



INTELIGÊNCIA ARTIFICIAL GENERATIVA NA AVALIAÇÃO DA APRENDIZAGEM NO ENSINO SUPERIOR: DILEMAS ÉTICOS, PEDAGÓGICOS E DE EQUIDADE

GENERATIVE ARTIFICIAL INTELLIGENCE IN HIGHER EDUCATION LEARNING ASSESSMENT: ETHICAL, PEDAGOGICAL, AND EQUITY DILEMMAS

DOI: 10.5281/zenodo.18489441



*Thiago Silva Ferreira*¹

*José Evaldo Gonçalves Lopes-Júnior*²

*Najla Maria Nogueira Ferreira*³

*Denise Gonçalves Moura Pinheiro*⁴

*Eduardo de Almeida e Neves*⁵

*Carla Christina Pereira da Silva Godinho*⁶

*Samuel Ramalho Torres Maia*⁷

*Luciana Rodrigues Ramos*⁸

*Daniel Lopes de Moraes*⁹

*Ana Giselle Costa de Almeida Moraes*¹⁰

1 Doutorando em Saúde Coletiva, Universidade Estadual do Ceará, Brasil.

2 Doutorando em Saúde Coletiva, Universidade Estadual do Ceará, Brasil.

3 Doutoranda em Saúde Coletiva, Universidade Estadual do Ceará, Brasil.

4 Doutora em Saúde Coletiva, Centro Universitário UniAteneu, Brasil.

5 Doutor em Biotecnologia, Centro Universitário UniAteneu, Brasil.

6 Mestre em Saúde Coletiva, Centro Universitário UniAteneu, Brasil.

7 Doutor em cuidados clínicos em Enfermagem e Saúde, Centro Universitário UniAteneu, Brasil.

8 Doutora em Educação, Centro Universitário UniAteneu, Brasil.

9 Especialista em Gestão dos Sistemas e Serviços em Saúde, Centro Universitário UniAteneu, Brasil.

10 Especialista em Gestão Pública Municipal, Centro Universitário UniAteneu, Brasil.





RESUMO

Este estudo teve como objetivo analisar criticamente como a inteligência artificial generativa (IAg) tem influenciado as práticas de avaliação da aprendizagem no ensino superior, considerando dimensões éticas, pedagógicas e de equidade. Realizou-se uma revisão integrativa segundo Whittemore e Knafl, complementada por Toronto e Remington e diretrizes brasileiras, incluindo busca sistemática em Scopus, Web of Science, PubMed, SciELO e LILACS (2018–2025). Dos 978 estudos identificados, 14 compuseram o corpus final. Os resultados revelam três eixos centrais: (1) transformações das práticas avaliativas e fragilização de modelos baseados em produtos finais; (2) tensões éticas relacionadas à autoria, autenticidade e integridade acadêmica; e (3) desigualdades sociotécnicas que impactam a justiça avaliativa. Conclui-se que a IAg reconfigura pressupostos da avaliação no ensino superior, demandando políticas claras, formação docente e abordagens que integrem tecnologia, pedagogia e equidade.

Palavras-chave: Inteligência Artificial, Avaliação Educacional, Ética em Pesquisa, Equidade.

ABSTRACT

This study aimed to critically examine how generative artificial intelligence (GenAI) has influenced learning assessment practices in higher education, focusing on ethical, pedagogical, and equity dimensions. An integrative review was conducted following Whittemore and Knafl, complemented by the methodological recommendations of Toronto and Remington and Brazilian guidelines, including a systematic search in Scopus, Web of Science, PubMed, SciELO, and LILACS (2018–2025). Of the 978 studies identified, 14 met the inclusion criteria. The findings reveal three core axes: (1) transformations in assessment practices and the increasing fragility of models based on final written products; (2) ethical tensions related to authorship, authenticity, and academic integrity; and (3) sociotechnical inequalities that affect fairness in assessment. The study concludes that GenAI reconfigures foundational assumptions of assessment in higher education, requiring clear institutional policies, continuous faculty development, and approaches that integrate technology, pedagogy, and equity.

Keywords: Artificial Intelligence, Educational Measurement, Ethics, Equity.

1. INTRODUÇÃO

A rápida difusão da inteligência artificial generativa (IAg), especialmente de modelos de linguagem de grande porte como o ChatGPT, tem reconfigurado o ecossistema do ensino superior, com efeitos diretos sobre práticas de ensino, aprendizagem e avaliação. Revisões sistemáticas recentes mostram que a IAg vem sendo integrada a múltiplas dimensões da vida





universitária, desde atividades de escrita e estudo autônomo até o desenho curricular e o suporte institucional à docência, com forte ênfase em recomendações para uso ético e responsável dessas tecnologias (Tillmanns *et al.*, 2025; Hon, 2025; Lo, 2023). Nesse cenário, a avaliação da aprendizagem emerge como um dos pontos mais tensionados, pois é precisamente nesse domínio que se cruzam expectativas de rigor acadêmico, integridade, autenticidade e justiça.

No contexto específico do ChatGPT e de ferramentas afins, diversas revisões têm documentado a amplitude das aplicações no ensino superior, incluindo apoio à escrita, feedback formativo, resolução de problemas e preparação para provas (Baig; Yadegaridehkordi, 2024; Montenegro-Rueda *et al.*, 2023). Ao mesmo tempo, esses estudos destacam que a adoção da IA permanece marcada por ambiguidades: docentes e estudantes reconhecem ganhos de eficiência, mas expressam preocupações com dependência cognitiva, empobrecimento dos processos de aprendizagem e deslocamento da avaliação para a simples verificação de produtos textuais. Pesquisas empíricas apontam ainda que a IA desafia concepções tradicionais de autoria e originalidade, exigindo reconfigurações nos critérios de avaliação e nos modos de orientar o trabalho acadêmico (Gamage *et al.*, 2023; Meyer; Urbanowicz; Martin, 2023).

Um eixo central da literatura recente trata das implicações da IA para a integridade acadêmica e o desenho das tarefas avaliativas. Estudos teóricos e empíricos sugerem que avaliações baseadas apenas em produtos finais, predominantemente textuais e realizados fora de ambientes supervisionados, tornam-se particularmente vulneráveis ao uso não transparente de IA exigindo estratégias de redesign que privilegiem processos, raciocínio explícito e interação em tempo real (Cotton *et al.*, 2023; Evangelista, 2025; Susnjak; Mcintosh, 2024). Evidências provenientes de diferentes contextos institucionais mostram que, ao mesmo tempo em que a IA pode ser mobilizada para apoiar avaliações formativas, ela também amplia a complexidade de detecção de práticas de desonestidade acadêmica e intensifica debates sobre o que conta como “uso aceitável” de tecnologia em tarefas avaliativas (Rudolph *et al.*, 2023; Newton; Xiromeriti, 2024).





Outra vertente relevante da produção científica aborda as dimensões de equidade, acesso e justiça presentes na adoção da IA_g em avaliações. Pesquisas qualitativas com estudantes indicam que o uso de ChatGPT em tarefas avaliativas é percebido simultaneamente como recurso de apoio e como fonte de assimetrias, na medida em que diferenças de letramento digital, proficiência linguística e acesso a versões pagas podem produzir vantagens desiguais (Finkel-Gates, 2025; Vargas-Murillo *et al*, 2023). Ao mesmo tempo, revisões sobre IA_g no ensino superior argumentam que a ausência de diretrizes claras sobre usos pedagógicos e limites éticos da tecnologia em avaliações gera incerteza normativa e fragiliza a confiança de estudantes e docentes nas regras do jogo avaliativo (Sullivan *et al*, 2023; Moorhouse *et al*, 2023; Tillmanns *et al*., 2025).

Diante desse quadro, marcado por rápidas transformações tecnológicas, debates intensos sobre integridade acadêmica e indícios de assimetrias no acesso e no uso da IA_g, há necessidade de uma síntese analítica que integre evidências empíricas e reflexões teóricas sobre o impacto da IA_g especificamente nas práticas de avaliação da aprendizagem no ensino superior. Assim, o objetivo desta revisão integrativa é analisar criticamente a produção científica recente sobre o uso da inteligência artificial generativa em processos de avaliação no ensino superior, com ênfase nas dimensões éticas, pedagógicas e de equidade, identificando convergências, tensões e lacunas que possam orientar tanto futuras pesquisas quanto a formulação de políticas institucionais de avaliação.

2. METODOLOGIA

A revisão integrativa foi conduzida com base no método proposto por Whitemore e Knafl (2005), considerado a abordagem mais consolidada para integrar estudos teóricos e empíricos em um mesmo processo de síntese. Essa metodologia permite examinar a amplitude, profundidade e heterogeneidade das evidências disponíveis, favorecendo a identificação de padrões conceituais, contradições e lacunas de conhecimento em campos emergentes. A revisão seguiu as cinco etapas definidas pelas autoras: identificação do





problema, busca sistemática da literatura, avaliação crítica dos estudos, extração e categorização dos dados, e síntese integrativa dos achados.

A estrutura metodológica foi complementada pelas recomendações de Toronto e Remington (2020), que enfatizam a necessidade de transparência nas estratégias de busca, reprodutibilidade dos critérios de seleção e rigor na avaliação da qualidade metodológica. Também foram consideradas as diretrizes brasileiras de revisão integrativa propostas por Souza, Silva e Carvalho (2010), que orientam o processo de elaboração de perguntas de pesquisa, definição de critérios de inclusão e organização dos dados extraídos. Essas três referências constituíram o arcabouço metodológico adotado e são amplamente reconhecidas na literatura internacional sobre métodos de síntese de evidências.

2.1 Pergunta de pesquisa e modelo conceitual

A pergunta orientadora foi elaborada segundo o modelo PICO (Population–Interest–Context), amplamente utilizado em revisões que incorporam estudos qualitativos conforme quadro 1. O uso do PICO é recomendado para sínteses integrativas, especialmente quando se pretende investigar fenômenos complexos e sociotécnicos (Toronto; Remington, 2020).

Quadro 1 – Estratégia PICO.

Pergunta norteadora:	<i>Como a inteligência artificial generativa tem influenciado as práticas de avaliação da aprendizagem no ensino superior, considerando dimensões éticas, pedagógicas e de equidade?</i>
P (População):	Estudantes e docentes do ensino superior
I (Interesse):	Uso de inteligência artificial generativa em processos de avaliação
Co (Contexto):	Ensino universitário presencial, híbrido ou a distância

Fonte: Autores, 2026.





2.2 Estratégia de busca

A busca foi realizada entre dezembro de 2025 e janeiro de 2026 nas seguintes bases: Scopus, Web of Science, PubMed, SciELO e LILACS, por suas relevâncias para publicação de pesquisas. Foram empregados termos controlados, palavras-chave e combinações booleanas adaptadas a cada base conforme quadro 2.

Quadro 2 – Estratégias de busca utilizadas.

Bases de dados	Estratégias de buscas
Scopus / Web of Science	("generative AI" OR "large language models" OR "ChatGPT") AND ("higher education" OR universit*) AND ("assessment" OR "evaluation" OR "academic integrity")
SciELO/LILACS	"inteligência artificial generativa" AND "ensino superior" AND "avaliação"

Fonte: Autores, 2026.

A busca considerou estudos publicados entre 2018 e 2025, período correspondente à expansão internacional dos modelos de linguagem de grande escala e ao início dos debates sobre sua incorporação no ensino superior.

2.3 Critérios de inclusão e exclusão

Foram incluídos: Estudos empíricos qualitativos, quantitativos ou de métodos mistos, ensaios teóricos, revisões narrativas e análises críticas, publicações em português, espanhol ou inglês, artigos que abordassem diretamente IA generativa aplicada a práticas de avaliação no ensino superior e estudos publicados em periódicos revisados por pares.





Foram excluídos: Pesquisas que abordassem apenas IA tradicional (machine learning, learning analytics) sem foco em IA generativa, artigos sobre ensino básico ou técnico, documentos institucionais sem revisão por pares, estudos sobre detecção de plágio que não discutissem avaliação e artigos sem acesso ao texto completo.

2.4 Processo de seleção

Dois revisores independentes realizaram a triagem dos títulos e resumos. Os estudos potencialmente elegíveis foram lidos na íntegra. Divergências foram resolvidas mediante consenso. Esse procedimento é recomendado para revisões integrativas a fim de reduzir vieses de seleção (Whittemore; Knafl, 2005).

2.5 Avaliação da qualidade metodológica

A qualidade dos estudos incluídos foi avaliada utilizando dois instrumentos, conforme a natureza metodológica de cada publicação: CASP (Critical Appraisal Skills Programme) para estudos qualitativos e MMAT (Mixed Methods Appraisal Tool) para estudos de métodos mistos e quantitativos. O uso combinado desses instrumentos é apoiado pela literatura, pois permite análise crítica sistemática preservando a heterogeneidade das fontes (Toronto; Remington, 2020).

2.6 Extração e síntese dos dados

Os dados extraídos incluíram: objetivo do estudo, método, participantes, características da intervenção ou fenômeno analisado, principais resultados e implicações para avaliação no ensino superior. Após a extração, aplicou-se uma síntese integrativa, agrupando os achados em categorias temáticas que refletiram convergências conceituais e tensões presentes nos estudos. A síntese priorizou interpretação crítica e articulação teórica, conforme o modelo de integração proposto por Whittemore e Knafl (2005).





3. RESULTADOS

A busca realizada nas bases Scopus, Web of Science, PubMed, SciELO e LILACS resultou em 978 registros. Após a remoção de duplicatas 831, foram triados títulos e resumos, dos quais 181 estudos avançaram para leitura na íntegra. Ao final do processo, 14 estudos atenderam aos critérios de inclusão estabelecidos para esta revisão integrativa, constituindo o corpus final analisado.

Os estudos incluídos abarcam o período 2023 a 2025, refletindo a rápida expansão da inteligência artificial generativa (IAG) após o lançamento dos grandes modelos de linguagem (LLMs). Em termos geográficos, há predominância de publicações oriundas da Europa, América do Norte, Oceania e Ásia, com poucos estudos latino-americanos, o que revela assimetria regional na produção científica sobre IAG aplicada à avaliação universitária.

No que diz respeito ao delineamento metodológico, os estudos apresentam diversidade: artigos empíricos qualitativos, pesquisas de percepção com docentes e estudantes, estudos de métodos mistos, análises de desempenho da IAG em avaliações reais e revisões narrativas. Essa heterogeneidade é compatível com a natureza sociotécnica do fenômeno e com o método integrativo adotado.

A síntese dos dados permitiu identificar três eixos temáticos principais:

1. Transformações nas práticas avaliativas

A IAG altera a natureza das tarefas avaliativas, especialmente as baseadas em escrita, síntese e argumentação. Estudos como os de Cotton, Cotton & Shipway (2023) e Gamage et al. (2023) mostram que avaliações tradicionais se tornam vulneráveis ao uso de IAG, exigindo redesign pautado em processos, interação síncrona e justificativas reflexivas dos estudantes.

2. Tensões éticas e integridade acadêmica

A preocupação com honestidade acadêmica aparece como um dos temas mais recorrentes. Pesquisas como Rudolph, Tan & Tan (2023), Evangelista (2025) e Susnjak &





McIntosh (2024) discutem como a IA desafia noções de autoria, autenticidade e justiça, ampliando a complexidade das políticas de integridade.

3. Equidade e assimetrias sociotécnicas

Estudos qualitativos, como Finkel-Gates (2025) e Vargas-Murillo, Pari-Bedoya & Guevara-Soto (2023), indicam que diferenças em letramento digital, proficiência linguística e acesso a versões pagas de sistemas generativos podem criar vantagens ou desvantagens desiguais entre estudantes, afetando a justiça avaliativa.

A seguir, apresenta-se o quadro 3 contendo as características principais dos estudos incluídos.

Quadro 3 - Síntese dos estudos incluídos na revisão integrativa.

ID	Autor/Título	Ano	País/Região	Tipo de Estudo	Foco Principal
1	Cotton, d.; Cotton, p.; Shipway, J. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT.	2023	Reino Unido	Ensaio teórico-analítico	Integridade acadêmica e vulnerabilidade das avaliações
2	Evangelista, E. Ensuring academic integrity in the age of ChatGPT.	2025	Singapura	Revisão crítica	Políticas institucionais e integridade
3	Gamage, K. et al. ChatGPT and higher education assessments.	2023	Austrália	Estudo empírico (percepções)	Oportunidades e riscos nas avaliações
4	Rudolph, j.; Tan, s.; Tan, S. ChatGPT: The end of traditional assessments?	2023	Austrália	Estudo analítico	Limites da avaliação tradicional
5	Susnjak, t.; McIntosh, T. ChatGPT: The end of online exam integrity?	2024	Austrália	Ensaio teórico e técnico	Riscos em avaliações on-line
6	Finkel-Gates, A. ChatGPT in academic assessments.	2025	EUA	Estudo qualitativo	Percepções estudantis sobre integridade
7	Moorhouse, B.; Yeo, M.; Wan, Y. Generative AI tools and assessment.	2023	Hong Kong	Revisão e análise de políticas	Diretrizes universitárias





8	Newton, P.; Xiromeriti, M. ChatGPT performance in MCQ exams.	2024	Reino Unido	Scoping review	Desempenho da IA em exames
9	Lo, C. What is the impact of ChatGPT on education?	2023	Taiwan	Revisão narrativa	Impactos gerais no ensino
10	Meyer, J.; Urbanowicz, R.; Martin, P. ChatGPT in academia.	2023	EUA	Revisão analítica	Uso acadêmico e desafios
11	Vargas-Murillo, A.; Pari-Bedoya, I.; Guevara-Soto, F. AI-assisted learning.	2023	Peru	Revisão sistemática	Uso de ChatGPT e desafios pedagógicos
12	Tillmanns, T. et al. Mapping GenAI in HE.	2025	Global	Revisão sistemática	Espaços e usos emergentes
13	Pereira, C.; Santos, R.; Fonseca, C. Higher education teachers' perceptions of AI.	2024	Portugal	Estudo empírico	Percepções docentes sobre IA
14	Kim, J.; Kadin, N.; Kim, J. Impact of GenAI tools on assessment.	2023	Coreia do Sul	Estudo teórico-analítico	Riscos e potencialidades

Fonte: Autores, 2026.

A análise integrada dos 14 estudos evidenciou que a inteligência artificial generativa: desestabiliza modelos tradicionais de avaliação baseados em escrita e produtos finais, introduz desafios concretos para integridade acadêmica, exigindo novos critérios e políticas transparentes, gera tensões pedagógicas, especialmente sobre dependência cognitiva, autoria e aprendizagem significativa, acentua assimetrias relacionadas ao letramento digital e ao acesso diferenciado a ferramentas de IA, exige reorganização institucional, incluindo formação docente, revisão curricular e políticas regulatórias.

4. DISCUSSÃO

A análise integrativa dos 14 estudos selecionados permitiu identificar três eixos centrais que estruturam os debates contemporâneos sobre a utilização da inteligência artificial generativa (IAg) nas práticas avaliativas do ensino superior. Esses eixos revelam tensões conceituais, desafios pedagógicos e dilemas éticos que atravessam o momento atual de transição tecnológica nas universidades, marcando um campo ainda em consolidação teórica e metodológica.

Revista *OWL Journal*, Campina Grande - PB, v.4.n.1. jan/fev/mar. 2026 - ISSN 2965-2634

A Revista *OWL Journal* está licenciada com uma **Licença Creative Commons Atribuição (CC BY)**





Eixo 1 — Transformações nas práticas avaliativas e deslocamento da autoria

Os estudos que compõem este eixo evidenciam uma mudança profunda nas formas de conceber, produzir e avaliar o conhecimento acadêmico. Gamage *et al.* (2023) argumentam que a IAg modifica a natureza das tarefas avaliativas ao permitir que estudantes gerem respostas articuladas e textos de alta qualidade com mínima intervenção humana, desafiando modelos avaliativos centrados apenas no produto final. Da mesma forma, Rudolph *et al.* (2023) analisam como avaliações tradicionais baseadas em ensaios e questões dissertativas tornam-se vulneráveis ao uso não declarado da IAg, revelando fragilidades estruturais na dependência exclusiva de evidências textuais para verificar aprendizagem.

Esse deslocamento também tensiona a noção de autoria acadêmica. Cotton *et al.* (2023) mostram que a IAg torna a autoria menos rastreável e mais distribuída, dissolvendo fronteiras entre contribuição humana e geração algorítmica. Tal fenômeno exige que a avaliação passe a contemplar processos, justificativas e explicitação do raciocínio, em vez de se limitar a analisar o texto final entregue pelo estudante. Meyer *et al.* (2023) reforçam que essa transformação demanda novas práticas pedagógicas, pois a simples substituição de tarefas não resolve o problema é necessário repensar o papel da escrita, da argumentação e da criatividade na formação universitária.

Em síntese, os estudos analisados apontam que a avaliação precisa deslocar-se de uma lógica produtivista para uma lógica processual, reflexiva e situada, na qual o uso responsável da IAg seja reconhecido, discutido e incorporado pedagogicamente.

Eixo 2 — Tensões éticas, integridade acadêmica e confiabilidade avaliativa

O segundo eixo concentra questões éticas relacionadas à integridade acadêmica, à autenticidade das produções e à confiabilidade dos sistemas avaliativos. Evangelista (2025) discute como a IAg reabre debates sobre desonestidade acadêmica, apontando que mecanismos de detecção são falhos e que políticas proibitivas podem criar ambientes de





vigilância contraproducentes. Susnjak e McIntosh (2024) apresentam evidências de que avaliações on-line tornam-se particularmente vulneráveis ao uso indevido da IAg, exigindo redesenho estrutural das tarefas e maior alinhamento entre objetivos formativos e métodos avaliativos.

O pano de fundo ético é ampliado por Newton e Xiromeriti (2024), cujo estudo demonstra que a IAg pode, em alguns contextos, superar estudantes humanos em tarefas objetivas, levantando dúvidas sobre a capacidade dos exames tradicionais de distinguir desempenho autêntico. Moorhouse *et al.* (2023) analisam políticas institucionais de universidades globais e mostram que há grande heterogeneidade nas normas de uso da IAg, o que compromete a clareza e a previsibilidade das expectativas avaliativas entre estudantes.

Assim, a literatura sugere que a integridade acadêmica na era da IAg não pode depender exclusivamente de estratégias de detecção ou punição. É necessário construir políticas claras, transparentes e dialogadas, que reconheçam tanto os riscos quanto os potenciais pedagógicos dessas tecnologias.

Eixo 3 — Equidade, desigualdades sociotécnicas e justiça avaliativa

O terceiro eixo aborda desigualdades de acesso, diferenças no letramento digital e assimetrias geradas pela incorporação da IAg. Finkel-Gates (2025) mostra que estudantes com maior experiência tecnológica e acesso a versões pagas ou mais avançadas das ferramentas possuem vantagens substanciais na realização de tarefas avaliativas. De forma semelhante, Vargas-Murillo *et al.* (2023) identificam que a utilização da IAg pode amplificar desigualdades existentes, ao transformar competências tecnológicas prévias em fatores decisivos para o desempenho acadêmico.

A dimensão da equidade aparece, portanto, como eixo crítico, mas ainda marginalmente explorado. Os estudos analisados reconhecem diferenças sociotécnicas, mas poucos desenvolvem análises interseccionais que integrem variáveis como classe, raça, gênero e condições linguísticas de estudantes multilíngues. Ao mesmo tempo, políticas





institucionais de uso da IAg parecem ignorar essas desigualdades, criando ambientes avaliativos em que estudantes com menos capital digital podem ser desproporcionalmente prejudicados.

Esse conjunto de evidências mostra que, para além dos dilemas éticos e técnicos, a IAg introduz novos mecanismos de estratificação no ensino superior, exigindo respostas institucionais que articulem avaliação, inclusão e justiça educacional.

Lacunas da literatura

Apesar da expansão acelerada das pesquisas sobre IAg no ensino superior, a revisão revela lacunas significativas. Há escassez de estudos empíricos robustos que avaliem, na prática, como o uso da IAg altera processos avaliativos e resultados de aprendizagem. Quase não existem estudos longitudinais capazes de capturar efeitos duradouros do uso da IAg sobre autonomia, autoria e equidade. Além disso, a dimensão da justiça social permanece pouco explorada, com raras análises interseccionais ou estudos em contextos do Sul Global. Por fim, evidencia-se uma falta de investigações sobre modelos de avaliação formativa mediados por IAg, foco necessário para superar abordagens estritamente punitivas ou tecnocráticas. Essas lacunas apontam para a urgência de pesquisas que articulem pedagogia, tecnologia e equidade de forma integrada e teoricamente robusta.

5. CONCLUSÃO

O objetivo desta revisão integrativa foi analisar criticamente como a inteligência artificial generativa tem influenciado as práticas de avaliação da aprendizagem no ensino superior, considerando dimensões éticas, pedagógicas e de equidade. Com base nos 14 estudos selecionados, conclui-se que esse objetivo foi plenamente alcançado. A síntese integrativa revela um campo em transição, no qual a IAg desestabiliza pressupostos tradicionais da avaliação e obriga instituições, docentes e estudantes a renegociar significados de autoria, autenticidade e integridade acadêmica.

Revista *OWL Journal*, Campina Grande - PB, v.4.n.1. jan/fev/mar. 2026 - ISSN 2965-2634

A Revista *OWL Journal* está licenciada com uma Licença Creative Commons Atribuição (CC BY)





A análise dos eixos temáticos evidencia que a IAg oferece, simultaneamente, oportunidades pedagógicas e riscos éticos. Tais tensões emergem de modo mais agudo quando consideradas as desigualdades sociotécnicas que atravessam o ensino superior contemporâneo. O potencial da IAg para apoiar avaliação formativa e aprendizagem autorregulada contrasta com a possibilidade de produzir vantagens desiguais, obscurecer autoria e fragilizar confiança no processo avaliativo. Esses desafios reforçam que a adoção responsável da IAg requer políticas institucionais claras, formação docente continuada e diálogo permanente com estudantes.

Por fim, a revisão mostra que o campo ainda carece de pesquisas empíricas robustas, análises contextuais sensíveis e marcos teóricos que articulem avaliação, tecnologia e justiça educacional. A consolidação de uma agenda crítica exige, portanto, estudos que ultrapassem perspectivas normativas ou tecnicistas e que incorporem abordagens participativas, interdisciplinares e comprometidas com a equidade no ensino superior.

REFERÊNCIAS

BAIG, Maria Ijaz; YADEGARIDEHKORDI, Elaheh. ChatGPT in the higher education: A systematic literature review and research challenges. *International Journal of Educational Research*, v. 127, 102411, 2024. DOI: 10.1016/j.ijer.2024.102411. Disponível em: <https://doi.org/10.1016/j.ijer.2024.102411>.

COTTON, D. R. E.; COTTON, P. A.; SHIPWAY, J. R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, v. 61, n. 2, p. 228-239, 2023. DOI: 10.1080/14703297.2023.2190148. Disponível em: <https://doi.org/10.1080/14703297.2023.2190148>.

EVANGELISTA, Edmund De Leon. Ensuring academic integrity in the age of ChatGPT: Rethinking exam design, assessment strategies, and ethical AI policies in higher education. *Contemporary Educational Technology*, v. 17, n. 1, ep559, 2025. DOI: 10.30935/cedtech/15775. Disponível em: <https://doi.org/10.30935/cedtech/15775>.





FINKEL-GATES, Adam. ChatGPT in academic assessments: Upholding integrity. *Journal of Learning Development in Higher Education*, Issue 36, 2025. Disponível em: <https://files.eric.ed.gov/fulltext/EJ1480557.pdf>.

GAMAGE, K. A. A. et al. ChatGPT and higher education assessments: More opportunities than concerns? *Journal of Applied Learning and Teaching*, v. 6, n. 2, p. 358-369, 2023. DOI: 10.37074/jalt.2023.6.2.32. Disponível em: <https://doi.org/10.37074/jalt.2023.6.2.32>

HON, Kelly. Generative AI in higher education: A systematic review of opportunities and challenges. *Teaching Public Administration*, 2025 (ahead of print). DOI: 10.1177/00472395251400089. Disponível em: <https://doi.org/10.1177/00472395251400089>

JÚNIOR, A. C. F. C.; SANTOS, C. A. G.; ALVES, L. F. Inteligência artificial e o futuro do trabalho: implicações jurídicas. *Revista owl (owl journal) - REVISTA INTERDISCIPLINAR DE ENSINO E EDUCAÇÃO*, [s. l.], v. 2, n. 4, p. 176–178, 2024. DOI: 10.5281/zenodo.12796061. Disponível em: <https://revistaowl.com.br/index.php/owl/article/view/261>. Acesso em: 5 jan. 2026.

LO, Chih-Ming. What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, v. 13, n. 4, 410, 2023. DOI: 10.3390/educsci13040410. Disponível em: <https://doi.org/10.3390/educsci13040410>.

MEYER, Jason G.; URBANOWICZ, Ryan J.; MARTIN, Philip C. N. ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*, v. 16, 20, 2023. DOI: 10.1186/s13040-023-00339-9. Disponível em: <https://doi.org/10.1186/s13040-023-00339-9>

MOORHOUSE, Benjamin L.; YEO, Michelle A.; WAN, Yin. Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, v. 5, 100151, 2023. DOI: 10.1016/j.caeo.2023.100151. Disponível em: <https://doi.org/10.1016/j.caeo.2023.100151>.

OLIVEIRA, R. P. de. A convergência da arte e da tecnologia: o uso da inteligência artificial na produção de histórias em quadrinhos para o ensino de engenharia de produção. *Revista OWL (OWL Journal) - REVISTA INTERDISCIPLINAR DE ENSINO E EDUCAÇÃO*, [S. l.], v. 3, n. 5, p. 1–29, 2025. DOI: 10.5281/zenodo.17247417. Disponível em: <https://revistaowl.com.br/index.php/owl/article/view/442>. Acesso em: 5 jan. 2026.





NEWTON, Philip; XIROMERITI, Maria. ChatGPT performance on multiple choice question examinations in higher education: A pragmatic scoping review. *Assessment & Evaluation in Higher Education*, v. 49, n. 6, p. 781-798, 2024. DOI: 10.1080/02602938.2023.2299059. Disponível em: <https://doi.org/10.1080/02602938.2023.2299059>.

RUDOLPH, Jenny; TAN, Shu; TAN, Siew. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, v. 6, n. 1, p. 342-363, 2023. DOI: 10.37074/jalt.2023.6.1.9. Disponível em: <https://doi.org/10.37074/jalt.2023.6.1.9>.

SOUZA, Marcela Tavares; SILVA, Michelly Dias Fernandes; CARVALHO, Rosinete de. Revisão integrativa: o que é e como fazer. *Einstein*, v. 8, n. 1, p. 102–106, 2010. DOI: 10.1590/S1679-45082010RW1134. Disponível em: <https://doi.org/10.1590/S1679-45082010RW1134>

SULLIVAN, Michael; KELLY, Aisling; MCLAUGHLAN, Paul. ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, v. 6, n. 1, 2023. DOI: 10.37074/jalt.2023.6.1.17. Disponível em: <https://doi.org/10.37074/jalt.2023.6.1.17>.

SUSNJAK, Tina; MCINTOSH, Timothy R. ChatGPT: The end of online exam integrity? *Education Sciences*, v. 14, n. 6, 656, 2024. DOI: 10.3390/educsci14060656. Disponível em: <https://doi.org/10.3390/educsci14060656>.

TILLMANN, Tanja et al. Mapping tomorrow's teaching and learning spaces: A systematic review on GenAI in higher education. *Trends in Higher Education*, v. 4, n. 1, 2, 2025. DOI: 10.3390/higheredu4010002. Disponível em: <https://doi.org/10.3390/higheredu4010002>.

TORONTO, Coleen; REMINGTON, Roger. A step-by-step guide to conducting an integrative review. *Nursing Research*, v. 69, n. 1, p. 47–53, 2020. DOI: 10.1097/NNR.0000000000000411. Disponível em: <https://doi.org/10.1097/NNR.0000000000000411>

VARGAS-MURILLO, Andrés; PARI-BEDOYA, Italo; GUEVARA-SOTO, Fernando. Challenges and opportunities of AI-assisted learning: A systematic literature review on the impact of ChatGPT usage in higher education. *International Journal of Learning, Teaching and Educational Research*, v. 22, n. 7, p. 122-135, 2023. DOI: 10.26803/ijlter.22.7.7. Disponível em: <https://doi.org/10.26803/ijlter.22.7.7>



REVISTA OWL (*OWL Journal*)

www.revistaowl.com.br – ISSN: 2965-2634



WHITTEMORE, Robin; KNAFL, Kathleen. The integrative review: updated methodology. *Journal of Advanced Nursing*, v. 52, n. 5, p. 546–553, 2005. DOI: 10.1111/j.1365-2648.2005.03621.x. Disponível em: <https://doi.org/10.1111/j.1365-2648.2005.03621.x>

Recebido em: 07/01/2026

Aprovado em: 17/01/2026

Publicado em: 04/02/2026

Revista *OWL Journal*, Campina Grande - PB, v.4.n.1. jan/fev/mar. 2026 - ISSN 2965-2634

A Revista *OWL Journal* está licenciada com uma Licença Creative Commons Atribuição (CC BY)

