HEA | An tÚdarás um Ard-Oideachas
The Higher Education Authority

Generative AI in Higher Education Teaching & Learning

AI-Resilient Assessment Practices

Contributors

James O'Sullivan
Colin Lowry
Ross Woods
Tim Conlon

HEA Generative AI Policy Framework
https://hub.teachingandlearning.ie/genai/policy-framework

HEA Generative AI Resource Portal
https://hub.teachingandlearning.ie/genai/

Version 1.0, December 2025
Higher Education Authority, Dublin
DOI: 10.82110/Z7MD-RM66

How to cite:
O'Sullivan, James, Colin Lowry, Ross Woods & Tim Conlon. *Generative AI in Higher Education Teaching & Learning: AI-Resilient Assessment Practices*. Dublin: Higher Education Authority, 2025. DOI: 10.82110/Z7MD-RM66.

This document, and all original content contained within, is licensed under the Creative Commons Attribution-ShareAlike 4.0 International Public License (CC BY-SA 4.0).

Context

The emergence of sophisticated generative artificial intelligence tools, particularly large language models capable of producing coherent, well-structured text on virtually any topic, represents a fundamental disruption to established assessment practices across Irish higher education. This document provides guidance to higher education institutions on developing assessment strategies that maintain academic integrity while acknowledging the permanence of AI in educational and professional contexts.

While the Higher Education Authority's policy framework for the ethical adoption of generative AI in higher education[1] upholds the rights of academics to reject the use of generative AI in their coursework, the unfortunate reality is that creating 'AI-proof' assessments presents extremely difficult pedagogical challenges. Such efforts assume that institutional practices can outmanoeuvre technology that is increasingly embedded in everyday digital tools and professional environments. Instead, these guidelines advocate for AI-resilient assessment practices, approaches that acknowledge AI's inevitable presence while maintaining the validity, reliability, and integrity of educational evaluation.

Resilience in this context implies institutional capacity for adaptation and flexibility while maintaining core educational purposes. The impossibility of AI-proof assessment stems from multiple factors that institutions and teaching staff must acknowledge. The pace of AI development ensures that technical barriers erected today may be obsolete within relatively short timeframes. Students' inevitable use of generative AI as a tool for both learning and plagiarism, as well as the increasing expectation AI literacy and fluency form part of transdisciplinary graduate attributes, makes prohibition both impractical and pedagogically questionable. Current detection tools remain imperfect and raise serious concerns about false accusations that could damage institutional credibility and student trust.[2] Critically, resources devoted to technological countermeasures divert investment from meaningful educational enhancement.

---

[1] https://hub.teachingandlearning.ie/genai/policy-framework

[2] Debora Weber-Wulff et al., 'Testing of Detection Tools for AI-Generated Text,' *International Journal for Educational Integrity* 19 (2023), https://doi.org/10.1007/s40979-023-00146-z; Jahna Otterbacher, 'Why Technical Solutions for Detecting AI-Generated Content in Research and Education Are Insufficient,' *Patterns* 4, no. 7 (2023), https://doi.org/10.1016/j.patter.2023.100796; Weixin Liang et al., 'GPT Detectors Are Biased against Non-Native English Writers,' *Patterns* 4, no. 7 (2023), https://doi.org/10.1016/j.patter.2023.100779; Chaka Chaka, 'Reviewing the Performance of AI Detection Tools in Differentiating between AI-Generated and Human-Written Texts: A Literature and Integrative Hybrid Review,' *Journal of Applied Learning and Teaching* 7, no. 1 (2024): 115–26, https://doi.org/10.37074/jalt.2024.7.1.14.

The disruption introduced by generative AI forces Irish higher education back towards assessment principles that educational theorists have promoted for decades. Put simply, institutions now need assessment that can still function as a credible measure of learning when high-quality text, code, and other artefacts can be produced or refined at negligible effort. That requires tasks that make students' judgement visible, treat knowledge as contextual rather than generic, and generate evidence that cannot be reduced to a single polished submission.

A central implication is the renewed importance of contextualisation. Assessment tasks need to be embedded in plausible, consequential scenarios that require students to navigate complexity, uncertainty, and competing considerations. The aim is to elicit critical and professional judgement: students must diagnose what matters, select an approach, justify trade-offs, and adapt to constraints, rather than follow a formula. Even so, staff should be clear-eyed about a persistent vulnerability, as a carefully designed scenario does not, by itself, protect validity if the primary evidence remains a written response that can be generated, polished, or substantively reworked by generative AI systems.

If assessment is to remain trustworthy, it must capture how students arrive at conclusions, not just what they submit at the end. Design should require traces of thinking and development across time, making visible decision rationales, intermediate versions, annotations, and other forms of working that can be interrogated. This can reduce straightforward substitution by shifting attention from product to evidential trail, but institutions should also acknowledge the limit of this protection, and that sophisticated use of gen AI can simulate development, particularly when 'process evidence' is itself submitted as unverified prose.

Tasks that draw on students' observations, situated experiences, and emerging disciplinary identities can introduce a particularity that generic outputs struggle to reproduce convincingly. When students must apply learning to their own contexts, or to contexts encountered in supervised settings, assessment becomes less interchangeable. The caveat is familiar: where claims cannot be checked, or where plausible fiction would satisfy the criteria, gen AI can still generate 'personal' material with little difficulty.

Reflective elements remain valuable as a way of eliciting how students understand their own learning and developing competence. On their own, written reflections are extremely vulnerable to AI generation, but their value increases when they are paired with modes of assessment that allow claims to be probed against observable practice such as oral questioning, supervised demonstrations, in-class problem-solving, annotated decision logs, or other evidence that ties an account of learning to

something that can be tested.

These guidelines ask institutions to shift from a posture of control to one of credible evidence. In practice, this means designing assessment ecosystems that combine modes, time, and interaction, so that what is being judged is not just a finished artefact but the learner's capacity to reason, choose, and justify within disciplinary contexts. It also means making institutional responsibilities explicit: providing staff with workable designs that do not depend on constant policing, ensuring students understand expectations for responsible AI use and disclosure, and investing in professional development, quality assurance, and consistent procedures when concerns arise.

Generative AI will continue to evolve, and higher education will not win by chasing each technical turn. It can, however, protect the integrity of awards by foregrounding judgement and making learning legible, and treating assessment as a public commitment to standards rather than a contest between educators and their tools.

Practical approaches to AI-resilient assessment

Invigilated examinations

In-person, invigilated examinations represent one of the most immediately available AI-resilient assessment methods. The controlled environment of traditional examination halls, where students lack access to digital devices and work under direct supervision, provides robust protection against AI assistance during the assessment itself. Invigilated examinations, while resource-intensive, offer important advantages including demonstration of individual capability under controlled conditions, standardised evaluation across large cohorts, and maintenance of comparability of standards.

The effectiveness of in-person examinations as an AI-resilient approach depends critically on their design and security. Institutions must acknowledge emerging threats, particularly from wearable technologies that may become increasingly sophisticated and difficult to detect. Smartwatches, augmented reality glasses, and other wearable devices pose growing challenges to examination integrity, therefore institutions should implement clear policies on wearable technology, including requirements for removal or disabling of all such devices, and detection capabilities as these technologies evolve.

Remote examinations, by contrast, have become fundamentally compromised. Despite proctoring software, students can readily access AI tools through secondary devices or collaborative platforms. Remote examinations should not be considered AI-resilient and should be avoided where assessment integrity is paramount. Where circumstances require remote assessment, institutions should employ alternative methods outlined in this framework rather than attempting to maintain traditional examination formats online.

To maximise the AI-resilience of in-person examinations, institutions should evolve question design beyond formats that can be addressed through pre-examination AI preparation. Students may use AI extensively in revision, potentially memorising AI-generated responses to anticipated questions, so examinations should prioritise assessment of capabilities that cannot be pre-prepared. In person, open-book examinations requiring analysis of previously unseen materials provided during the examination maintain strong authenticity. Questions requiring integration of multiple sources presented during the examination, or application of concepts to novel scenarios revealed only in the examination itself, resist prior AI preparation.

Case-based examinations where students receive detailed scenarios during the assessment and must provide time-limited analysis under controlled conditions demonstrate particular resilience. Such formats might include requiring students to critique arguments presented for the first time in the examination, identify errors in provided solutions, or evaluate competing interpretations of data supplied during the assessment. The key principle is that examinations should test capability to work with new information under controlled conditions rather than ability to reproduce pre-prepared content.

Maintaining large-scale in-person examinations has resource implications, including space requirements, invigilation costs, and accessibility provisions. Such investments should be considered against the fundamental compromise to assessment integrity posed by alternatives, and most institutions have experience in this form of examination and have already devised mature systems for exam coordination and invigilation.

Reimagined written assessment

Authentic assessment, grounded in situated, experiential, and inquiry-based learning, has long sought to align assessment with the kinds of intellectual work students are expected to perform beyond the university. In this tradition, assessment is a structured encounter with context and process. In the current context of widespread access to generative AI, this pedagogical lineage acquires renewed relevance.

While written assessment remains valuable for developing and evaluating sustained argumentation, institutions must acknowledge its fundamental vulnerability to AI assistance. No written assessment submitted outside controlled conditions can be considered fully resistant to generative AI use, but assessment designs that emphasise situated engagement and visible intellectual labour can increase authenticity and make inappropriate AI use more difficult and detectable.

Scaffolded writing processes that document intellectual development over time exemplify this approach. Rather than treating writing as a single performative act, they position it as an inquiry unfolding through stages of formulation, critique, and revision. By requiring proposal development, annotated bibliographies, draft sections, peer review engagement, and systematic revision, institutions create multiple points of engagement that resist wholesale AI substitution. Each stage should require response to specific feedback, demonstrating evolution of thinking. Teaching staff should recognise that determined students could use AI at each stage, making this approach resilient but not impervious.

Assessment tasks that require engagement with specific, locally relevant, or temporally bounded materials further anchor student work in lived contexts. Requirements for engagement with specific, locally relevant, or very recent sources provide additional authenticity. When students must analyse institutional documents unique to their institution, engage with guest speaker content from their specific modules, or respond to events occurring during the assessment period, generic AI responses become less applicable. Personalised prompts that explicitly require integration of personal experience, observation, or data collection create assessments that are inherently unique to each student.

Rather than asking students to ‘discuss sustainability in business’, they might instead be required to analyse the sustainability practices of three businesses within walking distance of their home, drawing on direct observation and interviews with staff. Rather than asking students to ‘analyse themes of migration in contemporary Irish literature’, they might be required to examine how three texts from the course relate to migration stories within their own family or community, incorporating interviews with

two people who have experienced migration and analysing how their narratives confirm, complicate, or challenge the literary representations they have studied, with specific attention to the language and metaphors used in those narratives. The requirement to identify specific language and metaphors used by real people in their own community makes AI substitution technically challenging as well as more transparent, as students would need to fabricate entire family histories and interview transcripts. Such specificity aligns assessment with experiential and inquiry-based learning, requiring genuine fieldwork that resists synthetic substitution. It is still possible that a student would use gen AI to fabricate entire family histories and interview transcripts, but it should be more apparent to an examiner.

Living documents that evolve through documented iterations reinforce this emphasis on process rather than product. Living documents that evolve through documented iterations provide process visibility that resists single-point AI intervention. Systems that capture iterative development create assessment artefacts that reveal engagement over time.

In all instances, institutions should implement mechanisms through which students can be called upon to discuss and defend their work, both pre- and post-submission. Such practices reaffirm the pedagogical purpose of authentic assessment: the cultivation of accountable, situated intellectual judgement, rather than the pursuit of illusory AI-proofing.

Oral and dialogic assessment

Oral assessment provides significantly greater resilience to AI substitution than written formats, though it places additional demands on staff time and institutional capacity, and presents practical implementation constraints.

Structured professional conversations that probe understanding through dynamic questioning provide robust assessment authenticity. The conversational format allows assessors to pursue unexpected avenues, challenge assumptions, and require real-time application to novel scenarios. This format particularly suits professional programmes where communication of expertise represents a core competency. However, institutions must ensure assessors are trained to conduct such assessments effectively and to recognise where students may be reciting memorised AI-generated content.

Combined written and defence formats balance prepared demonstration with adaptive expertise. While students may use AI in preparation of written materials, substantial defence components requiring response to unexpected questions maintain authenticity. The defence must be sufficiently substantial

and unpredictable that surface knowledge becomes apparent. Institutions should develop question banks that test genuine understanding rather than recall, and ensure defence components represent significant proportions of overall marks.

Peer teaching sessions where students must explain concepts to peers and respond to confusion in real-time create complex performance requirements resistant to AI substitution. The need to adapt explanations based on audience response and provide alternative examples when initial explanations fail requires deep understanding. Institutions must structure these sessions carefully to ensure all students engage meaningfully rather than allowing confident students to dominate.

Dialogic assessment refers to a family of approaches in which the credibility of student learning is established through interaction rather than through the retrospective inspection of a submitted artefact. Its value in the context of generative AI is straightforward: where a take-home product can no longer function as reliable evidence of authorship or understanding, live dialogue enables assessors to verify reasoning as it unfolds. The emphasis shifts from whether a piece of work 'looks authentic' to whether a student can explain, justify, adapt, and self-correct under questioning in ways that align with the module's learning outcomes.

Dialogic approaches should not be understood as a return to purely 'oral' traditions in a narrow sense, and they can be paired with written or practical artefacts, but those artefacts function as prompts for defence rather than as the primary evidential object. In practice this might take the form of whiteboard problem-solving with probing questions; mathematical proof discussion focused on logical structure and the consequences of altered assumptions; oral defence of experimental design choices and interpretation of anomalous results; case study defence in professional programmes; studio critique in creative disciplines; patient consultation performance and simulated client consultation using structured scenarios; field research presentations that foreground methodological adaptation and reflexive awareness; historical role-playing exercises that require students to inhabit constraints and period-specific reasoning; mock trial or moot court formats; policy simulations; structured academic debates; Socratic seminars; and peer code review in computing contexts. The common thread is not the format but the evidential mechanism: 'process visibility' produced through human interaction.

Where dialogic assessment is adopted, it benefits from explicit guidance on what counts as high-quality performance. Students should know, in advance, that what matters is traceable reasoning, principled justification, and the capacity to respond to challenge without collapsing into either bluff or deference. Clear criteria also help mitigate known equity risks associated with spoken assessment, particularly

where students differ in prior exposure to disciplinary discourse norms. This is one area where careful facilitation and inclusive participation design are not add-ons but the means by which the assessment remains both valid and fair.

Performance-based assessment

Performance-based assessment, while resource-intensive, provides strong resilience to AI substitution. Institutions should prioritise development of performance assessment where feasible, particularly in professional programmes.

Live problem-solving under observation, whether individual or collaborative, captures process as well as outcome. Case clinics, design charrettes, or technical challenges completed under time constraints require demonstration of approach as well as solution. The observational nature makes AI assistance impractical, though institutions must ensure problems are sufficiently novel that pre-prepared solutions are not applicable. Workplace simulations requiring navigation of complex scenarios in real-time provide authentic assessment of professional capabilities. However, institutions must regularly update simulation scenarios to prevent circulation of standard responses.

Laboratory work and field studies in scientific disciplines maintain strong authenticity when students must collect and analyse their own data. The physical dimension and unpredictability of empirical work require adaptive expertise. Nevertheless, institutions should be aware that data analysis and interpretation phases remain vulnerable to AI assistance unless conducted under controlled conditions.

Community-engaged assessment

Assessment connected to genuine community needs creates natural authenticity through stakeholder accountability. However, institutions must establish clear frameworks for community partnership and ensure appropriate support for both students and community partners.

Service learning projects addressing real community needs create assessment experiences that resist AI substitution through their relational and situated nature. The requirement for genuine stakeholder engagement and adaptation to evolving circumstances provides natural authenticity. Written outputs from such projects remain vulnerable to AI enhancement, and institutions should combine service learning with oral presentation or defence components.

Consultancy projects with actual organisational clients require tailored solutions accounting for specific constraints and contexts. The iterative nature of consultancy, with client feedback and evolving requirements, resists standardisation. Institutions must ensure clients understand assessment requirements and are prepared to verify student engagement.

Assessment in partnership with AI

Graduates will work in AI-augmented environments and should develop appropriate capabilities during their education. Strategic incorporation of AI into assessment can develop critical AI literacy while maintaining authenticity.

Structured AI engagement requiring critical evaluation of AI outputs develops essential graduate capabilities. Students might generate initial content using AI, then document their verification, enhancement, and critical analysis processes. Such assignments must include mechanisms to verify that critical engagement is genuine rather than performative.

AI literacy assessment can evaluate students' ability to craft effective prompts, assess outputs critically, and understand system limitations. Institutions must ensure such assessments do not simply test technical proficiency but develop critical understanding of AI's role in professional practice.

Implementation considerations

AI-resilient assessment has implications for institutional capacity, capability, and delivery models. Approaches such as oral assessment, process-oriented feedback, and portfolio-based systems may increase demands on staff time, infrastructure, and administrative coordination. These considerations highlight the need for institutions to take a strategic view of assessment design and delivery within existing resource and capacity considerations.

Dialogic assessment brings the 'question of scale' into the open. Many of the most robust approaches require time for interactive questioning and, therefore, smaller group sizes or a reallocation of staff effort across a module. Institutions can mitigate this by using calibrated teaching assistants, rotating short defence formats, for instance, ten-minute defences attached to a written task, and by designing dialogic checkpoints that sample understanding across the semester rather than concentrating

verification at the end.

It is also worth stating explicitly that, where the validity of a dialogic assessment derives from situated human judgement in real time, automated grading of the dialogue is conceptually misaligned with the assessment's purpose, even if AI tools may still have limited roles in administration, accessibility support, and feedback drafting where appropriately governed.

Institutions should explore efficient implementation models, including peer review systems that partially automate feedback processes, shared assessment resources across programmes, and technology platforms that support authentic assessment at scale, but efficiency must never compromise authenticity.

Movement towards AI-resilient assessment must not create new barriers to student success. Institutions should ensure multiple pathways for demonstrating achievement, recognising diverse student circumstances and capabilities. Students who struggle with oral presentation due to documented conditions must have equivalent alternative assessments that maintain authenticity. Technological requirements must not disadvantage students lacking personal access to equipment or high-speed internet.

The principles of Universal Design for Learning should inform assessment development. Cultural considerations must be addressed, recognising that different educational traditions value different forms of knowledge demonstration. Institutions must provide appropriate support for students unfamiliar with performance-based or oral assessment formats.

Traditional academic integrity policies require fundamental revision to address AI. Rather than prohibition-focused approaches, institutions should develop frameworks that acknowledge legitimate AI use while maintaining clear boundaries. Assessment declarations should evolve to require detailed attribution of AI assistance, including specific tools used, prompts employed, and verification processes undertaken.

AI literacy must be developed as a graduate attribute rather than treating AI solely as an integrity threat. Students should truly understand appropriate AI use in their disciplines, including when AI enhances versus substitutes for learning. Clear rubrics should specify acceptable AI use for each assessment, recognising that different tasks warrant different approaches.

Comprehensive professional development programmes are essential for successful implementation. HEIs should provide regular assessment design workshops with discipline-specific examples, create

communities of practice for sharing innovations, allocate time for assessment redesign within workload models, and recognise assessment innovation in promotion criteria.

Professional development must address both technical and emotional dimensions of transition. Many academics require support in understanding AI capabilities and limitations, designing authentic assessments, and maintaining confidence in their essential role in education.

Institutions should establish robust mechanisms for monitoring assessment effectiveness and adapting to evolving AI capabilities. Regular review cycles should evaluate whether assessments maintain integrity, support intended learning outcomes, and prepare students for professional practice. Student feedback should inform ongoing refinement, recognising that students often have sophisticated understanding of both AI capabilities and assessment vulnerabilities.

Collaborative, sector-wide evaluation of AI-resilient assessment approaches can support shared learning, including the exchange of both successes and challenges. Discipline-based networks and communities of practice may be well placed to contribute to the development and sharing of guidance, while respecting institutional autonomy and contextual differences.

References

Chaka, Chaka. 'Reviewing the Performance of AI Detection Tools in Differentiating between AI-Generated and Human-Written Texts: A Literature and Integrative Hybrid Review.' *Journal of Applied Learning and Teaching* 7, no. 1 (2024): 115–26. https://doi.org/10.37074/jalt.2024.7.1.14.

Liang, Weixin, Mert Yuksekgonul, Yining Mao, Eric Wu, and James Zou. 'GPT Detectors Are Biased against Non-Native English Writers.' *Patterns* 4, no. 7 (2023). https://doi.org/10.1016/j.patter.2023.100779.

Otterbacher, Jahna. 'Why Technical Solutions for Detecting AI-Generated Content in Research and Education Are Insufficient.' *Patterns* 4, no. 7 (2023). https://doi.org/10.1016/j.patter.2023.100796.

Weber-Wulff, Debora, Alla Anohina-Naumeca, Sonja Bjelobaba, Tomáš Foltýnek, Jean Guerrero-Dib, Olumide Popoola, Petr Šigut, and Lorna Waddington. 'Testing of Detection Tools for AI-Generated Text.' *International Journal for Educational Integrity* 19 (2023). https://doi.org/10.1007/s40979-023-00146-z.