Jurnal Inovasi Pendidikan
dan Sosial Humaniora

SAH PUBLISHER

Vol. 1, No. 1, August 2026

Pedagogical Agency of Teachers as Algorithm Users in AI-Supported Education

Tedi Gandara[1*], **Muhammad Adi Pratama**[2]

[1]STIT Al Azami Cianjur

[2]Universitas Sugeng Hartono

**Corresponding author. Email: alwigandara999@gmail.com*

Abstract. This paper examines how the spread of artificial intelligence in education changes the professional position of the teacher. Its purpose is to clarify what pedagogical agency means when a growing share of classroom judgement is mediated by algorithmic systems, and to specify the conditions under which teachers remain authors of pedagogical decisions rather than executors of machine outputs. The method is conceptual analysis and framework construction. Three bodies of literature are brought into dialogue: ecological and sociocultural accounts of teacher agency, research on artificial intelligence in education, and scholarship on algorithmic accountability, human oversight and automation in human factors research. The analysis shows that the teacher's role is shifting from user of technology to user of algorithmic outputs, a shift that relocates agency from operating tools to interpreting, contesting and taking responsibility for recommendations, scores and predictions produced elsewhere. Building on this diagnosis, the paper proposes the Algorithmic Pedagogical Agency Framework, an original five-dimensional model comprising algorithmic literacy, interpretive discretion, override and contestation rights, data stewardship, and collegial and institutional conditions, together with six propositions that link the dimensions to pedagogical agency. The framework identifies where agency is gained and lost in recommendation systems, automated grading and learning analytics dashboards, and specifies the conditions that convert nominal human oversight into real agency. Its novelty lies in treating agency as a property of sociotechnical arrangements rather than of individual teachers, and in naming accountability without authority as the central failure mode of current deployment. The paper concludes with implications for teacher education and for the procurement of educational AI.

Keywords: teacher agency; artificial intelligence in education; algorithmic decision making; human oversight; learning analytics; pedagogical agency

Received July 2026 / Revised August 2026 / Accepted August 2026

This work is licensed under a Creative Commons Attribution 4.0 International License.

INTRODUCTION

Artificial intelligence has moved from the periphery of schooling into the centre of pedagogical decision making. Systems that recommend learning resources, sequence curriculum content, score student work, predict which learners are at risk and summarise classroom performance now participate in judgements that were once the exclusive province of teachers. Influential accounts of artificial intelligence in education present this development as an augmentation of the teacher, in which machines absorb routine cognition and release professional time for relational and creative work [1]. Early programmatic statements argued that intelligent systems could personalise feedback, open the black box of learning and support teachers in addressing variability among learners, provided that educators remained the designers of the learning experience [2].

Yet the empirical literature has grown faster than its account of the teacher's own position. A systematic review of research on artificial intelligence applications in higher education found that studies concentrate overwhelmingly on prediction, assessment and profiling of students, while the educators who are expected to act on these outputs are largely absent from the research itself [3]. The teacher appears in this literature as an implementation variable, a beneficiary or a risk, but rarely as the subject whose agency is being reconfigured.

Critical scholarship has warned against accepting the augmentation story at face value. The figure of the teacher displaced or disciplined by intelligent machinery has a long history, and contemporary promises of personalisation carry older logics of behaviourism, surveillance and market reform into new technical forms [4]. Histories of the field show that artificial intelligence in education has repeatedly been shaped by funding structures, commercial interests and policy fashions that marginalise pedagogical questions [5]. At the same time, state of the art reviews caution that most deployed systems remain narrow, that evidence of learning benefit is uneven, and that the strongest claims concern administrative efficiency rather than educational value [6]. What is missing across both the optimistic and the critical literatures is a precise account of what happens to teacher agency when the immediate object of professional work is no longer a tool the teacher operates but an output the teacher receives. Teachers increasingly confront recommendations they did not generate, scores they did not assign and risk labels they did not choose, produced by models whose reasoning they cannot inspect, trained on data they did not collect, and procured through contracts they did not negotiate. The now standard reassurance that a human remains in the loop says nothing about what that human can actually do inside it.

This paper addresses that gap. Its purpose is to conceptualise the teacher as an algorithm user, a professional whose agency is exercised in relation to algorithmic outputs, and to construct a framework that specifies the dimensions on which such agency depends. The paper makes three contributions. First, it distinguishes the position of the algorithm user from the established figure of the technology user and shows why the distinction matters for agency. Second, it integrates ecological theories of teacher agency with scholarship on algorithmic accountability and the human factors of automation, two literatures that have developed largely in isolation from one another. Third, it proposes the Algorithmic Pedagogical Agency Framework, comprising five dimensions and six propositions, which can guide teacher education, institutional policy and the procurement of educational artificial intelligence. The argument proceeds as follows. The methods section explains the conceptual approach. The discussion then analyses the shift from technology user to algorithm user, reconstructs teacher agency as an ecological achievement, examines three sites where agency is gained and lost, presents the framework and its propositions, considers the risks of automation bias and accountability without authority, and draws out implications for teacher education and procurement policy.

METHODS

This is a conceptual paper. It reports no new empirical data, and its claims are theoretical rather than statistical. The method is conceptual analysis combined with framework construction from the literature, following the convention that a conceptual contribution clarifies contested terms, exposes the assumptions carried by existing usage and reorganises a field around a more adequate set of distinctions. The analysis draws on three bodies of scholarship selected for their direct bearing on the research problem. The first is

the literature on teacher and professional agency, including ecological, sociocultural and relational accounts. The second is research and critical commentary on artificial intelligence in education, covering intelligent tutoring, recommendation, automated assessment, learning analytics and generative systems. The third is scholarship on algorithmic decision making and automation outside education, including work on algorithmic accountability, opacity, fairness and the human factors of automated systems, together with official guidance from international and governmental bodies.

The construction of the framework proceeded in four steps. The first step was concept clarification: the meaning of agency was reconstructed from its major theoretical treatments, and the respects in which algorithmic mediation differs from earlier forms of educational technology were specified. The second step was site mapping: the concrete settings in which teachers encounter algorithmic outputs were identified from the education literature, and the gains and losses of agency at each site were recorded. The third step was dimensional synthesis: the conditions that recurred across sites as enabling or constraining agency were grouped into a small set of dimensions, each defined so that it could in principle be observed, supported or withdrawn by identifiable actors. The fourth step was proposition formulation: the relations among the dimensions were stated as explicit propositions, so that the framework makes claims that subsequent empirical research can examine, refine or reject. Sources were restricted to peer reviewed scholarship from established publishers, scholarly monographs, and official publications of bodies such as UNESCO, the European Union and national education authorities. The framework offered here should be read as an analytical instrument. Its value will be settled by its usefulness in research and policy, not by the analysis that produced it.

RESULT AND DISCUSSION

1. From Technology User to Algorithm User: The Nature of the Shift

The established image of the teacher and technology is that of a user operating a tool. A teacher using an interactive whiteboard, a spreadsheet or a word processor confronts an instrument whose behaviour is, in principle, transparent and reversible. The tool does what it is told; its operations can be inspected as they happen; errors are attributable to the operator or to visible malfunction; and the decision being supported remains entirely the teacher's own. Agency in this arrangement is a matter of skill and confidence. The competent user extracts more from the tool, and training that builds operational competence genuinely enlarges professional freedom.

The algorithm user occupies a different position. The immediate object of work is not an operation but an output: a ranked list of recommended resources, a predicted grade, a probability that a student will fail, a generated summary of class performance. Three differences follow. The first is epistemic. An algorithmic output is not a function performed but a claim advanced, a probabilistic judgement about a learner or a text, produced by a model whose internal reasoning is not available to the teacher. Accepting, discounting or rejecting such a claim is an act of interpretation, not of operation. The second difference is temporal and spatial. Outputs arrive from elsewhere: from models trained on other populations, configured by vendors, and embedded in platforms selected by institutions or governments. By the time a recommendation reaches the classroom, many decisions that shape it have already been taken, and the teacher's judgement is exercised downstream of choices made far away. The third difference is institutional. Because outputs carry the authority of data and of the organisation that purchased the

system, departing from them can require justification in a way that following them does not. The default settings of the system become the path of least resistance, and professional discretion is quietly re-priced. Selwyn's warning that intelligent systems import new forms of control over teachers, and not only new forms of assistance for them, captures this asymmetry [4]. The shift from technology user to algorithm user is therefore not a change of device but a change in the structure of professional judgement, and it requires an account of agency suited to that structure.

2. Teacher Agency as an Ecological Achievement

Agency theory supplies the necessary starting point. Emirbayer and Mische define agency as a temporally embedded process of social engagement, composed of three elements: an iterational element oriented to past patterns of thought and action, a projective element oriented to imagined future possibilities, and a practical evaluative element through which actors deliberate and respond within the contingencies of the present moment [7]. On this account, agency is never a possession or a trait. It is something actors do, with varying success, as these temporal orientations are brought to bear on concrete situations.

Priestley, Biesta and Robinson develop this insight into an ecological model of teacher agency. Teachers achieve agency, or fail to achieve it, through the interplay of their iterational resources, including professional knowledge, beliefs and life histories; their projective orientations toward short and long term purposes; and their practical evaluative judgements in the present, which are enabled or constrained by cultural conditions such as dominant discourses and values, structural conditions such as relationships, roles and power, and material conditions such as resources and the physical and technical environment [8]. Their related empirical work shows that teachers' beliefs matter precisely because they shape what teachers take to be possible and worth doing within these conditions, so that agency varies between teachers in the same school and for the same teacher across contexts [9]. Sociocultural accounts of professional agency converge on the same conclusion from another direction, emphasising that agency is exercised by subjects who interpret their work environments through professional identities and who act, resist or comply within conditions they did not choose [10].

Edwards adds the relational dimension that algorithmic mediation makes indispensable. Relational agency is the capacity to recognise what matters to other actors, to build common knowledge of those motives, and to align one's own expertise with that of others in pursuit of shared objects of activity [11]. In algorithmically mediated settings, the relevant others include not only colleagues, students and families but also the designers, data and defaults of technical systems, whose implicit priorities a teacher must learn to read before joint work of any kind is possible. The ecological conclusion is decisive for what follows. If agency is achieved within an environment, then an algorithmic system is not external to teacher agency but part of the environment in which agency is achieved or lost. It enters the material conditions of work directly, the structural conditions through the authority it lends to some decisions over others, and the cultural conditions through the discourse of data driven improvement that accompanies it. Any framework for pedagogical agency in artificial intelligence supported education must therefore specify features of this environment, not merely attributes of teachers.

3. Three Sites Where Agency Is Gained and Lost

The framework is grounded in the sites where teachers actually meet algorithmic outputs. Three sites dominate the literature, and each displays a characteristic pattern of gain and

loss. Artificial intelligence in education has been organised around three paradigms: systems that direct learning while the learner receives, systems that support learning while the learner collaborates, and systems that empower learning while the learner leads [12]. In all three paradigms the teacher's position is underdescribed, yet it is the teacher who is held responsible for what the paradigm does to a class.

The first site is the recommendation system, which selects and sequences content, resources and tasks. The gains are real. Recommendation can widen the repertoire available to a busy teacher, surface materials the teacher would not have found, and adapt practice activities at a granularity no individual could sustain. Commentators on artificial intelligence in higher education have long noted this promise of personalisation and of relief from routine selection work [13]. The losses are quieter. A recommender narrows as it assists: what it does not surface tends not to be taught, its sequencing becomes the de facto curriculum, and its optimisation target, usually engagement or predicted performance, substitutes for the teacher's educational purposes. Because recommendations arrive as defaults, declining them costs time and confidence, and the teacher's curricular agency contracts to the margin of an externally defined path.

The second site is automated grading and feedback, now extended by generative systems that draft comments and whole responses. The gains are speed, consistency on routine items and the possibility of feedback at a frequency that manual marking cannot match. The losses concern the construct being assessed and the teacher's knowledge of students. Automated scoring rewards what models can detect, which pressures assessment toward surface features, and a teacher who no longer reads student work closely loses the diagnostic intimacy on which planning depends. The risks are compounded by bias: algorithmic systems in education can encode and amplify the inequities of their training data, and their errors are not randomly distributed across student groups [14].

The third site is the learning analytics dashboard, which aggregates behaviour traces into indicators, alerts and predictions. The gains are early warning, a shared evidential basis for intervention and visibility of patterns invisible at classroom scale. The losses are framing and surveillance. Datafication recasts learners as profiles and teachers as managers of indicators, a transformation that extends beyond the school into policy and commercial platforms [15]. The ethical literature on learning analytics has insisted from the outset that the data collected about students are not neutral, that consent, obligation and the purposes of surveillance require explicit governance, and that learners and teachers alike can be harmed by well intentioned systems [16]. Behind all three sites stand two general conditions. Algorithmic systems raise epistemic concerns, because their evidence is often inconclusive, inscrutable or misguided, and normative concerns, because their effects distribute advantage and burden unequally [17]. Opacity is not a single problem but several: intentional secrecy, technical illiteracy among users, and the intrinsic complexity of machine learning models each block understanding in different ways [18]. Ananny and Crawford show that transparency alone cannot deliver accountability, because seeing a system is not the same as being able to govern it, and because the ideal of the knowing human overseer obscures the institutional arrangements that accountability actually requires [19]. Kroll and colleagues similarly argue that accountability for algorithmic decisions must be built into institutional and technical design, through procedures that make decisions reviewable, rather than being delegated to the vigilance of individual operators [20]. These conclusions set the design problem for

the framework: agency must be specified as a set of conditions that institutions create or fail to create for teachers.

4. The Algorithmic Pedagogical Agency Framework

The Algorithmic Pedagogical Agency Framework defines pedagogical agency in artificial intelligence supported education as the achieved capacity of a teacher, within a specific sociotechnical arrangement, to form educational purposes, to interpret and evaluate algorithmic outputs in the light of those purposes and of local knowledge, and to act on that evaluation, including by departing from, contesting or reconfiguring the output, while remaining answerable for the decision taken. The definition is deliberately ecological. It locates agency neither in the teacher alone nor in the system alone but in the relation between them as that relation is organised by institutions. The framework specifies five dimensions, each of which names a condition that can be strengthened or withdrawn, a characteristic form of agency it supports, and a characteristic failure that follows when it is absent.

The first dimension is algorithmic literacy. This is the teacher's working understanding of what algorithmic systems are and do: that outputs are probabilistic inferences from data rather than observations of students, that models inherit the purposes and limitations of their training data and design, that accuracy varies across groups and contexts, and that confidence displays and scores are conventions rather than measurements of truth. Algorithmic literacy is not computer science. It is the professional knowledge that allows a teacher to treat an output as a claim requiring appraisal instead of a fact requiring compliance. Without it, interpretive work cannot begin, because the teacher has no basis on which an output could be doubted. Its characteristic failure is credulity: the dashboard is believed because it is a dashboard.

The second dimension is interpretive discretion. This is the recognised professional space in which a teacher reads an algorithmic output against knowledge that the system does not hold: knowledge of the student's circumstances, of the class as a community, of the subject matter and of the educational purposes at stake. Interpretive discretion converts data into judgement. It is exercised when a teacher decides that a risk alert reflects a family crisis rather than disengagement, or that a recommended resource, however well matched to a performance profile, is wrong for these students now. Such discretion is the practical evaluative heart of agency in the ecological model, and it depends on institutional recognition as much as on individual confidence. Where performance management treats deviation from system indications as error, discretion exists on paper only. Its characteristic failure is displacement, in which the output's reading of the student silently replaces the teacher's.

The third dimension is override and contestation rights. Interpretation without consequence is commentary. This dimension names the formal and practical entitlements to act against an output: to change a grade, to ignore a recommendation, to remove a label, to require human review of an automated decision, and to contest the system itself by reporting errors, challenging procurement and feeding professional judgement back into design. Rights of this kind vary in their cost of use. An override that requires a written justification to a manager, exposes the teacher to blame if the alternative judgement proves wrong, or simply takes longer than a crowded timetable allows, will predictably fall into disuse. The dimension therefore includes the procedural economics of dissent: agency depends not on the existence of an override button but on whether pressing it is

cheap, safe and consequential. Its characteristic failure is the hollow loop, in which a human nominally oversees decisions that in practice ratify themselves.

The fourth dimension is data stewardship. Teachers stand inside a data lifecycle they rarely govern: data about students are collected through platforms, combined across contexts, retained for periods teachers do not set and reused for purposes, including the evaluation of teachers themselves, that were not part of the original bargain. Data stewardship names the teacher's role and rights within that lifecycle: knowing what is collected, being able to correct errors at source, participating in decisions about retention and reuse, and being protected against the repurposing of support data as surveillance. Stewardship matters for agency because upstream data decisions prefigure downstream outputs. A teacher excluded from data governance meets the consequences of that exclusion later, in the classroom, in the form of outputs built on data she had no part in shaping. Its characteristic failure is exposure, in which teachers and students supply the raw material of systems whose judgements then return to govern them.

The fifth dimension is collegial and institutional conditions. Agency is practised in departments, schools and systems, and its conditions are collective: leadership that defends professional judgement, timetabled space for teachers to interrogate systems together, union and professional association involvement in technology decisions, procurement processes that include practitioners, and accountability regimes that reward warranted deviation rather than mere compliance. Collegial conditions matter because individual resistance to a platform default is costly and often futile, whereas a department that collectively examines a dashboard, agrees thresholds for action and shares the risk of override makes agency a normal professional act. This dimension is the framework's acknowledgement that the first four dimensions are not personal virtues but organisational provisions. Its characteristic failure is isolation, in which each teacher faces the system alone and the path of compliance is the path of prudence.

5. Propositions Linking the Dimensions to Agency

The dimensions do not contribute to agency independently, and the framework's explanatory content lies in the relations among them. Six propositions state those relations.

Proposition 1. Algorithmic literacy is necessary but not sufficient for pedagogical agency. Literacy enlarges agency only when it is joined to interpretive discretion and usable override rights. Where literacy grows without rights, the result is informed compliance: teachers who understand a system's limitations in detail and follow its outputs anyway, a condition more corrosive of professional identity than uninformed compliance because the teacher's own knowledge testifies against her actions.

Proposition 2. Interpretive discretion is the mediating dimension of the framework. Algorithmic outputs become pedagogical resources only when they pass through a teacher's contextual interpretation; every other dimension affects agency principally by widening or narrowing the space in which that interpretation can occur and can count. Interventions that bypass interpretation, however efficient, reduce agency by definition, because they relocate the practical evaluative moment of teaching from the teacher to the system.

Proposition 3. Override rights decay toward nominal oversight as their procedural cost rises. Time penalties, justification burdens and asymmetric blame for deviating function as a tax on agency, and rational teachers respond by reserving override for extreme cases.

Systems that audit the frequency of teacher overrides as a performance indicator accelerate the decay, because they convert a right into a risk. Real oversight therefore requires protected, low cost and routinely exercised override, including periodic review of cases in which the teacher and the system disagreed.

Proposition 4. Agency is lost upstream before it is exercised downstream. Decisions about data collection, model procurement and default configuration, taken before any classroom encounter, determine the range within which classroom discretion operates. Teacher participation confined to the point of use, through training on a finished system, arrives after the consequential decisions have closed. Data stewardship and institutional conditions are therefore not supplements to agency but prior conditions of it.

Proposition 5. The dimensions are moderated jointly by collegial and institutional conditions. The same literacy, discretion and rights produce substantial agency in a school whose leadership defends judgement and whose staff interrogate systems collectively, and negligible agency in a school organised around compliance metrics. Agency is consequently an ecological property that cannot be raised by training individuals alone, and evaluations of educational artificial intelligence that measure only teacher attitudes or skills will systematically misdiagnose its absence.

Proposition 6. Where accountability exceeds authority, agency collapses into defensive compliance. If teachers are answerable for outcomes produced by systems they cannot inspect, adjust or refuse, the rational strategy is to follow the output and keep the record clean, since the output provides an alibi that independent judgement does not. Accountability without authority thus manufactures the automation bias it later deplores, and any governance arrangement that assigns responsibility for algorithmic decisions to teachers must assign commensurate authority over those decisions, or accept that it has made the system, and not the teacher, the effective decision maker.

6. Risks: Automation Bias, Deskilling, and Accountability Without Authority

The human factors literature explains why the failure modes named in the framework are not hypothetical. Bainbridge's classic analysis of the ironies of automation showed that automating a process does not remove the human operator but changes her task into monitoring a system she no longer practises, so that when the automation reaches its limits the operator is asked to intervene with skills that the automation itself has allowed to decay [21]. Parasuraman and Riley mapped the resulting pathologies of human automation interaction as misuse, disuse and abuse: overreliance on automation, its rejection after salient errors, and its deployment by designers and managers in ways that disregard the human operator's situation [22]. Automation bias, the tendency to accept automated advice and to stop searching for disconfirming evidence, has been demonstrated experimentally in decision making tasks, including among experienced operators who had been explicitly instructed that the aid could err [23]. Lee and See's account of trust in automation adds the calibration problem: appropriate reliance requires trust matched to actual capability, while both overtrust and undertrust destroy the benefits of the partnership, and trust is shaped by how systems present themselves as much as by how they perform [24]. Clinical research reaches similar conclusions, finding automation bias across professional groups and settings and identifying workload, time pressure and the framing of the automated advice as recurring mediators [25]. Teachers meet algorithmic outputs under precisely the conditions this literature identifies as dangerous:

high workload, severe time scarcity, outputs presented with numerical confidence, and personal accountability for the final decision.

Deskilling is the slow form of the same process. A teacher who no longer marks extended writing, because a system scores it, gradually loses the evaluative judgement that marking sustained; a teacher whose curriculum arrives pre sequenced loses the design capability that sequencing exercised. Generative artificial intelligence sharpens both risks and opportunities. Large language models can support lesson preparation, feedback drafting and resource creation, and their educational potential is genuine, but their fluent and confident outputs are unusually effective at recruiting automation bias, and their use in assessment raises integrity questions that institutions have often answered with detection and surveillance rather than with pedagogical redesign [26]. In early years and school settings, the ethical challenges of artificial intelligence, including bias, privacy, transparency and the quiet reshaping of teacher and learner roles, have been documented as practical governance problems rather than distant speculations [27]. The common thread is that risk does not reside in teacher incompetence. It resides in arrangements that demand vigilance while removing the conditions of vigilance: practice, time, information and authority.

These risks converge in the condition this paper names accountability without authority. The teacher remains the legally and professionally answerable figure for grades awarded, interventions triggered and opportunities allocated, while the effective grounds of those decisions sit inside proprietary models, vendor contracts and institutional defaults beyond her control. The arrangement is attractive to institutions because it combines the efficiency claims of automation with the reassurance of a responsible human, and it is corrosive for exactly the reason Ananny and Crawford identify in the oversight ideal generally: the human in the loop becomes a liability sink, present to absorb blame rather than empowered to govern [19]. The framework treats this not as an implementation blemish but as a predictable outcome whenever Proposition 6 is ignored, and it implies that the honest alternatives are only two: enlarge teacher authority to match teacher accountability, or openly assign accountability to the institutions and vendors that hold the authority.

7. Implications for Teacher Education and Procurement Policy

The framework yields distinct implications for those who prepare teachers and those who purchase systems. For teacher education, the central implication is that algorithmic literacy belongs in the professional core rather than in optional technology modules, and that its content should follow the framework rather than vendor documentation. Beginning teachers need a working model of inference from data, of training data and bias, of accuracy variation across groups, and of the difference between a score and a judgement, taught through the actual classes of system they will meet: recommenders, scoring engines, dashboards and generative tools. They also need supervised practice in interpretation and override, including case work in which the system is wrong and the professional task is to notice, to reason and to act, since a capability exercised only in lectures will not survive first contact with a timetable. Ethical and privacy principles for learning analytics, including consent, data minimisation and the rights of the data subject, provide a ready curricular foundation for the stewardship dimension [28]. Teacher education must also prepare teachers for the relational reality of artificial intelligence in schools, in which systems arrive through policy and commerce and must be read politically as well as technically [29]. The so called learnification and datafication of

education proceed partly by redefining what teachers take their work to be, and professional formation is one of the few sites where that redefinition can be consciously contested [30]. Finally, integrity in the generative era is better served by redesigning assessment toward authentic tasks and process evidence than by escalating detection, which places teachers in the role of police and students in the role of suspects [31].

For procurement policy, the framework converts directly into purchasing criteria. Regulation is beginning to supply a floor. The European Union's Artificial Intelligence Act treats most educational uses that determine access, evaluate learning outcomes or monitor students as high risk, and its human oversight provisions require that overseers be enabled to understand a system's capacities and limitations, to remain aware of the tendency to over rely on its outputs, to interpret those outputs correctly, and to disregard, override or reverse them and to intervene in or stop the system's operation [32]. Its deployer provisions go further in the direction of this framework, requiring that the persons entrusted with oversight have the necessary competence, training, authority and support. Procurement should treat these requirements as specifications rather than aspirations. Contracts should require demonstrable override and correction functions, intelligible explanations pitched to professional users, audit rights over model changes, data governance terms that prohibit the repurposing of classroom data for unrelated evaluation, and teacher representation in selection and review. International guidance supports the same direction: UNESCO's guidance on generative artificial intelligence insists on human agency, teacher capacity building and age appropriate, validated deployment as preconditions of adoption [33], and the United States Department of Education's report on artificial intelligence and the future of teaching and learning places humans in the loop at the head of its recommendations, insisting that educators be involved in design and governance rather than consulted after the fact [34]. Fairness deserves specific attention in specifications, because institutions often cannot lawfully or practically collect the sensitive data that standard fairness testing requires; procurement should therefore ask vendors how discrimination is detected and mitigated under that constraint, including through trusted third parties and pedagogically interpretable models [35]. More generally, the governance literature on algorithmic ethics warns that principles without institutional mechanisms change nothing, and that actionable governance must assign responsibility, create review procedures and fund them [36]. A procurement process governed by the framework would evaluate a system not by asking whether it keeps a teacher in the loop, but by asking what the loop entitles the teacher to know, to decide, to refuse and to change.

CONCLUSION

This paper has argued that artificial intelligence changes teaching not primarily by giving teachers new tools but by giving them a new object of work: the algorithmic output, which arrives carrying claims about students that the teacher must interpret, may contest and will be held answerable for. Reconstructing teacher agency as an ecological achievement shows that such agency is made or unmade by the arrangements surrounding the teacher, and the Algorithmic Pedagogical Agency Framework names the five dimensions on which it depends: algorithmic literacy, interpretive discretion, override and contestation rights, data stewardship, and collegial and institutional conditions. Its propositions warn that literacy without rights produces informed compliance, that override decays under procedural cost, that agency is lost upstream in data and procurement decisions before it is exercised in classrooms, and that accountability without authority collapses into

defensive compliance and automation bias. The measure of an educational artificial intelligence system is therefore not only what it predicts or produces, but what it does to the professional who must answer for its outputs, and systems worthy of education are those procured, governed and taught in ways that leave teachers more capable of judgement, not less.

REFERENCES

[1] W. Holmes, M. Bialik, and C. Fadel, *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Boston, MA, USA: Center for Curriculum Redesign, 2019.

[2] R. Luckin, W. Holmes, M. Griffiths, and L. B. Forcier, *Intelligence Unleashed: An Argument for AI in Education*. London, UK: Pearson, 2016.

[3] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education: Where are the educators?" *International Journal of Educational Technology in Higher Education*, vol. 16, art. 39, 2019, doi: 10.1186/s41239-019-0171-0.

[4] N. Selwyn, *Should Robots Replace Teachers? AI and the Future of Education*. Cambridge, UK: Polity Press, 2019.

[5] B. Williamson and R. Eynon, "Historical threads, missing links, and future directions in AI in education," *Learning, Media and Technology*, vol. 45, no. 3, pp. 223-235, 2020, doi: 10.1080/17439884.2020.1798995.

[6] W. Holmes and I. Tuomi, "State of the art and practice in AI in education," *European Journal of Education*, vol. 57, no. 4, pp. 542-570, 2022, doi: 10.1111/ejed.12533.

[7] M. Emirbayer and A. Mische, "What is agency?" *American Journal of Sociology*, vol. 103, no. 4, pp. 962-1023, 1998, doi: 10.1086/231294.

[8] M. Priestley, G. Biesta, and S. Robinson, *Teacher Agency: An Ecological Approach*. London, UK: Bloomsbury Academic, 2015.

[9] G. Biesta, M. Priestley, and S. Robinson, "The role of beliefs in teacher agency," *Teachers and Teaching*, vol. 21, no. 6, pp. 624-640, 2015, doi: 10.1080/13540602.2015.1044325.

[10] A. Eteläpelto, K. Vähäsantanen, P. Hökkä, and S. Paloniemi, "What is agency? Conceptualizing professional agency at work," *Educational Research Review*, vol. 10, pp. 45-65, 2013.

[11] A. Edwards, *Being an Expert Professional Practitioner: The Relational Turn in Expertise*. Dordrecht, The Netherlands: Springer, 2010, doi: 10.1007/978-90-481-3969-9.

[12] F. Ouyang and P. Jiao, "Artificial intelligence in education: The three paradigms," *Computers and Education: Artificial Intelligence*, vol. 2, art. 100020, 2021, doi: 10.1016/j.caeai.2021.100020.

[13] S. A. D. Popenici and S. Kerr, "Exploring the impact of artificial intelligence on teaching and learning in higher education," *Research and Practice in Technology Enhanced Learning*, vol. 12, art. 22, 2017, doi: 10.1186/s41039-017-0062-8.

[14] R. S. Baker and A. Hawn, "Algorithmic bias in education," *International Journal of Artificial Intelligence in Education*, vol. 32, no. 4, pp. 1052-1092, 2022, doi: 10.1007/s40593-021-00285-9.

[15] B. Williamson, *Big Data in Education: The Digital Future of Learning, Policy and Practice*. London, UK: SAGE Publications, 2017.

[16] S. Slade and P. Prinsloo, "Learning analytics: Ethical issues and dilemmas," *American Behavioral Scientist*, vol. 57, no. 10, pp. 1509-1528, 2013, doi: 10.1177/0002764213479366.

[17] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data & Society*, vol. 3, no. 2, 2016, doi: 10.1177/2053951716679679.

[18] J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," *Big Data & Society*, vol. 3, no. 1, 2016, doi: 10.1177/2053951715622512.

[19] M. Ananny and K. Crawford, "Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability," *New Media & Society*, vol. 20, no. 3, pp. 973-989, 2018, doi: 10.1177/1461444816676645.

[20] J. A. Kroll, J. Huey, S. Barocas, E. W. Felten, J. R. Reidenberg, D. G. Robinson, and H. Yu, "Accountable algorithms," *University of Pennsylvania Law Review*, vol. 165, no. 3, pp. 633-705, 2017.

[21] L. Bainbridge, "Ironies of automation," *Automatica*, vol. 19, no. 6, pp. 775-779, 1983.

[22] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human Factors*, vol. 39, no. 2, pp. 230-253, 1997, doi: 10.1518/001872097778543886.

[23] L. J. Skitka, K. L. Mosier, and M. Burdick, "Does automation bias decision-making?" *International Journal of Human-Computer Studies*, vol. 51, no. 5, pp. 991-1006, 1999.

[24] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," *Human Factors*, vol. 46, no. 1, pp. 50-80, 2004, doi: 10.1518/hfes.46.1.50.30392.

[25] K. Goddard, A. Roudsari, and J. C. Wyatt, "Automation bias: A systematic review of frequency, effect mediators, and mitigators," *Journal of the American Medical Informatics Association*, vol. 19, no. 1, pp. 121-127, 2012, doi: 10.1136/amiajnl-2011-000089.

[26] E. Kasneci *et al.*, "ChatGPT for good? On opportunities and challenges of large language models for education," *Learning and Individual Differences*, vol. 103, art. 102274, 2023, doi: 10.1016/j.lindif.2023.102274.

[27] S. Akgün and C. Greenhow, "Artificial intelligence in education: Addressing ethical challenges in K-12 settings," *AI and Ethics*, vol. 2, pp. 431-440, 2022, doi: 10.1007/s43681-021-00096-7.

[28] A. Pardo and G. Siemens, "Ethical and privacy principles for learning analytics," *British Journal of Educational Technology*, vol. 45, no. 3, pp. 438-450, 2014, doi: 10.1111/bjet.12152.

[29] C. Perrotta and N. Selwyn, "Deep learning goes to school: Toward a relational understanding of AI in education," *Learning, Media and Technology*, vol. 45, no. 3, pp. 251-269, 2020, doi: 10.1080/17439884.2020.1686017.

[30] J. Knox, B. Williamson, and S. Bayne, "Machine behaviourism: Future visions of 'learnification' and 'datafication' across humans and digital technologies," *Learning, Media and Technology*, vol. 45, no. 1, pp. 31-45, 2020, doi: 10.1080/17439884.2019.1623251.

[31] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," *Innovations in Education and Teaching International*, vol. 61, no. 2, pp. 228-239, 2024, doi: 10.1080/14703297.2023.2190148.

[32] European Union, "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)," *Official Journal of the European Union*, Jul. 12, 2024.

[33] F. Miao and W. Holmes, *Guidance for Generative AI in Education and Research*. Paris, France: UNESCO, 2023, doi: 10.54675/EWZM9535.

[34] U.S. Department of Education, Office of Educational Technology, *Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations*. Washington, DC, USA: U.S. Department of Education, 2023.

[35] M. Veale and R. Binns, "Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data," *Big Data & Society*, vol. 4, no. 2, 2017, doi: 10.1177/2053951717743530.

[36] A. Tsamados, N. Aggarwal, J. Cowls, J. Morley, H. Roberts, M. Taddeo, and L. Floridi, "The ethics of algorithms: Key problems and solutions," *AI & Society*, vol. 37, no. 1, pp. 215-230, 2022, doi: 10.1007/s00146-021-01154-8.