

Implementing Generative AI Assessment Guidelines in a Mexican University

**Melchor Sánchez-Mendiola, Enrique Buzo-Casanova, Elibidú Ortega-Sánchez,
Manuel García-Minjares, and Adriana Cruz Juárez**

Coordination of Evaluation, Innovation and Educational Development (CEIDE)

National Autonomous University of Mexico (UNAM), Mexico City, Mexico

Correspondence: melchorsm@unam.mx

Abstract

Generative artificial intelligence (GenAI) complicates authorship and the interpretation of evidence of learning, yet institutional responses often remain limited to prohibition, detection, or broad principles. This work-in-progress case study examines how the National Autonomous University of Mexico (UNAM) developed and launched three coordinated guidelines for GenAI use in assessment across upper-secondary, undergraduate, and graduate education. The corpus comprised 120 pages of guidelines and project records. Structured document analysis, process tracing, and framework mapping identified shared governance principles, level-specific adaptations, and a pathway from institutional guidance to course-level assessment decisions. The documents share a learning-centered core of validity, fairness, transparency, academic integrity, privacy, bias awareness, metacognition, and human oversight. Their emphasis differs by level: gradual, guided use in upper-secondary education; assessment-cycle redesign and rubrics at undergraduate level; and AI literacy, research integrity, curated-source tools, and oral dialogue at graduate level. Development occurred through five collaborative phases over approximately 11 months. The documents received 5,276 downloads during their first month, indicating early reach rather than adoption or effectiveness. We propose a coherence-with-adaptation model and an implementation evaluation plan addressing acceptability, appropriateness, feasibility, adoption, fidelity/adaptation, reach, and sustainability.

1 Introduction

GenAI can draft and revise text, synthesize information, translate, create images and code, simulate dialogue, generate assessment tasks, and provide feedback. These capabilities make

polished products increasingly ambiguous as evidence of unaided learning. A written submission may represent independent reasoning, legitimate assistance, close human-AI collaboration, or substantial substitution of the learner's work. The validity question is therefore not simply whether GenAI was used, but whether the evidence supports the intended interpretation of learning under the assessment conditions specified (American Educational Research Association et al., 2014).

Institutional responses initially emphasized misconduct and detection. Those concerns remain important, but prohibition-only approaches can obscure legitimate educational uses, inequitable access, privacy risks, and preparation for AI-mediated professional work (Dabis & Csáki, 2024; UNESCO, 2023). Reviews of university guidance show a shift toward course-level communication, assessment redesign, responsible use, and stakeholder support (An et al., 2025; Moorhouse et al., 2023). Effective adoption also requires clear roles, infrastructure, implementation supports, and continuing revision rather than a static policy document (Chan, 2023; Jin et al., 2025; Robert & McCormack, 2024).

Assessment governance must protect tasks in which unaided performance is part of the construct while permitting or requiring GenAI when critical use of the technology is itself a learning outcome. It must communicate a common institutional position without erasing differences among disciplines, educational levels, assessment purposes, and faculty judgment. It must also translate broad principles into usable assignment instructions, rubrics, disclosure rules, and evidence-collection strategies.

This paper examines three guidelines for GenAI use in educational assessment at the upper-secondary, undergraduate, and graduate levels in a public university in Mexico (Universidad Nacional Autónoma de México [UNAM], 2026a, 2026b, 2026c). We ask: (1) How can guidance remain coherent across levels while preserving

developmental and disciplinary relevance? (2) What process supported collaborative guideline development? (3) What implications follow for assessment validity and implementation?

2 Institutional Context and Development Process

UNAM is a large public university spanning upper-secondary, undergraduate, and graduate education across heterogeneous disciplines and organizational settings. Assessment decisions are distributed among academic bodies, programs, courses, and faculty. Institutional survey work before publication also indicated uneven GenAI uptake and limited assessment use in some sectors, including graduate education (Benavides-Lara et al., 2025). A single prescriptive rule would therefore be difficult to enact and pedagogically undesirable. The project pursued coordinated guidance rather than a uniform binding regulation. The university Educational Evaluation Council and the Coordination of Evaluation, Innovation and Educational Development (CEIDE) convened three collaborative networks representing the three educational levels. Their shared objective was to provide practical, ethical, and pedagogically grounded guidance for assessment of, for, and as learning. The documents are advisory, they complement existing norms and faculty decisions rather than replacing them. All were published in Spanish under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 license (UNAM, 2026a, 2026b, 2026c).

Project records document five phases. Start-up and agreements established the objective, structure,

references, writing conventions, schedule, and shared workspace. Content development used approximately biweekly meetings and distributed writing. Internal integration and review merged contributions, checked coherence and coverage, and standardized references. Validation and final revision incorporated coordinator review, stylistic editing, and feedback. Closure and delivery included approval, production, and public release. The process lasted approximately 11 months.

The networks worked from a common architecture but adapted examples and emphasis. This operationalized the project's central design principle: coherence need not mean uniformity. A shared core defined values and decision requirements; each guide translated them into developmentally and disciplinarily appropriate recommendations.

3 Methods

3.1 Design and Corpus

We conducted an implementation case study using document analysis, process tracing, and framework mapping. The primary corpus consisted of the final guidelines: 32 pages for upper-secondary, 57 for undergraduate, and 31 for graduate education, totaling 120 pages. Supporting materials included the collaborative work plan, network presentations, integrated review materials, and the April 2026 launch presentation. No student records, faculty performance data, or individual-level survey responses were analyzed. Table 1 summarizes the corpus and the orientation of each guide.

Table 1: Documentary corpus and level-specific orientation of the three UNAM GenAI assessment guidelines.

Level and document	Scope and audience	Distinctive emphasis	Representative operational guidance
Upper-secondary (32 pages)	Faculty, academic staff, assessment designers, and students; developmental context includes many learners below legal adulthood.	Gradual, guided use; verification; metacognition; privacy and equitable access.	Guided self-assessment and generative quizzes in early semesters; cases and simulations in intermediate stages; integrative projects, source checking, use declarations, and reflective learning records in later stages.
Undergraduate (57 pages)	Diverse degree programs, disciplines, course formats, and faculty assessment practices.	Full assessment-cycle redesign from planning through metaevaluation; explicit local rules and rubrics.	Specify AI-use conditions; diversify evidence; communicate criteria; retain process evidence where relevant; use rubrics for critical and creative GenAI use; review validity, bias, and assessment consequences.
Graduate (31 pages)	Specialization, master's, and doctoral education across 42 programs and research-intensive contexts.	AI literacy, prompt quality, research integrity, disciplinary application, and authorship verification.	Combine no-AI, AI-assisted, and AI-integrated tasks; use curated-source/RAG environments; require transparent declaration; use dialogue or oral defense when undeclared use is suspected.

3.2 Document Analysis and Framework Mapping

The unit of analysis was a substantive principle, recommendation, warning, example, or implementation support. A comparison matrix used seven a priori domains from assessment and AI-governance literature: learning purpose, validity, fairness and equity, transparency, academic integrity, privacy and data protection, and human oversight. Metacognition, AI literacy, and implementation support emerged during review. The analysis was interpretive rather than frequency based; it sought functional similarities and meaningful adaptations, not topic counts.

Recommendations were then mapped to two complementary frameworks. The first distinguished assessment of, for, and as learning (Black & Wiliam, 2009; Sánchez Mendiola & Martínez González, 2024). The second distinguished no-AI, AI-assisted, and AI-integrated assessment (Digital Education Council & Pearson, 2025). The relevant condition depends on the construct: no AI when independent performance is the target; assistance when the tool supports but does not replace reasoning; and integration when responsible, critical GenAI use is part of the intended competence.

Process tracing reconstructed the development phases and governance tensions. Written products were compared with project records to distinguish common commitments from level-specific adaptations.

3.3 Early Reception and Planned Implementation Evaluation

Available reception evidence comprised the public launch and aggregate download counts. Downloads were interpreted only as reach and interest; they do not show that faculty read, accepted, adopted, or implemented the guidance, or that assessment quality improved.

For the next phase, we adapted implementation outcomes proposed by Proctor et al. (2011), informed by Nilsen's (2015) distinction among process, determinant, and evaluation frameworks. Planned outcomes are acceptability, appropriateness, feasibility, adoption, fidelity and local adaptation, reach, and sustainability. This distinction matters because dissemination can be extensive without meaningful practice change.

4 Results

4.1 A Common Assessment-Governance Core

Across the documents, GenAI is framed as pedagogical augmentation rather than a substitute for learner agency or professional judgment. The shared core combines three functions. First, it protects validity: learning purpose and construct should precede technology choice; task conditions, evidence, and interpretation should align; and AI-generated instruments or feedback require human review. Second, it establishes ethical and procedural safeguards: equitable access, advance communication of permitted uses, disclosure of material AI assistance, protection of sensitive data, verification of claims, bias awareness, and human responsibility for consequential judgments. Third, it treats metacognition and AI literacy as educational aims: learners should explain how they used a tool, what they accepted or rejected, and how they checked the result.

Table 2 translates these commitments into operational safeguards. Collectively, they shift the governing question from "Is AI allowed?" to "What use condition produces defensible evidence for this learning purpose?" The documents therefore combine protective and developmental aims: preventing substitution, bias, privacy loss, and invalid inference while supporting feedback, self-regulation, critical evaluation, and professional preparation.

Table 2: Common assessment-governance principles identified across the three guidelines.

Governance principle	Operational translation	Primary safeguard
Learning centrality and validity	Define the construct, learning outcome, task, and evidence before selecting an AI tool.	Prevents technology-driven design and construct drift.
Explicit AI-use condition	State whether AI is prohibited, permitted as assistance, or integrated as part of the competence.	Clarifies administration conditions and score interpretation.
Transparency and integrity	Communicate rules in advance and require disclosure of material AI assistance.	Supports authorship, accountability, and consistent expectations.
Fairness and access	Consider devices, connectivity, accessibility, AI literacy, and non-AI alternatives.	Limits opportunity gaps and exclusion.

Governance principle	Operational translation	Primary safeguard
Privacy and data protection	Avoid sensitive inputs; anonymize information and prefer institutionally approved tools.	Protects learners, faculty, and confidential academic material.
Verification and bias awareness	Corroborate claims, sources, and calculations; examine cultural and algorithmic bias.	Reduces error, confabulation, and inequitable consequences.
Human oversight	Retain faculty responsibility for consequential interpretation, feedback, and adjudication.	Preserves professional accountability and contextual judgment.
Metacognition and AI literacy	Ask learners to explain prompts, choices, corrections, rejected outputs, and limitations.	Makes critical human contribution and responsible use visible.

4.2 Level-Specific Adaptation

The upper-secondary guide emphasizes developmental readiness and guided progression. Because many learners are below legal adulthood, privacy and protection receive particular attention. It recommends gradual, guided, and pedagogically intentional use. Early applications include guided self-assessment, generative quizzes, and concept maps; later stages introduce cases, simulations, interdisciplinary problems, integrative projects, and more independent critique. Student guidance emphasizes source verification, citation, privacy, reflection on the system's contribution, and learning records. Rubrics may assess recognition of AI limitations, verification of information, and explanation of how the tool supported learning (UNAM, 2026a).

The undergraduate guide treats assessment as a full decision cycle: planning, evidence collection, interpretation, feedback, decision making, and metaevaluation. This prevents AI from being reduced to item generation or misconduct detection and accommodates examinations, projects, portfolios, performance assessments, self-assessment, and peer assessment. Faculty are advised to specify the AI-use condition for each task, communicate criteria in advance, preserve process evidence when relevant, and use rubrics that evaluate critical and creative AI use rather than polished output alone. Metaevaluation asks whether instruments were valid and reliable, criteria clear, bias reduced, and deep learning and justice promoted (UNAM, 2026b).

The graduate guide addresses specialization, master's, and doctoral education across 42 programs. It emphasizes AI literacy, prompt quality, disciplinary application, research integrity, and preservation of higher-order analytical and creative work. It combines no-AI tasks for foundational competence, AI-assisted work in which students remain responsible for validation and synthesis, and AI-integrated tasks that assess

critical and ethical use. For evidence-intensive activities, it recommends retrieval-augmented or curated-source environments grounded in faculty-selected scientific documents. AI use should be declared explicitly. When undeclared use is suspected, dialogue and an oral response or defense are recommended before reliance on AI-detection software, treating authorship verification as an educational and evidentiary problem (UNAM, 2026c).

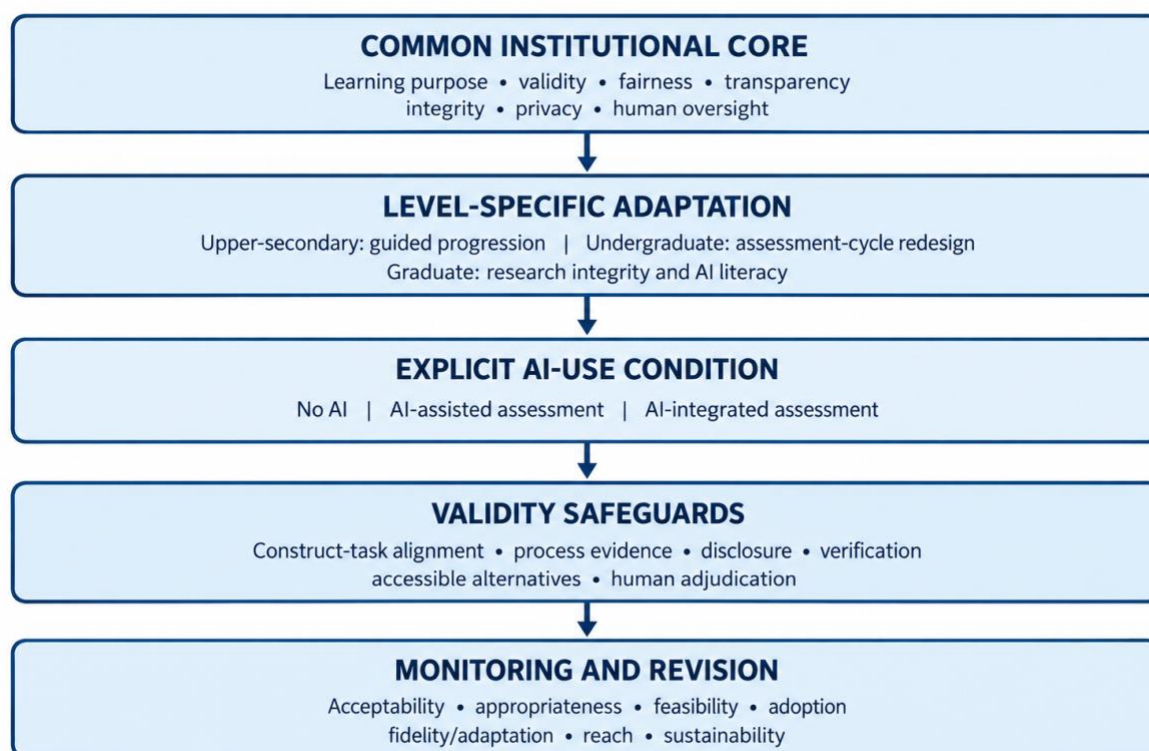
4.3 From Principles to Course-Level Assessment Decisions

Cross-document comparison yielded a five-layer policy-to-practice pathway (Figure 1): shared institutional core; level and disciplinary adaptation; explicit task-level AI-use condition; validity safeguards; and monitoring and revision. The safeguards include construct-task alignment, proportionate process evidence, disclosure, source verification, human adjudication, and accessible alternatives.

The pathway places tool rules downstream from construct decisions. An unaided explanation may support an inference about independent understanding; responsible GenAI-assisted synthesis requires tool access and criteria for verification, attribution, and judgment. Neither condition is inherently more authentic; each is defensible only in relation to the intended claim. Similarly, drafts, prompt logs, annotations, oral explanations, or reflective statements are useful only when they illuminate construct-relevant reasoning without disproportionate burden or privacy risk.

At minimum, an assessment that involves GenAI should communicate the learning outcome, AI-use condition, work that must remain human, disclosure requirement, possible process evidence, verification expectations, access and privacy safeguards, and locus of final judgment.

Figure 1: Coherence-with-adaptation pathway from institutional principles to task-level assessment conditions and implementation monitoring.



4.4 Early Reach and Current Implementation Status

The guidelines were launched publicly on April 29, 2026 and released as open-access documents. During their first month, they accumulated 5,276 downloads, indicating substantial early visibility for Spanish-language, assessment-specific GenAI guidance. The metric does not show full reading, perceived usefulness, adoption, or practice change. Current evidence therefore supports conclusions about development, documentary coherence, level-specific adaptation, and dissemination reach, not implementation fidelity, effectiveness, or learner outcomes.

5 Discussion

5.1 Coherence Without Pedagogical Uniformity

The main contribution is a governance design that combines a stable institutional core with differentiated operational guidance. Existing analyses repeatedly identify the need for clear course-level communication, ethical guidance, assessment redesign, and stakeholder responsibilities (An et al., 2025; Chan, 2023; Jin et al., 2025). The UNAM case adds a cross-level

perspective: upper-secondary, undergraduate, and graduate education share values but differ in developmental readiness, disciplinary complexity, research intensity, and learner autonomy. One undifferentiated document could obscure these differences; unrelated documents could create inconsistency.

The distinction between policy and guidance is consequential. Formal rules can establish mandatory boundaries for privacy, integrity, or high-stakes decisions, while flexible guidance helps faculty interpret those boundaries for particular constructs and tasks (Robert & McCormack, 2024). The documents primarily perform the second role. Their adaptability is a strength, but optional guidance may be enacted unevenly; publication must be followed by implementation support.

5.2 Implications for Measurement and Validity

GenAI makes the assessment condition part of the evidentiary argument. A product created under unspecified conditions has reduced interpretability because the contributions of learner, tool, and other sources are unclear. Explicit AI-use conditions improve interpretive clarity, but validity still

depends on alignment among construct, task, scoring criteria, and consequences.

Three implications follow. Institutions should preserve selected no-AI assessments when independent recall, reasoning, communication, or performance is essential. When assistance is permitted, tasks should elicit verification, selection, correction, and explanation rather than reward unexamined output fluency. When AI use is part of the construct, criteria should assess problem framing, source judgment, bias awareness, transparency, and accountability, not prompt mechanics alone.

Richer evidence systems should not imply surveillance of every interaction. Excessive monitoring can create privacy risks and construct-irrelevant burden. A defensible approach samples the minimum additional evidence needed: selected drafts, a short use declaration, annotated sources, an oral explanation, or a reflective account. The measurement contribution is therefore a decision framework for specifying when and why additional evidence is warranted, with human judgment retained for contextual interpretation.

5.3 From Dissemination to Implementation

It is important to emphasize that institutional reach is not implementation. Exposure must be followed by evidence that intended users find the guidance acceptable and appropriate, can feasibly apply it, and incorporate it into assessment design. Faculty development is central: educators need support to define AI-use conditions, redesign tasks, construct criteria for critical AI use, verify generated material, protect data, and respond fairly to suspected misuse.

The next phase should combine surveys and interviews with review of syllabi, assignment instructions, rubrics, and assessment plans. Disciplinary case studies can examine how faculty redesign tasks and how students experience the expectations. Program-level monitoring can assess reach and unequal access to training or approved tools. Recurring questions and implementation problems should feed revision, while successful adaptations can become reusable templates. This feedback loop treats the guidelines as living governance infrastructure rather than a finished publication.

5.4 Transferability

The design problem extends beyond UNAM. Multi-level, multi-campus, or professionally diverse institutions often need common expectations without erasing local expertise. The coherence-with-adaptation model may be especially useful where faculty autonomy is strong and central guidance must coexist with disciplinary variation.

The case also broadens the geographical and linguistic base of GenAI governance work, much of which has examined English-language policies from highly ranked institutions in high-income settings (An et al., 2025; Moorhouse et al., 2023). Spanish-language, open-access guidance from a large Latin American public university foregrounds digital inequality, developmental protection, and local faculty capacity. Transfer requires adaptation rather than replication, but the sequence (common principles, contextual translation, task-level rules, and implementation evaluation) is broadly applicable.

6 Limitations and Next Steps

The authors participated in developing or coordinating the guidelines, providing process access but creating a risk of favorable interpretation. We therefore distinguish documentary findings from impact claims and make institutional positionality explicit. Future evaluation should include non-participant users and, where feasible, independent analysis.

The study contains no faculty or student experience data, observed adoption, fidelity, assessment-quality ratings, or learning outcomes; downloads cannot establish use. The documents also differ in length, style, and operational detail because distinct networks produced them, and the interpretive analysis does not estimate theme prevalence. Finally, rapidly changing technologies and regulations may make tool-specific recommendations obsolete even when governing principles remain relevant.

The planned mixed-method evaluation will operationalize seven implementation outcomes (Table 3). Where redesigned assessments are implemented, later studies should also examine validity evidence, fairness, faculty workload, student learning processes, and unintended consequences.

Table 3: Planned evaluation of implementation outcomes, adapted from Proctor et al. (2011).

Implementation outcome	Key question	Candidate evidence for the next phase
Acceptability	Do intended users consider the guidance clear and useful?	Faculty and student surveys; interviews; recurring questions.
Appropriateness	Does it fit the educational level, discipline, and assessment purpose?	Ratings by level/program; disciplinary case studies; expert review.
Feasibility	Can users apply it with available time, expertise, access, and infrastructure?	Workload estimates; barriers/enablers; privacy and access audits.
Adoption	Are recommendations used in real assessment practice?	Syllabus statements; assignment instructions; rubrics; program guidance.
Fidelity and adaptation	Which core principles are retained, and how is guidance modified locally?	Document comparison; faculty narratives; assessment-design audits.
Reach	Which units and user groups encounter and use the resources?	Downloads, event participation, training coverage, unit-level distribution.
Sustainability	Are responsibilities, updates, support, and version control maintained?	Review schedule; governance ownership; updated editions; continuing faculty development.

7 Conclusion

The UNAM case demonstrates an assessment-governance approach that moves beyond a binary choice between permitting and prohibiting GenAI. Its central principle is coherence with adaptation: a common learning-centered and ethical core translated for upper-secondary, undergraduate, and graduate education. The policy-to-practice pathway begins with the construct, specifies the AI-use condition, identifies proportionate evidence and safeguards, and ends with monitoring and revision.

The project remains a work in progress. Published guidance and strong early reach demonstrate institutional capacity to communicate a framework, not practice change or improved assessment. The next test is whether faculty and programs use it to produce more interpretable, fair, transparent, and educationally meaningful evidence of learning. In the GenAI era, institutions must specify the human reasoning, judgment, and responsibility their assessments are intended to reveal.

Ethical Considerations and Researcher Positionality

This phase analyzed public institutional documents, project records, and aggregate download counts. It involved no individual-level student or faculty data. All authors participated in the initiative, and this dual role is acknowledged as a potential source of interpretive bias.

Acknowledgments

The authors thank the Educational Evaluation Council; the upper-secondary, undergraduate, and graduate collaborative networks; the faculty and academic staff who drafted and reviewed the guidelines; and CEIDE teams supporting editing, design, publication, and dissemination.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Educational Research Association. <https://www.testingstandards.net/open-access-files.html>
- An, Y., Yu, J. H., & James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, 22, Article 10. <https://doi.org/10.1186/s41239-025-00507-3>
- Benavides-Lara, M. A., Rendón Cazales, V. J., Escalante Rivas, N., Martínez Hernández, A. M. del P., & Sánchez Mendiola, M. (2025). Presencia y uso de la inteligencia artificial generativa en la Universidad Nacional Autónoma de México. *Revista Digital Universitaria*, 26(1). <https://doi.org/10.22201/ceide.16076079e.2025.26.1.10>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5-31. <https://doi.org/10.1007/s11092-008-9068-5>

- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20, Article 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Dabis, A., & Csáki, C. (2024). AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI. *Humanities and Social Sciences Communications*, 11, Article 1006. <https://doi.org/10.1057/s41599-024-03526-z>
- Digital Education Council & Pearson. (2025). The next era of assessment: A global review of AI in assessment design. <https://www.digitaleducationcouncil.com/post/the-next-era-of-assessment-a-global-review-of-ai-in-assessment-design>
- Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, 8, Article 100348. <https://doi.org/10.1016/j.caeai.2024.100348>
- Moorhouse, B. L., Yeo, M. A., & Wan, Y. W. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, Article 100151. <https://doi.org/10.1016/j.caeo.2023.100151>
- Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science*, 10, Article 53. <https://doi.org/10.1186/s13012-015-0242-0>
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(2), 65-76. <https://doi.org/10.1007/s10488-010-0319-7>
- Robert, J., & McCormack, M. (2024). 2024 EDUCAUSE action plan: AI policies and guidelines. EDUCAUSE. <https://library.educause.edu/resources/2024/5/2024-educause-action-plan-ai-policies-and-guidelines>
- Sánchez Mendiola, M., & Martínez González, A. (Eds.). (2024). Evaluación y aprendizaje en educación universitaria: Estrategias e instrumentos (2.ª reimp.). Universidad Nacional Autónoma de México, Coordinación de Evaluación, Innovación y Desarrollo Educativos. <https://www.ceide.unam.mx/index.php/evaluacion-y-aprendizaje-en-educacion-universitaria/>
- UNESCO. (2023). Guidance for generative AI in education and research. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- Universidad Nacional Autónoma de México. (2026a). Guía de uso de inteligencia artificial generativa en la evaluación educativa en bachillerato [Guidelines for the use of generative artificial intelligence in educational assessment at the upper-secondary level]. Consejo de Evaluación Educativa; Red Colaborativa del Bachillerato; Coordinación de Evaluación, Innovación y Desarrollo Educativos. <https://www.ceide.unam.mx/index.php/informes/#CEE>
- Universidad Nacional Autónoma de México. (2026b). Guía de uso de inteligencia artificial generativa en la evaluación educativa en licenciatura [Guidelines for the use of generative artificial intelligence in undergraduate educational assessment]. Consejo de Evaluación Educativa; Red Colaborativa de Licenciatura; Coordinación de Evaluación, Innovación y Desarrollo Educativos. <https://www.ceide.unam.mx/index.php/informes/#CEE>
- Universidad Nacional Autónoma de México. (2026c). Guía de uso de inteligencia artificial generativa en la evaluación educativa en posgrado [Guidelines for the use of generative artificial intelligence in graduate educational assessment]. Consejo de Evaluación Educativa; Red Colaborativa de Posgrado; Coordinación de Evaluación, Innovación y Desarrollo Educativos. <https://www.ceide.unam.mx/index.php/informes/#CEE>