions are judged too consequential, rights-sensitive or difficult to reverse for autonomous delegation.	Meaningful human judgment and accountability are mandatory.



The ladder is not intended as a maturity model in which higher levels are more advanced or desirable. A school should not aspire to move every process toward Level 4. In fact, institutional maturity may sometimes be demonstrated by deliberately keeping a decision at Level 1 or Level 2 despite the technical capacity to automate it further. The appropriate level depends upon the nature of the action rather than the sophistication of the technology.

At **Level 1**, AI expands visibility without exercising authority. A system might identify timetable anomalies, summarize curriculum coverage or consolidate assessment information. The professional remains responsible for interpretation. At **Level 2**, the system moves from describing circumstances to recommending a response. It might suggest possible interventions, resource allocations or scheduling alternatives, but the consequential choice remains human.

Level 3 introduces coordination. Here an agent may interact with several systems to arrange low-risk activity—reconciling calendars, identifying available rooms, assembling meeting materials or coordinating routine communication. The system possesses more operational freedom, but the underlying decision space remains deliberately narrow.

The most significant transition occurs at **Level 4: Execute Within Bounds**. Here the institution has authorized the agent to act without obtaining separate approval for every step. A system might automatically reschedule non-instructional meetings within predetermined constraints, initiate routine reminders, reorder approved resources below a financial threshold or complete defined administrative workflows. This is genuine, delegated agency. Its legitimacy therefore depends upon explicit permissions, observable actions, escalation rules and reliable human override.

Level 5 is deliberately different. It does not describe a more capable form of AI. It identifies a boundary around forms of authority that an institution determines should remain human. Decisions involving serious safeguarding judgments, disciplinary sanctions, exclusion, high-stakes assessment, substantial changes to educational pathways, consequential psychological classification or significant restriction of student opportunity are obvious candidates. Legal requirements will vary across jurisdictions, but the governance principle should remain consistent: some decisions carry forms of moral and public accountability that cannot be discharged simply by demonstrating that an algorithm performs them accurately.

UNESCO's September 2026 ministerial statement reinforces this broader orientation, emphasizing that AI integration in education must remain subject to public accountability and that learning is not served when AI substitutes indiscriminately for human relationships and professional judgment. The implication for agentic schools is important: the objective is not maximum autonomy. It is **appropriate autonomy**.

Moving up the ladder should therefore become harder as the stakes increase. Three variables are especially useful: **stakes, reversibility and contestability**. A low-stakes action that can be easily reversed and challenged may tolerate considerable autonomous execution. A decision that alters a learner's opportunities, becomes difficult to undo or cannot be meaningfully contested should require substantially stronger human authority.

This principle protects schools from a seductive but dangerous assumption: that once technology becomes reliable enough, greater delegation naturally follows. Reliability matters, but it is not the sole determinant of legitimate authority. A system may be technically excellent and still be inappropriate as the final decision-maker.

The Educational Agency Ladder therefore relocates the conversation from technological capability to institutional authorization. The question is no longer simply whether an agent can perform an action competently, but whether the school is prepared to grant it the **right to act**.

> ***Capability does not confer authority. Authority must be deliberately assigned.***

5. HIGH STAKES, LOW REVERSIBILITY

The Educational Agency Ladder becomes most important where decisions are consequential and difficult to undo. In such cases, the central question is not whether an AI system can reach a defensible conclusion. It is whether the institution can justify allowing that conclusion to acquire authority over another person.

Education contains many decisions whose consequences extend well beyond the moment in which they are made. A disciplinary sanction may alter a student's record and relationship with school. A course-placement decision can change access to future study. A safeguarding judgment can trigger formal intervention in family life. An assessment decision may influence progression, certification or admission. A recommendation about special educational provision can affect how a learner is subsequently taught and understood. Even when such decisions are later reversed, their effects may not disappear completely.

This is why **reversibility** deserves far more attention in AI governance. Technology discussions often concentrate on accuracy: how often did the system make the correct recommendation? Accuracy matters, but it cannot capture the full ethical weight of an error. Incorrectly scheduling a meeting and incorrectly restricting a learner's educational pathway are both mistakes, yet their consequences are incommensurable. One can be repaired with another calendar invitation. The other may reshape opportunity, expectations and institutional records.

The same distinction applies to agentic systems. As consequences become more serious, the institution should require more than model reliability. It should demand stronger human judgment, clearer evidence, greater transparency, meaningful opportunities for challenge and an identifiable person who possesses both the authority and responsibility to alter the decision.

UNESCO's guidance on generative AI in education makes this principle explicit by warning against ceding human accountability to AI in high-stakes decisions and by emphasizing that institutions should validate systems for bias, pedagogical appropriateness and potential harm before adoption. The OECD's 2026 review similarly finds that education systems are increasingly emphasizing accountability, transparency and human oversight rather than treating AI adoption as unrestricted automation.

These principles become even more important with agents because the risk may arise not from a single recommendation but from a **sequence of actions**. An agent might classify a case, select an intervention, notify relevant staff and update the institutional record before a professional has meaningfully examined the underlying judgment. Individually, each action may appear procedural. Collectively, they can convert an uncertain inference into an institutional reality.

Schools therefore need to distinguish between **decision support** and **decision displacement**. Decision support enhances professional judgment by surfacing evidence, identifying patterns, testing alternatives or reducing administrative burden. Decision displacement occurs when human professionals remain nominally responsible while the meaningful judgment has already been made by the system and is routinely accepted.

This is one reason "human approval" is not always an adequate safeguard. If a principal receives dozens of algorithmically generated recommendations each day, approval can easily become a procedural click rather than an independent act of judgment. Automation bias—the tendency to over-rely on computerized recommendations—can make technically supervised systems function in practice as autonomous ones. The issue is therefore not whether a human signature exists, but whether the decision process preserves enough time, information and professional authority for genuine scrutiny.

A useful governance is:

> ***The higher the stakes, the lower the permissible autonomy. The lower the reversibility, the stronger the required human authority.***

Contestability forms the third part of this equation. A learner affected by a consequential AI-supported decision should be able to understand that the decision occurred, know the significant basis on which it was made, provide relevant context and access a meaningful route for review. As AI becomes increasingly embedded in real-world decision environments, the OECD's 2026 AI literacy framework similarly emphasizes the capacity to understand, critically evaluate and make informed judgments about AI-mediated outcomes.

Schools should consequently identify a category of **non-delegable institutional authority** before agentic systems become operationally attractive. Safeguarding determinations, serious disciplinary decisions, exclusion, final high-stakes assessment judgments, consequential psychological classifications and decisions substantially constraining educational opportunity should ordinarily remain at Level 5 of the Educational Agency Ladder.

AI may still contribute at earlier stages. It may retrieve evidence, identify anomalies, organize records, suggest questions or present alternatives. But contributing intelligence is not equivalent to possessing final authority.

This distinction protects an essential feature of educational institutions: consequential judgments remain attached to people who can explain them, reconsider them, hear another person's account and accept responsibility for what follows. The more difficult a decision is to reverse, the less appropriate it becomes for institutional accountability to dissolve into the machinery of an autonomous workflow.



6. THE MULTI-AGENT SCHOOL

The most consequential version of the agentic school may not involve one highly capable system. It may involve **many specialized agents operating simultaneously across the institution**.

A school could eventually deploy one agent to monitor attendance patterns, another to support instructional planning, another to coordinate scheduling, another to manage routine family communication, another to identify professional-learning needs, and another to assist with resource allocation. Each might be individually useful. The harder problem emerges when their decisions begin interacting.

This possibility is no longer purely theoretical. Research on educational multi-agent systems is expanding rapidly. A 2026 IEEE survey describes multi-agent learning environments as systems in which distinct agents assume different pedagogical or operational roles and interact with one another, while identifying safety, privacy, fairness, transparency and human–agent coordination as unresolved trustworthiness challenges. A systematic review published the same year similarly found that educational agents are moving from isolated support tools toward multi-role collaborative systems and identified multi-agent ecosystems as a major research frontier.

Recent experimental designs illustrate what this could mean. A 2026 multi-agent framework for personalized education uses autonomous agents to interpret learner context and coordinate adaptive decisions rather than simply generating static recommendations. Other work has explored teacher, learner and evaluator agents collaborating to generate differentiated instructional materials, demonstrating how specialist agents can divide responsibility for analysis, adaptation and quality assurance.

The institutional significance lies in **interaction effects**. A single agent can be assessed against a relatively clear objective. A network of agents introduces the possibility that multiple legitimate objectives begin competing.

Consider a student whose attendance has declined while academic performance is also weakening. An attendance agent may prioritize increased presence. A wellbeing agent may identify signs suggesting reduced workload. An instructional-support agent may recommend additional intervention. A scheduling agent may seek the most efficient arrangement of available support periods. A resource-allocation agent may prioritize students whose predicted benefit from intervention is greatest.

None of these systems necessarily behaves irrationally. The difficulty is that they may be optimizing **different definitions of what is desirable**.

Schools already manage these tensions through professional deliberation. Educators routinely balance attainment against wellbeing, consistency against individual circumstance, efficiency against inclusion, and immediate intervention against long-term independence. These are not engineering defects waiting to be eliminated. They are constitutive tensions within education.

A multi-agent school therefore raises a governance question more difficult than determining whether an individual agent is accurate:

> ***Who arbitrates when intelligent systems pursue different versions of the educational good?***

One answer might be to create a coordinating "super-agent" responsible for resolving competing recommendations. Technically, such an arrangement may become feasible. Institutionally, however, it does not solve the underlying problem. It merely relocates the judgment. Someone must still determine which objectives receive priority, how competing values are weighted, which exceptions matter and when optimization should be interrupted altogether.

This exposes an important limitation in the language of efficiency. Schools are **multi-goal institutions**. They are expected simultaneously to promote learning, equity, wellbeing, safeguarding, belonging, inclusion, citizenship and responsible stewardship of finite resources. A system designed to optimize one objective may unintentionally degrade another. Fernando Reimers and colleagues make a related systems argument in their 2026 analysis of AI and education, emphasizing that schools operate as living, adaptive, multi-goal environments in which technological interventions can generate nonlinear and second-order effects across interconnected parts of the system.

The governance architecture of a multi-agent school must therefore extend beyond permissions for individual agents. Leaders will need visibility into **agent-to-agent interaction**: which systems share information, which agent can trigger another, where priorities conflict, how disputes are escalated, and whether cumulative actions remain aligned with educational purpose.

This is also where institutional values must become unusually explicit. Agentic systems do not remove values from decision-making. They operationalize them. Every objective function, escalation rule, ranking mechanism and permission structure encodes assumptions about what matters enough to optimize.

A future-ready school should consequently avoid constructing an invisible bureaucracy of interacting AI systems whose collective behaviour no individual leader fully understands. Multi-agent architecture requires institutional orchestration, not merely technical interoperability.

The more agents a school authorizes to act, the more important it becomes to preserve a human capacity to see the system as a whole—and to intervene when several individually rational actions begin producing an educationally irrational result.

7. ACCOUNTABILITY CANNOT BE DELEGATED

Agentic systems create an accountability problem that conventional automation rarely exposes so sharply. When a system can plan, select tools, initiate actions and continue working across several steps, responsibility can become diffused across designers, vendors, school leaders, staff and the system itself. Yet educational institutions cannot allow consequential authority to disappear into that chain.

This is increasingly recognized beyond education. UNESCO's October 2026 analysis of agentic workflows warns that as autonomous systems act more quickly and across increasingly complex processes, conventional "human-in-the-loop" safeguards can deteriorate into passive approval. Meaningful oversight requires architectural limits: explicit permissions, prohibited actions, escalation rules and mechanisms through which humans can genuinely interrupt or reverse what an agent is doing.

For schools, this means that every agentic system should have an identifiable **human owner**. Ownership does not require that one person approve every routine action. It means someone holds institutional responsibility for the domain in which the agent operates, understands the scope of its authority, can inspect its behaviour and has the power to suspend or modify that authority when necessary.

Accountability also requires **traceability**. If an agent changes a schedule, initiates communication, recommends an intervention or triggers another system, the institution should be able to reconstruct what happened: what information was used, which permissions were exercised, what action was taken and where human intervention occurred. UNESCO's broader AI ethics framework similarly emphasizes auditability, traceability, impact assessment and the principle that AI systems must not displace ultimate human responsibility.

This becomes particularly important when agentic systems operate at speed. A nominal opportunity for human review has little value if actions occur faster than professionals can meaningfully examine them. Oversight must therefore be designed into the workflow rather than added cosmetically at the end. Some actions may proceed automatically; others should pause at predefined thresholds. Uncertainty, conflicting evidence, rights-sensitive decisions, unusual circumstances or actions with low reversibility should trigger escalation.

Schools will also need to resist **accountability laundering**: the tendency to attribute an undesirable outcome to "the system" when the institution itself selected the technology, configured its permissions and authorized its use. An agent does not acquire moral responsibility merely because it possesses operational autonomy. The institution remains responsible for deciding where that autonomy begins and where it must stop.

This aligns with UNESCO's 2026 argument that education systems must retain agency over the terms on which AI is integrated, rather than allowing technological capability to determine institutional practice. Public accountability, deliberative governance and the preservation of education as a common good remain central precisely because AI systems are becoming more capable of shaping real-world decisions.

The practical implication is straightforward: an agentic school needs more than policies describing acceptable AI use. It needs a **governance layer** capable of identifying every authorized agent, the purpose it serves, the data it can access, the actions it may initiate, its escalation thresholds, its human owner and the conditions under which its authority expires.

The governing principle should be unambiguous:

Autonomy may be delegated. Accountability may not.

The next task, therefore, is practical. If schools accept that principle, how do they begin building agentic capacity without surrendering institutional control? That requires a transition pathway—not from "non-AI school" to "AI school," but from fragmented experimentation toward deliberately governed, bounded agency.

The Agent Accountability Architecture

If accountability cannot be delegated, schools need a practical mechanism for ensuring that every agentic system remains institutionally governable. The **Agent Accountability Architecture**, conceptualized by Javeria Rana, provides six controls that should be established before an AI agent is authorized to act within a school.

Accountability Control	Leadership Question	Practical Requirement
Mandate	What is the agent authorized to achieve?	Define a narrow purpose, operating domain and prohibited uses.
Permission	What may the agent see and do?	Specify data access, tools, systems, action rights and financial or operational limits.
Ownership	Who remains accountable for this agent?	Assign a named human role with authority to inspect, restrict, suspend or retire the system.
Traceability	Can we reconstruct what the agent did?	Maintain records of significant inputs, recommendations, actions, approvals, overrides and system-to-system interactions.

Escalation	When must autonomous action stop?	Establish triggers for uncertainty, conflict, high stakes, exceptional circumstances or low-reversibility decisions.
Review and Expiry	Does the agent still deserve the authority it was given?	Set review dates, evaluate outcomes and harms, revise permissions and allow authorization to expire.

These six controls shift accountability from a vague promise of "human oversight" toward a visible governance structure. An agent should never possess open-ended authority merely because its earlier performance has been satisfactory. Its mandate should remain bounded, its permissions should be no broader than necessary, and its autonomy should always be attached to an identifiable institutional owner.

The architecture also prevents a common governance failure: **permission creep**. Agentic systems may begin with a narrow administrative purpose and gradually acquire access to additional data, applications or decision domains because integration is convenient. Periodic review forces leaders to ask whether the expansion remains educationally justified rather than assuming that previously granted trust should automatically extend to new forms of action.

Escalation is particularly important. Schools should identify circumstances in which an agent must stop rather than improvise. Conflicting information, unusual learner circumstances, safeguarding implications, rights-sensitive decisions, substantial resource consequences or uncertainty beyond an agreed threshold should trigger human judgment. Effective escalation therefore becomes a sign of intelligent system design, not system failure.

Finally, authorization should have an **expiry condition**. An agentic system should not retain institutional authority indefinitely simply because nobody has formally withdrawn it. Schools should periodically re-establish that the original purpose still exists, the risks remain acceptable, the permissions remain proportionate and the educational benefits continue to justify deployment.

The Agent Accountability Architecture therefore establishes a simple institutional discipline:

No agent without a mandate.

No mandate without limits.

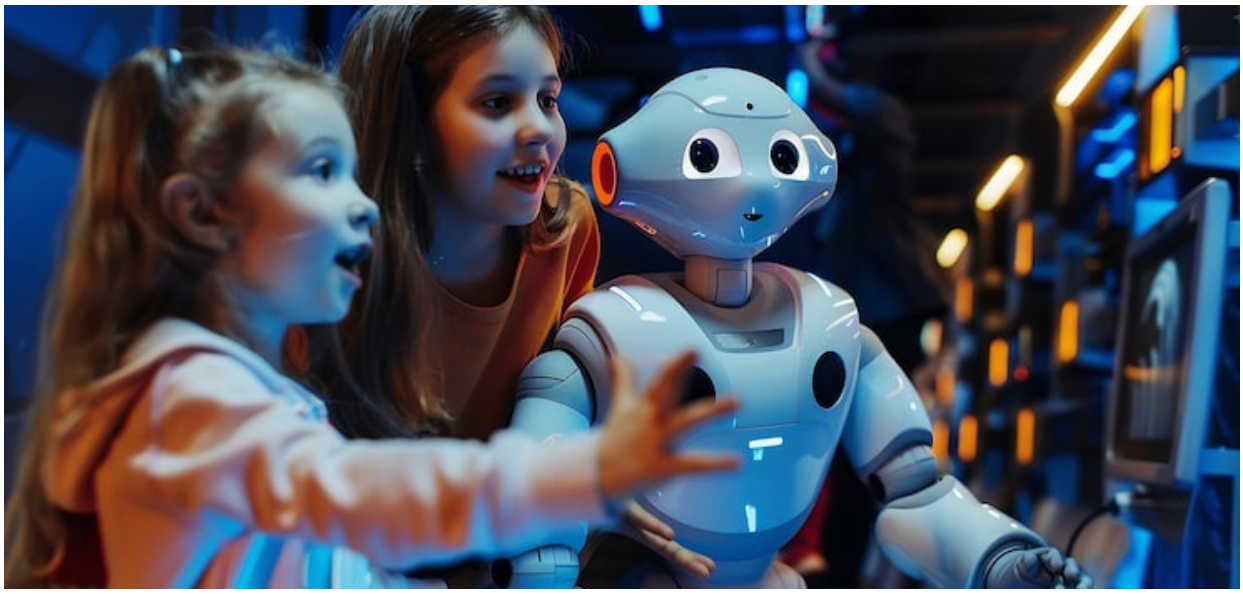
No autonomy without ownership.

No action without traceability.

No uncertainty without escalation.

No authority without review.

Used alongside the **Educational Agency Ladder**, the two frameworks answer different questions. The Agency Ladder determines **how much authority should be delegated**; the Agent Accountability Architecture determines **how that delegated authority remains governable once it has been granted**.



8. BECOMING AN AGENTIC SCHOOL

Becoming an agentic school does not begin with purchasing an autonomous AI platform. It begins with understanding how authority already moves through the institution. Schools contain hundreds of recurring decisions, but relatively few have been deliberately classified according to their stakes, reversibility, required professional judgment or suitability for delegation. Agentic readiness therefore requires organizational redesign before technological expansion.

This principle aligns with UNESCO's current advice to begin AI design with the educational problem rather than the algorithm. Its 2026 work with policymakers and educators deliberately reversed the common sequence of technology adoption: identify the problem, stakeholders and educational purpose first, and only then determine whether AI has a legitimate role. Similarly, the U.S. Institute of Education Sciences' October 2026 AI Planning Protocols ask school and district leaders to clarify decision rights, assess privacy and tool risk, evaluate readiness, communicate decisions and establish monitoring processes.

For an agentic school, this can be translated into six practical moves:

MAP DECISIONS BEFORE DEPLOYING AGENTS

The first task is a **decision inventory**. Leaders should map recurring processes across curriculum, assessment, attendance, student support, scheduling, communication, professional learning, procurement and administration. The purpose is not to document every workflow exhaustively, but to identify where judgment actually resides.

A timetable conflict, a curriculum recommendation and a safeguarding concern are all "decisions," yet they carry radically different forms of authority. Mapping them allows leaders to distinguish routine information processing from recommendation, coordination and consequential judgment.

Assign the Appropriate Level of Agency

Each prospective AI use case should then be located on the **Educational Agency Ladder**. Some systems should simply **Inform**. Others may **Recommend**. Low-risk workflows may permit agents to **Coordinate** or **Execute Within Bounds**. Decisions involving significant rights, welfare or educational opportunity should remain within **Reserved Human Authority**.

This prevents a common form of technological drift in which autonomy expands merely because the system proves capable of doing more. The correct level of agency should be determined by the nature of the decision, not by the maximum capability of the product.

Establish Institutional Red Lines

Schools should explicitly identify decisions they will not delegate.

These are likely to include serious safeguarding determinations, exclusion, major disciplinary sanctions, final high-stakes assessment judgments, consequential psychological classifications and decisions that substantially restrict a learner's educational pathway or opportunity.

The exact boundary will vary by jurisdiction and institutional context. What matters is that the boundary exists **before convenience begins negotiating it downward**.

A school without explicit non-delegable authority eventually risks allowing technical capability to determine ethical permission.

Begin with Bounded Pilots

Agentic adoption should begin where stakes are comparatively low, consequences are reversible and system behaviour can be observed closely. Suitable early environments may include meeting coordination, routine reminders, document collation, approved resource management or constrained administrative workflows.

The objective of an early pilot is not simply to prove that an agent saves time. It is to discover how delegated agency behaves within the institution: where unexpected actions occur, when humans override recommendations, which permissions prove excessive, how staff respond, and whether the agent generates new dependencies or unintended work.

This is where **Experimental Governance** becomes more valuable than enthusiastic adoption. Schools should determine the criteria for **stop, adapt or scale** before the pilot begins rather than interpreting every successful automation as justification for expansion.

Create an Agent Governance Register

Every authorized agent should appear in an institutional register containing, at minimum, its purpose, Agency Ladder level, permitted data access, authorized actions, prohibited actions, escalation conditions, human owner, review date and method of override.

The register transforms AI governance from a policy document into an operational architecture. Leaders can see not merely which AI products the institution owns, but **which systems possess permission to act**.

Combined with the Agent Accountability Architecture, this also makes permission creep easier to detect. If an administrative agent originally authorized to coordinate appointments later gains access to sensitive student-support information, that expansion should constitute a new governance decision rather than an invisible software update.

BUILD AGENCY LITERACY IN PEOPLE

The final transition is professional rather than technological. Staff in an agentic school need more than prompt-writing skills. They need **agency literacy**: the ability to understand what an AI system is authorized to initiate, recognize the difference between recommendation and delegated judgment, inspect its actions, challenge its outputs, trigger escalation and intervene when automated behaviour no longer serves educational purpose.

Leadership development must therefore move beyond "how to use AI." School leaders will increasingly need to understand permissions, decision rights, automation bias, escalation, auditability and the institutional consequences of allowing systems to operate across interconnected environments.

This is why becoming agentic should not be confused with becoming more automated. A school may automate extensively while exercising very little strategic agency of its own. The more mature institution deliberately decides **where intelligence assists, where agency is bounded and where human authority remains non-transferable**.

The transition can be summarized in three imperatives:

Automate tasks.

Delegate bounded agency.

Retain institutional authorship.

That last principle is the most important. An agentic school is not defined by how independently its technologies can operate. It is defined by whether the institution remains capable of determining the purposes those systems pursue, the authority they possess and the conditions under which they must stop.

The practical work of accomplishing that transition—mapping decision rights, identifying institutional red lines, designing bounded pilots and constructing an agent-governance architecture—requires more than a technology policy. It requires deliberate institutional design. That is where the conversation must ultimately move from understanding agentic AI to **governing the agentic school**.

CONCLUSION: SCHOOLS MUST REMAIN AUTHORS OF THEIR OWN AGENCY

The arrival of agentic AI does not simply give schools more powerful tools. It changes the nature of the institutional question. When systems can observe, plan, coordinate and act, the issue is no longer only whether artificial intelligence is accurate, efficient or pedagogically useful. It is whether schools have deliberately decided **which forms of authority may be delegated, which must remain contestable and which should never be surrendered at all**.

That distinction will define the maturity of the agentic school.

A technologically advanced institution may still be poorly governed if it allows autonomous systems to expand their permissions without clear boundaries, assigns consequential decisions to opaque workflows, or retains nominal human oversight without meaningful human authority. Conversely, a genuinely future-ready school may use highly capable agents while preserving strict limits around high-stakes judgment, maintaining transparent escalation pathways, auditing system behaviour and ensuring that every delegated action remains attached to an identifiable human responsibility.

The most important leadership shift is therefore from **AI adoption to agency governance**.

Schools will need to ask not simply what an agent can do, but what the institution is willing to authorize it to do on its behalf. They will need to distinguish execution from judgment, efficiency from legitimacy, assistance from authority, and autonomy from accountability. The Educational Agency Ladder helps determine how much decision authority may be delegated. The Agent Accountability Architecture ensures that delegated authority remains governed. The practical pathway toward agentic schooling begins with decision mapping, red lines, bounded pilots, governance registers and the development of agency literacy across staff.

The deeper challenge, however, is not technical. It is institutional authorship.

> ***Every agent acts toward an objective. Every permission structure encodes priorities. Every escalation rule contains assumptions about what matters. Every automated decision carries some conception of educational value. Agentic systems therefore do not remove institutional judgment; they embed it into infrastructure. The school of the future must remain capable of recognizing those judgments as its own.***

This is why the defining principle of agentic education should be neither maximum autonomy nor permanent human micromanagement. It should be **bounded agency under deliberate institutional authority**.

Future-ready schools will undoubtedly automate more. They will delegate more. They will increasingly rely on systems capable of acting across complex environments. But the measure of progress should not be how completely people can be removed from the process.

It should be whether the institution remains the author of the purposes, limits and responsibilities through which its intelligent systems act.

Automate tasks. Delegate bounded agency. Retain institutional authorship.

FURTHER READING

Stuart Russell — *Human Compatible: Artificial Intelligence and the Problem of Control*
Russell examines one of the foundational problems beneath agentic AI: how increasingly capable systems can pursue objectives while remaining aligned with human purposes and subject to meaningful control. His work is especially relevant to educational leaders thinking about delegated authority, bounded autonomy and the danger of allowing technical optimization to substitute for institutional judgment.

Ajay Agrawal, Joshua Gans and Avi Goldfarb — *Power and Prediction: The Disruptive Economics of Artificial Intelligence*
Agrawal, Gans and Goldfarb distinguish AI prediction from the wider decision systems in which predictions acquire value. Their analysis is useful for schools because it shifts attention away from isolated AI capabilities toward the redesign of workflows, decision rights and organizational structures that occurs once intelligent systems become embedded in institutional processes.

Brian Christian — *The Alignment Problem: Machine Learning and Human Values*
Christian explores the difficulty of translating human goals, preferences and values into computational systems. For education, the book raises an essential question: when an AI system optimizes an objective such as attainment, attendance or efficiency, what educational values have been encoded—and what important purposes may have been left outside the model?

Ben Williamson — *Big Data in Education: The Digital Future of Learning, Policy and Practice*
Williamson examines how algorithms, data infrastructures and digital platforms reshape educational governance. Although written before the current wave of agentic AI, the book provides an important conceptual foundation for understanding how technological systems can acquire institutional authority and influence what schools recognize, measure and prioritize.

Tom Davenport and Nitin Mittal — *All-in on AI: How Smart Companies Win Big with Artificial Intelligence*
Although written primarily for organizational and business contexts, this book is useful for understanding AI as an institutional capability rather than a collection of isolated tools. Educational leaders can read it critically alongside the governance arguments in this article, particularly when considering why technological integration requires changes in organizational design, leadership and decision-making structures.

Ethan Mollick — *Co-Intelligence: Living and Working with AI*
Mollick offers an accessible examination of how people and organizations can work alongside increasingly capable AI systems. For educational leaders, its value lies in considering how human roles may change as AI becomes more proactive, while also reinforcing the importance of experimentation, professional judgment and deliberate boundaries around delegation.

Taken together, these works help shift the conversation from **what AI can generate** toward the more consequential questions of **what systems are permitted to decide, how authority is distributed, whose values are being operationalized, and how human institutions retain authorship as artificial intelligence becomes increasingly capable of acting**.

Continue the Conversation



The move toward agentic education will not be solved by adding another AI policy to the staff handbook. Schools will need to examine how authority actually moves through their systems: which decisions can be automated, where bounded agency is appropriate, which forms of judgment must remain human, and how accountability is preserved when intelligent systems begin acting across institutional workflows.

These questions form the foundation of my professional-learning masterclass:

Governing the Agentic School

Decision Rights, Human Authority and AI Systems That Act

The facilitated masterclass is designed for school leaders, system leaders, academic teams and education organizations seeking to move from fragmented AI experimentation toward deliberate institutional governance. Participants examine their own decision structures, identify suitable and unsuitable domains for delegation, establish non-negotiable red lines, and design bounded agentic pilots aligned with educational purpose rather than technological novelty.

The masterclass extends the ideas introduced in this article through practical tools including decision-rights mapping, the Educational Agency Ladder, agent accountability analysis, institutional red-line design, governance registers, escalation protocols and **stop–adapt–scale** planning for bounded AI pilots.

The aim is not to help schools become more automated for its own sake. It is to help them become more intentional about **where intelligence assists, where agency may be delegated, and where institutional authorship must remain unmistakably human**.

For facilitated masterclasses, leadership consultation or professional-learning partnerships, connect with me through **Café Learning** or emailing ***info@cafelearning.org***. The accompanying ***Governing the Agentic School*** masterclass brief provides a practical introduction to this work and a pathway for leadership teams ready to begin designing their own agentic governance architecture.

CAFE LEARNING MASTERCLASS BRIEF

PUBLIC EDITION / LEADERSHIP SERIES 02

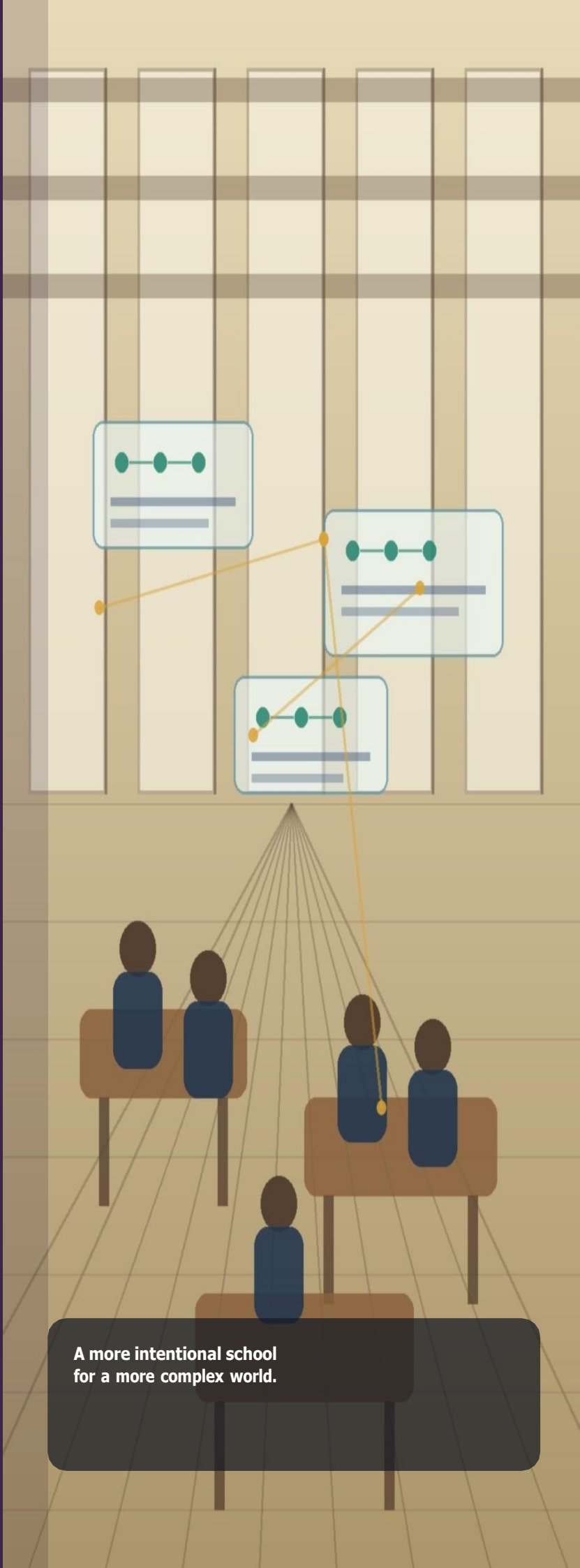
Governing the Agentic School

Decision Rights, Human Authority and AI Systems That Act

A practical masterclass for school and system leaders ready to design, govern and implement agentic AI with confidence.

RESEARCH-INFORMED

PRACTICE-READY



Javeria Rana

Educationist / Academic Leader / Researcher

SELECTED PUBLIC LEARNING RESOURCE

Full facilitated masterclass available separately.

www.cafelearning.org

MEET THE AUTHOR

Leadership at scale, grounded in learning

This public brief translates research and field experience into a compact governance resource for agentic schools.



Javeria Rana

Director Academics, UIES
Founder & CEO, Cafe Learning

Academic leadership across a network of 500+ schools in Pakistan. Cambridge Master Trainer, Scientix Ambassador, and Global Schools Program Mentor & Advocate.

WHY THIS MASTERCLASS

Agentic AI can help schools coordinate work, reduce friction and support learners more responsively. It also introduces a new governance problem: systems can increasingly act, not merely generate.

CLARIFY

1 Which decisions may be delegated - and which must remain human.

GOVERN

2 Build a practical architecture for bounded agency.

PILOT

3 Move from experimentation to safe, observable implementation.

LEAD

4 Develop agency literacy for an agentic future.

PUBLIC EDITION BOUNDARY

This brief provides selected concepts and field tools. The full facilitated masterclass adds contextual cases, team diagnostics, decision-rights mapping, escalation design, pilot protocols and implementation support.

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

WHY NOW

Agentic AI changes the governance question

The shift is from what AI can generate to what institutions permit AI to do.

33
Empirical studies in a 2026 systematic review of GenAI-agents in education.

4
Core educational agent roles identified: cognitive, self-regulatory, affective and human-AI interaction.

37%
of lower-secondary teachers reported using AI for their job in 2024.

WHAT THE EVIDENCE SAYS
The educational evidence base is growing, but remains concentrated in higher education. Recent reviews repeatedly identify trust, governance, contextual adaptation and multi-agent ecosystems as unresolved priorities.

THE SYSTEM-LEVEL QUESTION

UNESCO's 2026 global consultation asks what it takes for education systems to exercise their own agency in the face of rapid AI diffusion - rather than allowing technological capability to define institutional practice.

The challenge is not adoption alone. It is institutional authorship.

Sources: Wang et al. (2026); OECD Digital Education Outlook 2026; UNESCO global consultation on education in the age of AI (2026).

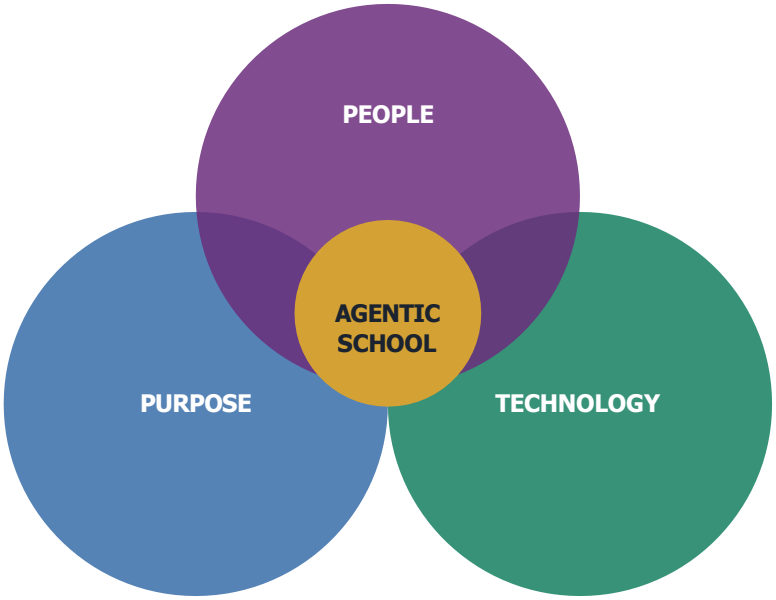
© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

CORE CONCEPT

What Is an Agentic School?

More than technology. A deliberate institution.

An agentic school is not one in which AI acts everywhere. It is a school sophisticated enough to decide where agency may be delegated, under whose authority, toward which purposes, and where human authorship must remain unmistakable.



01 Intentionality
We decide where agency is appropriate.

02 Accountability
Delegated authority remains governable.

03 Institutional authorship
Human authority remains visible.

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

FROM TOOLS TO AGENCY

From AI Tools to Agentic Systems

A clear distinction for clearer decisions.

1 Traditional software

Follows predefined instructions.

e.g. attendance report

2 Generative AI

Produces content in response to a prompt.

e.g. lesson plan or summary

3 Automated system

Executes predefined rules or triggers.

e.g. send low-attendance alert

4 Agentic system

Can observe, plan, use tools and act toward an objective.

e.g. coordinate support and follow up

Human initiates every action → Increasing autonomy and delegated agency

AUTOMATION TRANSFERS LABOUR. DELEGATED AGENCY TRANSFERS DISCRETION.

A timetable generator and an agent that decides how to resolve timetable conflicts may use similar technical infrastructure, but they do not carry the same institutional authority. The governance question is: what kind of decision right is embedded in the action?

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

JAVERIA RANA FRAMEWORK

The Educational Agency Ladder

Five levels of delegated authority - higher capability does not mean higher entitlement.

Higher stakes
Lower reversibility
Stronger human authority

5 Reserved Human Authority

Decisions too consequential, rights-sensitive or difficult to reverse for autonomous delegation.

4 Execute Within Bounds

Acts autonomously inside a narrow mandate with explicit limits, escalation and override.

3 Coordinate

Organizes low-risk activity across systems under established rules.

2 Recommend

Proposes an action or set of alternatives; humans accept, reject or modify.

1 Inform

Detects, retrieves, summarizes or organizes information for human interpretation.



Lower stakes
More reversible
More delegation possible

RULE OF THUMB

The higher the stakes and the lower the reversibility, the stronger the required human authority. Contestability should increase as consequences become more serious.

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

JAVERIA RANA FRAMEWORK

The Agent Accountability Architecture

Six controls for governable agency.

M

Mandate

Define purpose, domain and prohibited uses.

P

Permission

Specify data access, tools, systems and action limits.

O

Ownership

Assign a named human role with authority.

T

Traceability

Record significant inputs, actions, approvals and overrides.

E

Escalation

Define when autonomous action must stop.

R

Review & Expiry

Re-evaluate outcomes, permissions and continued authority.

No agent without a mandate. No mandate without limits. No autonomy without ownership.
No action without traceability. No uncertainty without escalation. No authority without review.

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

INSTITUTIONAL ORCHESTRATION

The Multi-Agent School

When several rational systems optimize different versions of the educational good.



THE ORCHESTRATION QUESTION

Who arbitrates when the attendance agent prioritizes presence, the wellbeing agent recommends reduced demand, the instructional agent increases intervention and the resource agent optimizes scarce capacity? Multi-agent architecture requires institutional orchestration, not merely technical interoperability.

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

IMPLEMENTATION PATHWAY

Becoming an Agentic School

Six moves from fragmented experimentation to deliberately governed agency.

01 Map decisions before deploying agents

Inventory where judgment actually resides across curriculum, assessment, attendance, support, scheduling and operations.

02 Assign the appropriate level of agency

Place each use case on the Educational Agency Ladder.

03 Establish institutional red lines

Name decisions that remain Reserved Human Authority.

04 Begin with bounded pilots

Start low-stakes, reversible and observable; define stop-adapt-scale rules first.

05 Create an Agent Governance Register

Record purpose, permissions, prohibited actions, owner, escalation, override and review.

06 Build agency literacy in people

Develop staff capacity to inspect, challenge, escalate and intervene.

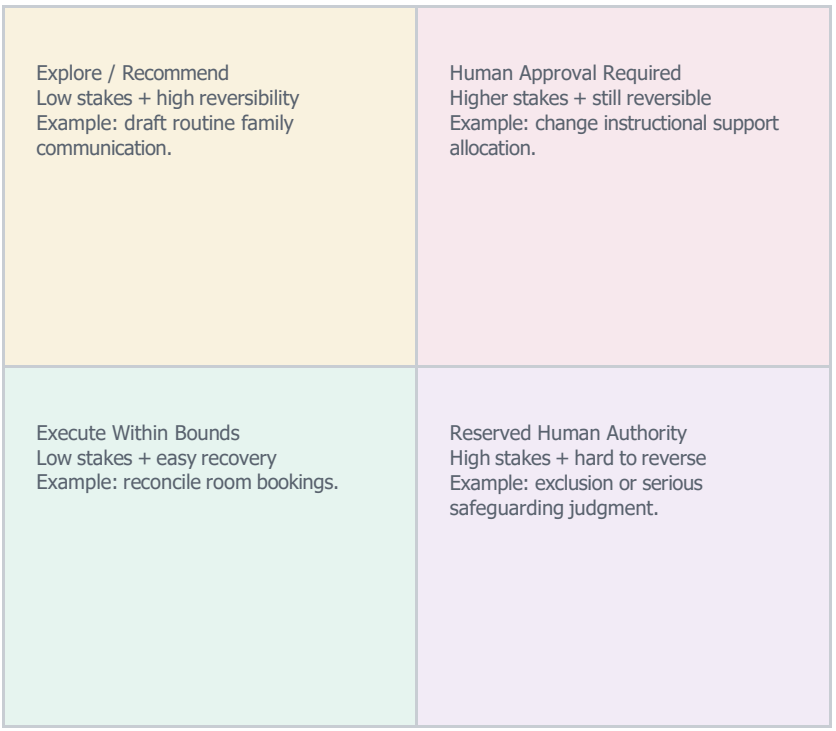
Automate tasks. Delegate bounded agency. Retain institutional authorship.

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

SELECTED FIELD TOOLS

Decision Rights Before Deployment

Use the nature of the decision - not the maximum capability of the technology - to set the level of agency.



FULL MASTERCLASS TOOLKIT

- Decision-rights mapping template
- Educational Agency Ladder worksheet
- Agent Governance Register
- Institutional Red-Line Register
- Risk & impact assessment tool
- Escalation-trigger checklist
- Bounded pilot planning guide
- Stop-adapt-scale review protocol
- Leadership reflection tools

Agent Governance Register - minimum fields
Purpose · Agency level · Data access · Authorized actions · Prohibited actions · Escalation · Human owner · Override · Review date

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

TURN INSIGHT INTO ACTION

Your 30-60-90 Day Agentic Readiness Plan

Choose one bounded use case, one governance question and one review point.

30

MAP + BOUND

Map decision flows
Identify one low-risk use case
Set red lines
Assign a human owner

60

PILOT + OBSERVE

Launch a bounded pilot
Track overrides and exceptions
Test escalation
Review evidence weekly

90

DECIDE + EMBED

Stop, adapt or scale
Update the governance register
Build agency literacy
Define next review date

OUTCOMES FOR YOUR SCHOOL

A shared language for agentic AI · clearer decision rights and institutional red lines · a practical governance architecture · stronger leadership confidence · one bounded pilot ready for responsible implementation

WHO SHOULD ATTEND

School leaders · system leaders · curriculum and academic teams · operational/support leaders · digital transformation and AI governance teams

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.

RESEARCH AND ATTRIBUTION

Selected References

This public brief draws on current research, policy guidance and Javeria Rana's author-developed governance frameworks.

Wang, C. et al. (2026). GenAI-agents in education: A systematic review of macro-level trends, educational roles, theoretical foundations and future agendas. Interactive Learning Environments. https://doi.org/10.1080/10494820.2026.2672562

OECD. (2026). OECD Digital Education Outlook 2026: Exploring Effective Uses of Generative AI in Education. OECD Publishing. https://doi.org/10.1787/062a7394-en

UNESCO. (2026). Global consultation on education in the age of AI: Sustaining education as a common good through deliberative governance.

Odrozek, K. (2026). Who answers for the machine? Closing the accountability gap in agentic workflows. UNESCO Global AI Ethics and Governance Observatory.

Busselle, B., Davenport, J., & Mendenhall, J. (2026). Helping District Leaders Make Informed Decisions about AI / AI Planning Protocols. Regional Educational Laboratory Northwest, Institute of Education Sciences.

Luo, J., Tang, Z., & Sun, M. (2026). Generative artificial intelligence in education from a technology research perspective: A scoping review of empirical studies from leading educational technology conferences. AI, Brain and Child, 2, 7.

Rana, J. (2026). Educational Agency Ladder and Agent Accountability Architecture. Author-developed frameworks for governing agentic AI in education.

CONTINUE WITH CAFE LEARNING

FACILITATE

Invite Javeria Rana to lead the complete masterclass with your school or leadership team.

CONSULT

Request bespoke leadership consultation on decision rights, red lines and agent governance.

PARTNER

Explore professional-learning, research, curriculum and system-improvement collaboration.

info@cafelearning.org · www.cafelearning.org

© 2026 Javeria Rana and Cafe Learning. Personal professional-learning use only.



***Future-Ready Schools** is an exclusive feature by Javeria Rana on The Worthy Educator. Check back regularly for new insights on education transformed!*