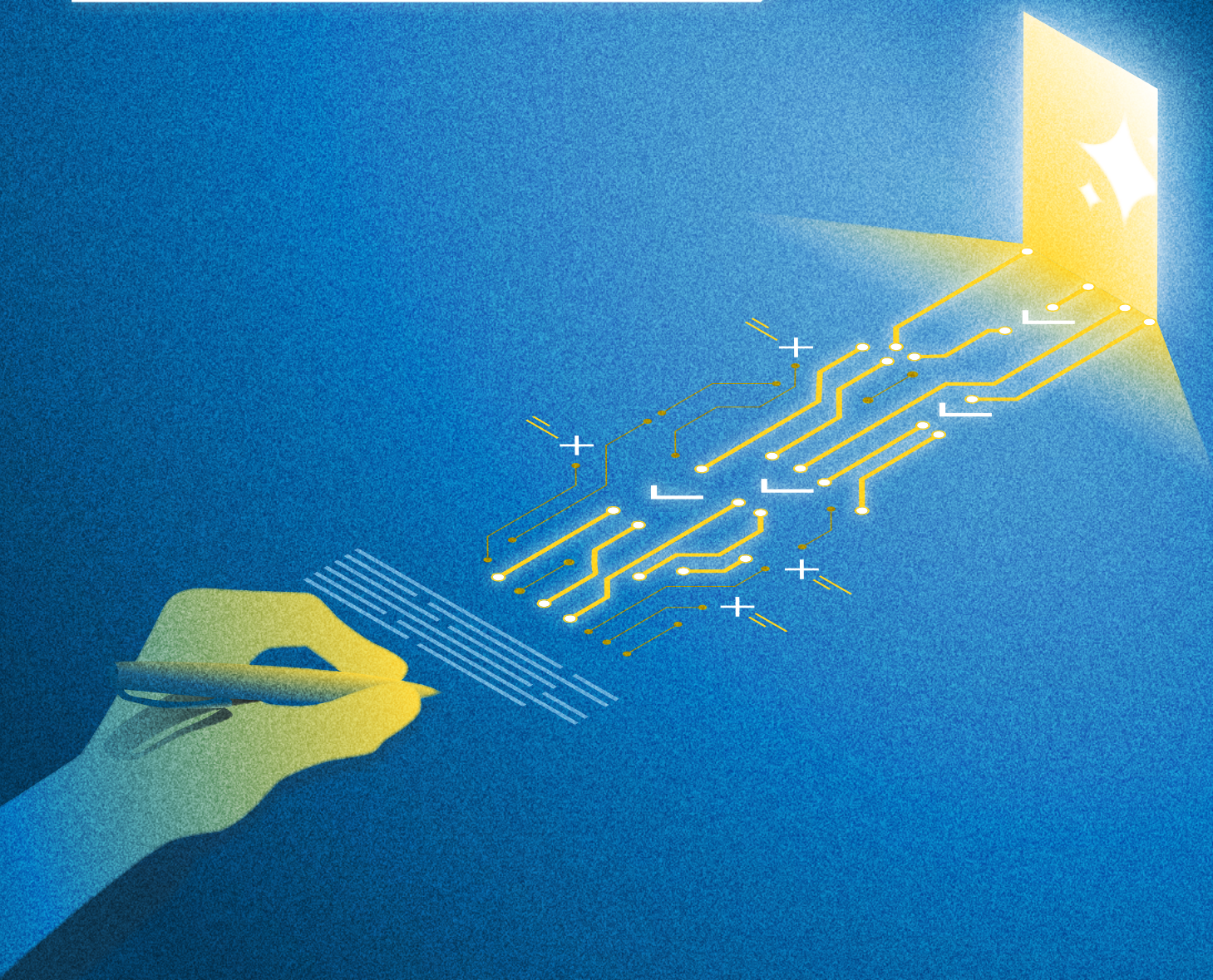
Assessing Teacher Learning in the Age of AI



European Schoolnet® | Academy

Publisher: European Schoolnet (EUN Partnership AISBL)
Rue de Trèves, 61
1040 Brussels, Belgium

Authors: Mary Richardson, UCL Institute of Education
Ev Abbott, UCL Institute of Education
Hertz Benjamin, European Schoolnet
Saltidou Efi, European Schoolnet

Design: Mattia Gentile, European Schoolnet

Please cite this publication as:

Richardson, M., Abbott, E., Hertz, B. & Saltidou, E. (2026). *Assessing Teacher Learning in the Age of AI*. European Schoolnet, Brussels, Belgium.

Published in February 2026, this report is based on the 6th European Schoolnet Academy Thematic Seminar organised in February 2026.

Cover design: assessment as a guided journey, where human evaluation evolves into structured pathways supported by Artificial Intelligence, shaped by informed choices rather than automatic outcomes.

Disclaimer: the drafting of this document was supported through the use of AI tools for the purposes of organization, clarity, conciseness, grammar, and spelling.

CONTENTS

AUTHORS

Mary Richardson, Professor of Educational Assessment, UCL Institute of Education

Mary is Professor of Educational Assessment at UCL Institute of Education in London. She has over 20 years' teaching and research experience including education practice and policy, education history and testing and assessment practice. Her doctoral candidates focus on a wide range of topics related to assessment including social media representation of testing, ethics and anxiety in testing, assessment of simulated learning in medicine, and assessment of creativity.

She has a background in educational assessment development and awarding through work as a Senior Research Officer in the dept for Research and Statistics at AQA, a national examination board in England. Her current research is anchored in assessment with specialist expertise in student experiences and how assessment shapes learner identity, public understanding of testing, the ethics of assessment and currently, what generative AI might mean to all our assessment futures.

She sits on the Research Board for Qualifications Wales, for AQA Exam Board's Research Department and is the co-convenor of the British Educational Research Association's Special Interest Group (SIG) for Curriculum and Assessment. She works on a range of international and national research projects and has conducted work in Georgia, Kazakhstan, Finland, Spain, Portugal, Ireland, Czech Republic and in Canada. She is the co-founder of IOE's Educational Assessment research group and runs monthly seminars and other public events to promote understanding of assessment in education.

Ev Abbott, Doctoral researcher, UCL Institute of Education

Ev Abbott is an experienced educator with 20 years in the field, working across the UK, Malaysia and Egypt. She has contributed extensively to in-school and regional initial teacher training and professional development. Ev is currently a doctoral researcher at the UCL Institute of Education, where her research focuses on assessment and creativity.

Benjamin Hertz, Senior Pedagogical Manager, European Schoolnet

Benjamin Hertz is a senior pedagogical manager at European Schoolnet. He is responsible for the pedagogical management of the European Schoolnet Academy, the first European MOOC platform dedicated to K12 education professionals. He is also in charge of the professional development offer on the European School Education Platform, the European Commission's website for the school education community. Benjamin has a background in educational technology with a particular interest in online learning and teacher professional development. Prior to his work for European Schoolnet he was teaching history at one of the leading sixth form colleges in the UK where he drove forward the implementation of technology-enhanced learning and supported the roll-out of a new learning management system with training and advice. Benjamin holds an MPhil degree in European Politics and a teacher qualification from the University of Oxford.

Efi Saltidou, Pedagogical Operational Manager, European Schoolnet

Efi Saltidou is a pedagogical operational manager at European Schoolnet. She oversees the coordination and pedagogical review of courses on European Schoolnet Academy, the first European MOOC platform for K12 education professionals, as well as courses on the European School Education Platform, the European Commission's portal for the school education community. She also supports the management of projects focused on digital education, innovative learning, and teacher training. Efi began her career as a primary school teacher in Greece, specialising in the integration of new technologies in teaching through a master's degree. She then became a teacher trainer at the Future Classroom Lab and an instructional designer of both online and onsite courses for educators. Over time, her passion shifted towards adult learning and teacher professional development, experimenting with innovative teaching approaches and contributing to research in online learning and educational technology.

INTRODUCTION

Assessing teacher learning - both during initial teacher education and throughout in-service professional development - is crucial to ensure that teachers acquire and continue to refine the competencies needed for effective teaching. Artificial intelligence (AI) is changing both *what* competencies teachers need to learn and *how* we assess that learning.

Teachers now need to understand how to use AI tools in the classroom, how to teach *with* and *about* AI, and how to address new ethical questions arising from AI in education and society in general. Teacher educators need to consider carefully how to assess these new competencies and what type of evidence is credible to show teacher learning and growth in this area.

However, AI is also challenging traditional assessment methods and tools while at the same time offering opportunities for more ongoing and formative processes of assessment. For example, some traditional metrics of teacher learning – reflection pieces, lesson plans, tests of content knowledge – are being thrown into question. If a student teacher can ask an AI to write a reflection paper, or a veteran teacher can use AI to design a lesson plan, then evaluating a teacher's growth based only on those artifacts is no longer reliable.

At the same time, AI offers new opportunities that could help make teacher assessment more data-driven, personalized, and continuous. For example, AI-supported lesson observations or AI-supported pedagogical coaching could offer teachers previously difficult to access formative assessment processes that are embedded in their daily practice.

The report explores these issues and is organised into three parts. First, it synthesises how teacher learning is currently assessed in initial teacher education (ITE) and continuous professional development (CPD). Second, it considers how AI is reshaping the knowledge, skills and attitudes teachers need to develop, and what this implies for assessment. Third, it explores how AI is prompting changes to assessment methods and tools. It closes with a set of practical recommendations for teacher educators and those involved in the organization of teacher education.

Current approaches to assessing teacher learning

Assessment focus: what is assessed

Assessing those learning to teach is distinctive and complex (see Figure 1). Becoming an educator requires the integration of subject knowledge with pedagogical, interpersonal and professional skills, alongside classroom management, planning and strategic decision-making. Global professional standards such as those developed by UNESCO & Education International (2024) show that teacher learning is best understood as multi-dimensional and developmental, rather than as a single body of knowledge and skills to be tested.

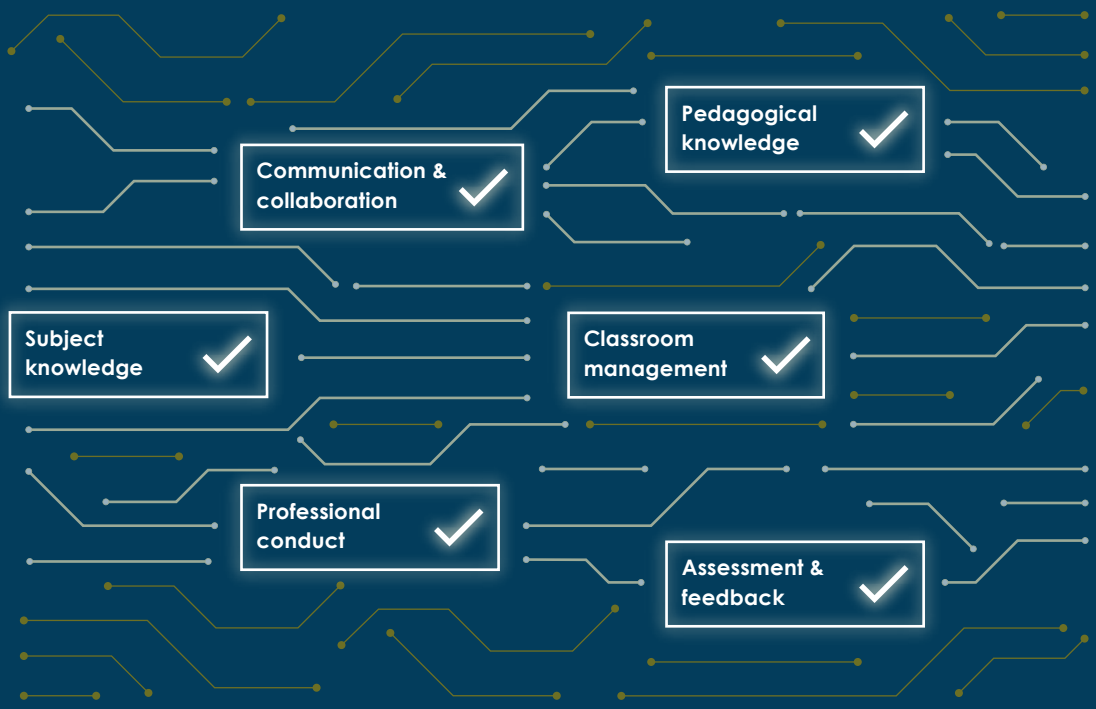


Figure 1: Summary of key skills necessary for teaching

What is being assessed in teacher learning is also highly contextual and is shaped by national and institutional policy. In England, for instance, trainee teachers must demonstrate that they meet government-defined Teachers' Standards (Department for Education, 2021). Similar national standards frame assessment in other systems (e.g., Estonia – (Ministry of Education and Research, 2025) Spain – (Gomez-Gonzalvo & Lores-Gomez, 2025) whereas in some contexts (e.g., Sweden – European Commission, 2024) programmes and individual teacher educators have greater autonomy in what to focus on.

Progression is a central tenet of all learning: understanding where a learner has come from, where they are, and where they are going. In teacher education, assessment of progression encompasses not only the development of the teacher but also evidence of pupil learning in their classrooms. Reviewing pupil outcomes - such as test results,

coursework, or classroom engagement – is therefore also sometimes taken into account when assessing teacher learning and it has long been integral to reflective dialogue between mentors and student teachers. This developmental view helps explain why mentoring, observation and reflective dialogue feature so prominently in teacher assessment (Walkington et al., 2025).

In initial teacher education, programmes increasingly adopt competence-based approaches that target teachers' theoretical knowledge, practical skills and professional dispositions (Caena, 2014; European Commission, 2015). In addition to subject content knowledge and pedagogical knowledge, student teachers are typically assessed on their ability to plan and enact lessons, manage classrooms, use instructional strategies, assess pupil learning, and reflect on their practice. Many programmes also treat reflective and ethical dispositions - such as critical self-reflection and a commitment to continuous learning - as important outcomes to be evidenced in assessment (Kelchtermans, 2009).

In continuous professional development, assessment tends to focus less on academic grading and more on evidencing professional growth: whether teachers have developed new knowledge or skills and are applying them effectively in their contexts. This may include changes in classroom practice (e.g., adopting formative assessment techniques), increased capacity to teach diverse learners, or improved assessment literacy. In some settings, evidence about pupil learning or engagement is also used as part of professional dialogue about a teacher's development (Timperley et al., 2007).

Taken together, assessing teacher learning is inherently complex because it involves judging the integration of knowledge, practice and professional dispositions over time, rather than measuring a single, stable set of skills. What "counts" as progress is also highly contextual, shaped by national standards, programme choices, and the realities of classrooms, so evidence often combines observation, mentoring dialogue, and traces of pupil learning. This is why teacher assessment is best understood as multidimensional and developmental, requiring multiple forms of evidence (often through portfolio-based approaches) to capture growth fairly (Meeus et al., 2009).

Assessment methods: how it is assessed

Building on the principle that teaching competence is multi-dimensional, teachers are typically assessed using multiple sources of evidence. These may include academic tasks, practice-based evidence from placement, and structured reflection, often coordinated through mentor–student teacher dialogue. Figure 2, created by the authors, outlines common modes of assessment used in student teacher education.

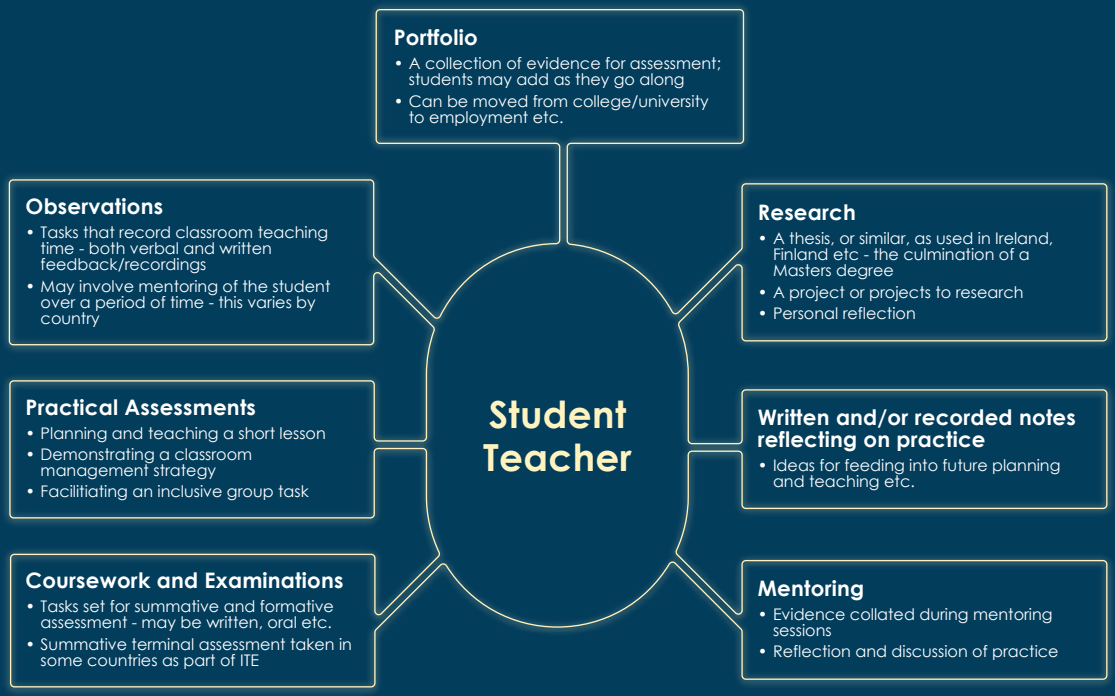


Figure 2: Common modes of assessment of student teachers

Nevertheless, Figure 2 is not intended as a comprehensive or prescriptive model. As Stobart (2008) reminds us, assessment design should begin with purpose: clarity about what is being assessed and why, and about what counts as credible evidence in a given context. In practice, programmes combine approaches in different ways to balance validity, feasibility, and fairness.

For example, ITE programmes typically combine academic assessment with performance-based evidence from practice placements and structured reflection. Common approaches include:

- **Written exams and coursework:** Many programmes still use written tests or exams to assess subject-matter expertise and knowledge of pedagogy. Coursework may include essays, case analyses or lesson plans that demonstrate understanding of theory and its application.
- **Classroom performance assessments:** Supervised teaching practice (practicum) is commonly assessed through classroom observation by mentors and/or university supervisors, using rubrics aligned to professional standards. Feedback discussions are often part of the assessment process.
- **Teaching portfolios:** Student teachers increasingly compile portfolios (often digital) that bring together artefacts such as lesson plans, assessment materials, pupil work samples, observation notes and reflections. Portfolios allow assessors to triangulate evidence across time and contexts.

- **Video analysis and reflection:** Video recordings of lessons are used in some programmes to support more granular analysis of classroom interaction and to anchor reflective dialogue. Video can also support peer feedback and self-assessment.
- **Reflective journals and narratives:** Some ITE programmes require reflective journals, essays or inquiry narratives to evidence growth in professional judgement, ethical reasoning and responsiveness to learners.
- **Peer and self-assessment:** In line with more participatory models of learning, some programmes include structured peer review (e.g., of lesson designs or video clips) and guided self-assessment against standards.
- **Oral presentations and professional dialogue:** To assess communication and professional reasoning, programmes may include oral examinations, presentations, or structured discussions (e.g., defending pedagogical choices made during practicum).

Across ITE, there is a strong trend toward combining multiple assessment forms to build a holistic picture of teacher candidates' competencies (Aparicio Herguedas et al., 2023). Blended designs typically pair evidence of practice (e.g., observation and artefacts) with analytic and reflective tasks, so that student teachers are assessed on what they know, what they can do, and how they reason about their decisions and learners' experiences.

Assessing in-service teacher learning is often less formalised than in ITE (see for example, Rodríguez-Robles et al., 2025), but it remains important for supporting and recognising ongoing professional growth. CPD is frequently assessed through cycles of goal setting, evidence collection, reflection and feedback. Evidence may be gathered by teachers themselves, by peers, mentors or school leaders, and increasingly through collaborative professional learning structures.

- **Teacher appraisal and observation:** Many systems include some form of appraisal or evaluation that draws on classroom observation, professional conversations and goal review. Where aligned to development (rather than only accountability), these processes can incorporate evidence of learning from CPD.
- **Self-assessment and reflective practice:** CPD often relies on teacher agency, so self-assessment tools, reflective journals and inquiry cycles are common ways of evidencing learning, especially when paired with feedback from peers or mentors.
- **Peer review and professional learning communities (PLCs):** Peer observation, lesson study and PLCs can embed peer assessment and feedback as part of collective improvement work.
- **Portfolios and artefacts of practice:** Some CPD routes require teachers to submit portfolios or artefacts demonstrating new practices, often with an accompanying rationale or reflection.
- **Workshops, courses and certifications:** When teachers participate in formal courses, assessment may include quizzes, assignments, lesson plans or applied projects.

- **Impact on teaching and learners:** In some models, evidence about changes in classroom practice and learner outcomes (e.g., engagement, progress on targeted goals) is brought into professional dialogue.

Overall, assessment in CPD tends to emphasise formative, ongoing self- or peer-reflection, supporting teachers to identify next steps, rather than one-off summative grading.

AI AND EVOLVING TEACHER LEARNING NEEDS

As AI and related technologies reshape aspects of teaching and learning, teacher education must reconsider not only how teachers use new tools, but also how they interpret AI-mediated evidence, uphold professional values, and make judgements in complex contexts. This section summarises emerging expectations for teacher learning in both ITE and CPD, and highlights what these shifts imply for assessment.

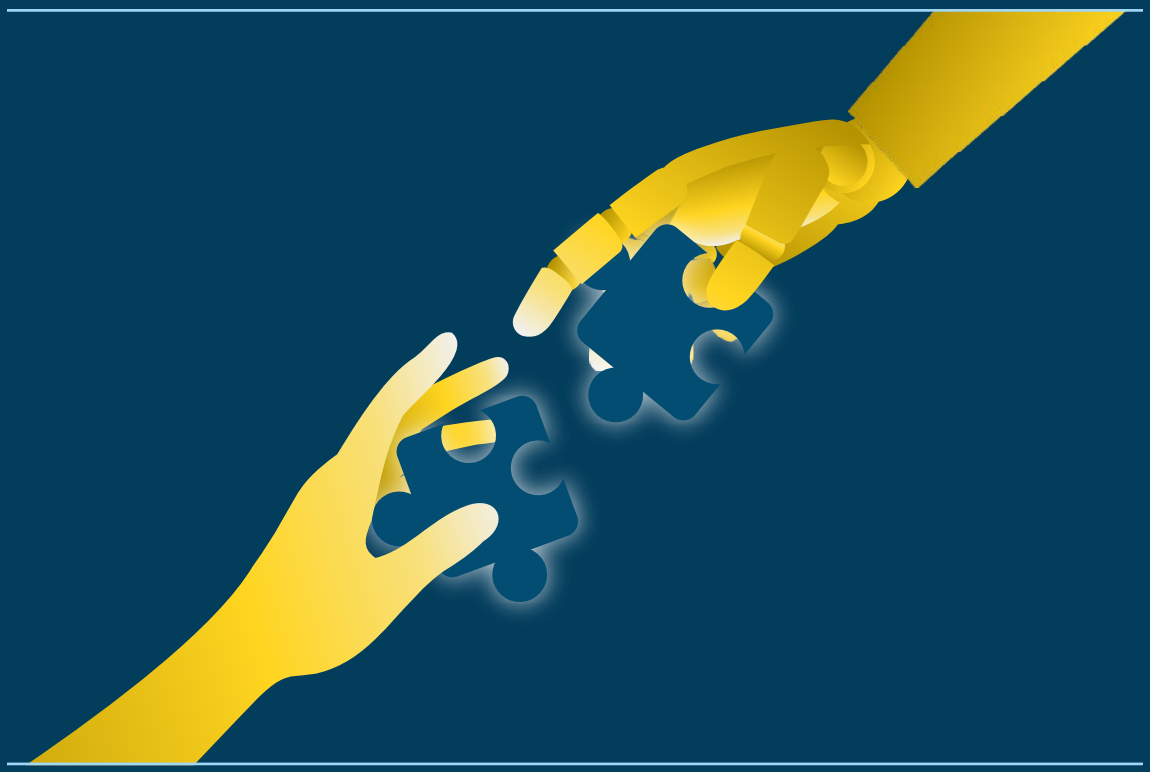
AI's growing presence in education demands that teachers acquire new kinds of knowledge and capabilities alongside traditional pedagogy. Recent frameworks converge on several core competency areas for teachers in an AI-driven world:

- **Foundational AI literacy:** Teachers need a basic understanding of how AI works – for example, being able to identify AI mechanisms and their operation (e.g. knowing how machine learning models are trained, and their limitations) (Filo et al., 2024). This foundational knowledge helps teachers grasp what AI can and cannot do, enabling them to critically evaluate AI tools. It also includes understanding key concepts like algorithmic bias, data privacy, and the difference between traditional software and adaptive, learning algorithms (Filo et al., 2024). In short, AI literacy is now essential for teachers, so that they can confidently navigate and supervise AI tools used in their classrooms (Lademann et al., 2026).

- **Skills for integrating AI in teaching:** Beyond theory, educators must learn how to effectively and informedly use AI in their professional practice (Filo et al., 2024). This involves practical skills like employing AI-powered tools for lesson planning, content creation, or providing adaptive learning supports. Teachers should be capable of selecting appropriate AI tools and integrating them into teaching and assessment in pedagogically sound ways. Critically, this competency also means maintaining human oversight, leveraging AI to enhance teaching while keeping human agency at the centre of educational decisions (European Digital Education Hub's squad on artificial intelligence in education, 2024). Effective integration skills ensure that teachers remain in control of the learning process, using AI as a support to better meet students' needs.
- **Ethical and responsible AI use:** The infusion of AI in schools raises complex ethical questions. Fütterer et al. (2025) urge for caution alongside the enthusiasm for employing AI in the classroom stating "…the role of AI [may] become a double-edged sword, promising advances in efficiency and effectiveness but posing important challenges to privacy and ethical standards in education''. Teachers, therefore, need knowledge and attitudes to use AI responsibly and ethically. Key considerations include issues of fairness, transparency, and student privacy (European Digital Education Hub's squad on artificial intelligence in education, 2024). Educators must be aware of AI biases and potential risks, for example, understanding that an AI-generated content might reflect certain cultural or gender biases. And they must know strategies to mitigate these in the classroom. Teachers also carry the responsibility of guiding students in the ethical use of AI (such as preventing plagiarism with generative AI and promoting digital integrity). Developing a strong ethical compass and critical mindset toward AI should be considered a core part of teacher professionalism (Kohout-Diaz, 2026).
- **Continuous learning and adaptability:** Finally, given the rapid evolution of AI, teachers must adopt an attitude of lifelong learning and adaptability. New AI tools and updates emerge frequently; what is state-of-the-art today may be outdated next year. Teachers, therefore, need the meta-skill of continuously updating their digital competencies. This involves staying informed about emerging AI in education trends and being willing to experiment and learn from practice.

Notably, these competencies encompass not only technical knowledge and skills but also values and attitudes. Successful integration of AI in schools depends heavily on educators' dispositions. Teachers need positive yet critical attitudes toward AI, seeing its benefits for teaching, while remaining mindful of its limitations and risks (Lademann et al., 2026).

Implications for assessing teacher learning



As AI alters what teachers need to learn, it also prompts a rethinking of *how we assess teacher learning outcomes*. Whether in initial teacher education or in-service training, the traditional metrics of teacher competence must be updated to include AI-related knowledge and skills. This has several implications for the design of teacher assessments, evaluation criteria, and certification processes:

- **Inclusion of AI competencies in teacher standards:** Teacher qualification frameworks and standards will likely incorporate explicit AI competencies. For example, a teaching standards body might add requirements such as "demonstrates understanding of AI tools relevant to teaching" or "able to use data from AI-driven assessments to inform instruction." This means that assessments leading to licensure or certification should test for these competencies. Some jurisdictions have started moving in this direction. For instance, Latvia is participating in the AI-Empowered Teaching project, which seeks to create an open online course comprising multiple modules designed to equip both pre-service and in-service teachers with the knowledge, skills, and competencies needed to integrate AI into the curriculum. Additionally, France, Ireland, Italy, Lithuania, Luxembourg, Slovenia, and Spain are involved in the Artificial Intelligence Digital Literacy (AI-DL) project, which fosters data literacy competences in post-primary education and supports collaborative groups of teachers and school leaders in incorporating generative AI into classroom practice. (European Schoolnet, 2025). UNESCO's new framework explicitly aims to "*guide the development of national AI com*petency frameworks… and help in designing assessment parameters" for

teachers (Miao & Cukurova, 2024). Concretely, this could involve adding AI-related questions or tasks in teacher licensing exams, or requiring preservice teachers to complete a project involving an educational AI application. By embedding AI competencies into the formal standards, we ensure that new teachers cannot qualify without demonstrating at least foundational AI literacy and integration skills.

- **Performance-based assessments and portfolios:** Given the complex, applied nature of many AI competencies, traditional pen-and-paper tests are often not sufficient. There is likely to be a greater emphasis on performance-based assess*ments* of teacher learning. For instance, a teacher might be assessed on how they integrate an AI tutoring system into a unit plan to personalize learning, or how they facilitate a classroom discussion on AI ethics. Teacher preparation programs could require creating an e-portfolio where candidates showcase evidence of their AI-related skills – perhaps including a sample of student feedback generated with an AI tool and the teacher's reflection on it. Such artifacts provide richer evidence of a teacher's ability to thoughtfully use AI in context, beyond what a multiple-choice test could capture. In Israel, researchers working on an AI competency framework have even involved teachers in co-creating assessment rubrics for AI literacy, noting the need for tools that allow teachers to *self-evaluate their proficiency* and growth in using AI (Filo, Rabin & Mor, 2024). This suggests that teacher assessments may include self-assessment and reflection components, acknowledging that mindset and ethical judgement (so critical in AI use) are best evaluated through guided reflection.
- **Assessing ethical reasoning and attitudes:** One of the more challenging implications is how to assess the attitudinal and ethical dimensions of teacher learning. It is relatively straightforward to test factual knowledge of AI or to observe a concrete skill. But evaluating a teacher's ethical stance or their disposition toward AI-assisted teaching can be subtle. This highlights the importance of using more qualitative assessment methods. These might include structured debates or scenario analyses as part of coursework – for example, giving a case where a student misuses an AI tool and asking the teacher candidate to discuss how they would respond, evaluating the answer for awareness of fairness, student agency, and so on. Feedback from mentors and supervisors is also crucial: in practicum or induction years, mentors can observe how new teachers incorporate (or resist) AI and provide evaluative feedback. While such qualities are harder to "score," they can be documented in narrative evaluations or rubric criteria (e.g. an indicator in an observation rubric like "Demonstrates critical and ethical use of educational technologies"). The inclusion of these aspects in assessment criteria signals to teachers that *how* they use AI – responsibly, transparently, and in a student-centered way – is as important as *whether* they can use it.
- **Continuous professional development:** For in-service teachers, the impact and ongoing development of AI further strengthens the idea that professional development isn't just about occasional workshops but ongoing competency development that might be tracked over time. Schools and districts may incorporate AI-related goals into teacher appraisal and evaluation systems. The broader implication is the need for a lifelong assessment of teacher competencies,

aligning with the idea that learning does not stop at certification, especially in a fast-changing domain like AI. Teachers may continually demonstrate their learning through periodic assessments, much like other professions (medicine, IT) require ongoing certification of new skills.

AI's growing role in education is broadening the content of teacher learning and necessitating parallel changes in how we assess teachers' knowledge and growth. Teachers now need to learn about AI – from technical basics to ethical and pedagogical nuances – and both their initial training and ongoing development must reflect this. For teacher educators and policymakers, this means updating curriculum and standards. It also means devising assessments that can validly measure these new competencies, whether through revised exams, performance tasks, or reflective evaluations. By ensuring that what we expect teachers to learn (and how we verify they've learned it) keeps pace with technological change, we prepare the teaching profession to make effective use of AI's potential while safeguarding the human values at the heart of education (European Digital Education Hub's squad on artificial intelligence in education, 2023; Miao & Cukurova, 2024).

AI AND HOW TEACHER LEARNING IS ASSESSED

AI, and especially generative AI, challenges conventional assessment practices in teacher education because it can produce fluent texts, lesson materials and analyses that resemble high-quality student work. Assessment approaches that focus primarily on final products (e.g., essays, unit plans or portfolios) risk conflating a teacher candidate's learning with the capabilities of the tools they used: are assessments capturing teachers' professional reasoning, judgement and growth, or primarily the outputs of AI systems? Some institutions are therefore simply replacing coursework with supervised examinations settings where any use of AI is prohibited. However, such an approach risks narrowing assessment to what is easiest to secure rather than what is most educationally meaningful.

While AI poses challenges, it also provides opportunities to enhance how we assess and support teacher learning. AI technologies have the potential to make teacher evaluation more objective, comprehensive, and formative, if used appropriately. Traditional observations and paper assessments capture only snapshots of teacher performance and sometimes carry subjectivity or bias. In contrast, AI systems can continuously analyze a wealth of data – from classroom video and audio, to student feedback and outcomes – to build a multifaceted picture of a teacher's practice. As Almubarak et al. (2025) note, "*AI enables the analysis of large volumes of data from various sources… facilitating a more comprehensive understanding of teaching effectiveness.*" By spotting patterns in

teaching methods and engagement, AI can "*more accurately and objectively assess educators' strengths and areas for growth*", and even provide immediate feedback for improvement. However, with these opportunities come new responsibilities and it is essential that AI is used in ethical and fair ways in teacher assessment.

This section therefore explores how AI requires a shift from product-focused assessment to a more process-focused approach and also explores the opportunities for more comprehensive, data-driven, and formative evaluation that AI offers as well as the risks and challenges associated with using AI in this way.

From product to process: assessing learning-in-practice

Written assessments focused on final products (e.g. lesson plans, unit plans, reflection journals) risk being undermined when AI can produce polished outputs with minimal human effort. AI complicates inferences about independent understanding or reflection when authorship becomes shared or opaque.

Although still commonly used, AI detection programmes are often unreliable. As AI technology develops and user prompts become more sophisticated, the avoidance of detection becomes easier. AI detectors pose the added risk of false positives, potentially damaging the relationship and trust between the course leader and student (Ahmed, Elsaid, & Almeer, 2023). Adapting traditional assessments to avoid the risk of AI collaboration, with strategies such as invigilated assessment, limiting access to technology, risks deprioritizing the kind of skills teachers ultimately need. They also limit accessibility for those with additional needs where technology can be of crucial assistance.

While completely AI resistant assessment is difficult to attain, dialogic strategies strengthen the AI resistance of a task by involving the teacher in the interrogation of their work and reflection on their teaching practice. Socratic discussions, group presentations, oral examinations, podcasts and video tasks all offer opportunities for critical and creative thinking about teaching practice in action. The value of being open to a range of ways to reflect on teaching is evidenced in the post-16 sector with examples from five European partners exploring digitalization and use of AI tools by student teachers and teacher educators (Attwell et al., 2020). Attwel et al.'s study found that the use of portfolios, reflective or process journals lessens the reliance on a single end point assessment. Such a shift to using process-oriented evidence in assessment, for example by asking teachers to document how AI was used, how outputs were evaluated, and how decisions were adapted to context, means that learning to use AI ethically and transparently becomes part of the assessment target.

Another example comes from the European Schoolnet Academy where *process-oriented assessment* was introduced in early 2025 with the aim to assess "*the entire journey of [an output's] development*" rather than only the submitted product (*European Schoolnet Academy, 2025)*. Participants must now document their steps, reflect on decisions, and disclose how they used any AI tools. By evaluating "*the steps taken to develop [the work], ensuring that learners actively engage with the material and demonstrate their*

understanding," this approach *"reduces over-reliance on AI for final outputs"* while encouraging reflection and critical thinking. In practice, teacher-learners submit their work *with* written reflections explaining how they brainstormed ideas, how they organized their plan, which resources or AI tools they used, and how they adjusted AI-generated suggestions to make the work their own. Peer feedback is still used, but final evaluation is *"contextualised within the learner's documented process,"* so that both the end product *and* the skills demonstrated along the way are assessed. This shift toward process not only protects academic integrity but also turns AI from a cheating tool into a learning support, as teachers-in-training must consciously integrate AI outputs and justify their use.

Opportunities enabled by AI

One area where AI could offer opportunities for the assessment of teacher learning is AI-driven classroom observation and analytics. For example, natural language processing and computer vision can analyze videos of a lesson to record key aspects of teaching. An AI model can detect, say, how often a teacher asks open-ended questions, whether all student groups get equal attention, or if the teacher's movements engage the whole room. Research prototypes have used video data to identify specific teaching behaviors (like *"Teacher-Explains"* or *"Student-Raises-Hand"*) and measure student participation, with a notable level of accuracy (Almubarak et al., 2025). Such systems can consistently track complex classroom dynamics that human observers might miss or only sample briefly. For instance, an AI using computer vision can follow how a teacher circulates among groups or responds to students' facial expressions, providing insights into classroom management and inclusivity (Almubarak et al., 2025). Likewise, audio and natural language processing analysis can assess discourse quality – who speaks when, the types of questions posed, the sentiments or confusion signals in student responses, etc. All this granular data can be synthesized into feedback for the teacher. Notably, these technologies address long-standing limitations of traditional observations: they can operate at scale (every lesson, not just one visit), reduce bias coming from personal relationships, and give consistent metrics to guide improvement. Early studies indicate that such AI-supported observations can make evaluations more reliable and fairer, while generating *"actionable insights for personalized professional development"* (Almubarak et al., 2025).

A concrete example is the recent *M-Powering Teachers* project at Stanford, which developed an AI feedback system for instructors. This tool uses natural language processing to analyze classroom transcripts and automatically measure how teachers interact with student ideas (a practice known as "uptake"). In a controlled study, the automated feedback significantly improved instructors' teaching behaviors: teachers who received AI feedback were more likely to acknowledge and build on student contributions in subsequent classes. Their students, in turn, completed more assignments and reported higher satisfaction. Crucially, the AI could provide detailed, timely feedback after each class – highlighting, for example, specific moments where the teacher's response followed up on a student's idea, or showing the balance of teacher vs. student talk time (Spector,

2023). This kind of immediate formative feedback is rarely possible at scale with human coaches. One professor noted that "*timely, specific feedback can improve teaching, but it's just not feasible for someone to sit in every class… an automated tool*" can fill that gap in a scalable and cost-effective way. Indeed, teachers in the study likened the AI feedback to a "fitness tracker" for teaching – a private aid to self-improve, not a surveillance tool. The tool gave nonjudgmental, constructive feedback that teachers could comfortably use to adjust their practice, without the pressure of an evaluator in the room (Spector, 2023).

Another area of AI application in assessing teacher learning is AI-powered simulations. In training, practice is vital, but novice teachers have limited chances to experiment in real classrooms. AI offers new ways to assess and develop teaching skills through realistic simulations. For instance, advances in AI-driven avatars and virtual students mean a trainee can teach a "class" of computerized students and get instant feedback on their performance. Small-scale and exploratory studies are starting to provide some evidence that AI-powered rehearsal of teaching scenarios can boost teacher learning: educators could teach the *same* virtual class multiple times and gradually improve by learning from mistakes in a low-stakes environment (Caudwell & Mallaband, 2025). While simulations can provide controlled opportunities to practise decision-making and receive feedback, they raise important questions about authenticity and what aspects of teaching can be validly assessed outside real classrooms.

Within educational research, there are still relatively few studies focusing on simulated teaching situations as part of the learning and assessment journey for student teachers. Sailer et al. (2023), for example, explored the development of diagnostic reasoning skills—those required to navigate the diversity and heterogeneity of classroom contexts, balancing collective and individual learning needs. Their study used written scenarios combined with either static or adaptive automated feedback. The latter, which provided interactive and personalised prompts, was described as 'promising' in supporting real-time learning, although the authors highlight the significant preparation and financial investment required. Subsequent studies (e.g. Bauer et al., 2025) reinforce these findings, showing that small, adaptive 'micro-feedback moments' can support student teachers in justifying pedagogical decisions. From an assessment perspective, such simulations align with emerging performance-based approaches, as they foreground decision-making, professional judgement, and reflection on action. While they lack the full authenticity of real classrooms, they may serve as valuable formative assessment opportunities, particularly for practising rare or high-risk scenarios that are difficult to observe consistently in placement settings.

Another area of potential is the use of AI for virtual coaching for teachers and consequently formative assessment. For example, AI can guide reflective practice by acting as an intelligent tutor for self-reflection. This could take the form of a digital reflective journal that doesn't just passively record thoughts, but actually prompts the teacher to dig deeper – "What made you choose that strategy? How did students react? How else might you approach it?" AI can provide those cues, reducing the mental burden of figuring out how

to reflect and letting teachers focus on the substance of their experiences (Caudwell & Mallaband, 2025). AI virtual coaching tools are already emerging such as Iris Connect, Riffbot or Edthena AI Coach, which provide a structured reflection cycle where teachers describe a problem or upload a video of their lesson, receive guided observation prompts, reflect on their practice, set goals, and get curated resources for improvement. Some of these platforms provide dashboards summarizing teachers' progress throughout these reflection cycles (e.g. which competencies improved), giving a new kind of data for evaluating professional growth. Importantly, some of these tools market themselves as operating as part of a human-based mentoring or coaching relationship by offering additional data and guidance that can be the basis for discussion between a mentor and a mentee, rather than entirely replacing that human relationship.

The above examples illustrate that AI can turn assessment into a continuous, embedded part of teacher learning – with feedback loops that operate in real time or close to it, and with personalized mentoring and support at scale.

Risks and safeguards

It is essential to note that while offering significant potential there are also inherent dangers and risks when using AI tools in this way and we must ensure ethical and fair use of AI in teacher assessment. Data privacy is paramount when using videos or tracking performance; teachers (and students) must consent to any AI monitoring. Moreover, it is essential that student teachers understand how data are generated and interpreted, and assessment of that data should prioritise professional growth over surveillance or accountability alone.

Bias in AI algorithms is another concern – if an AI model is not trained on diverse classrooms or teaching styles, its feedback could be skewed. For instance, a computer vision tool might misread cultural communication styles, or a natural language processing feedback system might falter with regional accents or multilingual settings (Spector, 2023). Addressing these issues requires ongoing human oversight, transparency in how AI evaluations are generated, and, importantly, involving teachers in designing and refining the tools.

Moreover, the goal should be for AI-driven assessment to be a supportive mirror for teacher growth, not a punitive measure. When implemented with care, AI can enhance the *formative* aspect of teacher assessment – helping teachers identify what to improve – while leaving summative judgments ultimately to human professionals who can account for context.

It is also important to consider that over-reliance on efficiency-driven tools can risk distorting assessment validity if the purpose of assessment is not kept central to design decisions. As Coe et al. (2014) argue, no single measure can adequately capture teaching quality; AI-informed assessment should therefore integrate multiple sources of evidence, keep humans firmly 'in the loop', and support holistic judgement of a student teacher's developing competence.

The relational dimension of teacher education further complicates the use of AI in assessment. Research on student–teacher relationships highlights their significance for engagement and achievement (Hattie, 2009; Roorda et al., 2011), with clear implications for feedback practices in ITE and CPD. AI-powered lesson observation tools, chatbot mentors and automated feedback systems may scale support and provide timely input informed by pedagogical research. However, principles of authentic and ethical assessment emphasise that feedback should remain dialogic, relational and embedded in trust, rather than reduced to one-way, automated judgement. For example, Jensen et al. (2025) found, that although students engaged with chatbots, their need for sustained, in-depth dialogue meant that the role of the supervisor and mentor remained essential. Evidence from studies also suggests that while chatbots may offer short-term assistance, over-reliance can increase social isolation (Crawford et al., 2024), and learners often value personalised feedback from human educators over AI-generated responses (Nazaretsky et al., 2024). While AI may scaffold reflection, the uniquely human qualities of judgement, emotion and contextual understanding remain central to any meaningful engagement with teaching practice. Assessment tools that make use of AI therefore need to preserve space for ambiguity, professional discretion and relationship-based feedback. These findings reinforce the need to ensure that AI-enhanced assessment practices strengthen rather than erode any human relationships in the context of teacher education.

AI is reshaping teacher assessment by undermining traditional product-focused approaches and prompting a shift toward process-oriented methods that emphasize transparency, reflection, and ethical use. At the same time, AI offers new opportunities for continuous, data-driven evaluation through tools like simulations and virtual coaching - though these innovations require careful attention to privacy, bias, and the preservation of human relationships in feedback.

RECOMMENDATIONS

In light of the issues examined throughout this report and building on the analysis of how AI is reshaping teacher learning, professional practice, and the validity of existing assessment methods, this final section sets out a series of recommendations. These recommendations are grounded in the evidence and arguments presented above and are intended to support policy-makers, teacher educators, and school leaders in moving towards assessment approaches that are fit for purpose in an AI-rich educational landscape. Taken together, they translate the report's key insights into practical and strategic directions, aimed at strengthening authenticity, ethical practice, and sustained professional growth in the assessment of teacher learning.

1. Update teacher standards and frameworks

- Embed AI competencies into national and institutional teacher standards, including foundational AI literacy, integration skills, and ethical use. This ensures that both initial teacher education and continuous professional development reflect the needs of teachers.

2. Redesign assessment approaches

- Shift from product-focused to process-oriented assessment methods that require teachers to document their decision-making, use of AI tools, and reflections on ethical considerations. This reduces over-reliance on polished outputs and promotes transparency and critical thinking.
- Incorporate performance-based tasks and portfolios that showcase authentic integration of AI in teaching, rather than relying solely on written exams or static artifacts.

3. Address ethical and attitudinal dimensions

- Include structured activities such as debates, case analyses, and reflective journals to assess teachers' ethical reasoning and attitudes toward AI use. These should be part of both formative and summative assessment processes.

4. Leverage AI for formative assessment

- Explore AI-supported observation tools, virtual coaching, and simulations to provide timely, personalized feedback and scalable professional development. These innovations can make assessment more continuous and embedded in practice.
- Use AI analytics to complement - not replace - human judgment, ensuring that feedback remains dialogic and relational.

5. Safeguard privacy, fairness, and human oversight

- Establish clear data privacy protocols for any AI-driven assessment involving classroom video, audio, or student data. Obtain informed consent and prioritize professional growth over surveillance.
- Mitigate algorithmic bias by ensuring diverse training data and maintaining human oversight in interpreting AI-generated evaluations.

6. Promote lifelong learning and continuous assessment

- Develop systems for ongoing competency tracking in CPD, recognizing that AI-related skills require regular updating. Consider periodic micro-credentials or certification aligned with evolving standards.

7. Preserve human relationships in assessment

- Ensure that AI-enhanced tools support - not replace - mentoring and professional dialogue, maintaining trust and relational feedback as central to teacher development.

CONCLUSION

Artificial intelligence is transforming both what teachers need to learn and how their learning is assessed. Traditional product-focused approaches such as essays, lesson plans, and static portfolios are increasingly insufficient in an era where AI can generate polished outputs with minimal effort.

At the same time, AI opens up new possibilities for continuous, data-driven, and formative assessment through tools like classroom analytics, simulations, and virtual coaching. These innovations can make teacher evaluation more authentic and embedded in practice, but they also raise critical challenges around privacy, bias, and the preservation of human relationships in feedback.

To ensure that assessment remains valid, ethical, and supportive of professional growth, teacher educators and policy makers must update standards, adopt process-oriented and performance-based methods, and integrate AI thoughtfully, always keeping human judgment and trust at the centre. By doing so, we can harness AI's potential to strengthen teacher learning while safeguarding the values that underpin quality education.

REFERENCES

Ahmed, M., Elsaid, K., & Almeer, S. (2023). *Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. International Journal for Educational Integrity, 19*, Article 17. https://link.springer.com/article/10.1007/s40979-023-00140-5

Almubarak, A., Alhalabi, W., Albidewi, I. & Alharbi, E. (2025). *An AI-powered framework for assessing teacher performance in classroom interactions: A deep learning approach. Frontiers in Artificial Intelligence, 8*, 1553051. https://doi.org/10.3389/frai.2025.1553051

Aparicio Herguedas, S., Pascual, I., & Molina, M. (2023). The effect of assessment procedures in the development of competences during initial teacher education: *A systematic review*. *Journal of Technology and Science Education, 13*(3), 2085–760. https://doi.org/10.3926/jotse.2085

Attwell, G., Deitmer, L., Perini, M., Roppertz, S., Attwell, G., Deitmer, L., Roppertz, M., & Tutlys, S. (2020). Digitalisation, artificial intelligence and vocational occupations and skills: what are the needs for training teachers and trainers? *Trends in Vocational Education and Training Research, Vol. III. Proceedings of the European Conference on Educational Research (ECER), Vocational Education and Training Network (VETNET)*, 30–42.

Bauer, E., Sailer, M., Niklas, F., Greiff, S., Sarbu-Rothsching, S., Zottmann, J. M., Kiesewetter, J., Stadler, M., Fischer, M. R., Seidel, T., Urhahne, D., Sailer, M., & Fischer, F. (2025). AI-Based Adaptive Feedback in Simulations for Teacher Education: An Experimental Replication in the Field [Article]. *Journal of Computer Assisted Learning*, *41*(1), n/a. https://doi.org/10.1111/jcal.13123

Caudwell, R. & Mallaband, P. (2025, 9 June). *AI in teacher professional development: The unnecessary and absolutely necessary cognitive load of developing as a teacher.* Chartered College Impact Journal. https://my.chartered.college/impact_article/i-in-teacher-professional-development-the-unnecessary-and-absolutely-necessary-cognitive-load-of-developing-as-a-teacher/

Caena, F. (2014). *Teacher competence frameworks in Europe: Policy-as-discourse and policy-as-practice*. *European Journal of Education, 49*(3), 311–331. https://doi.org/10.1111/ejed.12088

Coe, R., Aloisi, C., Higgins, S., & Major, L. E. (2014). *What makes great teaching? review of the underpinning research*. Sutton Trust. https://www.suttontrust.com/our-research/great-teaching/

Crawford, J., Allen, K.-A., Pani, B., & Cowling, M. (2024). When artificial intelligence substitutes humans in higher education: the cost of loneliness, student success, and retention [Article]. *Studies in Higher Education (Dorchester-on-Thames)*, *49*(5), 883–897. https://doi.org/10.1080/03075079.2024.2326956

Department for Education. (2021). Teachers' Standards. *Department for Education*,

2011(June), 1–11. https://www.gov.uk/government/publications/teachers-standards

Education International & UNESCO. (2024). *Global framework of professional teaching standards*. Education International / UNESCO.

European Comission. (2024). *Sweden: Teachers and education staff*. In *Eurydice* (Issue April 2010). https://eurydice.eacea.ec.europa.eu/national-education-systems/germany/teachers-and-education-staff

European Digital Education Hub's squad on artificial intelligence in education. (2024). *AI and ethics, human rights, law and educational data: Briefing report no. 6*. European Digital Education Hub. https://www.ai4t.eu/wp-content/uploads/2024/01/AI-squad-output_briefing-report-6.pdf

European Schoolnet Academy. (2025, 10 March). *AI usage policy and process-oriented assessment: Key updates for learners*. European Schoolnet Academy blog. https://blog.europeanschoolnetacademy.eu/2025/03/ai-usage-policy-and-process-oriented-assessment-key-updates-for-learners/

European Schoolnet (11/12/2025). Artificial Intelligence in School Education. *An overview of policy priorities and initiatives across 23 education systems*, Brussels, Belgium. http://www.eun.org/documents/411753/12100059/Agile+collection+of+Information+vol.6-11.12.25_Updated.pdf/d0e683fa-7c71-4913-ae1d-69acbb12598b

Filo, Y., Rabin, E. & Mor, Y. (2024). *An artificial intelligence competency framework for teachers and students: Co-created with teachers*. *European Journal of Open, Distance and E-Learning, 26*(S1), 93–106. https://doi.org/10.2478/eurodl-2024-0012

Fütterer, T., Goldberg, P., Bühler, B., Sikimić, V., Trautwein, U., Gerjets, P., Stürmer, K., & Kasneci, E. (2025). Artificial intelligence in classroom management: A systematic review on educational purposes, technical implementations, and ethical considerations. *Computers and Education: Artificial Intelligence, 6*. https://doi.org/10.1016/j.caeai.2025.100483

Gómez-Gonzalvo, F., & Lores-Gómez, B. (2025). Teacher Evaluation Patterns in Spain: Has Evaluation Become More Technical? *European Journal of Education*, *60*(4). https://doi.org/10.1111/ejed.70284

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement* (1st ed.) [Book]. Routledge. https://doi.org/10.4324/9780203887332

Jensen, L. X., Bearman, M., Boud, D., & Konradsen, F. (2025). Feedback encounters in doctoral supervision: the role of generative AI chatbots. *Assessment and Evaluation in Higher Education*, *0*(0), 1–14. https://doi.org/10.1080/02602938.2025.2478155

Kelchtermans, G. (2009). Who I am in how I teach is the message: Self-understanding, vulnerability and reflection. *Teachers and Teaching: Theory and Practice, 15*(2), 257–272. https://doi.org/10.1080/13540600902875332

Kohout-Diaz, M. (2026). *Making sense of AI in teacher education: A qualitative study of perceptions, practices and pedagogical tensions. Teaching and Teacher Education, 171*, 105342. https://doi.org/10.1016/j.tate.2025.105342

Lademann, J., Henze, J., Honke, N., Wollny, C. & Becker-Genschow, S. (2026). *Teacher training in the age of AI: impact on AI literacy and teachers' attitudes. Frontiers in Education, 10*, Article 1671306. https://doi.org/10.3389/feduc.2025.1671306

Meeus, W., van Petegem, P., & Engels, N. (2009). Validity and reliability of portfolio assessment in pre-service teacher education. *Assessment and Evaluation in Higher Education*, *34*(4), 401–413. https://doi.org/10.1080/02602930802062659

Miao, F. & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO. https://www.unesco.org/en/articles/ai-competency-framework-teachers

Ministry of Education and Research. (2025). *Estonian Education System*. https://www.hm.ee/en

Rodríguez-Robles, C., del Pino Rodriguez, L., Peña Trapero, N., Serván-Núñez, M. J., & Soto Gómez, E. (2025). Lesson study and the challenges of initial teacher education in Europe: keys to sustainability. A literature review [Article]. *International Journal for Lesson and Learning Studies*, *14*(2), 153–170. https://doi.org/10.1108/IJLLS-05-2024-0102

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The Influence of Affective Teacher-Student Relationships on Students' School Engagement and Achievement: A Meta-Analytic Approach [Article]. *Review of Educational Research*, *81*(4), 493–529. https://doi.org/10.3102/0034654311421793

Sailer, M., Bauer, E., Hofmann, R., Kiesewetter, J., Glas, J., Gurevych, I., & Fischer, F. (2023). Adaptive feedback from artificial neural networks facilitates pre-service teachers' diagnostic reasoning in simulation-based learning [Article]. *Learning and Instruction*, *83*, 101620. https://doi.org/10.1016/j.learninstruc.2022.101620

Spector, C. (2023, 8 May). *AI-driven feedback tool improves teaching practices, Stanford-led research finds*. *Stanford Report*. AI feedback tool improves teaching practices | Stanford Report

Stobart, G. (2008). *Testing times: the uses and abuses of assessment*. Routledge.

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. New Zealand Ministry of Education. https://www.educationcounts.govt.nz/publications/series/2515/15341

Walkington, H., Saunders, C., Helmbold, B., Schart, M., & Ebid, A. (2025). *Salient practices in mentoring teachers as researchers: The Goethe-Institute's German as a foreign language continuing professional development programme*. *Zeitschrift für Bildungsforschung*. https://doi.org/10.1007/s35834-025-00514-4

European Schoolnet®
Academy

www.eunacademy.org