ReGeo ISSN: 2177-3246

INTELIGÊNCIA ARTIFICIAL NA PRODUÇÃO DE TRABALHOS ESCOLARES: CRITÉRIOS DE AVALIAÇÃO DA AUTORIA E DA APRENDIZAGEM NO ENSINO MÉDIO

ARTIFICIAL INTELLIGENCE IN THE PRODUCTION OF SCHOOL ASSIGNMENTS: CRITERIA FOR EVALUATING AUTHORSHIP AND LEARNING IN SECONDARY EDUCATION

INTELIGENCIA ARTIFICIAL EN LA PRODUCCIÓN DE TRABAJOS ESCOLARES: CRITERIOS DE EVALUACIÓN DE LA AUTORÍA Y DEL APRENDIZAJE EN LA EDUCACIÓN MEDIA

https://doi.org/10.56238/revgeov17n11-006

Data de submissão: 18/09/2026 **Data de publicação:** 02/11/2026

André Vilhena Vieira
Mestrado em modelagem computacional do conhecimento
E-mail: andrevieira78@gmail.com
Lattes: http://lattes.cnpq.br/6456209533408278

João Fábio Rodrigues Braga
Mestre em Sociologia
Instituição: Universidade Federal do Amazonas (UFAM)
E-mail: jfabiobraga@gmail.com
Lattes: http://lattes.cnpq.br/6023866859046933

Gabriel Gonçalves Pessoa de Castro
Doutorado em Engenharia Metalúrgica
E-mail: Gabriel_mecanica@hotmail.com
Lattes: http://lattes.cnpq.br/8407541449956886

Lucas Araújo Conegundes
Especialista
E-mail: coachconegundes@gmail.com
Lattes: http://lattes.cnpq.br/9174966620436292

Cléuma de Melo Barbosa
Doutoranda em Ciências da Educação
Instituição: Universidad San Carlos (USC)
E-mail: cleuma10@yahoo.com.br
Lattes: http://lattes.cnpq.br/0351310423782928

Cíntia de Souza Adelino
Doutoranda em educação
E-mail: cintia-adelino@hotmail.com
Lattes: http://lattes.cnpq.br/8675113846123568

RESUMO
A disponibilidade de sistemas de inteligência artificial generativa colocou o professor do Ensino Médio diante de uma pergunta para a qual o seu instrumental de avaliação não foi desenhado: como atribuir

autoria a um trabalho escolar quando o texto pode ter sido produzido por uma máquina. Este artigo examina duas camadas do problema. A primeira é normativa e consiste no mapeamento lexical da Base Nacional Comum Curricular do Ensino Médio, documento que organiza o currículo e orienta a avaliação na etapa. A segunda é bibliométrica e consiste no levantamento da literatura indexada sobre detecção de texto gerado por máquina e sobre avaliação na educação secundária, com todas as consultas restritas aos campos de título e resumo após um teste de contaminação. O mapeamento revela que a Base emprega autoria nove vezes, e que oito dessas ocorrências designam algo que o estudante exerce, não algo que o professor verifica. A única ocorrência em sentido de verificação integra uma habilidade de checagem de notícias falsas. Onze expressões do vocabulário da verificação de autoria não ocorrem uma única vez, entre elas plágio, originalidade, autenticidade, integridade acadêmica e inteligência artificial. No plano bibliométrico, o recorte que combina inteligência artificial generativa, educação secundária e avaliação retorna 25 registros, dos quais 12 são preprints sem revisão por pares, 9 pertencem a contexto clínico e apenas 1 trata de autoria na escrita do Ensino Médio, sendo também um preprint. Conclui-se que o professor opera sem amparo curricular e sem base de evidência consolidada.

Palavras-chave: Inteligência Artificial Generativa. Autoria. Avaliação da Aprendizagem.

ABSTRACT
The availability of generative artificial intelligence systems has placed secondary school teachers before a question their assessment toolkit was not designed for: how to attribute authorship to a school assignment when the text may have been produced by a machine. This article examines two layers of the problem. The first is normative and consists of the lexical mapping of the Brazilian National Common Curricular Base for Secondary Education, the document that organizes the curriculum and guides assessment at that stage. The second is bibliometric and consists of a survey of indexed literature on machine generated text detection and on assessment in secondary education, with all queries restricted to title and abstract fields after a contamination test. The mapping reveals that the Base uses the word authorship nine times, and that eight of those occurrences designate something the student exercises, not something the teacher verifies. The single occurrence in a verification sense belongs to a fake news checking skill. Eleven expressions from the vocabulary of authorship verification do not occur even once, among them plagiarism, originality, authenticity, academic integrity and artificial intelligence. At the bibliometric level, the intersection combining generative artificial intelligence, secondary education and assessment returns 25 records, of which 12 are preprints without peer review, 9 belong to clinical contexts and only 1 addresses authorship in secondary school writing, itself also a preprint. It is concluded that teachers operate without curricular support and without a consolidated evidence base.

Keywords: Generative Artificial Intelligence. Authorship. Learning Assessment.

RESUMEN
La disponibilidad de sistemas de inteligencia artificial generativa ha colocado al profesor de Educación Secundaria ante una pregunta para la cual su instrumental de evaluación no fue diseñado: cómo atribuir autoría a un trabajo escolar cuando el texto pudo haber sido producido por una máquina. Este artículo examina dos capas del problema. La primera es normativa y consiste en el mapeo léxico de la Base Nacional Común Curricular de la Educación Secundaria, documento que organiza el currículo y orienta la evaluación en la etapa. La segunda es bibliométrica y consiste en el levantamiento de la literatura indexada sobre detección de texto generado por máquina y sobre evaluación en la educación secundaria, con todas las consultas restringidas a los campos de título y resumen tras una prueba de contaminación. El mapeo revela que la Base emplea autoría nueve veces, y que ocho de esas

ocurrencias designan algo que el estudiante ejerce, no algo que el profesor verifica. La única ocurrencia en sentido de verificación integra una habilidad de verificación de noticias falsas. Once expresiones del vocabulario de la verificación de autoría no ocurren ni una sola vez, entre ellas plagio, originalidad, autenticidad, integridad académica e inteligencia artificial. En el plano bibliométrico, el recorte que combina inteligencia artificial generativa, educación secundaria y evaluación retorna 25 registros, de los cuales 12 son preprints sin revisión por pares, 9 pertenecen a contexto clínico y apenas 1 trata de autoría en la escritura de la Educación Secundaria, siendo también un preprint. Se concluye que el profesor opera sin amparo curricular y sin base de evidencia consolidada.

Palabras clave: Inteligencia Artificial Generativa. Autoría. Evaluación del Aprendizaje.

1 INTRODUÇÃO

Um trabalho escolar entregue por um estudante do Ensino Médio chega ao professor como um texto. Até muito recentemente, a inferência que ligava esse texto ao estudante era tão automática que raramente se tornava objeto de reflexão: quem entregou, escreveu. A avaliação da aprendizagem apoiava-se inteiramente nessa inferência, porque avaliar o texto equivalia a avaliar o que o estudante havia aprendido. A inferência podia falhar, e falhava, nos casos conhecidos de cópia entre colegas ou de reprodução de fonte não declarada, mas falhava de modo detectável e excepcional.

A disponibilidade ampla de sistemas de inteligência artificial generativa rompeu essa automaticidade. Não porque tenha introduzido a possibilidade de fraude, que sempre existiu, mas porque desacoplou, de maneira barata e silenciosa, a produção do texto da realização da aprendizagem. Um estudante pode agora entregar um texto competente sobre um assunto que não compreende, sem copiar de ninguém, sem reproduzir fonte alguma de modo rastreável, e sem que o texto apresente as marcas que historicamente denunciavam a cópia.

Diante disso, o professor precisa de critérios. Precisa saber o que conta como autoria, o que conta como uso legítimo da ferramenta, o que conta como aprendizagem demonstrada e por qual procedimento essas distinções se estabelecem. A pergunta deste artigo é se tais critérios existem nos dois lugares onde seria razoável procurá-los: no documento normativo que organiza o currículo e orienta a avaliação no Ensino Médio brasileiro, e na literatura científica indexada sobre o tema.

A resposta é negativa nas duas frentes, e as razões da negativa são distintas e igualmente instrutivas. No plano normativo, a Base Nacional Comum Curricular do Ensino Médio trata a autoria exclusivamente como algo que o estudante exerce, associando-a a protagonismo e a criação. Não há, no documento, vocabulário para a autoria como objeto de verificação pelo professor. No plano bibliométrico, a literatura indexada que combina inteligência artificial generativa, educação secundária e avaliação é escassa, majoritariamente não revisada por pares e, em proporção substancial, nem sequer trata de educação.

É necessário delimitar o que o artigo não faz. Não se propõe um instrumento de avaliação, nem se recomenda o uso ou a proibição de ferramentas de detecção, nem se afirma que a Base Nacional Comum Curricular seja deficiente por não antecipar uma tecnologia que se difundiu anos após a sua homologação. O documento foi homologado em dezembro de 2018, e a difusão em massa de sistemas generativos de texto ocorreu a partir do final de 2022. A ausência que se descreve é cronologicamente explicável. O que o artigo sustenta é que essa ausência tem consequências práticas presentes, e que descrevê-la com precisão é condição para discuti-las.

O texto organiza-se em nove seções. A segunda descreve o procedimento metodológico, incluindo o teste de contaminação e as limitações de acesso encontradas. A terceira apresenta o mapeamento lexical da Base. A quarta examina as nove ocorrências de autoria no documento. A quinta trata da literatura sobre detecção de texto gerado por máquina. A sexta analisa o recorte da educação secundária. A sétima discute critérios possíveis a partir do que a literatura brasileira vem propondo. A oitava explicita limitações e a nona apresenta as considerações finais.

2 METODOLOGIA

O estudo combina análise documental com componente lexicométrico e levantamento bibliométrico. A coleta foi encerrada em 19 de outubro de 2026, data de referência de todas as consultas registradas nas referências.

2.1 CAMADA NORMATIVA

A fonte primária é o texto integral da Base Nacional Comum Curricular do Ensino Médio, obtido em formato de documento portátil no sítio oficial do portal da Base Nacional Comum Curricular, mantido pelo Ministério da Educação. O arquivo contém 154 páginas. O texto foi extraído página a página e normalizado quanto a espaços em branco, resultando em um corpo de 328.107 caracteres, sobre o qual foram realizadas as contagens.

Registra-se uma limitação de acesso relevante para a replicabilidade. Durante a janela de pesquisa, as tentativas de acesso aos domínios da Presidência da República, da Imprensa Nacional e do portal institucional do Ministério da Educação retornaram falha de conexão, sem resposta do servidor. O portal específico da Base Nacional Comum Curricular respondeu normalmente, e foi dele que o documento foi obtido. A conferência cruzada do texto com o veículo de publicação oficial da homologação não pôde, por isso, ser realizada.

Foram contadas, de forma insensível a maiúsculas e minúsculas, as ocorrências de vinte e uma expressões, organizadas em dois conjuntos conceituais. O primeiro reúne o vocabulário da autoria exercida, isto é, termos que designam a produção e a criação pelo estudante. O segundo reúne o vocabulário da autoria verificada, termos que designam a aferição, pelo professor ou pela instituição, de que o trabalho é efetivamente do estudante. A contagem é de ocorrências de cadeia de caracteres, não de lemas. Todas as ocorrências de autoria foram adicionalmente lidas em contexto e classificadas quanto ao sentido.

2.2 CAMADA BIBLIOMÉTRICA E TESTE DE CONTAMINAÇÃO

A base consultada foi a Europe PMC, por meio de sua interface programática. Essa base, por padrão, pesquisa o texto completo dos artigos de acesso aberto, e não apenas título e resumo, de modo que uma expressão mencionada de passagem é recuperada como se o artigo fosse sobre ela. Antes de qualquer análise, mediu-se a magnitude desse efeito no presente tema. Os resultados constam da Tabela 1.

Tabela 1. Teste de contaminação por texto completo na base Europe PMC

Expressão consultada	Busca livre	Título e resumo	Razão	Mecanismo provável da contaminação
"authorship"	702.225	9.242	76,0 vezes	Declarações de contribuição de autoria presentes em quase todo artigo
"generative AI"	187.860	5.559	33,8 vezes	Declarações obrigatórias de uso ou não uso de IA na redação
"AI-generated text"	1.042	126	8,3 vezes	Menção em discussões e limitações
"AI detector"	123	19	6,5 vezes	Menção em discussões e limitações
"ChatGPT"	77.626	12.690	6,1 vezes	Menção em discussões e limitações
"academic integrity"	3.707	864	4,3 vezes	Menção em seções de ética

A coluna Razão expressa o quociente entre a busca livre e a busca restrita a título e resumo. O mecanismo indicado é interpretativo e não foi verificado registro a registro.
Fonte: elaborada pela autora. Consultas realizadas em 19 out. 2026.

O caso de authorship é o mais extremo já encontrado pela autora em testes dessa natureza: 702.225 registros em busca livre contra 9.242 restritos a título e resumo, uma razão de 76,0 vezes. O mecanismo é transparente e merece registro porque revela algo sobre a própria infraestrutura da publicação científica. Praticamente todo artigo contém, em seu texto completo, uma seção de contribuições de autoria. A palavra está, portanto, em quase toda parte, e buscá-la em texto completo equivale a não filtrar nada.

O caso de generative AI, com razão de 33,8 vezes, tem mecanismo análogo e cronologicamente recente. A partir de 2023, periódicos passaram a exigir que os autores declarem se utilizaram sistemas generativos na redação do manuscrito. Essa declaração, inclusive quando é negativa, insere a expressão no texto completo. O resultado é que a prática de transparência quanto ao uso de inteligência artificial contaminou a busca em texto completo sobre inteligência artificial. Todas as consultas subsequentes foram restritas aos campos de título e resumo.

Sobre o corpus assim delimitado, construíram-se dois recortes. O recorte da detecção reúne os registros que combinam descritores de texto gerado por máquina com descritores de acurácia, falso positivo ou confiabilidade, totalizando 59 registros. O recorte da educação secundária reúne os

registros que combinam descritores de inteligência artificial generativa, descritores de etapa escolar secundária e descritores de avaliação, integridade acadêmica ou autoria, totalizando 25 registros. Ambos os conjuntos foram integralmente recuperados com seus resumos e classificados por leitura. Registros publicados em servidores de preprint foram identificados pelo prefixo do identificador digital de objeto.

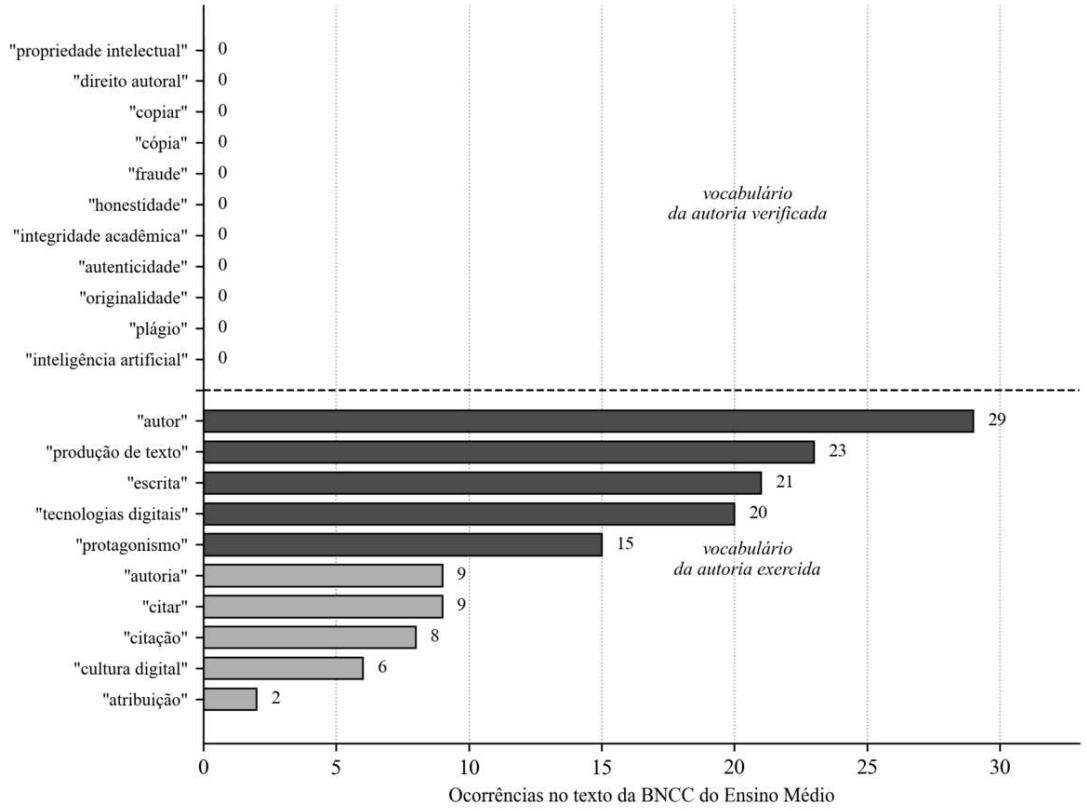
2.3 CAMADA BIBLIOGRÁFICA

As referências brasileiras foram localizadas em buscas na base Crossref restritas a periódicos nacionais e a publicações posteriores a 2024, com atenção aos periódicos Aracê, Lumen et Virtus e aos títulos indexados na SciELO. Cada identificador digital de objeto, nacional ou internacional, foi verificado individualmente na interface programática da Crossref, confirmando-se título, autoria, periódico, volume, fascículo, paginação e data. Nenhuma referência foi incluída sem essa verificação, e aplicou-se o corte de 19 de outubro de 2026.

3 O VOCABULÁRIO DA BASE NACIONAL COMUM CURRICULAR DO ENSINO MÉDIO

O Gráfico 1 apresenta o resultado do mapeamento lexical, separando o vocabulário da autoria exercida do vocabulário da autoria verificada.

Gráfico 1. Vocabulário da autoria no texto da BNCC do Ensino Médio



Fonte: elaborado pela autora. Contagem insensível a maiúsculas e minúsculas sobre o texto extraído do documento oficial, com 328.107 caracteres (BRASIL, 2018).

O contraste é categórico. O vocabulário da autoria exercida está presente e é recorrente: autor ocorre 29 vezes, produção de texto 23, escrita 21, tecnologias digitais 20, protagonismo 15, autoria 9, citar 9, citação 8, cultura digital 6 e atribuição 2. O vocabulário da autoria verificada é integralmente ausente: inteligência artificial, plágio, originalidade, autenticidade, integridade acadêmica, honestidade, fraude, cópia, copiar, direito autoral e propriedade intelectual registram, todas as onze expressões, zero ocorrência em um documento de 328.107 caracteres.

A ausência de inteligência artificial é, como já se disse, cronologicamente explicável e não constitui crítica. Mais significativa é a ausência simultânea de plágio, de originalidade, de autenticidade, de honestidade e de propriedade intelectual, todos conceitos anteriores à inteligência artificial generativa e disponíveis ao legislador curricular em 2018. O documento que organiza o currículo e orienta a avaliação de toda uma etapa da educação básica não contém vocabulário para tratar da autenticidade do trabalho escolar.

Convém ser rigoroso quanto ao alcance desse achado. A ausência de uma expressão no texto de um documento normativo não significa que a instituição escolar esteja desautorizada a tratar do assunto. Regimentos escolares, projetos político-pedagógicos e normas internas podem e

frequentemente disciplinam a questão, e o professor dispõe de autonomia didática. O que a ausência indica é que a Base não oferece ancoragem curricular para esses dispositivos locais, e que cada escola constrói seus critérios sem parâmetro nacional comum. Em um documento cuja finalidade declarada é estabelecer o que é comum a todos, essa é uma lacuna de natureza específica.

Dois termos presentes merecem comentário, porque poderiam ser lidos como contraexemplos. Citar e citação, com 9 e 8 ocorrências, pertencem ao vocabulário do trabalho com fontes, mas a leitura em contexto mostra que ocorrem em habilidades de análise de textos literários e argumentativos, no sentido de identificar e empregar citações como recurso de composição, e não no sentido de atribuir corretamente a autoria de material de terceiros. Integridade ocorre uma vez, e a leitura em contexto revela que se refere à integridade física, individual e coletiva, em habilidade da área de Ciências da Natureza sobre segurança em atividades cotidianas. Não há, portanto, nenhuma ocorrência de integridade em sentido acadêmico.

4 AS NOVE OCORRÊNCIAS DE AUTORIA

Como autoria é o termo central da pergunta deste artigo, todas as suas nove ocorrências foram lidas em contexto e classificadas. A Tabela 2 apresenta o resultado.

Tabela 2. Classificação das nove ocorrências de autoria na BNCC do Ensino Médio

Nº	Contexto da ocorrência	Sentido
1	Competência Geral 5: utilizar tecnologias digitais para "exercer protagonismo e autoria na vida pessoal e coletiva"	Autoria exercida
2	Área de Linguagens: foco na "ampliação da autonomia, do protagonismo e da autoria nas práticas de diferentes linguagens"	Autoria exercida
3	Culturas juvenis: participação "por meio da autoria de diversas produções que constituem as culturas juvenis"	Autoria exercida
4	Competência Específica 3 de Linguagens: exercer "protagonismo e autoria na vida pessoal e coletiva"	Autoria exercida
5	Reiteração da Competência Específica 3 de Linguagens	Autoria exercida
6	Práticas culturais: "incorporação de critérios técnicos e estéticos na análise e autoria das produções"	Autoria exercida
7	Habilidade EM13LP38: checagem de fatos noticiados, com verificação de "veículo, fonte, data e local da publicação, autoria, URL"	Autoria verificada, em notícia de terceiro
8	Práticas de leitura: socializar "obras da própria autoria (poemas, contos e suas variedades, roteiros...)"	Autoria exercida
9	Área de Ciências Humanas: reiteração da Competência Geral 5	Autoria exercida

Fonte: elaborada pela autora, por leitura integral de cada ocorrência em contexto (BRASIL, 2018).

Oito das nove ocorrências designam a autoria como algo que o estudante exerce. Ela aparece, com notável consistência, acoplada a protagonismo, autonomia e criação. A autoria é, na Base, uma conquista a ser promovida, um objetivo formativo, uma capacidade a desenvolver. Esse é um posicionamento pedagógico legítimo e, aliás, generoso: trata o estudante como produtor de cultura e não como suspeito.

A nona ocorrência é a única em que a autoria figura como objeto de verificação, e ela está na habilidade EM13LP38, cujo enunciado determina usar procedimentos de checagem de fatos noticiados e fotos publicadas, verificando ou avaliando veículo, fonte, data e local da publicação, autoria e endereço eletrônico, comparando diferentes fontes e consultando ferramentas e sítios checadores, de forma a combater a proliferação de notícias falsas.

A habilidade é pertinente e bem construída, mas o seu objeto é o texto alheio que circula na internet, não o texto do próprio estudante entregue ao professor. A operação que ela ensina é a de um leitor crítico diante de uma notícia, não a de um professor diante de um trabalho. É significativo que a única competência de verificação de autoria prevista na Base seja dirigida ao combate à desinformação jornalística, tema que em 2018 ocupava o debate público, e não à autenticidade da produção escolar.

Há, nessa distribuição, uma assimetria que organiza todo o problema prático. A Base ensina o estudante a desconfiar da autoria de um texto que ele lê, e não oferece ao professor nenhum parâmetro para avaliar a autoria de um texto que ele recebe. O estudante é formado como verificador de autoria alheia e tratado como autor presumido da própria produção. Enquanto a inferência entre entrega e escrita se mantinha automática, essa assimetria não produzia efeito. Com sistemas generativos disponíveis, ela passa a produzir.

5 A LITERATURA SOBRE DETECÇÃO DE TEXTO GERADO POR MÁQUINA

A resposta institucional mais imediata à ruptura descrita tem sido a adoção de ferramentas de detecção automática. Convém examinar o que a literatura indexada diz sobre a confiabilidade dessas ferramentas. O recorte da detecção, combinando descritores de texto gerado por máquina com descritores de acurácia, falso positivo ou confiabilidade, retornou 59 registros, distribuídos de modo revelador no tempo: 1 em 2020, 1 em 2023, 10 em 2024, 26 em 2025 e 21 até a data de corte em 2026. A literatura é, portanto, quase inteiramente posterior a 2023.

Entre os 59 registros, 44 estão em periódicos e 15 em servidores de preprint ou em plataformas de publicação imediata com revisão aberta posterior. Trinta e seis mencionam contexto educacional e 21 mencionam contexto biomédico. Apenas 8 mencionam explicitamente falsos positivos e apenas 4

mencionam viés contra escritores não nativos de inglês. Esse último número é digno de nota, porque o problema do viés é precisamente aquele com consequências mais graves para o estudante brasileiro.

O estudo de referência nesse ponto é o de Liang *et al.* (2023), publicado em periódico revisado por pares, que avaliou detectores de texto gerado por máquina e concluiu que eles são enviesados contra escritores não nativos de inglês. O mecanismo é compreensível: os detectores operam majoritariamente sobre medidas de previsibilidade estatística do texto, e a escrita de quem usa um repertório lexical e sintático mais restrito apresenta previsibilidade mais alta, exatamente como a escrita gerada por máquina. Um preprint recente localizado no corpus reafirma que auditorias empíricas expuseram viés sistemático contra escritores de inglês como segunda língua, mas trata-se de documento não submetido a revisão por pares, e por isso é aqui registrado como indício e não como evidência estabelecida.

A implicação para a escola brasileira é direta e grave, ainda que exija uma ressalva. O estudante do Ensino Médio brasileiro que escreve em português não está na mesma situação do estudante que escreve em inglês como segunda língua, e a transposição do achado de Liang *et al.* (2023) para o português não é automática: seria necessário estudo específico com detectores treinados em corpora de língua portuguesa. O que o achado autoriza a afirmar é mais modesto e ainda assim suficiente: o mecanismo de funcionamento dos detectores penaliza a escrita de baixa variabilidade lexical e sintática, e a escrita escolar de estudantes com menor repertório é justamente desse tipo. O risco de falso positivo concentra-se, previsivelmente, nos estudantes que mais dependem da escola.

Os estudos empíricos do corpus confirmam a fragilidade das ferramentas em contextos acadêmicos. Popkov e Barrett (2024) conduziram estudo experimental sobre a acurácia de detectores em escrita acadêmica da área de saúde comportamental. Doru *et al.* (2025) compararam, em estudo semirrandomizado, a detecção de redações de estudantes de medicina escritas por humanos e por sistemas generativos. Erol *et al.* (2025) avaliaram acurácia e limitações de detectores de produto de inteligência artificial, formulando no próprio título a pergunta sobre a confiabilidade dessas ferramentas. Hyatt *et al.* (2025) propuseram o uso agregado de resultados de múltiplos detectores especificamente como estratégia para eliminar falsos positivos na escrita de estudantes de áreas científicas, o que pressupõe, como ponto de partida, que o resultado de um detector isolado não é confiável o bastante para fundamentar uma decisão.

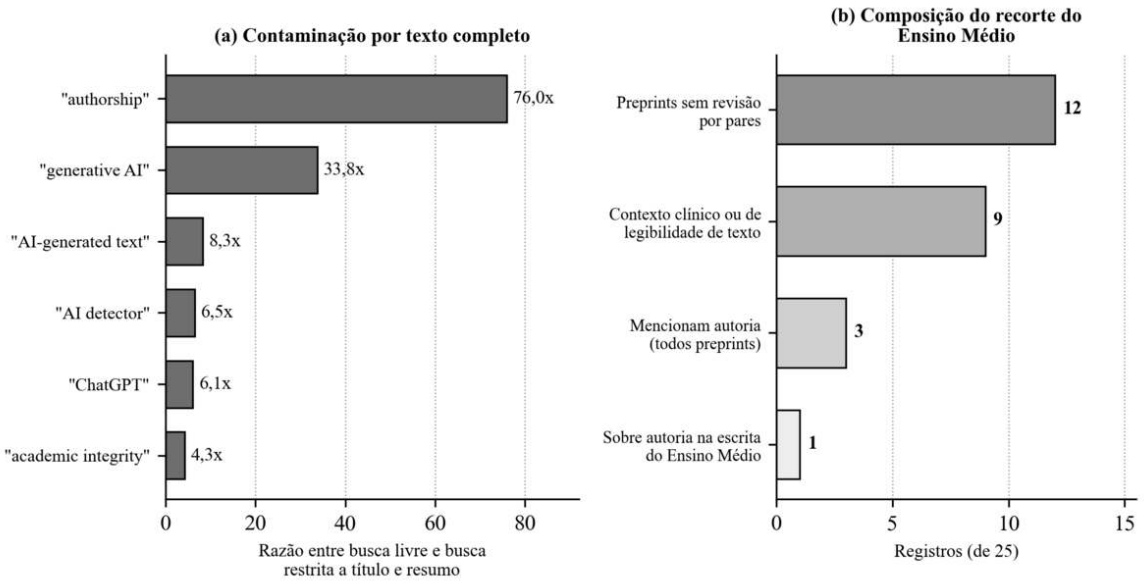
Uma linha distinta procura resolver o problema por atribuição estilométrica em vez de detecção binária. Mani e Arulanandam (2026) propõem um arcabouço de aprendizado de máquina estilométrico para atribuição de autoria entre textos de sistema generativo e textos de estudantes, e Hua e Yao (2024) investigam como a tokenização e o tamanho do conjunto de dados afetam a identificação de texto

gerado por máquina, em trabalho que situa explicitamente a questão no contexto da educação básica. A diferença de abordagem importa: a atribuição estilométrica compara o texto suspeito com textos anteriores do mesmo estudante, o que a aproxima de um raciocínio pedagógico, enquanto a detecção binária pergunta apenas se o texto parece de máquina, sem referência a quem o entregou.

6 O RECORTE DA EDUCAÇÃO SECUNDÁRIA

O recorte que mais diretamente responde à pergunta deste artigo combina inteligência artificial generativa, etapa escolar secundária e avaliação, integridade acadêmica ou autoria. Ele retorna 25 registros. O painel (b) do Gráfico 2 apresenta a sua composição, e o painel (a) retoma visualmente o teste de contaminação discutido na metodologia.

Gráfico 2. Contaminação por texto completo e composição do recorte da educação secundária



Fonte: elaborado pela autora. No painel (b), as categorias não são excludentes. Consultas realizadas em 19 out. 2026.

Três características desse conjunto de 25 registros desautorizam qualquer leitura otimista do número. A primeira é a proporção de literatura não revisada por pares: 12 dos 25 registros, isto é, 48,0%, são preprints hospedados em servidores como Research Square, Preprints.org e repositórios de ciência aberta. Quase metade do que a busca recupera não passou por avaliação editorial.

A segunda é a presença de literatura que não trata de educação. Nove dos 25 registros, 36,0% do conjunto, pertencem a contexto clínico ou de legibilidade de texto. São estudos sobre compreensão de tratamentos de doença glomerular, sobre legibilidade de folhetos para pacientes revisados por sistema generativo, sobre respostas de chatbot a perguntas sobre analgesia de parto, sobre educação de pais em oftalmologia pediátrica, sobre folhetos a respeito de dirigir após acidente vascular cerebral,

sobre respostas a respeito de paralisia cerebral e sobre legibilidade de respostas acerca de reconstrução de ligamento cruzado anterior.

O mecanismo que os traz para o recorte merece explicitação, porque é uma polissemia análoga à do teste de contaminação, mas operando dentro do campo semântico correto. A literatura biomédica avalia a legibilidade de materiais destinados a pacientes usando como parâmetro o nível de escolaridade necessário para compreendê-los, e expressa esse parâmetro em frases como nível de leitura de ensino médio. A expressão que, em pesquisa educacional, designa uma etapa de ensino, designa ali uma métrica de dificuldade textual. Uma busca sobre inteligência artificial e ensino médio recupera, por isso, uma literatura clínica substancial que nada diz sobre escolas.

A terceira característica é a mais decisiva. Apenas 3 dos 25 registros mencionam autoria, e os três são preprints. Dois deles são revisões de escopo sobre criatividade e aprendizagem criativa com inteligência artificial generativa na educação básica, nas quais a autoria aparece como tema correlato. Apenas 1 registro, intitulado PRISM e dedicado a perfis de uso de inteligência artificial, criatividade e autoria na escrita do Ensino Médio, trata diretamente do objeto deste artigo. Esse único registro é também um preprint.

A formulação precisa do achado é, portanto, a seguinte: na literatura biomédica indexada acessível pela estratégia empregada, e dentro do corte temporal adotado, não foi localizado nenhum estudo revisado por pares que trate especificamente de autoria e uso de inteligência artificial na escrita de estudantes do Ensino Médio. O único registro que trata do tema não passou por revisão por pares. O professor que procurar fundamento científico consolidado para decidir como avaliar autoria nesse contexto não o encontrará.

Tabela 3. Composição dos 25 registros do recorte da educação secundária

Característica do registro	Registros	Proporção	Observação
Preprints sem revisão por pares	12	48,0%	Identificados pelo prefixo do identificador digital de objeto
Publicados em periódico revisado por pares	13	52,0%	Inclui os registros de contexto clínico
Contexto clínico ou de legibilidade de texto	9	36,0%	Recuperados pela expressão de nível de leitura, não por tema educacional
Mencionam autoria	3	12,0%	Todos os três são preprints
Tratam de autoria na escrita do Ensino Médio	1	4,0%	Preprint, não revisado por pares

As categorias não são excludentes.

Fonte: elaborada pela autora, por leitura de títulos e resumos. As categorias não são excludentes. Consulta em 19 out. 2026.

7 CRITÉRIOS POSSÍVEIS: O QUE A LITERATURA BRASILEIRA VEM PROPONDO

A escassez identificada na literatura indexada não corresponde a silêncio na produção brasileira, que tem discutido o problema com densidade conceitual, ainda que predominantemente em chave ensaística e não experimental. Almeida e Barbosa (2026) rediscutem a autoria na produção de textos na escola em tempos de inteligência artificial generativa, e Amaral (2026) formula a questão como a tensão entre o eu e o algoritmo. Batista e Vasconcelos (2026) tratam do que chamam de autoria híbrida, e Brito Rocha, Brandão e Panfile (2025) examinam a avaliação da aprendizagem digital considerando explicitamente a hipótese de autoria híbrida entre humano e máquina.

A noção de autoria híbrida é conceitualmente produtiva porque desloca a pergunta. Em vez de perguntar se o texto é do estudante ou da máquina, pergunta qual foi a contribuição de cada um e se a contribuição do estudante evidencia aprendizagem. Essa reformulação tem uma vantagem prática considerável: ela não depende de detecção. Um professor pode avaliar a contribuição do estudante sem precisar estabelecer, com certeza probatória, a origem de cada sentença.

Dessa reformulação derivam critérios que não dependem de ferramenta de detecção e que a literatura brasileira vem indicando. O primeiro é o deslocamento do produto para o processo: avaliar versões sucessivas, registros de planejamento, escolhas descartadas e justificativas, em vez de avaliar apenas o texto final. Zucolar (2025), ao reler a psicogênese da escrita na era da inteligência artificial, situa essa discussão na tradição construtivista, para a qual o processo sempre foi o objeto privilegiado de observação.

O segundo é a arguição oral sobre o texto entregue. Se a finalidade da avaliação é verificar aprendizagem, e não verificar digitação, perguntar ao estudante sobre as escolhas do seu texto resolve diretamente o problema de atribuição, sem recorrer a inferência estatística sobre o estilo. Gonçalves e Bodolay (2025) realizaram análise comparativa entre textos produzidos com e sem auxílio de inteligência artificial no Ensino Médio, trabalho que, no conjunto localizado, é dos poucos a produzir material empírico brasileiro diretamente sobre a etapa aqui discutida.

O terceiro é a explicitação prévia das regras de uso. Costa e Garcia (2025) examinam o olhar de futuros professores sobre avaliação educacional e inteligência artificial, e Silva, Alarcão e Faria (2026) discutem possibilidades e limitações da inteligência artificial generativa na prática pedagógica brasileira. O ponto comum é que a ausência de regra explícita transfere ao estudante o ônus de adivinhar o que é permitido, e transfere ao professor o ônus de julgar depois do fato aquilo que não definiu antes.

Há ainda a dimensão regulatória. Muniz, Ferreira e Ferreira (2026) realizam análise comparativa dos marcos regulatórios de inteligência artificial na educação no Brasil e na União

Europeia, e Hermann (2025) discute o que inquieta na relação entre ética e inteligência artificial. Valença Ribeiro (2024), em estudo sobre autoria e estilo em diários de leituras de estudantes do Ensino Médio, oferece um ponto de apoio metodológico anterior ao problema: se é possível identificar estilo autoral em produções escolares, então existe uma base empírica para a abordagem estilométrica comparativa, que compara o texto suspeito com a produção anterior do mesmo estudante, em vez de compará-lo com um modelo genérico de escrita humana.

Essa convergência entre a proposta estilométrica internacional de Mani e Arulanandam (2026) e o estudo brasileiro de estilo autoral escolar de Valença Ribeiro (2024) sugere um caminho que nenhum dos dois formula explicitamente. Se o professor dispõe de produções anteriores do estudante, a pergunta deixa de ser se este texto parece de máquina e passa a ser se este texto parece deste estudante. A segunda pergunta é pedagogicamente mais defensável, porque admite resposta negativa sem acusação: um texto que não se parece com a produção anterior pode indicar uso de ferramenta, mas pode igualmente indicar progresso, mudança de repertório ou simples variação. A detecção binária não comporta essa ambiguidade; a comparação longitudinal, sim.

Resta registrar uma limitação dessas propostas. Todas elas aumentam a carga de trabalho docente. Avaliar processo exige acompanhar versões; arguir oralmente exige tempo por estudante; comparar com produção anterior exige arquivo organizado. Em turmas numerosas do Ensino Médio público brasileiro, essas exigências não são triviais, e qualquer recomendação que as ignore será, na prática, inaplicável. Pinninti (2025), ao estudar usos e percepções de estudantes de graduação sobre o ChatGPT para fins de escrita acadêmica, lembra que a adoção pelos estudantes precede e supera a capacidade institucional de resposta.

8 LIMITAÇÕES

A primeira limitação é de base bibliométrica. Utilizou-se uma única fonte de indexação, a Europe PMC, de cobertura predominantemente biomédica. Literatura educacional publicada em periódicos de educação não indexados nessa base não foi alcançada pelos recortes, e é razoável supor que parte substancial da produção sobre inteligência artificial e avaliação escolar esteja precisamente ali. Os achados de escassez devem ser lidos como relativos à literatura biomédica indexada, e não à totalidade da produção científica sobre o tema. Essa limitação é particularmente relevante neste artigo, dado que o objeto é educacional, e a sua consequência é que o número de 25 registros subestima a produção existente.

A segunda é documental. A análise recai sobre a Base Nacional Comum Curricular do Ensino Médio, e não sobre o conjunto da normatização educacional brasileira. Pareceres e resoluções do

Conselho Nacional de Educação, diretrizes curriculares estaduais, regimentos escolares e eventuais normas posteriores sobre inteligência artificial na educação não foram examinados, em parte porque a indisponibilidade de acesso aos domínios da Presidência da República, da Imprensa Nacional e do portal do Ministério da Educação durante a janela de pesquisa impediu a sua obtenção em fonte oficial. A afirmação de que falta vocabulário é uma afirmação sobre a Base, não sobre o ordenamento educacional brasileiro.

A terceira é lexicométrica. A contagem registra ocorrências de cadeias de caracteres e não captura paráfrases nem institutos designados por outras palavras. A Base poderia tratar de autenticidade do trabalho escolar sem empregar nenhuma das onze expressões contadas. A leitura em contexto de todas as ocorrências de autoria mitiga parcialmente esse risco para o termo central, mas não o elimina para os demais.

A quarta é interpretativa. A coluna de mecanismo provável da contaminação, na Tabela 1, apresenta hipóteses explicativas plausíveis que não foram verificadas registro a registro. Afirmar que a razão de 76,0 vezes para authorship decorre das declarações de contribuição de autoria é uma inferência sobre a causa, não uma medição dela. O mesmo vale para a atribuição da razão de 33,8 vezes de generative AI às declarações obrigatórias de uso de inteligência artificial.

A quinta é de transposição linguística, já enunciada na seção 5. O achado de viés de detectores contra escritores não nativos refere-se à língua inglesa e não foi replicado, no corpus localizado, para o português. A extensão do argumento ao estudante brasileiro apoia-se no mecanismo de funcionamento dos detectores, não em evidência direta, e foi apresentada com essa ressalva.

A sexta é de escopo. Nenhuma avaliação de qualidade metodológica, de risco de viés ou de tamanho de efeito foi conduzida sobre os registros recuperados. O artigo descreve a disponibilidade e a composição da literatura, não a sua solidez, e nenhuma inferência sobre a eficácia de qualquer ferramenta de detecção decorre dos dados apresentados.

9 CONSIDERAÇÕES FINAIS

Este artigo perguntou se existem critérios para avaliar autoria e aprendizagem em trabalhos escolares do Ensino Médio diante da disponibilidade de inteligência artificial generativa, procurando-os no documento curricular nacional e na literatura indexada.

No plano normativo, a Base Nacional Comum Curricular do Ensino Médio emprega autoria nove vezes, e oito dessas ocorrências designam algo que o estudante exerce, associado a protagonismo, autonomia e criação. A única ocorrência em sentido de verificação integra uma habilidade de checagem de notícias falsas, dirigida ao texto alheio que circula na internet e não ao trabalho entregue ao

professor. Onze expressões do vocabulário da verificação de autoria registram zero ocorrência em um documento de 328.107 caracteres, entre elas plágio, originalidade, autenticidade, integridade acadêmica e honestidade, todas disponíveis ao legislador curricular em 2018.

No plano bibliométrico, o recorte que combina inteligência artificial generativa, educação secundária e avaliação retorna 25 registros, dos quais 48,0% são preprints sem revisão por pares e 36,0% pertencem a contexto clínico, recuperados porque a expressão que designa uma etapa de ensino funciona, na literatura biomédica, como métrica de legibilidade. Apenas um registro trata de autoria na escrita do Ensino Médio, e também ele é um preprint.

Soma-se a esses um achado metodológico de alcance mais geral. O teste de contaminação revelou razão de 76,0 vezes para authorship e de 33,8 vezes para generative AI entre busca livre e busca restrita a título e resumo. O segundo caso é particularmente instrutivo: as declarações obrigatórias de uso de inteligência artificial, instituídas pelos periódicos como medida de transparência, tornaram a busca em texto completo sobre inteligência artificial praticamente inútil. Uma prática de integridade produziu um artefato metodológico, e quem pesquisar o tema sem conhecê-lo superestimará o volume de literatura em mais de trinta vezes.

A conclusão prática não é de desamparo, mas de deslocamento. A pergunta este texto é do estudante ou da máquina é mal formulada, porque exige prova de origem que nenhuma ferramenta disponível fornece com confiabilidade suficiente, e porque o risco de falso positivo recai, pelo mecanismo mesmo dos detectores, sobre os estudantes de menor repertório, isto é, sobre aqueles que mais dependem da escola. A pergunta que a literatura brasileira vem formulando, sobre qual foi a contribuição do estudante e se ela evidencia aprendizagem, é respondível com instrumentos que o professor já domina: avaliação de processo, arguição oral, comparação com a produção anterior do próprio estudante e explicitação prévia das regras de uso.

Essa reformulação tem um custo que não deve ser omitido: todos esses instrumentos exigem mais tempo docente, em turmas que raramente o permitem. Reconhecer o custo faz parte de propor a solução com honestidade. O que este artigo acrescenta é a constatação de que o professor que hoje enfrenta esse problema não encontra, nem no documento que organiza o seu currículo nem na literatura indexada acessível, o parâmetro que lhe permitiria decidir com segurança. Enquanto esse parâmetro não existir, cada escola e cada professor constroem o seu, e a desigualdade de critérios entre escolas torna-se mais uma dimensão da desigualdade educacional brasileira.

REFERÊNCIAS

ALMEIDA, Eduardo de Moura; BARBOSA, Jacqueline Peixoto. Produção de textos na escola: (re)discutindo autoria em tempos de inteligência artificial generativa. Trabalhos em Linguística Aplicada, v. 65, 2026. DOI: 10.1590/01031813v65020268684625. Disponível em: https://doi.org/10.1590/01031813v65020268684625. Acesso em: 19 out. 2026.

AMARAL, Mirian Maia do. Entre o eu e o algoritmo: a autoria em tempos de inteligência artificial generativa (IAGen). Acta Scientiarum. Education, v. 48, p. e78638, 2026. DOI: 10.4025/actascieduc.v48i1.78638. Disponível em: https://doi.org/10.4025/actascieduc.v48i1.78638. Acesso em: 19 out. 2026.

BATISTA, Patrícia; VASCONCELOS, Alana. Cultura digital e inteligência artificial generativa: desafios da autoria híbrida e da ética no uso da IA. Revista Acadêmica de Divulgação Científica, v. 4, n. 2, p. e4217, 2026. DOI: 10.17564/2965-6222.2026.v4.n2.e4217. Disponível em: https://doi.org/10.17564/2965-6222.2026.v4.n2.e4217. Acesso em: 19 out. 2026.

BRASIL. Ministério da Educação. Base Nacional Comum Curricular: Ensino Médio. Brasília, DF: MEC, 2018. Disponível em: http://basenacionalcomum.mec.gov.br/images/historico/BNCC_EnsinoMedio_embaixa_site_110518.pdf. Acesso em: 19 out. 2026.

BRITO ROCHA, Telma; BRANDÃO, Cleyton; PANFILE, Williams. Avaliação da aprendizagem digital segundo o ChatGPT-3.5 e a autoria híbrida humano-IA. AtoZ: novas práticas em informação e conhecimento, v. 13, p. 1-11, 2025. DOI: 10.5380/atoz.v13i0.92219. Disponível em: https://doi.org/10.5380/atoz.v13i0.92219. Acesso em: 19 out. 2026.

COSTA, Nicolas Valverde; GARCIA, Paulo Sérgio. Avaliação educacional e inteligência artificial: o olhar de futuros professores. ARACÊ, v. 7, n. 2, p. 9731-9746, 2025. DOI: 10.56238/arev7n2-297. Disponível em: https://doi.org/10.56238/arev7n2-297. Acesso em: 19 out. 2026.

DORU, Berin; MAIER, Christoph; BUSSE, Johanna Sophie; LÜCKE, Thomas; SCHÖNHOFF, Judith; ENAX-KRUMOVA, Elena; HESSLER, Steffen; BERGER, Maria; TOKIC, Marianne. Detecting artificial intelligence-generated versus human-written medical student essays: semirandomized controlled study. JMIR Medical Education, v. 11, p. e62779, 2025. DOI: 10.2196/62779. Disponível em: https://doi.org/10.2196/62779. Acesso em: 19 out. 2026.

EROL, Gökberk; ERGEN, Anıl; GÜLŞEN EROL, Büşra; KAYA ERGEN, Şebnem; BORA, Tevfik Serhan; ÇÖLGEÇEN, Ali Deniz; ARAZ, Büşra; ŞAHIN, Cansel; BOSTANCI, Günsu; KILIÇ, İlayda; MACIT, Zeynep Birce; SEVGI, Umut Tan; GÜNGÖR, Abuzer. Can we trust academic AI detective? Accuracy and limitations of AI-output detectors. Acta Neurochirurgica, v. 167, n. 1, 2025. DOI: 10.1007/s00701-025-06622-4. Disponível em: https://doi.org/10.1007/s00701-025-06622-4. Acesso em: 19 out. 2026.

GONÇALVES, Sirlene da Silva; BODOLAY, Adriana Nascimento. Análise comparativa entre textos produzidos com e sem auxílio de inteligência artificial no Ensino Médio. ARACÊ, v. 7, n. 10, p. e8677, 2025. DOI: 10.56238/arev7n10-003. Disponível em: https://doi.org/10.56238/arev7n10-003. Acesso em: 19 out. 2026.

HERMANN, Nadja. Ética e inteligência artificial: o que inquieta? Educação & Sociedade, v. 46, 2025. DOI: 10.1590/es.296750. Disponível em: https://doi.org/10.1590/es.296750. Acesso em: 19 out. 2026.

HUA, Haowei; YAO, Co-Jiayu. Investigating generative AI models and detection techniques: impacts of tokenization and dataset size on identification of AI-generated text. Frontiers in Artificial Intelligence, v. 7, 2024. DOI: 10.3389/frai.2024.1469197. Disponível em: https://doi.org/10.3389/frai.2024.1469197. Acesso em: 19 out. 2026.

HYATT, Jon-Philippe K.; BIENENSTOCK, Elisa Jayne; FIRETTO, Carla M.; WOODS, Elizabeth R.; COMUS, Robert C. Using aggregated AI detector outcomes to eliminate false positives in STEM-student writing. Advances in Physiology Education, v. 49, n. 2, p. 486-495, 2025. DOI: 10.1152/advan.00235.2024. Disponível em: https://doi.org/10.1152/advan.00235.2024. Acesso em: 19 out. 2026.

LIANG, Weixin; YUKSEKGONUL, Mert; MAO, Yining; WU, Eric; ZOU, James. GPT detectors are biased against non-native English writers. Patterns, v. 4, n. 7, p. 100779, 2023. DOI: 10.1016/j.patter.2023.100779. Disponível em: https://doi.org/10.1016/j.patter.2023.100779. Acesso em: 19 out. 2026.

MANI, P.; ARULANANDAM, Adaikalam. A stylometric machine learning framework for authorship attribution of ChatGPT and student generated texts. Frontiers in Artificial Intelligence, v. 9, 2026. DOI: 10.3389/frai.2026.1921847. Disponível em: https://doi.org/10.3389/frai.2026.1921847. Acesso em: 19 out. 2026.

MUNIZ, Aline Barbosa Xavier; FERREIRA, Fernanda Vernieri Vivacqua da Costa; FERREIRA, Valdivina Alves. Inteligência artificial na educação: uma análise comparativa dos marcos regulatórios no Brasil e na União Europeia. Ensaio: Avaliação e Políticas Públicas em Educação, v. 34, n. 132, 2026. DOI: 10.1590/s0104-40362026003405346. Disponível em: https://doi.org/10.1590/s0104-40362026003405346. Acesso em: 19 out. 2026.

PINNINTI, Lakshmana Rao. Undergraduate ESL students' use and perceptions of ChatGPT for academic writing purposes. Texto Livre, v. 18, 2025. DOI: 10.1590/1983-3652.2025.58320. Disponível em: https://doi.org/10.1590/1983-3652.2025.58320. Acesso em: 19 out. 2026.

POPKOV, Andrey A.; BARRETT, Tyson S. AI vs academia: experimental study on AI text detectors' accuracy in behavioral health academic writing. Accountability in Research, v. 32, n. 7, p. 1072-1088, 2024. DOI: 10.1080/08989621.2024.2331757. Disponível em: https://doi.org/10.1080/08989621.2024.2331757. Acesso em: 19 out. 2026.

SILVA, Alessandro Siqueira da; ALARCÃO, Davi Taveira Alencar; FARIA, Syd Pereira. Inteligência artificial generativa na educação brasileira: possibilidades e limitações para a prática pedagógica. Revista de Geopolítica, v. 17, n. 7, p. e2846, 2026. DOI: 10.56238/revgeov17n7-086. Disponível em: https://doi.org/10.56238/revgeov17n7-086. Acesso em: 19 out. 2026.

VALENÇA RIBEIRO, Vinicius. Autoria e estilo em diários de leituras de estudantes do Ensino Médio. Linguagem em (Dis)curso, v. 24, 2024. DOI: 10.1590/1982-4017-24-35. Disponível em: https://doi.org/10.1590/1982-4017-24-35. Acesso em: 19 out. 2026.

ZUCOLAR, Ionice. De Emília Ferreiro ao ChatGPT: psicogênese da escrita na era da inteligência artificial. ARACÊ, v. 7, n. 12, p. e10928, 2025. DOI: 10.56238/arev7n12-104. Disponível em: https://doi.org/10.56238/arev7n12-104. Acesso em: 19 out. 2026.