Channel Islands
CALIFORNIA STATE UNIVERSITY

White Paper: AI-Assisted Grading in Higher Education

Ethical Considerations and Institutional Guidance

AUTHOR(S)	ISSUE DATE	VERSION	STATUS
Cynthia Flores, AI Ethics Board Co-Chair; Brendan Cline, AI Ethics Board Co-Chair; Jeff Benedetti-Coomber; Dani Rae; Lorna Gonzalez (on behalf of the 2025-2026 AI Ethics Board)	May 2026	1.0	For Review

Executive Summary

This white paper examines the ethical implications of AI-assisted grading in higher education, focusing on the central question: *Is it morally acceptable to use AI-assisted grading, and if so, under what conditions can it be used without compromising equity, transparency, trust, and student learning?*

Through structured ethical deliberation and stakeholder engagement, several key tensions emerged. These include the balance between efficiency and meaningful pedagogical engagement, consistency and scalability versus bias and loss of nuance, and innovation versus institutional responsibility and student rights. Concerns were also raised about the potential outsourcing of learning and cognitive development when AI tools are over-relied upon in assessment.

Rather than offering prescriptive policy recommendations, this paper surfaces key ethical considerations and identifies emerging directions for responsible practice. These include framing AI as a support tool rather than a replacement for human judgment, prioritizing transparency and accountability, and adopting a cautious, iterative approach to implementation supported by faculty development and institutional oversight.

Keywords: AI Ethics · Higher Education · AI-Assisted Grading · Pedagogy

Table of Contents

Background and Context

Artificial Intelligence (AI) systems now have the ability to automate the assessment of many forms of student coursework. There are reasons to be excited about the opportunities opened by this development, but there are also reasons to be cautious about overly hasty adoption in higher education. These tools can support tasks such as rubric generation, feedback drafting, and automated scoring, and may offer potential efficiencies for faculty managing large or writing-intensive courses. Furthermore, these types of AI-assisted grading features and functionalities are making their way into existing, campus-supported digital learning infrastructure (e.g., on the roadmap for and integrated into learning management systems).

At the same time, AI systems remain opaque in their internal decision-making processes and are known to produce errors, particularly in complex or open-ended assessments. Despite claims of objectivity, these systems make probabilistic and often subjective determinations about quality, coherence, and correctness, especially in interpretive or discipline-specific work.

Within the institutional context, several realities shape how AI-assisted grading may be adopted. Faculty training related to AI tools is voluntary under the California Faculty Association Collective Bargaining Agreement, leading to uneven engagement and adoption. AI tools are widely available and evolving rapidly, yet institutional guidance and oversight structures are still emerging.

At present, there is no formal policy at CSUCI or most other schools directly governing faculty use of AI for grading. A growing set of policies and governance frameworks inform this landscape, including CSU Academic Senate *guidelines* on AI in instruction, CSUCI Senate Policy 24-07 on syllabus disclosures, and federal regulations such as FERPA. These frameworks emphasize transparency, academic integrity, and the protection of student data and intellectual property.

This is an ethically complex choice that raises deep questions about student-faculty relationships. This report surveys some of the nearby ethical terrain to help inform thoughtful decisions about whether—and, if so, how—to adopt AI for grading. Our core concern is in assessing both (a) whether it is morally acceptable to use AI to assist grading in higher

education, and (b) if it is used, how can faculty minimize the risks of compromising equity, transparency, trust, and student learning.

Stakeholder Perspectives

Stakeholder engagement revealed important differences and areas of convergence.

Students

Students emphasized the importance of transparency and consistency, particularly when expectations around AI use differ between faculty and students. Many expressed concern that AI-generated feedback can feel impersonal or generic, potentially weakening the developmental and relational aspects of feedback. At the same time, some students recognized the potential value of AI as a support tool for improving consistency and reducing grading fatigue.

Faculty

Faculty expressed interest in AI-assisted grading as a means of managing workload but emphasized that grading is not merely a technical task. Rather, it is part of a pedagogical relationship involving care, recognition, and responsiveness to students. Concerns centered on maintaining the ability to “see” individual student effort and provide meaningful feedback.

Staff & Administrators

Staff and administrators highlighted issues related to compliance, data protection, and governance. Questions emerged about whether institutions should adopt vetted tools rather than relying on ad hoc use, as well as the need for careful piloting and evaluation prior to broader implementation.

Ethical Considerations

The ethical implications of AI-assisted grading are multifaceted and can be understood through considerations of outcomes, responsibilities, and values. Additional salient ethical dimensions of AI Assessment include:

1. Efficiency: AI assessment offers some significant benefits for both faculty and students, and perhaps the most significant of these is the efficiency that they offer.
2. Feedback Quality: If adopting AI in assessment reduced the quality of feedback, this would count strongly against it. However, if it is likely that feedback students receive from AI will indeed be of very high quality, then it will be worth consideration.

3. Modern Education: Eventually, resisting the use of AI for things like grading might appear quaint and outdated—like someone today expressing reservations about using email or using web browsers to read news articles.
4. Fairness: Students deserve to be graded fairly. Fair grading involves at least (a) applying similar evaluative standards across students, and (b) ensuring that students get the grade that they deserve—that their grade reflects the quality of their work in the context of the course. AI grading systems need to be quality checked for known biases.

Consequences

From a consequential perspective, potential benefits include increased consistency, reduced administrative burden, and the possibility of reallocating faculty time toward higher-impact teaching and mentoring. However, these benefits are accompanied by risks, including bias against non-native English speakers, loss of nuance and creativity in evaluation, and the potential weakening of instructor-student relationships.

Responsibility

From a responsibility-based perspective, several obligations emerge. Students have a right to understand how their work is evaluated, and faculty and institutions remain accountable for grading outcomes, even when AI tools are involved. Questions of data ownership, intellectual property, and adherence to learning outcomes further complicate the ethical landscape.

Values

From a values-based perspective, AI-assisted grading raises questions about what higher education seeks to prioritize. Tensions arise between care and efficiency, transparency and opacity, and respect for students as individuals versus the risk of reducing evaluation to a purely instrumental process.

Equity & Vulnerability

Equity is a central concern in AI-assisted grading. Since faculty ought to take special care to ensure that their pedagogical choices do not further harm students from marginalized or vulnerable groups, considerations of fairness seem to favor an attitude of skepticism and caution when choosing whether to rely on AI systems for grading. AI systems may disproportionately impact students whose forms of expression do not align with standardized assumptions embedded in training data. Neurodiverse students and non-native English speakers may be particularly vulnerable to misinterpretation.

In addition, uneven access to AI literacy may limit some students' ability to critically engage with or challenge AI-generated evaluations. Concerns about data use, consent, and intellectual

property further underscore the need for careful consideration of how student work is processed and stored. Even when used with good intentions, AI systems can reflect and amplify existing inequities if their limitations are not clearly understood or addressed. Because grades influence students' academic progress, confidence, and sense of belonging, institutions have an ethical responsibility to consider how AI-assisted grading practices may disproportionately impact students who are already navigating structural barriers.

Additional considerations should be taken in regards to the following areas:

Transparency

At the very least, students deserve to know the general details of how they are being evaluated so that they can organize their efforts to achieve the best possible grade. It is unfair to hold students' work to standards that are kept secret.

Interpretability

It is possible to convey information about things like the way the instructor has programmed the AI, what system(s) are being used, and what the instructor does with the outputs of the AI system. However, given the limits on interpretability described above, it is just not possible at present to know precisely how modern AI systems assess student work. Even the expert engineers who develop modern AIs do not fully understand how they work.

Privacy

If one thinks that student privacy concerns are very serious—ethically and/or legally—then privacy provides a very strong reason to officially designate approved AI grading systems that appropriately protect student information.

Hypocrisy

Some students feel that it is hypocritical for faculty to use AI to grade their work if they are not permitted to use AI to produce their work (or if they are limited in how they may use AI to produce their work). "Why do they get to use AI but I don't?"

The Value of Attention to Grading

Grading students' work expresses respect for their time, although it is worth noting that much undergraduate work in the US is graded by teaching or grading assistants (this is a less common practice at CSUCI); and in courses with multiple choice assessments (e.g., quantitative courses), some types of grading were offloaded to machine-scoring long ago.

Guidance and Recommended Practices

AI-assisted grading presents opportunities for improving efficiency and supporting timely feedback; however, its ethical use depends on preserving trust, accountability, and meaningful human engagement. Based on stakeholder input and existing research, the following practices are recommended as considerations for responsible use:

Use AI as a support tool, not a replacement for human judgement.

AI may assist with summarizing patterns or drafting feedback but should not replace instructor interpretation or decision-making.

Ensure transparency in AI use.

Students should be informed when and how AI contributes to grading or feedback, as transparency is essential for maintaining trust and agency.

Maintain human accountability.

Instructors and institutions remain responsible for grading outcomes and any associated harms.

Limit use in high-stakes assessment contexts.

AI-assisted grading should be applied cautiously in high-stakes settings without robust testing and oversight, particularly given evidence of persistent bias.

Adopt a gradual, iterative implementation approach.

Piloting and evaluation should precede broader adoption, with attention to equity and pedagogical impact.

Provide faculty development and support.

Faculty should be equipped to understand AI limitations, risks, and ethical responsibilities.

Protect student data and intellectual property.

Student work should remain within approved systems that ensure compliance with privacy regulations.

Monitor equity impacts continuously.

Fairness should be actively assessed through ongoing review rather than assumed based on aggregate performance.

Open Questions

Several questions remain unresolved and warrant continued exploration:

- Under what conditions, if any, should AI-assisted grading be used in high-stakes assessments?
- How can institutions ensure consistency in expectations for AI use across faculty and students?
- What forms of oversight are necessary to ensure accountability without limiting pedagogical flexibility?
    - What are some preconditions for success (i.e., what do faculty need to know and be able to do in order to use AI-assisted grading features effectively)?
- How can institutions support equitable access to AI literacy for all students?
- What role should locally developed or institutionally controlled AI systems play in assessment?
- How do students—and, especially, students at CSUCI—perceive the use of AI in grading? Will this change over time?

Readings and References

1. MIT Open Learning. (2024). *Ethics and AI-powered learning and assessment*.
2. Farazouli, A. (2024). *Automation and Assessment: Exploring Ethical Issues of Automated Grading Systems from a Relational Ethics Approach*.
3. Andersen, N., Mang, J., Goldhammer, F., & Zehner, F. (2025). *Algorithmic Fairness in Automatic Short Answer Scoring*.
4. CSU Academic Senate. *Guidelines on AI in Instruction*.
5. CSUCI Senate Policy 24-07. *Syllabus Policy (AI disclosure)*.
6. CSUCI Senate Policy 14-014.(Policy for Online Teaching and Learning)
7. FERPA (Family Educational Rights and Privacy Act)

Supplemental Suggested Readings

1. Awomosu, A. (2026). *They built Stepford AI and called it "agentic."* How Not to Use AI.
2. Center on Privacy & Technology at Georgetown Law. (2026). *An open letter to Georgetown students, in response to recent announcements about "generative AI."*
3. Conijn, R., Kahr, P., & Snijders, C. (2023). *The effects of explanations in automated essay scoring systems on student trust and motivation.* Journal of Learning Analytics, 10(1), 37–53.Hirsch, A. NIU Center for Innovative Teaching and Learning. (2024). *The digital red pen: Efficiency, ethics, and AI-assisted grading.*
4. Emirtekin, E. (2025). *Large language model-powered automated assessment: A systematic review.* Applied Sciences, 15(10), 5683.

5. Furze, L. (2024). *Don't use GenAI to grade student work.* LeonFurze.com.
6. Gao, Y. (2025). *AI and auto-grading in higher education: Capabilities, ethics, and the evolving role of educators.* ASC Office of Distance Education, The Ohio State University.
7. Guest, O., Suarez, M., Müller, B. C. N., van Meerkerk, E., et al. (2025). *Against the uncritical adoption of "AI" technologies in academia.*
8. Heil, J., & Ifenthaler, D. (2024). *Ethics in AI-based online assessment in higher education.* In *Ethics in Online AI-Based Systems* pp. 55–70. Elsevier.
9. Huang, Y., & Wilson, J. (2025). *Evaluating LLM-based automated essay scoring: Accuracy, fairness, and validity.* Proceedings of the Artificial Intelligence in Measurement and Education Conference: Works in Progress, 71–83.
10. Johnson, M. S. (2025). *Responsible AI for measurement and learning: Principles and practices.* ETS Research Report No. RR-25-03.
11. Knight, L. (2025). *BBC harnesses AI to create writing classes given by Agatha Christie.* The Guardian.
12. Lindsay, E. D., Zhang, M., Johri, A., & Bjerva, J. (2023). *The responsible development of automated student feedback with generative AI.* arXiv.
13. Pack, A., Barrett, A., & Escalante, J. (2024). *Large language models and automated essay scoring of English language learner writing: Insights into validity and reliability.* Computers and Education: Artificial Intelligence, 6, 100234.
14. Purser, R. (2025). *AI is destroying the university and learning itself.* Current Affairs.
15. Schaller, N.-J., Ding, Y., Horbach, A., Meyer, J., & Jansen, T. (2024). *Fairness in automated essay scoring: A comparative analysis of algorithms on German learner essays from secondary education.* Proceedings of the 19th Workshop on Innovative Use of NLP for Building Educational Applications, 210–221.
16. Tan, L. Y., Hu, S., Yeo, D. J., & Cheong, K. H. (2025). *A comprehensive review on automated grading systems in STEM using AI techniques.* Mathematics, 13(17), 2828.
17. Whitham, R., Stockton, G., Richards, D., Lindley, J., Jacobs, N., & Coulton, P. (2023). *Implications of generative AI on learning and assessment in higher education and design research practice.*
18. Zhong, Y., Hao, J., Fauss, M., Li, C., & Wang, Y. (2024). *AI-generated essays: Characteristics and implications on automated scoring and academic integrity.* arXiv.