



**INTELIGÊNCIA ARTIFICIAL NA EDUCAÇÃO:
AUTORIA, INTEGRIDADE ACADÊMICA E AVALIAÇÃO
EM XEQUE**

*ARTIFICIAL INTELLIGENCE IN EDUCATION:
AUTHORSHIP, ACADEMIC INTEGRITY AND ASSESSMENT
UNDER THREAT*

Jean Piton-Gonçalves¹

Universidade Federal de São Carlos – UFSCar

RESUMO

A Inteligência Artificial Generativa (IAGen) instalou-se no cotidiano acadêmico antes de qualquer resposta institucional ou docente consistente. Este ensaio teórico, fundamentado em revisão narrativa de literatura recente sobre IA na Educação (IAEd), argumenta que a IAGen não representa apenas uma nova ferramenta pedagógica, ela compromete, de forma simultânea e estrutural, três pressupostos sobre os quais a educação formal está organizada. O primeiro pressuposto é que o texto escrito é evidência de pensamento próprio; o segundo, que aprender exige esforço cognitivo do aprendiz; e o terceiro, que o avaliador é capaz de distinguir o que o estudante sabe do que ele produziu. Os três pressupostos são examinados nas dimensões conceitual, empírica, epistemológica e ética, a partir de marcos normativos internacionais e nacionais e de evidências sobre autoria, integridade acadêmica e atrofia cognitiva. Conclui-se que a incorporação acrítica da IAGen, sem que o campo nomeie o que está sendo comprometido, é mais perigosa do que a ferramenta em si.

¹ Doutor em Ciências da Computação e Matemática Computacional pela Universidade de São Paulo (USP). Professor Associado da Universidade Federal de São Carlos (UFSCar). ORCID: <https://orcid.org/0000-0002-7392-2001>. E-mail: jpiton@ufscar.br

Palavras-chave: Inteligência Artificial na Educação. Inteligência Artificial Generativa. Autoria na Inteligência Artificial. Produção Acadêmica.

ABSTRACT

Generative Artificial Intelligence (GenAI) has become embedded in everyday academic practice before any consistent institutional or pedagogical response has been established. This theoretical essay, grounded in a narrative review of recent literature on Artificial Intelligence in Education (AIED), argues that GenAI does not merely represent a new pedagogical tool; rather, it simultaneously and structurally compromises three assumptions upon which formal education is organized. The first assumption is that written text constitutes evidence of one's own thinking; the second is that learning requires the learner's cognitive effort; and the third is that the evaluator is able to distinguish what the student knows from what the student has produced. These three assumptions are examined across conceptual, empirical, epistemological, and ethical dimensions, drawing on international and national normative frameworks as well as evidence concerning authorship, academic integrity, and cognitive atrophy. It is concluded that the uncritical incorporation of GenAI, without the field explicitly naming what is being compromised, is more dangerous than the tool itself.

Keywords: Artificial Intelligence in Education. Generative Artificial Intelligence. Authorship in Artificial Intelligence. Academic Production.

1 INTRODUÇÃO

O presente artigo caracteriza-se como um ensaio teórico fundamentado em revisão narrativa de literatura recente sobre Inteligência Artificial na Educação (IAEd) (do inglês *Artificial Intelligence in Education*, AIED), com foco nas implicações para o trabalho docente. A seleção das fontes considerou três critérios principais: relevância temática para o debate sobre Inteligência Artificial (IA) (do inglês *Artificial Intelligence*, AI) e docência, atualidade das publicações (prioritariamente entre 2023 e 2026) e diversidade de perspectivas analíticas, incluindo estudos empíricos, revisões sistemáticas de literatura, levantamentos quantitativos e documentos normativos.

No âmbito nacional, destaca-se o Referencial para Desenvolvimento e Uso Responsáveis de Inteligência Artificial na Educação (Brasil, 2026), publicado pelo Ministério da Educação em fevereiro de 2026, sendo este o primeiro marco federal brasileiro dedicado ao tema, abrangendo todos os ciclos educacionais, da educação básica à pós-graduação. A análise seguiu uma abordagem interpretativa e comparativa, buscando identificar convergências e tensões na literatura sobre IAEd.

A revisão narrativa foi conduzida a partir de buscas exploratórias realizadas entre 2023 e 2026 nas bases SciELO-BRASIL, Google Scholar, revistas científicas da área de Educação e periódicos com extrato QUALIS CAPES, utilizando combinações de descritores como “*generative AI*”, “*artificial intelligence in education*” e “*academic integrity*”. O período selecionado corresponde à difusão massiva de modelos generativos de linguagem, com exceção de referências seminais anteriores mantidas por sua relevância conceitual. Cabe destacar que parte dos referenciais mobilizados são, eles próprios, revisões sistemáticas da literatura (como Long *et al.* (2026) e Gonçalves e Santos (2025)), de modo que o alcance empírico efetivo deste ensaio é mais amplo do que o número de referências diretas sugere.

O argumento central que orienta este ensaio é o de que a IAGen compromete, de forma simultânea e estrutural, três pressupostos sobre os quais a educação formal está organizada: (i) que o texto escrito é evidência de pensamento próprio, (ii) que aprender exige esforço cognitivo do aprendiz e (iii) que o avaliador é capaz de distinguir o que o estudante sabe do que ele produziu.

2 IA NA EDUCAÇÃO

Buscando compreender o impacto das ferramentas de IAEd, a literatura vem apontando uma tensão estrutural que é, ao mesmo tempo, tecnológica, pedagógica e ética. Ou seja, enquanto estudantes de diferentes níveis incorporam ferramentas de IA ao seu cotidiano acadêmico, a vasta maioria dos professores permanece sem orientação institucional, sem formação específica e, muitas vezes, sem posição definida sobre o “fenômeno” recente da IA (Farinosi; Melchior, 2025; Gentile *et al.*, 2023) generativa.

A revisão sistemática de Long *et al.* (2026), que analisou 73 estudos empíricos publicados entre 2015 e 2025 sobre IA no ensino superior, identificou que o engajamento cognitivo, comportamental e emocional dos estudantes é potencialmente fortalecido quando mediado pelas estratégias pedagógicas adotadas pelos docentes.

No contexto brasileiro, Gonçalves e Santos (2025), em revisão sistemática da literatura sobre *chatbots*² com IAEd, chegaram à conclusão que, embora as ferramentas ofereçam potencial real de personalização e suporte ao aprendizado,

² São sistemas de IA projetados para simular conversas com usuários por meio de linguagem natural.

sua eficácia depende diretamente da qualidade pedagógica com que são integradas pelo professor. Os referidos autores identificaram que instituições com infraestrutura robusta tendiam a relatar resultados mais positivos, enquanto contextos com recursos limitados enfrentavam obstáculos que precedem a própria adoção tecnológica (Gonçalves; Santos, 2025). Esse achado ressoa com o diagnóstico mais amplo da realidade brasileira, em que as condições materiais de acesso constituem variável tão determinante quanto a qualidade da mediação docente.

3 IA E IA GENERATIVA: DISTINÇÕES CONCEITUAIS PARA A EDUCAÇÃO

Até meados dos anos 2000, pesquisar um tema acadêmico exigia, por exemplo: (i) visita física à biblioteca e busca manual em fichários de catálogo; (ii) cópias xerográficas de páginas de livros como único meio portátil de registro e (iii) cálculos realizados à mão ou com calculadoras científicas de uso individual. Menos de duas décadas depois, esse universo foi reconfigurado, uma vez que buscadores, bases de dados *online* e *softwares* de cálculo simbólico e numérico transformaram, de forma estrutural, as formas de pesquisar, calcular e escrever.

A IA Generativa (IAGen) é um subcampo da IA capaz de produzir conteúdo novo e coerente a partir de padrões estatísticos “aprendidos” em grandes volumes de dados gerados por humanos (Cheng; Calhoun; Reedy, 2025). Sua expressão mais relevante no contexto acadêmico são os Grandes Modelos de Linguagem (em inglês *Large Language Models*, LLMs), como o ChatGPT³, que geram respostas linguisticamente adequadas a partir de associações entre sequências textuais, sem mecanismos internos de verificação de fatos ou rastreamento de fontes (Cheng; Calhoun; Reedy, 2025).

A IAGen representa uma “ruptura tecnológica” em que a ferramenta não apenas busca ou calcula, mas ela produz conteúdo. O lançamento do ChatGPT em novembro de 2022 representou uma inflexão histórica, não porque a IAGen fosse tecnicamente inédita, mas porque, pela primeira vez, uma ferramenta dessa natureza tornou-se acessível a qualquer usuário, sem custo inicial, sem instalação e sem conhecimento técnico prévio, de modo que a adoção se deu de forma veloz e em larga escala, anterior a qualquer resposta institucional ou docente.

³ É um *chatbot* desenvolvido pela *OpenAI*, capaz de gerar textos e responder perguntas em linguagem natural a partir de LLMs treinados com extensos conjuntos de dados textuais.

O termo IA é frequentemente utilizado de forma ampla para designar diferentes tipos de sistemas computacionais capazes de executar tarefas associadas à inteligência humana, como reconhecimento de padrões, tomada de decisão ou análise de dados. Contudo, essa categoria abrange uma diversidade de aplicações tecnológicas que operam em contextos bastante distintos. Em muitos setores, a IA está associada principalmente a processos de automação e otimização.

Sistemas de IA são amplamente utilizados na indústria, em logística, em sistemas de recomendação ou em processos de análise preditiva, nos quais algoritmos são empregados para classificar dados, detectar padrões ou automatizar decisões. Nessas aplicações, a interação direta com usuários humanos nem sempre é central, e os sistemas operam frequentemente em ambientes técnicos ou industriais.

No contexto educacional contemporâneo, grande parte das discussões e preocupações referem-se à IAGen, que é capaz de produzir textos, imagens, códigos e outros conteúdos em resposta a solicitações dos usuários. Ferramentas desse tipo permitem que o usuário formule perguntas em linguagem natural e obtenha respostas relativamente complexas, incluindo explicações conceituais, resoluções de exercícios ou redações completas, por exemplo.

A diferença entre IA em sentido amplo e IAGen é particularmente relevante para a Educação. Enquanto muitas aplicações tradicionais de IA operam em segundo plano em plataformas digitais⁴, os modelos generativos atuam diretamente na produção de conteúdo, que pode ser utilizado em atividades escolares e acadêmicas. Isso introduz novas questões relacionadas à autoria, à confiabilidade das informações produzidas e ao papel dessas ferramentas nos processos de ensino, aprendizagem e avaliação. Nesse sentido, os desafios educacionais associados à IA decorrem principalmente da capacidade dos sistemas generativos de produzir respostas textuais que simulam explicações e o comportamento humano. É essa característica que deve colocar a IA no centro das discussões atuais da Educação.

⁴ Exemplos: serviços bancários para detecção de fraudes e análise de crédito; serviços de *streaming*, como *Netflix* e *Spotify*, que empregam algoritmos de recomendação de conteúdo; grandes infraestruturas de computação e *data centers* de empresas como *Google*, *Amazon* e *Microsoft*, que utilizam IA para otimização de tráfego, armazenamento e gerenciamento de recursos computacionais, dentre outros.

O que distingue qualitativamente a IA das Tecnologias da Informação e Comunicação (TICs) anteriores é a combinação de três atributos:

1. *Estocasticidade*: os modelos generativos não determinam uma resposta única a partir de uma entrada, mas amostram sequências de *tokens* (que são os blocos de processamento), de modo que as mesmas entradas (do inglês *inputs*) podem produzir saídas (do inglês *outputs*) distintos a cada execução. Essa variabilidade não é ruído eliminável, mas uma propriedade estrutural dos modelos, com implicação direta para a Educação. Em suma, o mesmo enunciado de um exercício, por exemplo, submetido à IAGen em momentos diferentes, geralmente gera resoluções diferentes.
2. *Agentividade*: em que a própria ferramenta toma a decisão do que apresentar ao usuário. É a capacidade do sistema de executar ações ou tomar decisões de forma total ou parcialmente autônoma.
3. *Bidirecionalidade*: em que a interação é dinâmica e não linear. No contexto da interação humano-computador e da comunicação, o ciclo interativo constrói conteúdo.

No que se refere a uma categorização da IAEd, Long *et al.* (2026) propuseram uma classificação organizada em cinco categorias a saber:

1. *Aprendizagem personalizada e sistemas adaptativos*: utilizam algoritmos de IA para analisar dados comportamentais e cognitivos dos estudantes e recomendar trajetórias de aprendizagem individualizadas. Esses sistemas ajustam conteúdos, ritmo e nível de dificuldade conforme o desempenho do aluno, aumentando motivação e participação ativa. Essa abordagem não é nova, como mostra o trabalho de Piton-Gonçalves *et al.* (2004), que desenvolveram um Sistema Tutor Inteligente para a aprendizagem do inglês instrumental.
2. *IA Generativa*: envolve o uso de modelos generativos para apoiar a interação em sala de aula e fornecer retroalimentação imediata. Essas ferramentas podem ser utilizadas em simulações, atividades baseadas em estudo de casos ou diálogos tutorados, estimulando o pensamento crítico, a autonomia do estudante e a participação nas atividades acadêmicas.

3. *Gamificação e tecnologias imersivas*: combina elementos de jogos com algoritmos de IA capazes de adaptar as atividades ao desempenho dos estudantes. Ambientes imersivos, como realidade virtual e realidade aumentada, ampliam esse processo ao permitir experiências de aprendizagem baseadas em simulação.
4. *Acesso imediato a recursos e suplementação de conteúdos*: ferramentas de IA podem facilitar o acesso rápido a informações, resumos e materiais complementares, especialmente em ambientes digitais ou híbridos. Tais ferramentas permitem recuperação instantânea de conteúdos, apoio ao estudo assíncrono e interação entre pares em plataformas educacionais, reduzindo barreiras cognitivas e tecnológicas no processo de aprendizagem.
5. *Adaptação contextual em cenários educacionais específicos*: a IA também tem sido aplicada em contextos de ensino remoto, educação especial ou situações de aprendizagem flexível. Estratégias como microaprendizagem, vídeos curtos e ferramentas adaptativas podem apoiar diferentes perfis de estudantes e necessidades educacionais.

A classificação de Long *et al.* (2026) organiza a IA por função pedagógica, o que permite ao estudante selecionar ferramentas a partir das necessidades da aprendizagem, e não a partir das características técnicas dos sistemas. Nesse sentido, é importante que o professor conheça tais classificações para as tomadas de decisão quanto ao uso pedagógico das ferramentas.

4 A IAGEN COMO FENÔMENO DE MASSA

A dimensão empírica desse cenário é ilustrada pelos dados do trabalho de Farinosi e Melchior (2025), em levantamento realizado com 1.366 estudantes universitários italianos distribuídos em 24 instituições⁵, aponta que 69,2% declararam ter utilizado IAGen para projetos pessoais e 38,7% para tarefas acadêmicas. No mesmo estudo, constatou-se que 85% dos professores não demonstraram o uso explícito em relação ao uso de IA em sala de aula. Em uma relação entre atitude e comportamento, o referido trabalho encontrou que 58,3% dos

⁵ Estudo conduzido entre outubro e dezembro de 2023.

respondentes consideravam o uso de IAGen em trabalhos acadêmicos moralmente errados, mas, ainda assim, uma parcela substancial o praticava.

Nesse sentido, como argumentam Gentile *et al.* (2023), a chegada da IA de uso massivo criou uma situação inédita, em que docentes precisam mediar uma tecnologia que seus alunos já dominam operacionalmente, sem que eles próprios tenham recebido qualquer formação para isso.

A rapidez dessa incorporação indica a necessidade de que o tema seja debatido de forma sistemática e contínua em universidades, instituições de ensino e pesquisa, secretarias de educação e escolas. Soma-se a isso o fato de que as ferramentas computacionais baseadas em IA encontram-se em permanente evolução, com ampliação constante de funcionalidades, surgimento frequente de novas aplicações, que expandem suas capacidades e formas de utilização.

No plano prático, a discussão deve ser institucional, uma vez que essas tecnologias já passaram a integrar as rotinas acadêmicas, escolares e administrativas. Nesse contexto, a questão central desloca-se para a definição de diretrizes institucionais e pedagógicas capazes de orientar o uso dessas ferramentas nos processos de ensino e aprendizagem.

O uso massivo de IAGen por professores e estudantes é um fenômeno de abrangência global e internacional. Nessa direção, torna-se importante a criação de referenciais e/ou regras para a utilização da IAGen. No plano federal brasileiro, o Ministério da Educação publicou em março de 2026 o *Referencial para Desenvolvimento e Uso Responsáveis de Inteligência Artificial na Educação* (Brasil, 2026), primeiro documento de política pública de abrangência nacional sobre o tema. O Referencial parte da premissa de que a adoção de IAEd somente é legítima quando subordinada ao projeto pedagógico das instituições e ao princípio constitucional do direito à educação, e propõe diretrizes para gestores, docentes e estudantes em todos os níveis de ensino (Brasil, 2026).

No cenário internacional apenas sete países no mundo contavam com marcos referenciais nacionais de competências em IA para professores (UNESCO, 2025), o que evidencia que a velocidade de adoção tecnológica superou, em muito, a capacidade das instituições de formar docentes para um uso crítico, ético e pedagógico das novas ferramentas.

Com o objetivo de preencher essa lacuna, a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) publicou, em 2025, o primeiro *Marco Referencial de Competências em IA para Professores* (MRCP-IA) de caráter universal, documento que será central para algumas das análises aqui desenvolvidas e está detalhado na Seção 5.

5 O MARCO REFERENCIAL DA UNESCO

Diante da constatação de que a maioria dos países necessitam de orientação sistemática para a formação docente em contextos de IA, o MRCP-IA (UNESCO, 2025) se constitui no primeiro marco de caráter universal do gênero, concebido não como prescrição obrigatória, mas como padrão de referência para orientar marcos nacionais e programas de formação continuada. Essencialmente, o MRCP-IA está organizado em uma matriz bidimensional, composta por cinco aspectos e três níveis de progressão, resultando em 15 blocos de competências.

Os cinco aspectos são: (i) mentalidade centrada no ser humano; (ii) ética da IA; (iii) fundamentos e aplicações de IA; (iv) pedagogia de IA; e (v) IA para o desenvolvimento profissional. Os três níveis seguem uma lógica de aprofundamento gradual, sendo: *Adquirir*, em que se desenvolvem a alfabetização básica e o uso crítico de ferramentas; *Aprofundar*, em que se dá a integração pedagógica com responsabilidade humana; e *Criar*, em que se espera uso inovador e contribuição para políticas e padrões éticos (UNESCO, 2025).

Um aspecto central da concepção do MRCP-IA é que a progressão entre os níveis não é linear nem homogênea. Por exemplo, um mesmo professor pode estar no nível *Criar* na dimensão de pedagogia e ainda no *Adquirir* na dimensão de ética, o que implica que programas de formação baseados no marco precisam partir de diagnósticos individualizados, e não de pacotes uniformes de capacitação (UNESCO, 2025).

O Referencial MEC (Brasil, 2026) pode ser lido como o complemento nacional a esse marco: enquanto o MRCP-IA oferece um padrão de referência universal, o documento brasileiro operacionaliza essa visão no contexto nacional, organizando as competências docentes em duas dimensões: (i) competências para o ensino, que abrangem letramentos digitais, segurança e privacidade e (ii) competências para o

ensino e aprendizagem mediadas por IA (Brasil, 2026). A convergência entre os dois documentos fortalece a validade do quadro analítico aqui adotado.

6 A IA NA VOZ DOS PROFESSORES E ESTUDANTES

A literatura sobre o que professores efetivamente realizam com IA em sala de aula ainda é escassa. A coletânea *La docencia en tiempos de IA* (Satorre Cuerda, 2024) documenta, em oito capítulos, experiências concretas de professores universitários espanhóis que integraram IA ao ensino, buscando cenários em que ela amplie (e não substitua) a atividade cognitiva do estudante. Portanto, pode-se dizer que a qualidade da integração da IA ao ensino não depende apenas do domínio técnico da ferramenta, mas da qualidade do “repertório pedagógico” do professor.

O relatório *Artificial Intelligence in Education: uses, opportunities, and risks in the Brazilian context* (UNESCO; Cetic.br, 2025) documenta como a IA tem sido utilizada em diferentes contextos educacionais no Brasil, destacando oportunidades de personalização do ensino, apoio aos professores em tarefas administrativas e ampliação do acesso a recursos digitais, sempre enfatizando a manutenção da mediação humana.

Nesse cenário de tensão entre potencial e resistência, o trabalho de Gonçalves e Santos (2025) documenta que, em escolas públicas urbanas brasileiras onde *chatbots* foram implementados, a resistência inicial dos professores decorreu, em grande parte, do temor de que a ferramenta viesse a substituir o papel docente. Os autores concluem que o treinamento continuado e a clareza sobre o papel complementar (e não substitutivo) do *chatbot* são condições necessárias para que a integração seja bem-sucedida (Gonçalves; Santos, 2025). Essa conclusão se alinha ao que o MRCP-IA prescreve como competência do nível *Aprofundar*, isto é, integrar IA ao ensino com responsabilidade humana, sem delegar ao sistema as decisões pedagógicas centrais.

Estudantes que utilizam, por exemplo, o *ChatGPT* cotidianamente podem não ter qualquer noção sobre como o modelo foi treinado, que vieses carrega ou por que produz alucinações, de modo que o analfabetismo funcional em IA produz dependências acríticas que, no longo prazo, são pedagógica e cognitivamente prejudiciais (Hassen, 2025; UNESCO, 2025). No que se refere aos professores:

(...) enquanto as gerações anteriores de ferramentas de TIC se concentravam mais em facilitar tarefas rotineiras, as ferramentas de IA tendem a substituir a tomada de decisões humanas, com base em previsões de padrões extraídos da análise de dados extensivos de exemplos passados. Consequentemente, existe o risco de que a dependência excessiva de IA possa levar à atrofia das competências essenciais dos professores. (UNESCO, 2025, p. 14).

Relativo aos impactos da IA na vida dos estudantes, resultados de Farinosi e Melchior (2025) mostram que 66,1% dos estudantes respondentes manifestaram alta preocupação social com os impactos da IA, ao passo que a preocupação pessoal com os riscos individuais era baixa, o que indica que os estudantes percebem o perigo no plano coletivo e abstrato, mas não o internalizam como risco para sua própria formação. Além disso, 81,4% defendiam alguma forma de normatização institucional e 81% eram contra a proibição total (Farinosi; Melchior, 2025), o que evidencia que os estudantes demandam orientação. Em contrapartida, os estudos supracitados mostram que os professores, em sua maioria, ainda não estão em posição de oferecer tal orientação.

Essa demanda por orientação institucional adquire contornos ainda mais complexos quando se examina a dimensão epistemológica do problema, pois a própria natureza técnica da IAGen impõe limites estruturais ao seu uso em contextos que exigem rigor certificatório.

No que se refere ao ensino de ciências exatas, por exemplo, a literatura aponta um risco adicional e específico, uma vez que a IAGen pode produzir passos de demonstração incorretos, erros matemáticos apresentados com confiança sintática impecável (Hassen, 2025). É exatamente esse modo de pensar que o uso acrítico de IA pode comprometer. Quando estudantes delegam a resolução de problemas à IA sem engajamento cognitivo, ocorre o que Hassen (2025) denomina *atrofia cognitiva*, isto é, a habilidade não exercitada enfraquece progressivamente e a dependência da ferramenta aumenta, de modo que, em ciências exatas, a atrofia em estágios iniciais produz um efeito cascata de longo prazo.

No plano epistemológico, o fenômeno das “alucinações da IAGen” contraria diretamente o princípio da verificabilidade que define as ciências formais (Hassen, 2025). Agrava esse quadro a estocasticidade dos modelos, visto que seu caráter aleatório e probabilístico impede previsões exatas, em que as mesmas entradas podem gerar saídas distintas em execuções diferentes, criando um cenário

estruturalmente problemático em contextos de demonstração formal, nos quais a validade lógica da prova é um requisito e não um ideal. Esse caráter estocástico também implica que, mesmo quando o usuário explicita integralmente os *prompts*⁶ utilizados para outro usuário, a reprodução dos resultados não é garantida, pois diferentes contas, contextos de interação ou históricos de uso (logs) podem influenciar o comportamento do sistema e produzir respostas distintas. Além disso, a chamada “caixa-preta” dos modelos de IA (UNESCO, 2025) torna impossível auditar um raciocínio computacional da forma como se audita uma prova matemática.

Essa preocupação é reconhecida institucionalmente pelo Referencial MEC (Brasil, 2026) que denomina essa dimensão como “Proteção da aprendizagem ativa e da autoria acadêmica frente à automação excessiva” (p. 54), alertando que a disponibilidade imediata de respostas prontas pode estimular a terceirização do esforço cognitivo e levar estudantes “a suprimir etapas fundamentais da aprendizagem, como o enfrentamento da complexidade, a resolução de ambiguidades e a construção ativa de sentidos” (Brasil, 2026, p. 15).

A análise de Long *et al.* (2026) oferece sustentação empírica a essa preocupação. Os autores identificaram que o uso acrítico do *ChatGPT* implicou na redução dos níveis de pensamento ativo em estudantes que priorizaram a conveniência sobre o engajamento cognitivo. Por outro lado, os sistemas de tutoria inteligente mostraram os maiores ganhos de engajamento cognitivo, embora apenas quando utilizados de forma consistente e sustentada.

7 RISCOS EPISTEMOLÓGICOS NA ACADEMIA

No plano da autoria e da integridade acadêmica, a IAGen coloca questões ainda sem resposta objetiva na literatura especializada. Monografias, dissertações e teses são, por definição legal, obras de autoria individual. Dessa forma, um texto gerado majoritariamente por máquina e entregue sob esse regime levanta questões jurídicas sobre a veracidade da autoria e configura, no plano ético, uma modalidade de plágio: não é o texto de outro ser humano que se apresenta como próprio, mas a produção intelectual de uma máquina (Farinosi; Melchior, 2025).

⁶ É a instrução textual fornecida pelo usuário a IAGen, que orienta a tarefa ou pergunta a ser respondida e condiciona o conteúdo da resposta produzida pelo sistema computacional.

A dificuldade prática na academia agrava o problema. Para um avaliador/examinador não especialista no tema, a distinção entre o que o autor sabe e o que a máquina gerou torna-se, prática e metodologicamente, impossível por meio da análise exclusiva do texto escrito.

No plano ético mais amplo, a tentativa de contornar processos avaliativos não constitui uma novidade histórica, pois práticas como cola, fraude e plágio sempre existiram nas instituições de ensino. A introdução da IA, contudo, altera a natureza desse fenômeno, ampliando sua escala, facilitando sua execução e tornando-o mais invisível. Nesse contexto, a contradição observada por Farinosi e Melchior (2025), em que estudantes reconhecem a imoralidade do uso acadêmico da IA, mas a utilizam mesmo assim, não pode ser interpretada como mera inconsistência ética, mas como um indicativo de lacunas institucionais na definição de normas claras de conduta acadêmica.

Nessa linha, Gonçalves e Santos (2025) sistematizam dois riscos estruturais dos *chatbots* com IA: o *viés algorítmico*, pelo qual modelos treinados em dados desequilibrados podem exprimir desigualdades e preconceitos implícitos; e as *questões de privacidade de dados*, nos quais estudantes (incluindo menores de idade) interagem com plataformas cujos termos de uso raramente são lidos ou compreendidos por seus responsáveis. Para Gonçalves e Santos (2025, p. 55) as “questões éticas, como a privacidade dos dados e o viés algorítmico, são cruciais para uma integração responsável da IA na educação”, o que requer supervisão humana contínua e auditoria dos algoritmos.

O Referencial do MEC endossa institucionalmente esse diagnóstico ao identificar, entre seus desafios estruturais, tanto a reprodução de preconceitos por sistemas treinados em bases de dados que refletem desigualdades históricas quanto à necessidade de proteção da privacidade e dos dados de estudantes e educadores com observância estrita da legislação vigente (Brasil, 2026).

Esse enquadramento em políticas públicas federais reforça a relevância institucional do argumento, uma vez que os riscos apontados pela literatura acadêmica internacional também passam a ser reconhecidos oficialmente no âmbito do Estado brasileiro.

Problemas de ordem jurídica e institucional tornam-se mais complexos quando se considera o impacto da IAGen sobre os fundamentos da autoria no texto

acadêmico. Como o texto científico também funciona como registro de atribuição intelectual, o uso de sistemas que geram conteúdo automaticamente dificulta distinguir o que corresponde ao pensamento do autor e o que foi produzido pela ferramenta, levantando questões sobre responsabilidade e integridade acadêmica.

8 AUTORIA E OS DESAFIOS ÉTICOS DO USO DA IAGEN

A disseminação de ferramentas de IAGen e, especificamente os Grandes Modelos de Linguagem (do inglês *Large Language Models*, LLMs), como o *ChatGPT*, introduziu uma ruptura nos pilares que sustentam a produção do conhecimento científico e acadêmico. Se no passado a integridade acadêmica pressupunha fronteiras relativamente claras entre autoria legítima, colaboração permitida e conduta indevida, a capacidade dos LLMs de gerar textos tecnicamente originais, coesos e convincentes tornou essas fronteiras estruturalmente mais tênues (Cheng; Calhoun; Reedy, 2025; Revell *et al.*, 2024).

No plano normativo nacional, o Referencial MEC nomeia explicitamente esse risco, ao afirmar que tecnologias generativas “tensionam noções consolidadas de autoria e integridade acadêmica” (Brasil, 2026, p. 17) e ao elevar a proteção dos direitos autorais, da propriedade intelectual e a prevenção do plágio ao estatuto de desafio estrutural da adoção de IAEd (Brasil, 2026).

Um dos efeitos do uso de LLMs na produção acadêmica reside na crescente dificuldade de rastrear a origem das ideias presentes em um texto e, por consequência, de atribuir crédito intelectual de forma fidedigna. Diferentemente de outras ferramentas tecnológicas que auxiliam na coleta ou análise de dados, os modelos generativos sintetizam, reordenam e reescrevem conteúdo extraído de vastos acervos de texto humano, sem que as fontes originais sejam identificadas ou referenciadas de maneira transparente (Cheng; Calhoun; Reedy, 2025). Os referidos autores alertam que o uso de LLMs pode incorporar ao texto final ideias provenientes de trabalhos anteriores sem que o autor humano sequer conheça as fontes originais, rompendo a cadeia tradicional de atribuição acadêmica. Os LLMs funcionam como modelos estatísticos de associação lexical, e não como sistemas de verificação de fatos ou de rastreamento de fontes:

(...) o modelo não utiliza explicitamente esse significado, mas gera cadeias de texto sintaticamente apropriadas com base em associações estatísticas. Não existem componentes de verificação de fatos nesse processo nos modelos atuais, o que constitui uma importante fonte das alucinações da IA e da propagação de vieses presentes no conjunto de dados de treinamento. (Cheng; Calhoun; Reedy, 2025, p. 5, tradução nossa⁷).

Essa característica estrutural dos LLMs tem implicações diretas para a integridade da produção científica: um texto gerado por IA pode reproduzir argumentos, enquadramentos teóricos ou achados de outros pesquisadores de forma altamente convincente, sem que isso se traduza em citação ou reconhecimento. Configura-se, assim, uma nova modalidade de apropriação indevida de produção intelectual, igualmente danosa à cadeia de reconhecimento que sustenta a ciência como empreendimento coletivo (Cheng; Calhoun; Reedy, 2025; Lund *et al.*, 2025).

9 AMBIGUIDADE DE AUTORIA E PLÁGIO MEDIADO POR MÁQUINA

Para compreender as implicações avaliativas da IAGen, é importante identificar *quem está sendo avaliado* quando uma pessoa (humana) entrega, por exemplo, um trabalho com assistência de IAGen. Nesse contexto, podemos distinguir três cenários:

- *Humano apoiado pela IAGen*: o humano é o principal autor do trabalho, utilizando-a como ferramenta de apoio. A avaliação recai principalmente sobre as competências do humano, enquanto a IAGen funciona como suporte. Exemplo: um estudante escreve seu ensaio sobre literatura e usa a ferramenta apenas para revisar ortografia e sugerir sinônimos.
- *IAGen operada pelo humano*: a ferramenta realiza a maior parte do trabalho sob instruções do humano, que orienta o processo, define objetivos, estabelece critérios e revisa o resultado final. A avaliação deve considerar tanto a operação da IAGen quanto a habilidade do humano em conduzir a ferramenta. Exemplo: um estudante usa a ferramenta para gerar

⁷ Do original “(...) the model is not explicitly utilising this meaning, but is rather generating syntactically appropriate strings of text based on statistical associations. There are no ‘fact-checking’ components to this process in current models, which is an important source of the AI hallucinations and propagation of biases within the training dataset”.

automaticamente gráficos e análises de dados, mas escolhe quais dados usar e revisa o relatório final antes de entregar.

- *IAGen autônoma, sem intervenção humana*: a máquina produz o trabalho de forma independente, sem contribuição significativa do humano. Nesse caso o trabalho não reflete as habilidades do humano. Exemplo: um estudante pede à IAGen para escrever todo o seu TCC, lendo pouco e sem compreender claramente o conteúdo do trabalho.

Uma consequência direta desse cenário de avaliação é a elevação da exigência sobre o avaliador/examinador, reafirmando o princípio pedagógico clássico de que, especialmente em bancas de Trabalho de Conclusão de Curso (TCC), Dissertação e Tese, deve-se ser comprovadamente especialista no tema avaliado, sendo capaz de formular perguntas que revelem se o autor domina o que “supostamente” ele mesmo escreveu.

No âmbito da produção estudantil, os limites tornam-se ainda mais difusos. Lund *et al.* (2025) verificaram, por meio de um levantamento aplicado a 401 estudantes de universidades norte-americanas, que muitos deles não reconhecem claramente se o uso de IA configura ou não fraude acadêmica. O estudo indica que estudantes tendem a tratar o uso de IA como uma categoria distinta e não como extensão direta do plágio convencional.

Essa percepção é problemática porque, do ponto de vista funcional, a submissão de textos integralmente gerados por IA como produção própria constitui o que Lund *et al.* (2025) denominam de *falsa autoria*. Yeo (2023, apud Lund *et al.*, 2025) observa que, embora o uso de IAGen para redigir ensaios possa não se enquadrar tecnicamente na definição canônica de plágio, é razoável classificá-lo como falsa autoria.

Do ponto de vista técnico, a situação é agravada pela capacidade que a IAGen tem de reescrever textos existentes de maneira altamente convincente, dificultando a detecção tanto por leitores humanos quanto por sistemas automatizados de verificação de plágio. Cheng, Calhoun e Reedy (2025) destacam que os modelos generativos podem reproduzir elementos de seus conjuntos de treinamento de forma virtualmente idêntica quando consultados de certas maneiras, introduzindo texto plagiado que de outra forma não figuraria no manuscrito. Esse fenômeno é especialmente preocupante porque a autoria real do conteúdo torna-se

opaca, enquanto o texto preserva aparência de originalidade (Cheng; Calhoun; Reedy, 2025). Ou seja, podemos inferir que se trata de uma situação de *plágio mediado por máquina*⁸.

Os riscos teóricos supracitados encontram respaldo em evidências empíricas. Revell *et al.* (2024) configura, em certo sentido, uma versão acadêmica do Teste de Turing⁹. Os referidos autores conduziram um estudo comparativo no qual ensaios produzidos por estudantes de graduação da Universidade de Oxford foram avaliados, em condições cegas, ao lado de ensaios gerados pelo GPT-4¹⁰. Os resultados indicaram que “os ensaios gerados por IA apresentaram desempenho comparável ao dos ensaios humanos (Revell *et al.*, 2024, p. 6, tradução¹¹ nossa). A capacidade da IAGen de produzir textos que se assemelham à produção universitária representa um desafio direto aos mecanismos tradicionais de verificação da autoria.

A convergência das evidências disponíveis aponta para um cenário em que as normas institucionais vigentes mostram-se insuficientes para enfrentar os desafios éticos impostos pela IAGen. Lund *et al.* (2025) demonstraram que o conhecimento das políticas institucionais sobre uso de IA não prediz significativamente as atitudes éticas dos estudantes nem seu comportamento efetivo. Esses dados sugerem que a simples formulação de políticas proibitivas (sem acompanhamento de formação ética) tem efeito limitado sobre a conduta estudantil.

Em síntese, a literatura científica internacional converge para o reconhecimento de que a IAGen cria novos problemas de autoria e atribuição intelectual, novas formas de plágio ou apropriação indevida (que pode ser até não intencional) e desafios éticos que exigem das instituições de ensino e pesquisa uma

⁸ A literatura recente descreve fenômenos próximos sob a denominação de *AI-assisted plagiarism* (plágio assistido por IA ou plágio mediado por IA). O trabalho de Ibrahim (2026) destaca que o surgimento da IAGen representa uma ameaça sem precedentes à integridade acadêmica em contextos de produção textual em L2 (são contextos de segunda língua), levando ao desenvolvimento dessa nova forma de violação, caracterizada pela geração ou reescrita de textos por sistemas de IA, cuja detecção por ferramentas automatizadas ainda apresenta limitações significativas, especialmente quando há tentativas de ocultação do conteúdo gerado.

⁹ Proposto pelo matemático Alan Turing em 1950, o Teste de Turing avalia se uma máquina é capaz de exibir comportamento inteligente indistinguível do de um ser humano. No teste original, um avaliador humano interage por escrito com um humano e com uma máquina sem saber qual é qual; se não conseguir distingui-los de forma confiável, considera-se que a máquina passou no teste.

¹⁰ É modelo de linguagem criado pela OpenAI.

¹¹ Do original: “*AI-generated essays performed comparably well to human essays*”.

reconfiguração profunda de seus paradigmas avaliativos, políticas de integridade e práticas pedagógicas (Revell *et al.*, 2024; Lund *et al.*, 2025).

10 CONSIDERAÇÕES FINAIS

A chegada da IAGen ao cotidiano acadêmico e estudantil vem redefinindo as práticas de estudo e produção de conhecimento. A diferença dessa tecnologia em relação às anteriores está na capacidade de produzir conteúdo assumindo uma “forma humana”, que pode ser indistinguível para um humano se a elaboração da produção foi *Humano apoiado pela IAGen, IAGen operada pelo humano ou IAGen autônoma*.

A produção automatizada de conteúdo (supervisionado ou não) coloca em xeque os três pressupostos sobre os quais a educação formal está organizada, anunciados na introdução deste ensaio e examinados ao longo de suas seções: que o texto escrito é evidência de pensamento próprio; que aprender exige esforço cognitivo do aprendiz; e que o avaliador é capaz de distinguir o que o avaliado sabe do que ele produziu. A análise desenvolvida demonstra que nenhum dos três permanece intacto diante da IAGen.

Se a IAGen pode gerar qualquer produto textual, o texto escrito deixa de ser instrumento exclusivo de verificação do saber. Não porque a escrita perca valor, mas porque ela passa a funcionar como um indício de que o estudante estudou, sem que possa, por si só, atestar o que ele sabe. O corolário é a valorização de formas avaliativas que a máquina não consegue simular, em especial a avaliação oral presencial conduzida por um examinador que conheça o tema com profundidade suficiente para formular perguntas imprevisíveis.

O problema, em última análise, não é a ferramenta. É a ausência de um quadro conceitual claro que permita ao campo educacional nomear o que está sendo comprometido e decidir, a partir daí, como agir. Enquanto essa nomeação não ocorrer, a resposta institucional tenderá a oscilar entre a proibição impraticável e a permissividade irrefletida.

Dois documentos analisados neste ensaio indicam que o campo começa a nomear esse problema. No plano universal, o MRCP-IA da UNESCO (2025) oferece um padrão de referência para orientar marcos nacionais e programas de formação continuada. No plano federal brasileiro, o Referencial do MEC (Brasil, 2026)

operacionaliza essa visão. A existência de ambos os documentos é significativa; a distância entre o referencial normativo e as práticas cotidianas das instituições é o nó central a ser enfrentado.

Cabe ressaltar que este ensaio não sustenta uma posição tecnofóbica nem propõe a exclusão da IA do ambiente educacional. A literatura revisada demonstra com clareza que, quando mediada por estratégias pedagógicas consistentes, (especificamente) a IAGen pode fortalecer o engajamento cognitivo, personalizar trajetórias de aprendizagem e ampliar o acesso a recursos. O argumento aqui é outro, o de que a incorporação acrítica da IAGen é mais perigosa do que a ferramenta em si.

REFERÊNCIAS

BRASIL. Ministério da Educação. **Referencial para Desenvolvimento e Uso Responsáveis de Inteligência Artificial na Educação**. Brasília: Ministério da Educação, fev. 2026. Disponível em: <https://www.gov.br/mec/pt-br/media/segape/referencial-oficial-pt.pdf>. Acesso em: 12 mar. 2026.

CHENG, A.; CALHOUN, A.; REEDY, G. Artificial intelligence-assisted academic writing: recommendations for ethical use. **Advances in Simulation**, v. 10, n. 22, p. 1–9, 2025. Disponível em: <https://doi.org/10.1186/s41077-025-00350-6>.

FARINOSI, M.; MELCHIOR, A. Adopt or ban? Students' perceptions and use of generative AI in higher education. **Humanities & Social Sciences Communications**, Londres, v. 12, n. 1, 2025. Disponível em: <https://doi.org/10.1057/s41599-025-05982-7>. Acesso em: 10 dez. 2025.

GENTILE, M.; CITTÀ, G.; PERNA, S.; ALLEGRA, M. Do we still need teachers? Navigating the paradigm shift of the teacher's role in the AI era. **Frontiers in Education**, Lausanne, v. 8, 2023. Disponível em: <https://doi.org/10.3389/educ.2023.1161777>. Acesso em: 7 mar. 2024.

GONÇALVES, W. A.; SANTOS, M. G. T. A. Do futuro dos Jetsons ao presente da educação: a inteligência artificial como ferramenta na aprimoração do ensino-aprendizagem. **Educte: Revista Científica do Ifal**, Maceió, v. 16, n. 1, p. 49–68, 2025. Disponível em: <https://periodicos.ifal.edu.br/educcte/article/view/2101>. Acesso em: 7 jan. 2026.

HASSEN, M. Z. The impact of AI on students' reading, critical thinking, and problem-solving skills. **American Journal of Education and Information Technology**, [S. l.], v. 9, n. 2, p. 82–90, 2025. Disponível em: <https://doi.org/10.11648/j.ajeit.20250902.12>. Acesso em: 7 dez. 2025.

IBRAHIM, K. Using AI to control AI-assisted plagiarism in L2 writing contexts: a critical review of the literature. **Language Testing in Asia**, v. 16, n. 21, p. 16–21, 2026. Disponível em: <https://doi.org/10.1186/s40468-026-00433-9>. Acesso em: 16 mar. 2026.

LONG, D. Y.; WANG, S.; MD RASHID, S.; LU, X. T. Artificial intelligence in higher education: a systematic review of its impact on student engagement and the mediating role of teaching methods. **Frontiers in Education**, Lausanne, v. 10, art. 1648661, 2026. Disponível em: <https://doi.org/10.3389/feduc.2025.1648661>. Acesso em: 10 mar. 2026.

LUND, B.; MANNURU, N. R.; TEEL, Z. A.; LEE, T. H.; ORTEGA, N. J.; SIMMONS, S.; WARD, E. Student perceptions of AI-assisted writing and academic integrity: ethical concerns, academic misconduct, and use of generative AI in higher education. **AI in Education**, v. 1, n. 1, art. 2, 2025. Disponível em: <https://doi.org/10.3390/aieduc1010002>. Acesso em: 7 dez. 2025.

PITON-GONÇALVES, J.; ALUISIO, S. M.; OLIVEIRA, L. H. M.; OLIVEIRA JÚNIOR, O. N. A learning environment for English for academic purposes based on adaptive tests and task-based systems. In: LESTER, J. C.; VICARI, R. M.; PARAGUAÇU, F. (ed.). **Intelligent Tutoring Systems**. Berlin: Springer, 2004. p. 1-11. (Lecture Notes in Computer Science, 3220). *E-book*. Disponível em: https://doi.org/10.1007/978-3-540-30139-4_1. Acesso em: 23 nov. 2004.

REVELL, T.; YEADON, W.; CAHILLY-BRETZIN, G.; CLARKE, I.; MANNING, G.; JONES, J.; MULLEY, C.; PASCUAL, R. J.; BRADLEY, N.; THOMAS, D.; LENEGHAN, F.. ChatGPT versus human essayists: an exploration of the impact of artificial intelligence for authorship and academic integrity in the humanities. **International Journal for Educational Integrity**, v. 20, n. 18, p. 1-19, 2024. Disponível em: <https://doi.org/10.1007/s40979-024-00161-8>. Acesso em: 3 dez. 2025.

SATORRE CUERDA, R. (coord.). **La docencia universitaria en tiempos de IA**. Barcelona: Octaedro, 2024. ISBN 978-84-1079-003-2. Disponível em: <https://dialnet.unirioja.es/servlet/libro?codigo=989667>. Acesso em: 7 mar. 2025.

UNESCO. **Marco referencial de competências em IA para professores**. Brasília: UNESCO Representação no Brasil, 2025. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000394280>. Acesso em: 20 dez. 2025.

UNESCO; CETIC.BR. **Artificial Intelligence in Education: uses, opportunities, and risks in the Brazilian context**. Brasília: Cetic.br/NIC.br, 24 nov. 2025. Disponível em: <https://cetic.br/en/publicacao/inteligencia-artificial-na-educacao-usos-oportunidades-e-riscos-no-cenario-brasileiro>. Acesso em: 20 dez. 2025.

YEO, M. A. Academic integrity in the age of artificial intelligence (AI) authoring apps. **TESOL Journal**, v. 14, art. 716, 2023. Disponível em: <https://doi.org/10.1002/tesj.716>. Acesso em: 12 mar. 2024.

Recebido em: 16/03/2026
Aceito em: 15/04/2026
Publicado em: 28/08/2026

