

REVIEW

Open Access



Artificial intelligence in higher education assessment as an ecosystem of pedagogy governance and legitimacy

Joshua King Obeng-Nyarko^{1*}

*Correspondence:
Joshua King Obeng-Nyarko
jkoben@essex.ac.uk
¹Essex Business School, University
of Essex, Colchester, Essex, UK

Abstract

Artificial intelligence (AI) is rapidly reshaping higher education assessment through applications such as automated feedback, adaptive testing, learning analytics, automated scoring, and generative support. Yet scholarship on AI-enabled assessment remains fragmented across technical, pedagogical, ethical, governance, and professional perspectives, limiting the field's capacity to explain how these dimensions interact in practice and how misalignment among them affects educational legitimacy. This paper develops the AI-Assessment Ecosystem Model, a conceptual ecosystem framework for analysing AI in higher education assessment as a sociotechnical and institutional phenomenon rather than a discrete technical intervention. Drawing on a conceptual synthesis of literature published primarily between 2019 and 2024, and informed by foundational work on assessment validity, feedback, institutional legitimacy, organisational change, inclusive design, and computer-supported collaborative learning, the paper identifies five interdependent domains that shape AI-enabled assessment: technological innovation, pedagogical alignment and learner needs, institutional strategy and governance, faculty and professional readiness, and ethical and inclusive implementation. The paper theorises legitimacy as a relational and contested process through which AI-supported assessment must be justified to students, faculty, institutions, regulators, and other stakeholders as educationally valid, professionally defensible, ethically acceptable, inclusive, and organisationally accountable. The framework argues that AI in assessment cannot be adequately evaluated by technical capability, implementation success, or ethical compliance alone; it must be understood through the relations, tensions, and potential misalignments among educational purpose, organisational conditions, professional practice, student agency, and technological design. The model is offered as a conceptual ecosystem framework for guiding future research, critique, and institutional decision-making.

Keywords Artificial intelligence, Higher education assessment, Educational technology, Sociotechnical systems, Ethical AI, Institutional governance, Legitimacy, Faculty development, Inclusive assessment



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

1 Introduction

Artificial intelligence (AI) is rapidly reshaping higher education, including how assessment is designed, delivered, interpreted, and governed. Across universities, AI-enabled systems are increasingly used to support feedback generation, adaptive testing, learning analytics, automated scoring, and a growing range of assessment-related decisions. These developments have intensified debate not only about efficiency and personalisation, but also about validity, fairness, academic integrity, transparency, and institutional accountability. As AI becomes more embedded in assessment processes, the central question is no longer whether higher education will engage with these technologies, but how such engagement can be organised in ways that are educationally sound, ethically defensible, inclusive, and institutionally coherent.

Recent scholarship has increasingly emphasised that AI in education must be understood not only in terms of technical capability but also in terms of its implications for pedagogy, governance, and the wider data infrastructures through which educational practice is organised [18, 57]. This question is particularly consequential in assessment. Assessment is not merely a technical mechanism for measuring performance; it is a high-stakes educational practice through which institutions make judgments about learning, progression, and achievement [30]. For that reason, AI in assessment raises issues beyond the operational. While AI may enable greater speed, scale, and responsiveness, it may also intensify standardisation, obscure how judgments are made, and displace human interpretation in ways that are difficult to justify on educational grounds. Ethical concerns such as opacity, data extraction, surveillance, bias, and contestability cannot, therefore, be treated as secondary constraints to be addressed after implementation. They are bound up with the design, interpretation, and governance of assessment systems themselves [46, 57].

Likewise, pedagogical value cannot be reduced to efficiency alone. Whether AI enhances assessment depends on how its use aligns with learning goals, disciplinary expectations, faculty judgment, and students' diverse experiences [20, 38]. Recent critical scholarship has also cautioned against framing AI in education primarily in terms of technical promise, emphasising instead the social harms, institutional distortions, and power asymmetries that may accompany AI adoption in educational settings [45].

Although scholarship on AI in education has expanded rapidly, much of it remains conceptually fragmented. Some studies focus primarily on technical capability, such as automated feedback, adaptive systems, explainable AI, or predictive analytics. Others foreground ethical issues such as bias, transparency, surveillance, and rights. A further body of work addresses classroom use, institutional policy, professional learning, or wider processes of digital transformation. These strands each provide valuable insight, but they are often treated in relative isolation from one another. In the context of higher education assessment, that is a significant limitation, because assessment is simultaneously pedagogical, social, ethical, institutional, and professional. When these dimensions are examined separately, AI adoption risks being shaped more by technical feasibility, institutional positioning, or policy pressure than by educational purpose.

A number of frameworks have sought to conceptualise AI's role in education, including models of human–AI collaboration and responsible AI implementation. These have made important contributions by moving beyond narrow techno-solutionist accounts [18, 28]. However, fewer frameworks address higher education assessment through a

single assessment-specific lens that connects pedagogical alignment, institutional governance, faculty readiness, student agency, and ethical inclusion. Existing work has not adequately theorised assessment as a sociotechnical and institutional process in which design, interpretation, implementation, contestation, and accountability are interdependent rather than sequential concerns.

Recent work in the *British Journal of Educational Technology* demonstrates that these issues are already being examined across several connected but still largely separate lines of inquiry. This includes research on assessment validity and learning analytics, teachers' digital assessment practices, the implications of generative AI for authentic assessment and academic integrity, and the development of AI literacy frameworks for educational contexts. Taken together, these contributions confirm both the relevance of AI to assessment and the growing importance of governance, pedagogy, and professional capacity. At the same time, they reveal a persistent conceptual limitation: existing work tends to address these issues through adjacent rather than integrated lenses, leaving underdeveloped an account of how technological design, educational purpose, institutional governance, professional practice, student agency, and ethical inclusion interact in shaping the legitimacy of AI-enabled assessment [8, 11, 26, 56].

This paper addresses that gap by developing the AI-Assessment Ecosystem Model, a conceptual ecosystem framework for analysing AI-enabled assessment in higher education as a sociotechnical and institutional arrangement rather than a discrete technical intervention. The paper's contribution lies not simply in identifying five relevant dimensions, but in theorising how their alignment, tension, and potential misalignment shape whether AI-enabled assessment becomes educationally meaningful, professionally defensible, ethically acceptable, inclusive, and institutionally accountable. The model brings together five interrelated domains: technological innovation, pedagogical alignment and learner needs, institutional strategy and governance, faculty and professional readiness, and ethical and inclusive implementation. In doing so, the paper shifts analysis from questions of capability and adoption to a relational account of how legitimacy is negotiated across multiple stakeholders, including students, faculty, programme teams, institutional leaders, professional bodies, technology providers, and regulators. AI in assessment, from this perspective, is not primarily a problem of optimisation but one of justification, coordination, and contestation across multiple interacting conditions of educational practice.

The remainder of the paper proceeds as follows. The next section reviews the literature on AI in higher education assessment and identifies the conceptual fragmentation that motivates the framework. The methodology section then outlines the conceptual synthesis used to develop the model. The paper subsequently presents the AI-Assessment Ecosystem Model and elaborates its five domains before providing an illustrative application, discussing implications and limitations, and identifying directions for future research. This article is therefore presented as a conceptual, theory-building contribution intended to organise a rapidly developing field rather than to report empirical validation of a single intervention.

1.1 Literature review

Research on AI in higher education assessment has expanded rapidly across educational technology, learning analytics, digital assessment, assessment design, and critical data

studies. One strand of this literature focuses on the technical possibilities of AI-enabled assessment, including automated scoring, adaptive testing, predictive analytics, multi-modal analysis, explainable AI, and feedback generation. This work demonstrates that AI is becoming an increasingly significant component of how institutions evaluate, support, and monitor student learning, particularly where scale, speed, and responsiveness are valued [25, 58].

A second strand foregrounds the ethical and political dimensions of AI-enabled educational systems. This literature raises concerns about bias, opacity, surveillance, data extraction, accountability, and the unequal distribution of risks and benefits when algorithmic systems shape decisions affecting learners. From this perspective, AI in education cannot be understood simply as a neutral toolset; it is embedded in wider relations of policy, infrastructure, institutional power, and commercial interest [4, 13, 46, 57].

A third strand examines AI in relation to pedagogy, learner support, and assessment design. This work considers how automated feedback, adaptive systems, intelligent tutoring, and analytics may support formative assessment, self-regulated learning, and more responsive teaching. At its strongest, this literature shows that questions of effectiveness are inseparable from questions of educational purpose: AI may improve speed or responsiveness, but its pedagogical value depends on whether it supports meaningful learning, valid judgment, and student agency [20, 38]. Assessment design scholarship is especially important here because it cautions against treating assessment solely as a mechanism for producing scores or feedback. Sustainable assessment emphasises students' longer-term capacity to make evaluative judgments beyond the immediate task, while feedback research stresses that feedback becomes educationally valuable only when learners can interpret, act on, and incorporate it into future work [5, 6, 7, 41]. This is particularly important for AI-enabled feedback, where apparent immediacy and personalisation may obscure whether students are developing the capacity to judge quality, understand standards, and participate actively in assessment processes. From this perspective, the legitimacy of AI-supported assessment depends not only on whether outputs are accurate or timely, but on whether assessment designs support student agency, evaluative judgment, and meaningful learning over time [15, 35].

A fourth strand addresses institutional strategy, governance, quality assurance, and organisational change. Here, scholars emphasise policy, accountability, leadership, regulatory expectations, and the conditions under which AI is authorised, regulated, and normalised within universities. This perspective is particularly important because AI in assessment is not implemented in a vacuum; it is shaped by procurement decisions, data governance, institutional priorities, quality assurance processes, and uneven distributions of expertise and authority [40, 54, 57]). The organisational change literature sharpens this point. Higher education institutions do not adopt technologies as neutral instruments; adoption is shaped by institutional culture, external pressures, resource dependencies, leadership strategies, and processes of organisational sensemaking [23, 24, 29]. Institutional theory further suggests that universities may pursue AI adoption partly to secure legitimacy by conforming to sectoral expectations, professional norms, or regulatory pressures, rather than because a given technology has been pedagogically justified in a particular assessment context [10, 42]. This matters because alignment across the model's domains may sometimes produce coherence, but it may also reflect isomorphic pressure, risk management, or market positioning. The framework,

therefore, treats alignment as an object of critical analysis rather than an automatic good.

Inclusive design and disability studies provide an additional foundation for the model's ethical and inclusive domain. Universal Design for Learning has been influential in arguing that learning environments should provide flexible means of engagement, representation, action, and expression, rather than retrofitting accommodations after barriers have already been created [31]. Disability studies extend this argument by showing how institutional systems often construct disability through inaccessible norms, assumptions, and infrastructures [12, 43]. Inclusive assessment scholarship similarly argues that inclusion should not be reduced to accommodations for individual students but should be understood in relation to the design of assessment systems and the agency of marginalised learners within academic communities [36, 37]. For AI-enabled assessment, this means that inclusion cannot be limited to technical accessibility or compliance. It requires attention to how AI systems classify learners, structure participation, accommodate varied forms of expression, and distribute both support and risk across different student groups. Student voice is therefore central: students are not merely data subjects or recipients of AI-mediated judgment, but participants whose experiences, contestations, and forms of agency shape whether assessment practices can be regarded as legitimate.

The computer-supported collaborative learning literature also offers relevant insights. CSCL research has long examined how digital systems mediate interaction, collaboration, knowledge-building, and group meaning-making rather than simply delivering content or automating instruction [9, 48, 51]. This literature is useful for AI assessment because many assessment practices are social, dialogic, and disciplinary rather than purely individual and computational. AI-supported assessment, therefore, needs to be understood not only as automated scoring or feedback, but also as a form of mediated educational interaction that may reshape peer learning, collaborative authorship, feedback dialogue, and the conditions under which students and teachers negotiate standards.

Within the British Journal of Educational Technology specifically, recent scholarship has highlighted several aspects of this evolving terrain. Divjak et al. [11] emphasise the importance of assessment validity in relation to learning analytics, while Viberg et al. [56] show how teachers' digital assessment practices are shaped by changing pedagogical and professional conditions. More recent work has extended these debates into the generative AI era, including discussions of authentic assessment and academic integrity in higher education, as well as the competencies required for AI literacy in educational settings [8, 26]. Collectively, this body of work demonstrates that BJET is already engaging questions central to AI-enabled assessment. However, these concerns are still often addressed through distinct lenses, leaving space for a framework that brings technological innovation, pedagogical alignment, governance, professional readiness, student agency, legitimacy, and ethical inclusion into closer conceptual relation.

These strands are individually valuable, but they are often only loosely connected. Technical studies tend to emphasise system performance and instructional efficiency; ethical studies foreground bias, privacy, transparency, and rights; pedagogical studies focus on feedback, learner needs, and classroom practice; and governance-oriented work addresses policy and institutional accountability. What remains underdeveloped

is not simply a summary of these concerns but an account of how they interact specifically in higher-education assessment. This matters because assessment is a high-stakes domain in which validity, progression, fairness, trust, contestability, and disciplinary judgment are especially acute [30]. A fragmented literature risks encouraging forms of AI adoption that are technically sophisticated yet pedagogically thin, ethically reactive, or weakly embedded in institutional strategy.

Several conceptual frameworks offer useful starting points for responding to this problem. Models of human–AI collaboration and responsible AI implementation move beyond narrow techno-solutionist accounts by foregrounding values, agency, and accountability [18, 19, 28]. However, they do not fully address higher education assessment through a single assessment-specific lens that connects pedagogical alignment, institutional governance, faculty readiness, student agency, legitimacy, and ethical inclusion. Nor do they adequately theorise what happens when these dimensions fall out of alignment. This gap motivates the development of the AI-Assessment Ecosystem Model, which synthesises these strands into a conceptual framework for explaining how technological innovation, pedagogy, governance, professional readiness, student agency, and ethical inclusion interact to shape the legitimacy of AI-enabled assessment.

1.2 Methodological approach

This paper uses a conceptual synthesis approach to develop a framework for understanding the integration of artificial intelligence in higher education assessment. The purpose of the study is not to produce an exhaustive systematic review of all relevant literature, but to identify, interpret, and synthesise key strands of scholarship bearing on the pedagogical, technological, institutional, professional, and ethical dimensions of AI-enabled assessment. The methodological orientation is therefore theory-building rather than aggregative, drawing on conceptual review traditions that integrate diverse bodies of literature into a coherent analytical model [53]. In this sense, the aim is not merely to organise themes in the literature but to generate a conceptual structure that explains how multiple dimensions of AI-enabled assessment interact, align, and come into tension within higher education settings.

To bound the review while retaining sufficient conceptual breadth, literature was searched across Scopus, Web of Science, ERIC, IEEE Xplore, and Google Scholar. Searches focused primarily on publications from 2019 to 2024, a period that captures the recent expansion of scholarship on AI in higher education assessment while also encompassing ethical, pedagogical, and governance debates that intensified prior to the widespread uptake of generative AI. Earlier foundational works were selectively included where necessary for conceptual interpretation, particularly regarding assessment validity, feedback, institutional legitimacy, organisational change, inclusive design, and computer-supported collaborative learning.

The search strategy combined terms relating to artificial intelligence, higher education, and assessment. Core search strings included combinations of the following terms: (“artificial intelligence” OR AI OR “machine learning” OR “learning analytics” OR “generative AI” OR “automated feedback” OR “automated scoring” OR “adaptive assessment”) AND (“higher education” OR university* OR college*) AND (assessment OR feedback OR grading OR evaluation OR “academic integrity”). Additional searches were conducted using more focused combinations, including “ethical AI in education,”

“AI governance in universities,” “AI assessment validity,” “AI feedback higher education,” “inclusive assessment,” “universal design for learning assessment,” “faculty readiness AI,” and “computer-supported collaborative learning assessment.” Searches were supplemented through backward and forward citation tracking of highly relevant sources.

The search and citation-following process produced a working pool of 107 potentially relevant publications. Because this was a conceptual synthesis rather than a systematic review, screening was used to establish a focused and theoretically useful corpus rather than to claim exhaustive coverage of the field. Overlapping records were identified by comparing available title, author, year, and DOI information, with Google Scholar results checked manually where metadata were inconsistent. The remaining sources were then reviewed for relevance to higher education assessment and to at least one of the paper’s core dimensions: technological innovation, pedagogical alignment, institutional governance, faculty or professional readiness, and ethical or inclusive implementation. Sources judged to make a substantive conceptual contribution were read in full or in substantial part and considered for inclusion in the synthesis. Through this process, 58 sources were retained for close analysis.

For the purposes of this conceptual synthesis, “conceptual relevance” was treated as an interpretive inclusion judgment guided by four criteria. A source was retained where it: (1) addressed AI, digital assessment, learning analytics, automated feedback, or related technologies in a higher education or clearly transferable assessment context; (2) contributed to understanding at least one of the five domains of the emerging model; (3) helped explain relationships, tensions, or misalignments among domains rather than offering only a narrow technical evaluation; or (4) provided foundational theoretical grounding for the framework, such as assessment validity, feedback literacy, institutional legitimacy, organisational change, inclusive design, or CSCL. Sources were generally excluded where they focused primarily on non-educational contexts, K–12 settings without clear transferability to higher education, technical system performance without assessment or institutional relevance, or general AI commentary without a substantive connection to assessment practice.

The analytical process was iterative and interpretive. Rather than using a formal systematic review coding framework, the synthesis used a structured note-taking protocol to record each source’s primary domain, secondary domains, key concepts, relevance to legitimacy, and contribution to understanding alignment or misalignment. Initial analysis focused on recurring themes such as technological affordances, automated feedback, adaptive assessment, validity, learner agency, governance, policy, professional judgment, faculty readiness, data ethics, bias, accessibility, inclusion, accountability, and trust. These themes were compared across sources and refined through repeated analysis in order to identify broader conceptual groupings. Sources that sat at the margins were retained only where they clarified a theoretical issue that the AI assessment literature itself left underdeveloped, such as institutional legitimacy, organisational isomorphism, sustainable assessment, or inclusive design.

Through this process, five interrelated domains were stabilised as the core components of the AI-Assessment Ecosystem Model: technological innovation, pedagogical alignment and learner needs, institutional strategy and governance, faculty and professional readiness, and ethical and inclusive implementation. The outcome of this process is not

a taxonomy of themes but a conceptual ecosystem framework in which the significance of each domain depends on its interaction with the others.

This approach was chosen because the literature on AI in higher education assessment is both rapidly evolving and dispersed across multiple fields. A conceptual synthesis is well-suited to this kind of problem because it allows diverse forms of scholarship to engage in dialogue to produce an integrative framework rather than a narrow summary of findings. In this study, the aim was therefore not to determine the effectiveness of particular AI tools, but to develop a conceptual framework to capture the interdependencies among technical, pedagogical, institutional, professional, and ethical considerations in assessment contexts.

This methodology has several limitations. First, the literature sample was selective rather than exhaustive. Second, the synthesis relied on interpretive judgment to identify and refine the model's domains. Third, the resulting model is conceptual and has not yet been empirically validated in institutional settings. These limitations are acknowledged not as weaknesses to be obscured, but as characteristics of a framework-building study whose purpose is to provide an analytical basis for future empirical and comparative research. To enhance transparency, Table 1 summarises the review design and analytical process; while Supplementary Table S1 lists the 58 retained sources along with their primary and secondary domains and the rationale for inclusion.

Table 1 Summary of review design and analytical process

Review component	Procedure	Decision rule
Review purpose	Conceptual synthesis for framework development	Sources were selected for their contribution to theory-building rather than exhaustive coverage.
Databases searched	Scopus, Web of Science, ERIC, IEEE Xplore, and Google Scholar	Databases were selected to cover education, social science, technology, and interdisciplinary AI scholarship.
Timeframe	2019–2024, with selective earlier foundational works	Earlier works were included when theoretically necessary, for example, assessment validity, legitimacy theory, feedback, organisational change, inclusive design, and CSCL.
Core search logic	AI-related terms AND higher education terms AND assessment-related terms	Search terms were adapted to each database while retaining the same conceptual structure.
Initial working pool	107 potentially relevant publications	Records were compared and manually checked for overlapping titles, DOIs, and author/year combinations.
Screening stages	Overlap checking; title/abstract relevance review; full-text or substantial-text conceptual screening	Sources had to address higher-education assessment or provide a clear, transferable theoretical grounding.
Final corpus	58 retained sources	Retained sources contributed to at least one model domain and to understanding alignment, tension, or legitimacy.
Inclusion criteria	Higher education assessment; AI or digital assessment; governance; professional readiness; ethics; inclusion; relevant foundational theory	Sources were included when they made a substantive conceptual contribution to framework development.
Exclusion criteria	Non-educational settings; K–12 without transferability; narrow technical performance without assessment relevance; general AI commentary	Sources were excluded where they did not contribute to the paper's assessment-specific conceptual argument.
Analytical process	Structured note-taking and iterative conceptual grouping of themes	Themes were compared iteratively until five interdependent domains were stabilised.

This table summarises the review design used to support the development of the framework. The process was designed as a conceptual synthesis rather than a systematic review; therefore, the table provides an audit trail of search, screening, and analytical decisions rather than a claim to exhaustive coverage of the field.

1.3 The AI-assessment ecosystem model

The AI-Assessment Ecosystem Model conceptualises AI-enabled assessment in higher education as a relational sociotechnical and institutional configuration rather than a discrete technological intervention. Its central claim is that the educational legitimacy of AI-enabled assessment depends not on the strength of any single element in isolation, but on the interaction among five interdependent domains: technological innovation; pedagogical alignment and learner needs; institutional strategy and governance; faculty and professional readiness; and ethical and inclusive implementation. The model is therefore neither linear nor additive. It is designed to explain how these domains shape one another, how tensions emerge when they fall out of alignment, and why technically capable systems may still prove pedagogically weak, ethically contested, or institutionally ungovernable. In this sense, the conceptual ecosystem framework advances current debates by treating AI in assessment not simply as a matter of adoption or implementation, but as a question of how educational purpose, organisational capacity, professional judgment, student agency, and technological design align, are contested, or become misaligned in practice.

This ecosystem perspective matters because assessment is a particularly high-stakes site of AI integration. Assessment is the process by which institutions translate educational values into judgments about performance, progression, and achievement [30]. For that reason, AI-enabled assessment cannot be evaluated solely on technical functionality, implementation success, or ethical compliance. A system may be technologically sophisticated yet pedagogically misaligned; it may be ethically ambitious yet institutionally poorly supported; or it may be strategically endorsed while faculty lack the expertise to use it critically and responsibly. The model therefore shifts attention from isolated variables to the conditions of interdependence through which AI-enabled assessment becomes educationally meaningful, ethically defensible, inclusive, and institutionally accountable [46, 47, 57].

Figure 1 presents the AI-Assessment Ecosystem Model. The figure should be read not as a linear implementation sequence, but as a representation of interdependent domains whose relationships shape the legitimacy of AI-enabled assessment.

The model should be read as a framework of mutual shaping rather than as a sequence of stages. The five domains are analytically distinct, but they do not operate independently. Technological innovation affects what forms of assessment become possible, but its value depends on pedagogical alignment and ethical scrutiny. Governance structures authorise and regulate AI use, but their effectiveness depends on faculty capacity and professional judgment. Ethical and inclusive implementation is not an external check applied after design, but a condition that shapes whether assessment practices can be justified in educational terms. The AI-Assessment Ecosystem Model is therefore intended to illuminate not only alignment, but also misalignment: the points at which technologies are adopted without pedagogical clarity, where governance outpaces professional readiness, or where efficiency is prioritised over fairness, transparency, and inclusion [20, 47].

Having introduced the model's overall structure, the next section clarifies its central organising concept: legitimacy. This is necessary because legitimacy is not treated here as simple alignment among the five domains, but as a relational and contested process of justification.

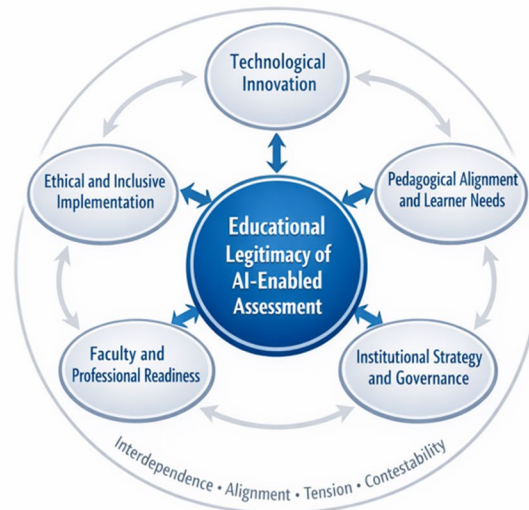


Fig. 1 AI-assessment ecosystem model

1.4 Theorising legitimacy in AI-enabled assessment

Legitimacy is the central organising concept of the AI-Assessment Ecosystem Model, but it should not be understood as a simple property of a technology or as a universal outcome that can be declared in advance. In this paper, legitimacy refers to the extent to which an AI-enabled assessment arrangement can be justified to relevant stakeholders as educationally valid, professionally defensible, ethically acceptable, inclusive, and institutionally accountable. It is therefore a relational and institutional condition rather than a purely technical attribute. A system may be accurate in narrow performance terms while still lacking legitimacy if students cannot contest its judgments, if faculty cannot interpret or challenge its outputs, if its use conflicts with disciplinary standards, or if governance structures cannot account for the risks it introduces.

This account draws on institutional theories of legitimacy. Suchman's distinction between pragmatic, moral, and cognitive legitimacy is useful because it shows that legitimacy may rest on different kinds of justification. Pragmatic legitimacy concerns whether stakeholders perceive an arrangement as serving their interests or solving a recognised problem. Moral legitimacy concerns whether the arrangement is judged to be normatively appropriate, fair, and aligned with educational values. Cognitive legitimacy concerns whether the arrangement becomes understandable, taken for granted, or accepted as part of normal institutional practice [50]. Scott's institutional pillars further clarify that legitimacy is shaped by regulative, normative, and cultural-cognitive conditions: rules and accountability structures, professional norms and values, and shared assumptions about what counts as acceptable practice [42].

Applied to AI-enabled assessment, these distinctions prevent legitimacy from becoming circularly defined as mere alignment among the model's domains. Alignment is not legitimacy itself. Rather, alignment is one condition through which legitimacy may be produced, challenged, or withheld. For example, an AI feedback system may have pragmatic legitimacy for institutional leaders if it enables timely, at-scale feedback, but it may lack moral legitimacy for students if it produces opaque or inequitable judgments. It may be normatively defensible to faculty if it supports formative learning, yet culturally

resisted if it conflicts with disciplinary expectations of human judgment. Conversely, a system may become cognitively normalised across a university while remaining ethically problematic or pedagogically thin.

The model, therefore, treats legitimacy as multi-level and contested. It operates at the micro level of student and teacher experience, the meso level of programme and institutional governance, and the macro level of regulatory, professional, and societal expectations. It is also stakeholder-relative: students, faculty, managers, professional bodies, technology providers, and regulators may assess legitimacy by different criteria. For this reason, misalignment across the five domains should not always be interpreted as a failure of coordination. It may reveal a genuine conflict of values, interests, or epistemic assumptions. The aim of the AI-Assessment Ecosystem Model is therefore not to eliminate all tension, but to make visible how tensions are produced, whose interests are served by proposed alignments, and whether AI-enabled assessment can be justified through transparent, participatory, and revisable institutional processes.

This theoretical clarification also indicates how the model could be studied empirically. Researchers might examine pragmatic legitimacy by investigating stakeholder perceptions of usefulness, burden, and benefit; moral legitimacy by analysing fairness, inclusion, validity, and contestability; cognitive legitimacy by studying how AI assessment practices become normalised in policy, discourse, and everyday professional routines; and institutional legitimacy by examining how regulative, normative, and cultural-cognitive conditions shape adoption. Legitimacy is therefore not treated as a static endpoint but as an ongoing process of justification, negotiation, and contestation.

A key contribution of the model is that it reframes AI-enabled assessment as a problem of institutional legitimacy rather than solely a matter of technical optimisation. Legitimacy here refers to the extent to which an AI-supported assessment arrangement can be justified to relevant stakeholders as educationally meaningful, professionally defensible, ethically acceptable, inclusive, and organisationally accountable. This legitimacy is not created by alignment alone, nor is alignment always desirable if it suppresses dissent, student agency, disciplinary difference, or ethical critique. Instead, legitimacy depends on whether the tensions among technological capability, educational purpose, institutional governance, professional judgment, and inclusion are made visible, debated, and governed in ways that remain accountable to students and other affected stakeholders. The framework is therefore not presented as a checklist or implementation tool. Its purpose is analytical: to provide a conceptual ecosystem framework for examining how interactions among multiple domains shape the promises, limits, and contestations surrounding AI in higher-education assessment.

What distinguishes the AI-Assessment Ecosystem Model from existing approaches is not simply the identification of technological, pedagogical, ethical, governance, and professional concerns, all of which are already visible in the literature. Its distinctive contribution lies in theorising how these domains become consequential through their interdependence and how their misalignment shapes the legitimacy of AI-enabled assessment. The framework, therefore, treats assessment not as a site where separate challenges accumulate, but as a sociotechnical and institutional domain in which educational purpose, institutional coordination, professional judgment, student agency, and technological design are mutually constitutive.

Having clarified the model's overall structure and its understanding of legitimacy, the following sections discuss the five interdependent domains of the AI-Assessment Ecosystem Model. These domains are elaborated separately for analytical clarity, but their significance lies in how they interact to shape the legitimacy of AI-enabled assessment.

1.5 Domain 1: technological innovation

The first domain concerns the technological affordances and limitations of artificial intelligence in higher education assessment. It addresses the ways AI systems can support assessment design, delivery, feedback, and interpretation through functions such as automation, adaptivity, multimodal analysis, generative support, and explainability. Recent developments in AI have expanded the range of assessment-related tasks that can be supported computationally, including automated scoring, intelligent feedback generation, adaptive question sequencing, predictive analytics, and analysis of large, complex datasets related to learner performance [25]. Generative AI has further extended these possibilities by enabling the production of prompts, exemplars, feedback suggestions, simulated dialogue, and draft evaluative comments that may assist both students and educators in assessment-related activities [22].

At the same time, technological innovation in assessment cannot be understood in purely instrumental terms. AI systems operate through data models, training assumptions, and probabilistic outputs that may not align neatly with the interpretive and context-sensitive nature of educational judgment [39, 57]. What appears technically efficient may still be pedagogically weak, insufficiently transparent, or poorly suited to disciplinary differences in how knowledge and performance are assessed. Explainability is therefore not merely a technical feature but a condition of educational intelligibility: students and faculty need to understand how AI-supported judgments are produced, what they can and cannot mean, and when they should be challenged. Within the AI-Assessment Ecosystem Model, technological innovation is therefore treated as an enabling but insufficient domain. It creates possibilities for assessment, but it does not determine their educational legitimacy. That legitimacy depends on whether technological capability is aligned with pedagogical purposes, governance arrangements, professional expertise, student agency, and ethical commitments.

1.6 Domain 2: pedagogical alignment and learner needs

The second domain concerns the relationship between AI-enabled assessment and educational purpose. Its central question is not simply whether AI can be used in assessment, but whether its use aligns with valid learning design, meaningful feedback, learner diversity, disciplinary standards, and the broader pedagogical aims of higher education. Assessment is not only a mechanism for measuring performance; it is also a process through which learning is shaped, interpreted, and supported. For this reason, the introduction of AI into assessment must be evaluated against pedagogical principles such as validity, formative value, student agency, evaluative judgment, and responsiveness to diverse learner needs [5, 30, 38]. This emphasis also resonates with recent BJET work highlighting the centrality of assessment validity in digitally mediated learning environments and the changing conditions of teachers' digital assessment practices [11, 56].

This domain is especially important because AI can create a misleading sense of educational value when speed, scale, or responsiveness are treated as ends in themselves. In

formative contexts, AI may provide timely, iterative, and personalised feedback, potentially helping students identify gaps, monitor progress, and engage in self-regulated learning [20, 35]. In summative contexts, however, the pedagogical stakes differ, especially where transparency, disciplinary judgment, standards, and validity are more acute. The model, therefore, treats pedagogical alignment not as a secondary consideration added after technical design, but as a central condition of legitimacy. A technologically advanced system that does not support meaningful learning, valid judgment, or inclusive participation cannot be regarded as educationally successful simply because it performs efficiently [19, 31].

A stronger emphasis on learner needs also requires attention to student agency and disciplinary differences. Students should not be positioned only as recipients of AI-generated feedback or as data points within analytics systems. Their capacity to understand, question, contest, and make use of assessment information is central to the pedagogical legitimacy of AI-enabled assessment. This is especially important because assessment practices vary significantly across disciplines: the standards used in fine arts, clinical education, laboratory science, business, humanities, and professional programmes are not reducible to a single model of performance or feedback. AI-supported assessment must therefore be evaluated in relation to the forms of knowledge, judgment, evidence, and participation that matter within particular disciplinary communities. Pedagogical alignment is consequently not a generic match between a tool and a task, but a situated judgment about whether AI supports meaningful learning, valid standards, student agency, and appropriate disciplinary forms of assessment.

1.7 Domain 3: institutional strategy and governance

The third domain addresses the institutional conditions through which AI-enabled assessment is made governable. It focuses on strategy, policy, leadership, coordination, quality assurance, and accountability, recognising that the adoption of AI in assessment is never simply a matter of individual educator choice or isolated technological experimentation. Assessment practices in higher education are shaped by institutional priorities, regulatory expectations, resourcing decisions, quality assurance processes, procurement arrangements, and organisational cultures [40, 57]. AI integration must therefore be understood as a governance issue as much as a pedagogical or technical one.

Within the AI-Assessment Ecosystem Model, governance serves as the connective infrastructure linking innovation to accountability. Decisions about procurement, data management, platform use, academic integrity, assessment policy, and acceptable AI use all shape whether AI-enabled assessment is introduced coherently or reactively. Governance is not limited to compliance; it involves creating the organisational conditions under which AI can be authorised, monitored, challenged, and revised in ways that remain accountable to educational goals [46]. This domain also makes visible a frequent source of misalignment: institutions may endorse AI strategically without establishing the policies, responsibilities, or support systems needed to govern its use responsibly in assessment practice.

The organisational change literature strengthens this domain by showing that institutional adoption may be driven by more than educational rationale. Universities may adopt AI in response to external competition, reputational pressures, regulatory

ambiguity, market positioning, or assumptions about sectoral modernisation [10, 23, 29]. Such pressures can generate pragmatic legitimacy for AI adoption while leaving moral or pedagogical legitimacy unresolved. Governance must therefore ask not only whether AI use is permitted, but how decisions are justified, who participates in them, how risks are monitored, and how assessment arrangements can be revised when harms, inequities, or pedagogical weaknesses emerge.

1.8 Domain 4: faculty and professional readiness

The fourth domain concerns the human capacity required to engage AI responsibly in assessment practice. It focuses on the knowledge, judgment, confidence, and support structures that educators, assessors, and academic staff need in order to interpret, use, and challenge AI-enabled systems in pedagogically sound and professionally defensible ways. While institutional strategy provides the governance context for AI adoption, faculty and professional readiness address whether those expected to work with AI in assessment are equipped to do so critically and competently.

This domain is crucial because AI in assessment does not remove the need for professional expertise; rather, it reconfigures the forms of expertise required. Educators need not only subject knowledge and assessment literacy, but also the capacity to interpret automated outputs, evaluate the quality of AI-generated feedback, recognise the risks of overreliance, identify bias or misalignment, and decide when human judgment should take precedence [20, 32]. In this sense, faculty readiness operates as the model's interpretive core. Technologies may be available, and governance structures may be in place, but without professional readiness, AI use in assessment risks becoming superficial, uncritical, or weakly aligned with educational purpose.

Recent AI literacy work strengthens this point by showing that readiness for AI in educational settings depends on identifiable competencies rather than generic digital confidence alone. AI literacy involves more than tool operation; it encompasses understanding data, models, ethics, limitations, human agency, and the capacity to critically evaluate AI outputs [8, 27, 34]. For faculty, this means that professional development must go beyond introductory workshops on generative AI tools. It must support educators in redesigning assessment, explaining AI use to students, interpreting AI-mediated evidence, preserving disciplinary judgment, and creating routes for student dialogue and contestation. Assessment in higher education often involves interpretation, contextual knowledge, disciplinary nuance, and ethical discretion that cannot be fully delegated to computational systems [30].

1.9 Domain 5: ethical and inclusive implementation

The fifth domain concerns the normative conditions under which AI-enabled assessment can be regarded as legitimate, fair, and inclusive. It addresses questions of transparency, bias, accessibility, participation, accountability, contestability, and trust. Whereas the governance domain focuses on the institutional mechanisms through which AI is managed, this domain focuses on the ethical principles and inclusion commitments that should shape how AI is designed, deployed, and evaluated in assessment contexts. Its central concern is not simply whether AI can be used efficiently, but whether its use can be justified in educational and social terms.

This domain is not supplementary. In assessment, where decisions affect progression, opportunity, recognition, and students' perceptions of fairness, ethical and inclusive implementation is a constitutive condition of legitimacy. AI systems may introduce efficiencies, but they may also reproduce existing inequities, obscure reasoning processes, and distribute harms unevenly across different student groups [4, 13]. Transparency, therefore, involves more than technical explainability; it also requires intelligibility, contestability, and meaningful opportunities for those affected to understand and challenge AI-supported judgments [25, 46]. Recent work on generative AI, authentic assessment, and academic integrity further underscores that ethical questions in assessment cannot be separated from design, legitimacy, and institutional responsibility [26]. Within the AI-Assessment Ecosystem Model, this domain cuts across all others by asking whether AI-enabled assessment can be defended not only as functional but as fair, inclusive, and educationally justifiable [47].

Ethical and inclusive implementation also requires participatory attention to students' lived experiences. Students differ in their access to technologies, prior exposure to AI tools, linguistic repertoires, cultural expectations, disability status, neurodiversity, and confidence in contesting institutional decisions. AI-enabled assessment may therefore redistribute advantage and disadvantage even where systems appear technically neutral. Inclusion should consequently be built into assessment design from the outset through accessible formats, alternative modes of expression, transparent criteria, clear routes for appeal, and opportunities for students to provide feedback on AI-mediated assessment practices. This positions students as agents and participants in assessment governance rather than as passive objects of automated judgment [31, 36, 43].

1.10 Synthesis: interdependence, tension, and legitimacy

Taken together, these five domains define the conceptual structure of the AI-Assessment Ecosystem Model. Their significance lies not merely in their presence but in their interdependence. Technological innovation without pedagogical alignment can produce systems that function efficiently yet lack educational purpose. Governance without faculty readiness risks formal oversight without meaningful practice. Ethical aspiration without institutional coordination remains declarative rather than operational. Conversely, pedagogically ambitious uses of AI may fail when technologies are poorly designed or when governance structures are weak. The model, therefore, explains AI-enabled assessment not through isolated factors but through the relations, tensions, and possible misalignments among domains that together shape whether assessment becomes legitimate, trustworthy, inclusive, and educationally meaningful.

The analytical value of the ecosystem perspective becomes particularly visible in conditions of misalignment. For example, a university may implement technically advanced AI systems for automated feedback and assessment at scale, yet fail to ensure alignment with pedagogical design, resulting in feedback that is efficient but educationally shallow or misdirected. Similarly, institutional policies may emphasise ethical AI use and transparency, while faculty lack the expertise or confidence to interpret and challenge AI-generated outputs, limiting meaningful accountability in practice. Conversely, pedagogically innovative uses of AI may struggle to gain traction where governance structures are underdeveloped or institutional strategy is unclear. These examples illustrate that AI-enabled assessment cannot be understood as the outcome of isolated improvements

within individual domains; rather, its effectiveness and legitimacy depend on how technological capability, pedagogical purpose, governance arrangements, professional judgment, student agency, and ethical considerations are brought into relation, and where tensions among them remain unresolved.

1.11 Illustrative application: AI-supported formative feedback

The analytical value of the AI-Assessment Ecosystem Model can be illustrated through a hypothetical case of AI-supported formative feedback in a large first-year undergraduate module. The institution introduces a generative AI tool that provides students with draft feedback on written assignments before final submission. The tool is attractive because it promises timely feedback at scale, reduces pressure on staff, and gives students repeated opportunities to revise their work. Considered only through the domain of technological innovation, the system appears beneficial: it is fast, adaptive, and capable of generating detailed comments.

The model, however, shows that technical functionality is only one part of the assessment arrangement. From a pedagogical alignment perspective, the key question is whether the AI feedback supports students' development of evaluative judgment or merely provides surface-level corrections. If students use the system to polish text without understanding criteria, standards, or disciplinary expectations, then the tool may increase efficiency while weakening assessment literacy. From the perspective of faculty and professional readiness, staff need to understand the system's limits, explain appropriate use, identify misleading feedback, and design assessment activities that keep human judgment central. Without such readiness, educators may either over-trust the tool or reject it without critical engagement.

Institutional governance introduces further questions. The university must specify what constitutes acceptable use, how data are stored, whether student work is used to train models, who is accountable for harmful or misleading feedback, and how the system aligns with the academic integrity policy. Ethical and inclusive implementation requires attention to whether all students can access the tool, whether multilingual or disabled students are affected differently, whether the feedback is culturally or linguistically biased, and whether students can challenge or discuss AI-generated advice.

This example also clarifies the role of legitimacy. The tool may have pragmatic legitimacy for institutional leaders because it addresses feedback workload, and for students because it provides immediate support. Yet it may lack moral legitimacy if it produces inequitable guidance, weakens student agency, or obscures accountability. It may gain cognitive legitimacy over time if its use becomes normalised across modules, but such normalisation should not be mistaken for educational defensibility. The model therefore helps researchers and institutions ask not simply whether the tool works, but for whom it works, under what conditions, and with what consequences for validity, professional judgment, student agency, and institutional accountability.

2 Discussion

The AI-Assessment Ecosystem Model is developed in response to a specific limitation in existing scholarship on artificial intelligence in higher education assessment: while prior work has examined technological capability, pedagogy, ethics, governance, and professional readiness in detail, it has not adequately theorised how these dimensions interact

to produce, or undermine, the legitimacy of assessment practices. Existing frameworks tend to treat these domains as analytically separable or sequentially addressed, leaving underdeveloped an account of how their misalignment shapes assessment outcomes. The contribution of this paper is therefore not simply integrative. It is to reconceptualise AI-enabled assessment as a relational and interdependent system in which technological design, pedagogical purpose, institutional governance, professional judgment, student agency, and ethical inclusion are mutually constitutive. From this perspective, the problem of AI in assessment is not primarily one of optimisation or adoption, but of justification: how these domains are brought into relation, where they come into tension, and how those tensions shape whether assessment becomes educationally meaningful, ethically defensible, inclusive, and institutionally legitimate.

Importantly, the model does not assume that alignment is always possible or desirable. Some forms of misalignment may indicate poor coordination, such as when institutional policy endorses AI use without staff development or data governance. Other forms of misalignment may reveal legitimate conflicts of value, such as tensions between efficiency and human judgment, standardisation and disciplinary diversity, innovation and accessibility, or institutional risk management and student agency. The model's purpose is therefore not to resolve all tensions into consensus, but to provide a conceptual language for analysing how such tensions are negotiated, whose interests are prioritised, and whether institutional arrangements remain open to critique, contestation, and revision.

In this respect, the model aligns with wider critical work in AI and education that treats digital systems not as neutral enhancements to teaching and learning, but as infrastructures that reshape institutional judgment, responsibility, and educational relations [44, 45, 57]. The analytical value of the ecosystem perspective becomes particularly visible in conditions of misalignment. A university may implement technically advanced AI systems for automated feedback and assessment at scale, yet fail to ensure alignment with pedagogical design, resulting in feedback that is efficient but educationally shallow or misdirected. Similarly, institutional policies may emphasise ethical AI use and transparency, while faculty lack the expertise or confidence to interpret and challenge AI-generated outputs, limiting meaningful accountability in practice. Conversely, pedagogically innovative uses of AI may struggle to gain traction where governance structures are underdeveloped or institutional strategy is unclear.

A key contribution of the model is that it reframes assessment as an ecosystem rather than a tool context. This shift matters conceptually because it resists linear accounts of AI adoption in which a new technology is introduced and then subsequently adjusted through policy, ethics, or professional development. Instead, the framework treats technological innovation, pedagogy, governance, faculty and professional readiness, and ethical and inclusive implementation as mutually constitutive domains. In practice, this means that no single dimension can be treated as sufficient on its own. Technically advanced systems may still be pedagogically misaligned; ethically ambitious policies may fail without institutional coordination; and governance structures may remain ineffective if staff lack the confidence or expertise to critically interpret AI outputs [20, 47].

The model also contributes to current debates in digital education by insisting that assessment is a particularly high-stakes site for AI integration. Assessment is the process by which educational values are translated into judgments of quality, progression, and

performance [30]. For that reason, decisions about AI in assessment carry implications not only for efficiency and innovation but also for validity, trust, fairness, inclusion, and institutional legitimacy. By situating assessment within a wider sociotechnical ecosystem, the framework highlights that AI adoption in this domain cannot be adequately evaluated by performance metrics or implementation success alone. It must also be judged in relation to educational purpose, transparency, contestability, student agency, and the distribution of benefits and risks across different learners and institutional contexts [13, 36].

A further contribution of the framework is that it provides a diagnostic language for analysing why AI initiatives succeed, stall, or become contested within institutions. In many settings, AI adoption may be driven primarily by technical possibility, efficiency pressures, or strategic positioning, while pedagogical purpose, staff readiness, and ethical inclusion are addressed only later or unevenly. The AI-Assessment Ecosystem Model clarifies why such sequencing is problematic by showing that legitimacy depends on the coordination and contestation of domains rather than progress within any one of them. In this respect, the model also aligns with critical work that cautions against reducing AI in education to questions of optimisation or innovation alone, instead foregrounding the social, political, and institutional harms that may follow when educational practices are reorganised around machine-readable logics [45]. By bringing into one framework concerns that recent British Journal of Educational Technology scholarship has often examined through separate lenses—including assessment validity and learning analytics, teachers' digital assessment practices, generative AI and authentic assessment, and AI literacy—the AI-Assessment Ecosystem Model is intended to contribute a more integrated conceptual language for ongoing debate in the journal [8, 11, 26, 56].

A further value of the framework lies in its heuristic potential. For researchers, the model provides a conceptual structure for examining how different dimensions of AI-enabled assessment interact, where tensions emerge, and how legitimacy is negotiated across institutional contexts. For higher education institutions, it offers a way to diagnose whether AI initiatives are shaped predominantly by technological possibility, policy pressure, institutional risk management, or educational rationale, and whether sufficient attention is given to governance, staff preparedness, student participation, and inclusion. The framework is therefore not presented as a validated implementation model or as a checklist that guarantees responsible practice. Its contribution is analytical rather than predictive: it is intended to support critique, interpretation, and comparative inquiry rather than prescribe a universal route to adoption.

The framework also has clear limitations. It is derived from conceptual synthesis rather than empirical testing, and its five domains reflect interpretive judgment about how the literature can be organised rather than fixed categories established through validation across multiple contexts [53]. The model is intended to support analysis, not to claim universal completeness. Its usefulness will depend on how well it travels across institutional settings, disciplinary traditions, regulatory environments, and learner populations. Further research is therefore needed to explore how the domains interact in practice, where misalignments are most consequential, and how the framework might be refined through empirical study. Comparative work across higher education contexts, as well as participatory research involving students and educators, would be especially valuable in testing the framework's relevance and limits.

3 Conclusion

This paper has developed the AI-Assessment Ecosystem Model, a conceptual framework for understanding the integration of artificial intelligence into higher education assessment. In contrast to approaches that treat AI primarily as a technical innovation or implementation problem, the framework argues that assessment must be understood as a sociotechnical, pedagogical, institutional, professional, and ethical phenomenon. To capture that complexity, the model brings together five interrelated domains: technological innovation, pedagogical alignment and learner needs, institutional strategy and governance, faculty and professional readiness, and ethical and inclusive implementation.

The paper's central argument is that responsible AI use in assessment cannot be secured solely through technological capability. Nor can it be achieved by addressing ethics, governance, pedagogy, professional development, or inclusion in isolation. AI becomes educationally legitimate only when these domains are considered together and when their tensions are made visible, debated, and governed. The ecosystem perspective developed here, therefore, offers a way to move beyond fragmented accounts of AI in assessment towards a more integrated understanding of how technologies, values, practices, students, and institutional conditions interact and, at times, come into tension in shaping assessment practice.

The paper's contribution is primarily conceptual, but not merely synthetic. It offers a framework for explaining why the legitimacy of AI-enabled assessment depends not on technical capability alone, but on the interaction of multiple interdependent domains whose tensions and misalignments shape assessment practice in consequential ways. By grounding legitimacy in institutional theory, the paper also clarifies that legitimacy is not the same as consensus or simple alignment. AI-enabled assessment may be pragmatically attractive to institutions, morally contested by students or faculty, and cognitively normalised through policy and routine, even when its educational justification remains uncertain. The model, therefore, reframes AI in higher education assessment as a problem of relational coordination, institutional legitimacy, and value contestation rather than implementation alone.

The model is not presented as a validated implementation tool or as a complete solution to the challenges of AI-enabled assessment. Rather, it is intended as a conceptual ecosystem framework for examining the tensions, dependencies, and responsibilities that shape how AI is adopted, interpreted, challenged, and revised in higher education assessment contexts. Future research should therefore investigate how the five domains interact in practice, where their tensions are most acute, and how legitimacy is negotiated across institutional, disciplinary, regulatory, and learner contexts. Empirical work should pay particular attention to student voice, faculty judgment, inclusive design, disciplinary differences, and governance arrangements that allow AI-supported assessment decisions to be explained, contested, and changed.

Ultimately, the value of the AI-Assessment Ecosystem Model lies in its effort to clarify what is too often treated as separate. AI in assessment is not only a matter of what systems can do, but of how educational institutions define purpose, distribute responsibility, exercise judgment, include diverse learners, and sustain legitimacy in increasingly 'datafied' environments. By framing assessment as an ecosystem rather than a technical application, this paper offers a more integrated basis for understanding not only the promises of AI in higher education, but also the tensions, dependencies, and conditions

under which its use can be educationally defensible, ethically legitimate, inclusive, and institutionally governable.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s44217-026-02225-y>.

Supplementary Material 1.

Author contributions

J O-N is responsible for all the contents of this paper.

Funding

Declaration.

I confirm that no funding is associated with this paper.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable. This study did not involve human participants.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 30 March 2026 / Accepted: 24 September 2026

Published online: 01 October 2026

References

1. Ajjawi R, Boud D. Researching feedback dialogue: An interactional analysis approach. *Assess Evaluation High Educ.* 2017;42(2):252–65. <https://doi.org/10.1080/02602938.2015.1102863>.
2. Bearman M, Dawson P, Ajjawi R, Tai J, Boud D. In: Re-imagining university assessment in a digital world. Springer; 2020. <https://doi.org/10.1007/978-3-030-41956-1>.
3. Bearman M, Nieminen JS, Ajjawi R. Designing assessment in a digital world: An organising framework. *Assess Evaluation High Educ.* 2023;48(3):291–304. <https://doi.org/10.1080/02602938.2022.2069674>.
4. Blodgett SL, Barocas S, Daumé III H, Wallach H. Language (technology) is power: A critical survey of bias in NLP. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020;5454–5476.
5. Boud D, Falchikov N. Aligning assessment with long-term learning. *Assess Evaluation High Educ.* 2006;31(4):399–413. <https://doi.org/10.1080/02602930600679050>.
6. Boud D, Molloy E. Rethinking models of feedback for learning: The challenge of design. *Assess Evaluation High Educ.* 2013;38(6):698–712. <https://doi.org/10.1080/02602938.2012.691462>.
7. Carless D, Boud D. The development of student feedback literacy: Enabling uptake of feedback. *Assess Evaluation High Educ.* 2018;43(8):1315–25. <https://doi.org/10.1080/02602938.2018.1463354>.
8. Chee H, Ahn S, Lee J. A competency framework for AI literacy: Variations by different learner groups and an implied learning pathway. *Br J Edu Technol.* 2025;56(5):2146–82. <https://doi.org/10.1111/bjet.13556>.
9. Dillenbourg P. What do you mean by collaborative learning? In: Dillenbourg P, editor. *Collaborative-learning: Cognitive and computational approaches*. Elsevier; 1999. pp. 1–19.
10. DiMaggio PJ, Powell WW. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *Am Sociol Rev.* 1983;48(2):147–60. <https://doi.org/10.2307/2095101>.
11. Divjak B, Svetec B, Horvat D, Kadoić N. Assessment validity and learning analytics as prerequisites for ensuring student-centred learning design. *Br J Edu Technol.* 2023;54(1):313–34. <https://doi.org/10.1111/bjet.13290>.
12. Dolmage JT. *Academic ableism: Disability and higher education*. University of Michigan Press; 2017.
13. Eubanks V. *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's; 2018.
14. European, Commission. Directorate-General for Education, Youth, Sport and Culture. *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*. Publications Office of the European Union; 2022. <https://doi.org/10.2766/153756>.
15. Hattie J, Timperley H. The power of feedback. *Rev Educ Res.* 2007;77(1):81–112. <https://doi.org/10.3102/003465430298487>.
16. Henderson M, Ajjawi R, Boud D, Molloy E, editors. (2019). *The impact of feedback in higher education: Improving assessment outcomes for learners*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-25112-3>
17. Henderson M, Ryan T, Phillips M. The challenges of feedback in higher education. *Assess Evaluation High Educ.* 2019;44(8):1237–52. <https://doi.org/10.1080/02602938.2019.1599815>.
18. Holmes W, Bialik M, Fadel C. *Artificial intelligence in education: Promises and implications for teaching and learning*. Centre for Curriculum Redesign; 2019.

19. Holstein K, Aleven V, Rummel N. (2020). A conceptual framework for human–AI hybrid adaptivity in education. In I. Bitencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán, editors, *Artificial intelligence in education* (pp. 240–254). Springer. https://doi.org/10.1007/978-3-030-52237-7_20
20. Ifenthaler D, Yau JY-K. Utilising learning analytics to support study success in higher education: A systematic review. *Education Tech Research Dev.* 2020;68:1961–90. <https://doi.org/10.1007/s11423-020-09788-z>
21. Jorre deS, Jorre T, Boud D, Johnson ED. Assessment for distinctiveness: Recognising diversity of accomplishments. *Stud High Educ.* 2021;46(7):1371–82. <https://doi.org/10.1080/03075079.2019.1689385>
22. Kasneci E, Sessler K, Küchemann S, Bannert M, Dementieva D, Fischer F, Gasser U, Groh G, Günemann S, Hüllermeier E, Krusche S, Kutyniok G, Michaeli T, Nerdel C, Pfeffer J, Poquet O, Sailer M, Schmidt A, Seidel T, Stadler M, Weller J, Kuhn J, Kasneci G. ChatGPT for good? On opportunities and challenges of large language models for education. *Learn Individual Differences.* 2023;103. <https://doi.org/10.1016/j.lindif.2023.102274>. Article 102274.
23. Kezar A. *How colleges change: Understanding, leading, and enacting change*. 2nd ed. Routledge; 2018.
24. Kezar A, Eckel PD. The effect of institutional culture on change strategies in higher education: Universal principles or culturally responsive concepts? *J High Educ.* 2002;73(4):435–60. <https://doi.org/10.1353/jhe.2002.0038>
25. Khosravi H, Shum SB, Chen G, Conati C, Tsai Y-S, Kay J, Knight S, Martinez-Maldonado R, Sadiq S, Gašević D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, Article 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
26. Kofinas AK, Tsay CH-H, Pike D. The impact of generative AI on the academic integrity of authentic assessments in higher education. *Br J Edu Technol.* 2025;56(6):2522–49. <https://doi.org/10.1111/bjet.13585>
27. Long D, Magerko B. What is AI literacy? Competencies and design considerations. *Proc 2020 CHI Conf Hum Factors Comput Syst.* 2020;1–16. <https://doi.org/10.1145/3313831.3376727>
28. Luckin R, Holmes W, Griffiths M, Forcier LB. *Intelligence unleashed: An argument for AI in education*. Pearson; 2016.
29. Marshall S. Change, technology and higher education: Are universities capable of organisational change? *Res Learn Technol.* 2010;18(3):179–92. <https://doi.org/10.1080/09687769.2010.529107>
30. Messick S. Validity. In: Linn RL, editor. *Educational measurement*. 3rd ed. Macmillan; 1989. pp. 13–103.
31. Meyer A, Rose DH, Gordon D. *Universal design for learning: Theory and practice*. CAST Professional Publishing; 2014.
32. Miao F, Cukurova M. AI competency framework for teachers. UNESCO; 2024.
33. Nazaretsky T, Ariely M, Cukurova M, Alexandron G. Teachers’ trust in AI-powered educational technology and a professional development program to improve it. *Br J Edu Technol.* 2022;53(4):914–31. <https://doi.org/10.1111/bjet.13232>
34. Ng DTK, Wu W, Leung JKL, Chiu TKF, Chu SKW. Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *Br J Edu Technol.* 2024;55(3):1082–104. <https://doi.org/10.1111/bjet.13411>
35. Nicol DJ, Macfarlane-Dick D. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Stud High Educ.* 2006;31(2):199–218. <https://doi.org/10.1080/03075070600572090>
36. Nieminen JH. Assessment for inclusion: Rethinking inclusive assessment in higher education. *Teach High Educ.* 2024;29(4):841–59. <https://doi.org/10.1080/13562517.2021.2021395>
37. Padden L, Tonge J, Moylan T, O’Neill G, editors. *Inclusive assessment and feedback: Universal Design case studies from IADT and UCD. UCD Access & Lifelong Learning*; 2019.
38. Panadero E, Broadbent J, Boud D, Lodge JM. Using formative assessment to influence self- and co-regulated learning: The role of evaluative judgement. *Eur J Psychol Educ.* 2019;34(3):535–57. <https://doi.org/10.1007/s10212-018-0407-8>
39. Perelman L. When the state of the art is counting words. *Assess Writ.* 2014;21:104–11.
40. Quality Assurance Agency for Higher Education. *Quality compass: Navigating the complexities of the artificial intelligence era in higher education*. QAA; 2024.
41. Sadler DR. Beyond feedback: Developing student capability in complex appraisal. *Assess Evaluation High Educ.* 2010;35(5):535–50. <https://doi.org/10.1080/02602930903541015>
42. Scott WR. *Institutions and organizations: Ideas, interests, and identities*. 4th ed. SAGE; 2014.
43. Seale J. *E-learning and disability in higher education: Accessibility research and practice*. 2nd ed. Routledge; 2014.
44. Selwyn N. Should robots replace teachers? AI and the future of education. Polity; 2019.
45. Selwyn N. On the limits of artificial intelligence (AI) in education. *Nordisk Tidsskrift Pedagogikk og Kritikk.* 2024;10(1):3–14. <https://doi.org/10.23865/ntpk.v10.6062>
46. Slade S, Prinsloo P. Learning analytics: Ethical issues and dilemmas. *Am Behav Sci.* 2013;57(10):1509–28. <https://doi.org/10.1177/0002764213479366>
47. Slade S, Prinsloo P, Khalil M. Trust us, they said. Mapping the contours of trustworthiness in learning analytics. *Inform Learn Sci.* 2023;124(9/10):306–25. <https://doi.org/10.1108/ILS-04-2023-0042>
48. Stahl G, Koschmann T, Suthers DD. Computer-supported collaborative learning: An historical perspective. In: Sawyer RK, editor. *The Cambridge handbook of the learning sciences*. Cambridge University Press; 2006. pp. 409–26.
49. Srijbos J-W. Assessment of computer-supported collaborative learning. *IEEE Trans Learn Technol.* 2011;4(1):59–73. <https://doi.org/10.1109/TLT.2010.37>
50. Suchman MC. Managing legitimacy: Strategic and institutional approaches. *Acad Manage Rev.* 1995;20(3):571–610. <https://doi.org/10.2307/258788>
51. Suthers DD. Technology affordances for intersubjective meaning-making: A research agenda for CSCL. *Int J Computer-Supported Collaborative Learn.* 2006;1(3):315–37. <https://doi.org/10.1007/s11412-006-9660-y>
52. Tai J, Ajjawi R, Boud D, Dawson P, Panadero E. Developing evaluative judgement: Enabling students to make decisions about the quality of work. *High Educ.* 2018;76(3):467–81. <https://doi.org/10.1007/s10734-017-0220-3>
53. Torraco RJ. Writing integrative literature reviews: Guidelines and examples. *Hum Resour Dev Rev.* 2005;4(3):356–67. <https://doi.org/10.1177/1534484305278283>
54. UNESCO. *Guidance for generative AI in education and research*. UNESCO; 2023.
55. Venter J, Coetzee SA, Schmulian A. Exploring the use of artificial intelligence (AI) in the delivery of effective feedback. *Assess Evaluation High Educ.* 2025. <https://doi.org/10.1080/02602938.2024.2415649>. Advance online publication.
56. Viberg O, Mutimukwe C, Grönlund Å, Andersson A. Exploring teachers’ (future) digital assessment practices in higher education: Instrument and model development. *Br J Edu Technol.* 2024;55(3):1186–208. <https://doi.org/10.1111/bjet.13462>
57. Williamson B, Eynon R. Historical threads, missing links, and future directions in AI in education. *Learn Media Technol.* 2020;45(3):223–35. <https://doi.org/10.1080/17439884.2020.1798995>

58. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *Int J Educational Technol High Educ*. 2019;16., Article 39. <https://doi.org/10.1186/s41239-019-0171-0>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.