



PRECISÃO OU INCERTEZA? UM ESTUDO SOBRE PLATAFORMAS DE VERIFICAÇÃO DE USO DA IA EM PRODUÇÕES ACADÊMICAS

José Carlos Lucio Maia¹
Ilana Murici Ayres²
Luciana Mamede da Silva³
Katiusse Kelle de Melo Soares⁴
Adenir José de Sousa⁵
Rafael Lima de Carvalho⁶
Rodrigo de Melo Brustolin⁷
Flávio Roldão de Carvalho Lelis⁸
Giovana Beatriz de Souza Lucio Maia⁹
Emylane Celi Mota Brito¹⁰

Resumo: A inteligência artificial (iA) reestrutura as lógicas de produção e de consumo da sociedade, aprimorando o conhecimento e favorecendo a transformação digital de operações condicionadas à atuação humana. Na produção textual, a iA pode ser aplicada em diversas etapas, desde a organização de ideias até a revisão linguística. Dados os riscos à integridade acadêmica, surgem aplicações para verificar o uso de iA na redação. Nesse contexto, questiona-se: essas plataformas oferecem um serviço efetivo? Como resposta, desenvolveu-se uma pesquisa quantitativa, não probabilística, com extração de 100 fragmentos textuais de trabalhos acadêmicos produzidos entre 1980 a 2025, disponíveis abertamente na *internet*, submetendo-os a uma plataforma escolhida aleatoriamente. Em síntese, o experimento mostrou que as expectativas quanto à assertividade do detector não se concretizaram. A aplicação analisada atribuiu até 100% de contribuição de iA para textos datilografados, publicados nas décadas de 1980 e 1990. Também foi realizado um Mapeamento Sistemático da Literatura (MSL) e criada uma nuvem de palavras a partir dos estudos selecionados. Nota-se irreversível o crescente uso das iAs. Os especialistas indicam que esta seja incorporada ao processo de aprendizado, estimulando seu uso ético e transparente, observando princípios pedagógicos. Conclui-se que educadores devem conhecer as limitações das ferramentas, evitando decisões com base em conclusões equivocadas, diante do risco danos aos direitos da personalidade de autores legítimos.

Palavras-chave: Inteligência artificial Autenticidade. Produção de texto. Direito autoral.

¹J. C. L. Maia (👤). Universidade Federal do Tocantins. Palmas, TO. Tribunal Regional Eleitoral de Goiás (TRE-GO). Goiânia, Goiás. Brasil. e-mail: lucio.maia@uft.edu.br.

²I. M. Ayres (👤). Universidade Federal do Tocantins. Palmas, TO. TRE-GO, Brasil.

³L. M. Silva (👤). Universidade Federal do Tocantins. Palmas, TO. TRE-GO, Brasil.

⁴K. K. M. Soares (👤). Universidade Federal do Tocantins. Palmas, TO. TRE-GO, Brasil.

⁵A. J. Sousa (👤). Universidade Federal do Tocantins. Palmas, TO. TRE-GO, Brasil.

⁶R. L. Carvalho (👤). Universidade Federal do Tocantins. Palmas, TO, Brasil.

⁷R. M. Brustolin (👤). Tribunal de Justiça de Goiás (TJ-GO), Goiânia, Goiás. TRE-GO, Brasil.

⁸F. R. C. Lelis (👤). Instituto Federal de Educação, Ciência e Tecnologia do Estado de Goiás. Goiânia, GO. Brasil.

⁹G. B. S. L. Maia (👤). Universidade Evangélica de Goiás. Anápolis, GO. Brasil.

¹⁰E.C.M. Brito (👤). Faculdade Anhanguera de Camaçari. Camaçari, BA. Brasil.

INTRODUÇÃO

A transformação digital promove uma realidade na qual a tecnologia tanto pode impulsionar o desenvolvimento da sociedade quanto servir para práticas ilícitas. Nessa lógica, os Sistemas de Gestão de *Compliance* tornam-se referências para as organizações manterem a conformidade do seu negócio, incutindo a *práxis* de seus colaboradores, não só em termos dos marcos legais, mas, também, em atenção aos padrões éticos de sua conduta. Por consequência natural, no meio acadêmico a preocupação com a integridade das ações (*do*) e não-ações (*do nothing*) se revela quando da criação de entidades como o *International Center for Academic Integrity* (ICAI) e nos crescentes debates acerca de integridade acadêmica.

Não se tem de forma uníssona a data exata do início da lógica hoje popularmente conhecida como inteligência artificial (iA). Alan Turing, um dos pioneiros da computação, em um artigo datado em 1950, apresentou a instigante reflexão se as máquinas poderiam pensar. Para tanto propôs uma solução denominada “*jogo da imitação*”, que também ficou conhecida como “*Teste de Turing*”. Em 1956, John McCarthy propôs o termo “*inteligência artificial*” para se referir “*à engenharia de criar máquinas inteligentes*”. Contudo, apesar do conceito existir há décadas, o uso em massa das iAs só se tornou viável após a exploração sistemática de dados pelas plataformas de redes sociais e o uso da computação em nuvem (Murer, 2025).

No contexto acadêmico, as novas tecnologias podem ser utilizadas para checar a autenticidade das produções textuais, visando, em tese, instigar transparência e fortalecer a confiança nos estudos científicos. Isso é o que se espera de ferramentas tecnológicas como as que se propõem a verificar o possível uso de inteligência artificial (iA) na elaboração de trabalhos acadêmicos. Porém, deve haver cautela quanto às suas potencialidades, assim como remeter um olhar crítico para as restrições de uso, por vezes não verbalizadas pelas empresas fornecedoras, e evitar que a utilização inadvertida se converta em violação de direitos dos autores cujas obras são submetidas à avaliação.

A transformação digital é um fenômeno complexo, multinível e dinâmico que gera uma nova realidade social, ambiental, tecnológica e econômica, na qual múltiplos atores interagem entre si, inseridos em diversificados ecossistemas digitais. Nessa linha, a gestão dos produtos digitais deve levar em consideração os elementos de orientação, cooperação e orquestração para agregar valor e vantagens competitivas. No âmbito da Educação, tais produtos podem estimular o pensamento crítico e o desenvolvimento humano-científico, mas, também, podem servir para realização de práticas antiéticas.

Assim, a integridade revela um amplo espectro de valores e crenças que devem arquitetar os esquemas estruturantes das organizações, o que inclui, sobretudo, todas as partes interessadas. Além de repelir as práticas reprováveis, ela valoriza e estimula a eficácia social, a responsabilidade, a excelência e a transparência. Não por menos, essas perspectivas devem também ser internalizadas nos ambientes escolares, na medida em que a integridade acadêmica pode ser entendida como sendo “*um compromisso, mesmo diante da adversidade, com seis valores fundamentais: honestidade, confiança, justiça, respeito, responsabilidade e coragem*” (ICAI, 2021, tradução livre).

Anteriormente ao uso das tecnologias digitais, seja para obter *insights*, organizar, revisar e ajustar textos, tinha-se um cenário no qual a escrita primitiva sofria interferências externas e, não raro, alterava substancialmente a construção original. Entre aqueles que protagonizavam essas modificações tinha-se o trabalho humano dos revisores/tradutores e daqueles tantos outros



que emitiam opiniões sobre o conteúdo e a forma do texto em construção. Estes agentes externos, por vezes agregando ou suprimindo elementos textuais que não se originaram no intelecto do autor original. Um exemplo amplamente conhecido é o dos textos bíblicos, que há centenas de anos possuem variadas versões, com conteúdos diferentes em uma mesma língua, muito em razão das diversas traduções realizadas.

Atualmente, têm-se um ecossistema de empresas que fornecem tecnologias que se propõem verificar a autenticidade de textos, estimando um percentual de contribuição das iAs. Há outras que oferecem a humanização de textos não humanos. Seja de uma ou de outra forma, o cenário é desafiador se todos os atores envolvidos não tiverem o pleno conhecimento dos benefícios e das restrições dessas tecnologias, que são capazes de alavancar o conhecimento, entendendo-as, sempre, na condição de ferramenta acessória para a atuação humana.

Fato é, os usuários podem criar uma falsa expectativa quanto ao funcionamento dessas tecnologias de verificação de autenticidade, cuja utilização indevida dos resultados pode gerar profundas consequências, tais como embasar sanções ou outros tipos de exposições que afetam o direito de personalidade dos autores e podem resultar em uma demanda judicial.

Sob a perspectiva jurídica, a utilização das iAs quando da produção de textos gera uma série de questionamentos. Por um lado, existem dúvidas sobre a existência e a titularidade de direitos autorais de obras produzidas por meio dessa tecnologia. Por outro vértice, não há limites claros referentes ao uso legítimo (*fair use*) de produções intelectuais para treinar as iAs. Há dúvidas em relação à caracterização dos textos elaborados pela tecnologia como derivações da obra original ou só como paráfrases ou paródias, que são livremente autorizadas pela legislação brasileira (Alves, 2024).

Nessa linha, não se pode ladear os possíveis desdobramentos jurídicos resultantes de falhas de funcionamento e/ou do manejo incorreto de ferramentas de verificação que ocasionem danos morais e materiais a legítimos autores de obras intelectuais. Atribuir indevidamente a alguém a prática de plágio ou de qualquer outra conduta análoga, constitui violação aos direitos da personalidade, atingindo a honra da pessoa ofendida e motivando o dever de reparação (Brasil, 2025).

Diante desse cenário, o presente estudo tem por objetivo examinar as implicações do uso de sistemas de detecção de escrita por inteligência artificial no contexto da integridade acadêmica, com ênfase na avaliação da autenticidade da produção textual. Para tanto, adota-se uma abordagem metodológica em duas etapas complementares.

Na primeira, para constatar se as plataformas de verificação de autenticidade de texto oferecem um serviço efetivo, efetuou-se análise empírica das respostas dadas por uma dessas tecnologias disponíveis gratuitamente na *internet*. Foram submetidos 100 fragmentos textuais à verificação do serviço escolhido. Os textos examinados eram trechos das introduções de trabalhos acadêmicos produzidos entre os anos de 1980 e 2025, os quais também estavam disponíveis de forma digital e gratuita na *internet*.

Entre os achados desta etapa, chamou a atenção o fato de a plataforma escolhida ter indicado altas taxas de contribuição de iA generativa para textos escritos décadas antes da popularização dessa tecnologia. Destaque-se que, entre os fragmentos textuais submetidos à análise, alguns foram redigidos com máquina de escrever mecânica. Essa constatação, por lógico, suscita dúvidas quanto à assertividade dos resultados de autenticidade sugeridos pela ferramenta utilizada. Isso evidencia limitações metodológicas dessas ferramentas e reforça a necessidade de sua utilização crítica e contextualizada em ambientes acadêmicos.



Em uma segunda etapa, realizou-se um mapeamento sistemático de literatura (MSL) na base *Scopus*, voltado à identificação das principais categorias conceituais e tendências de pesquisa sobre integridade, autoria e uso de IA na verificação de originalidade. A síntese dos estudos selecionados foi representada visualmente por meio de uma nuvem de palavras, construída a partir das palavras-chave extraídas e tratadas com técnicas de análise textual assistida por IA.

Os estudos que compuseram o *corpus* de análise do MSL evidenciam que o debate sobre integridade acadêmica diante do crescente uso das iAs ainda está fortemente ancorado em abordagens punitivas e de detecção. Contudo, os especialistas apontam a necessidade de reflexões estruturadas sobre governança, equidade e uso pedagógico responsável da tecnologia na educação.

REFERENCIAL TEÓRICO

Para aprofundar na busca de respostas à questão de pesquisa, este trabalho está fundamentado em estudos sobre: **a)** inteligência artificial, abrangendo seu histórico, conceitos, fundamentos, aplicações e limitações, atuais e futuras; **b)** *compliance* e integridade acadêmica; **c)** linguagem humana, tratando de sua complexidade e da dificuldade de diferenciar formas legítimas de intertextualidade de usos ilícitos de obras alheias; e **d)** aspectos jurídicos referentes aos direitos da personalidade e aos direitos autorais.

Autores como Abid Malik e Amjad (2025), Ardito (2025), Bordalejo *et. al.* (2025), Guan e Han (2025), Hau (2025), Kalmani *et. al.* (2025), Lund *et. al.* (2025), Miller *et. al.* (2025), Mondal (2025), Mulaudzi e Hamilton (2025), Poullet *et. al.* (2025), Popescu-Apreutesei *et. al.* (2025), Prykhodchenko *et. al.* (2025), Stillman (2025), Zhu *et. al.* (2025), Ullah, Bin Naeem, e Kamel Boulos (2024) e Elkhatat, Elsaid e Almeer (2023) tratam da relação entre os detectores do uso das iAs nas produções textuais no contexto da integridade acadêmica. Murer (2025), Suleyman e Bhaskar (2024), Aranha (2023) e Lee (2019) abordam o desenvolvimento da IA em suas obras, relatando suas origens, em termos conceituais, na década de 1950.

Ao longo da história, a IA se dividiu entre a abordagem “baseada em regras” e a de “redes neurais”, sendo esta última a adotada nos modelos tecnológicos aplicados às ferramentas que se popularizaram nos dias de hoje. Apesar de existir há várias décadas, até pouco tempo as iAs eram conhecidas e utilizadas por um número muito restrito de estudiosos da ciência da computação, nos poucos laboratórios com equipamentos avançados o suficiente para realização de experimentos nessa área. Era inviável o desenvolvimento e a utilização de iAs em larga escala, pois a criação destas exige equipamentos com alto poder computacional e grandes volumes de informação disponíveis para treinamento dos algoritmos.

Isso porque, no cerne dessa tecnologia, o algoritmo utiliza uma vasta quantidade de dados para reconhecer padrões e fazer correlações internas a fim de tomar uma decisão e entregar o que foi solicitado pelo usuário (Lee, 2019). O resultado pretendido pode ser tanto a elaboração de um texto com um estilo específico quanto o reconhecimento de padrões que indiquem a probabilidade de a origem deste não ser humana. No entanto, foram necessários avanços tecnológicos em várias frentes para tornar viável a ampla utilização das iAs.

Do ponto de vista científico, ocorreu o desenvolvimento de técnicas de processamento de linguagem natural (PLN), campo no qual se insere a IA generativa. Também houve progresso no setor industrial possibilitando a produção de equipamentos com maior capacidade de



processamento de dados. Além disso, os profissionais da computação passaram a dispor de vastas bases de dados textuais, fruto principalmente das redes sociais, para o desenvolvimento de aplicações de iA. Esse conjunto de fatores permitiu a criação de ferramentas dotadas de versatilidade, autonomia e capacidade de geração de conteúdo, que podem apresentar respostas diferentes para uma mesma questão, simulando reações humanas (Murer, 2025).

Ante a praticamente inevitável incorporação da iA a vários aspectos da vida social, especialmente no campo da educação, atualmente diversos debates têm sido realizados a respeito da legitimidade de aplicação dessa tecnologia. Estudos recentes tratam de potenciais usos legítimos da iA, que podem gerar aumento de produtividades para professores e alunos, tais como suporte para correção gramatical, aprendizado personalizado e geração de *insights* para aprofundamento dos estudos (Sarwanti *et al.*, 2024). Porém, por outro lado, pode haver indevida aplicação da tecnologia na produção de trabalhos inautênticos, ameaçando a credibilidade da produção acadêmica (Spinellis, 2025).

Ainda são poucas as instituições de ensino que possuem diretrizes claras sobre o uso das aplicações baseadas em iA (Ullah, Bin Naeem, e Kamel Boulos, 2024). Diante desta lacuna, com o alegado propósito de proteger a integridade acadêmica, emergem soluções tecnológicas voltadas a identificar o uso de iA em produções textuais.

Todavia estudiosos que analisam as aplicações da iA afirmam que, atualmente, nem pesquisadores humanos nem ferramentas tecnológicas são capazes de distinguir com segurança textos redigidos por humanos dos que foram produzidos por meio da tecnologia generativa (Hakam *et al.*, 2024). A detecção de uso das iAs na redação de trabalhos ainda é imprecisa, o que ressalta o risco de interpretações indevidas dos resultados apresentados, capazes de embasar decisões injustas quanto à autenticidade de trabalhos acadêmicos ou profissionais (Perkins *et al.*, 2024; Gosselin, 2025; Gautam, 2025).

Os resultados das ferramentas de detecção podem ser influenciados por diversos fatores como o estilo de escrita, o idioma, o tipo de *prompt* usado na geração dos escritos ou a aplicação de ferramentas de edição (Herath, Ode e Herath, 2025; Fariello, 2025). Além disso, há risco de homogeneização do discurso e enfiamento dos resultados (Gautam, 2025).

Dessa forma, os especialistas alertam sobre a necessidade de que as pessoas responsáveis pela avaliação de autenticidade de estudos levem em consideração as limitações dos detectores, que não devem ser utilizados isoladamente como base para penalizar estudantes ou profissionais, dado o alto risco de decisões injustas e causadoras de severos danos aos direitos da personalidade dos avaliados (Popescu-Apreutesei, 2025; Bordalejo, 2025).

Considerando as possibilidades de práticas de condutas ilegítimas e usos indevidos das aplicações que utilizam iA, tanto docentes como estudantes revelam preocupações quanto à ética à integridade acadêmica. Os estudiosos afirmam que é preciso elaborar políticas claras sobre os usos legítimos de iA (Lund *et al.*, 2025), da promoção da transparência e do investimento em tecnologia (Mulaudzi e Hamilton, 2025). Tratam ainda da necessidade de ampla capacitação em relação à iA (Ayanwale *et al.*, 2024), incorporando o uso da iA aos aspectos metodológicos das atividades desenvolvidas no ambiente escolar, conciliando o desenvolvimento tecnológico e a integridade acadêmica (Hau, 2025).

Nessa linha, é fundamental levar em consideração as dificuldades estabelecidas à identificação de autenticidade na escrita em razão da complexidade da comunicação humana. Dominici (2020) adverte que é ilusório supor que o avanço tecnológico, baseado em automação e algoritmos, elimine erros ou controle a comunicação, pois esta é complexa e não pode ser



reduzida a modelos lineares. Seus fluxos envolvem múltiplas variáveis e mediações. Os dados não falam por si, então, à medida que a tecnologia avança, a comunicação revela sua hipercomplexidade, assumindo papel decisivo na sociedade.

Diniz (2024) e Eco (2019) tratam da redação de textos acadêmicos e dos cuidados que os estudantes devem ter para não incorrer na prática de plágio. Reconhecem que raramente esse tipo de infração ocorre com transcrição literal de textos alheios, mas de formas sutis, como o pastiche ou falsas paráfrases. Ambos reconhecem que a dimensão do texto elaborado pode ser um bom indicador para quem escreve. Se o trecho redigido for mais curto do que o consultado, é maior a probabilidade de que não tenha ocorrido apropriação indevida de ideia alheia.

Há, ainda, a imperativa necessidade de distinguir entre formas ilícitas de composição textual, com uso fraudulento do trabalho de outra pessoa; de interações legítimas entre obras de diferentes autores, fazendo lembrar a inquietante questão apresentada em “Quase sem querer”: “*Mas quais são as palavras que nunca são ditas?*” (Manfredini Jr. *et al.*, 1986). Desta forma, observa-se que as análises de Nascimento e Brisolara (2019), quando tratam sobre a fronteira entre o plágio e a intertextualidade são pertinentes à temática sob análise.

Os avanços tecnológicos observados nas últimas décadas e a ampliação do uso da *internet* ampliaram o acesso a obras literárias e acadêmicas e o compartilhamento de informações. Isso permite a democratização do conhecimento e promove o desenvolvimento científico, mas também facilita práticas fraudulentas, como o plágio. Tornando a situação ainda mais complexa, observa-se que nem sempre são claros os limites entre o diálogo legítimo entre os textos, com aplicação ética e referenciadas das ideias de outros autores, e a apropriação indevida de obras alheias, que viola o direito autoral (Nascimento e Brisolara, 2019).

O crescente emprego das iAs generativas torna essas questões ainda mais preocupantes. Ocorre que essas tecnologias não utilizam estruturas textuais completamente novas. As aplicações usam uma vasta base de trabalhos desenvolvidos com o intelecto humano para estabelecer seus padrões de elaboração de textos. Logo, a verificação de autenticidade pode incorrer em erros de interpretação quanto ao texto analisado ser ou não de origem humana.

Todavia, mesmo diante das limitações demonstradas pela tecnologia, a forma como os resultados são apresentados pelas aplicações de detecção de iA, aliada à exibição de mensagens que exaltam a confiabilidade do serviço, pode transmitir uma falsa segurança aos usuários que não têm ciência de seu funcionamento, o que pode resultar em sérios danos morais e materiais a autores de trabalhos eventualmente sejam tratados injustamente como inautênticos.

Dessa forma, também são apresentadas reflexões sobre os possíveis desdobramentos jurídicos do tema sob análise. Em uma primeira vertente, analisa-se as implicações referentes a possíveis violações de direitos autorais, como a prática de plágio (Alves, 2024; Lopes e Almada, 2025). Outra abordagem diz respeito aos danos aos direitos da personalidade que podem ser causados por uma falsa atribuição de plágio ou de inautenticidade na produção de um texto profissional ou acadêmico, como a recente decisão do Superior Tribunal de Justiça (STJ) a respeito de falsa atribuição de plágio a autor de texto legítimo (Brasil, 2025), ressaltando o risco do uso das ferramentas de verificação sem plena ciência de seu funcionamento.

METODOLOGIA

Conforme abordado, este estudo tem como propósito instigar uma reflexão sobre o grau de assertividade das plataformas que se propõem a verificar a contribuição das iAs na produção



de textos acadêmicos, tendo como base uma análise longitudinal abrangendo o período de 1980 a 2025. Além disso, os pesquisadores buscaram identificar o atual estágio das discussões sobre a aplicação das iAs na educação, no contexto da integridade acadêmica, realizando um MSL para construir uma nuvem de palavras com os termos mais recorrentes nos estudos selecionados.

Essa proposta surge em resposta ao necessário questionamento acerca da efetividade dessas plataformas, sobretudo diante de situações em que se atribuem a participação de iA em textos elaborados em épocas nas quais tais tecnologias não estavam disponíveis ao público em geral. A adoção dessa abordagem justifica-se como forma de evitar a consolidação de uma falsa percepção de assertividade comprovatória destas plataformas. Ressalte-se que as respostas que são apresentadas podem levar à interpretação injusta e equivocada sobre a real contribuição das iAs quando do processo de escrita, redundando em falsos juízos de valor na análise e na avaliação de trabalhos acadêmicos.

Na primeira fase, esta pesquisa se apoia no método dedutivo, pois parte de uma premissa geral (os detectores de iA são efetivos), confrontando-a com as evidências empíricas apuradas para confirmar ou refutar sua validade. O objetivo é verificar a hipótese inicial de que tais plataformas podem induzir conclusões falsas sobre a real contribuição das iAs, gerando, assim, uma zona cinzenta quanto à extensão dessa eventual participação não humana.

Assim, quanto aos seus objetivos, caracteriza-se como descritiva e exploratória, na medida em que organiza e sistematiza dados numéricos conexos aos percentuais de contribuições das iAs em trabalhos acadêmicos, descrevendo as incidências ao longo do período-alvo. Não obstante, empreende-se, ainda, a explicitar traços da impossibilidade histórica de eventuais contribuições das iAs em datas anteriores à sua disponibilidade pública.

Em termos de sua natureza, pode ser classificado como sendo um estudo quantitativo não probabilístico. Neste estudo, foram coletados e analisados dados numéricos referentes aos percentuais de aplicação de iA registrados por uma plataforma de verificação de autenticidade textual. Para tanto, fragmentos textuais de 100 trabalhos acadêmicos (evidências), produzidos entre os anos de 1980 e 2025, foram submetidos ao processamento de um detector de uso de iA, aleatoriamente escolhido. Com base nos dados obtidos, identificou-se distribuições, padrões temporais e variações entre os diferentes tempos-alvo.

Quanto à plataforma de verificação utilizada na pesquisa, em atenção à grande diversidade disponível ao público em geral, por conveniência foi selecionada uma de acesso gratuito, capaz de realizar cruzamento com outras plataformas semelhantes. Em respeito aos princípios éticos e de confidencialidade aplicáveis à pesquisa científica, a identidade da ferramenta escolhida será preservada.

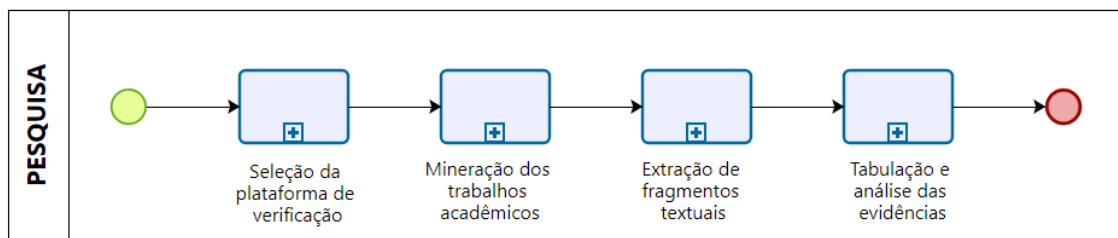
A mineração dos textos acadêmicos (evidências) deu-se por meio da extração de 100 dados a partir da consulta à plataforma de busca Google Acadêmico[®]. Optou-se por esta base para busca das amostras em razão da sua disponibilidade gratuita ao público em geral, sendo elegidas *strings* específicas para garantir a recuperação de evidências alinhadas a um tema-alvo (*dissertação, tese, integridade e pdf*). Não houve nenhuma espécie de filtro sobre disciplinas e temas, assim obteve-se estudos de diferentes áreas do conhecimento, o que forneceu um conjunto com grande variedade de estilos de linguagem.

Estrategicamente, de cada trabalho acadêmico verificado (dado) foi extraído um fragmento textual da seção composto por três parágrafos (evidências), como forma de compor o plano amostral, o qual privilegiou o conteúdo textual contido na subseção “Introdução” de cada publicação-alvo, garantindo, assim, padronização quando da composição da amostra.



Na Figura 1 tem-se a sistematização das etapas da pesquisa.

Figura 1 - Sistematização das etapas da pesquisa.



Fonte: Elaborado pelos autores via Bizagi Modeler®

Na segunda fase, realizou-se um MSL a fim de identificar e classificar as abordagens teóricas e empíricas existentes sobre integridade acadêmica, autenticidade textual e uso de inteligência artificial na avaliação da originalidade de textos. A pesquisa foi conduzida na base *Scopus* com a utilização da seguinte *string* de busca: (“*academic integrity*” and “*AI detection*”) restringindo a pesquisa ao *título*, ao *resumo* e às *palavras-chaves* de artigos de acesso aberto, publicados entre 2020 e 2025, em *inglês*, *português* e *espanhol*.

Foram selecionados 18 artigos científicos, que foram salvos em formato *.pdf*. Na sequência, com uso da ferramenta PDF 24, da GmbH, os arquivos *.pdf* foram agrupados e convertidos em um grande bloco de texto no formato *.txt*, com 991.610 caracteres. Com apoio da ferramenta de IA Chat GPT, da Open AI, efetuou-se uma primeira etapa de “limpeza” para remoção das chamadas *stop words* (palavras vazias). Com isso, busca-se remover do texto analisado palavras muito recorrentes que, por si, não agregam sentido, como artigos, conjunções e nomes próprios. Após, com uso da mesma ferramenta realizou-se a contagem das palavras, identificando as 100 mais recorrentes no texto.

A partir dos dados obtidos, utilizando a aplicação *online* WordArt.com, elaborou-se uma nuvem de palavras, representativa das tendências conceituais predominantes, permitindo a visualização dos eixos temáticos centrais. Essa segunda etapa reforça o arcabouço teórico necessário à compreensão crítica dos resultados empíricos, estabelecendo uma relação com os debates sobre transparência, responsabilidade e ética no uso das iAs para fins de verificação da produção acadêmica.

Demonstrados os aspectos metodológicos, passa-se na próxima seção a apresentar os resultados obtidos e as discussões a respeito destes, destacando a análise da distribuição dos percentuais de uso de IA atribuídos pela ferramenta, as limitações técnicas e as implicações éticas e jurídicas que decorrem de possíveis interpretações equivocadas quanto às respostas dadas pelas plataformas de verificação textual, especialmente quanto à autenticidade das produções acadêmicas e aos riscos de falsos positivos.

RESULTADOS E DISCUSSÕES

Preliminarmente, antes da apresentação dos dados apurados no período-alvo (1980 a 2025), importa destacar que a plataforma de verificação de contribuição de IA utilizada nesta pesquisa gera mensagens aos usuários após o processamento dos fragmentos de texto inseridos para análise (evidências).

No Quadro 1 têm-se mensagens extraídas da referida plataforma, resultantes da submissão de um fragmento textual de teste, inicialmente escrito de forma manuscrita em uma



folha de papel e, posteriormente, digitado na plataforma-alvo, obtendo-se mensagens padrão em duas versões linguísticas: Inglês (Inglaterra) e Português (Portugal).

Quadro 1 - Mensagens da plataforma de verificação de contribuição de iA

Inglês	Português
<i>“Detected AI writing patterns in your text”</i>	<i>“Detectámos padrões de escrita de IA no teu texto”</i>
<i>“71% of your text “Inteligência artificial” shows signs of AI generation”</i>	<i>“71% do seu texto «Inteligência artificial» parece ter sido gerado por IA”</i>
<i>“71% AI content”</i>	<i>“71% Conteúdo de IA”</i>

Fonte: adaptado pelos autores

Em atenção às mensagens padrão da plataforma utilizada nesta pesquisa, nada impede que os usuários sejam levados a acreditar que o fragmento textual em análise apresente padrões de escrita não humanos e que tenha sido gerado por uma iA, o que fragiliza, indiscutivelmente, a originalidade e a conduta dos autores no processo de construção textual.

Nessa linha, importa destacar que aqueles que recorrem a essas plataformas de verificação de contribuições de iA, especialmente quando do uso para fins avaliativos ou profissionais, devem conhecer minimamente os parâmetros lógicos dos componentes desses sistemas. Caso isso não ocorra, o que não se espera, corre-se o risco de que sejam realizados julgamentos equivocados sobre as atividades alheias com base em premissas frágeis.

Prosseguindo com a análise dos resultados do experimento conduzido, considerando o período-alvo e a sistemática de seleção de dados e extração das evidências, na Tabela 1 têm-se as quantidades de trabalhos acadêmicos (dados) que compuseram o campo de análise, com destaque para as suas categorias.

Tabela 1 - Quantidade de trabalhos acadêmicos por ano e por categoria

Categoria	1980	1985	1990	1995	2000	2005	2010	2015	2020	2025
Tese	-	1	-	1	1	1	1	1	1	2
Dissertação	10	6	6	8	4	1	4	3	5	5
Artigo	-	3	3	1	5	7	5	6	4	3
Monografia	-	-	1	-	-	-	-	-	-	-
Resenha	-	-	-	-	-	1	-	-	-	-

Fonte: elaborado pelos autores.

Quanto à escolha dos dados, destaca-se que foram selecionados os 10 primeiros trabalhos acadêmicos apresentados pelo Google Acadêmico®, sem preferência de categoria, de área do conhecimento ou de origem dos estudos. Dessa forma, obteve-se uma amostragem aleatória e não probabilística, com ampla variabilidade quanto aos estilos de redação dos textos submetidos à análise na plataforma adotada no presente estudo para verificar o uso das iAs.

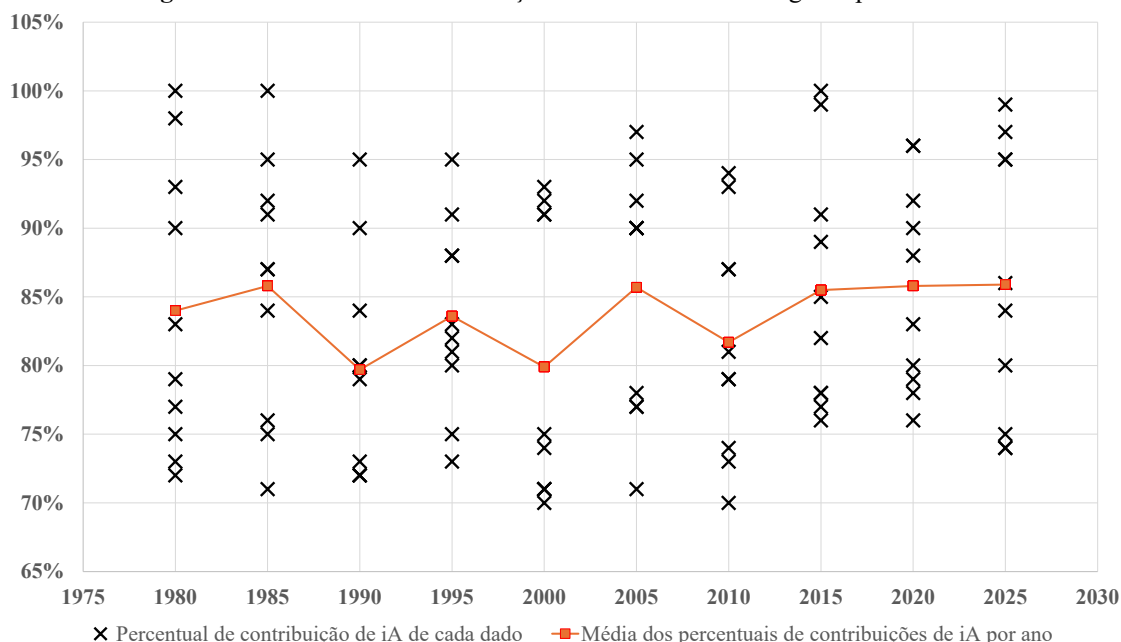
Cabe destacar que embora a iA tenha se popularizado de forma gratuita e de livre acesso apenas nessa última década, constatou-se por meio dos dados coletados que a plataforma de autenticidade adotada neste estudo atribuiu percentuais de contribuições não humanas a textos produzidos em décadas anteriores à sua disseminação. Em tese, tal indicação destoa da



realidade histórica, uma vez que, até então, essa tecnologia não estava disponível nos períodos analisados.

Na Figura 2 têm-se a dispersão dos percentuais das contribuições de iA atribuídos pela plataforma de verificação utilizada no estudo e o comportamento das respectivas médias por ano ao longo do período-alvo (1980 a 2025).

Figura 2 - Percentuais de contribuições de iA e média ao longo do período-alvo.



Fonte: Elaborado pelos autores

É conveniente mencionar que entre os dados anteriores a 1995, apenas um dos selecionados não correspondia a uma versão digitalizada de trabalho datilografado. Isso evidencia que os resultados não poderiam, em tese, ser interpretados como uma eventual “contribuição” de iA, mas, sim, como uma aderência formal da estrutura linguística do texto em análise ao padrão internalizado pela plataforma quando seu processo de aprendizado.

No período-alvo, constatou significativa dispersão dos percentuais de contribuições de iA quando da análise das evidências extraídas dos dados e submetidas à verificação na plataforma escolhida para esta pesquisa. A amplitude dos resultados variou desde 70% a 100%, destacando registros mais elevados nos anos de 1980, 1985, 2005 e 2015, com ocorrências que alcançaram 100%. Ainda que se observem valores que orbitam a faixa de 70% a 75%, não se pode assumir a homogeneidade dos indicadores, em razão de subsistirem *outliers*, que revelam desigualdade no fenômeno em estudo ao longo do período-alvo.

Os achados da pesquisa podem ser analisados sob as diversas perspectivas abordadas no referencial teórico, observando como essas temáticas se inter-relacionam. No que se refere ao desenvolvimento das tecnologias baseadas em iA, percebe-se que as plataformas de detecção de autenticidade de textos enfrentam limitações de desempenho, sobretudo devido à sua estrutura lógica subjacente. A tecnologia tenta distinguir textos humanos de conteúdos gerados pelas iAs utilizando métricas estatísticas derivadas de sua modelagem computacional, o que pode resultar em elevadas taxas de falso positivo ou falso negativo.

Como é comum às iAs, os algoritmos que determinam seu comportamento são desenvolvidos sob uma estrutura heurística, baseada em probabilidade. Esta é desenvolvida não



visando resultados perfeitos, mais sim apresentar respostas boas o suficiente às demandas dos usuários. No caso específico das plataformas que fazem a verificação de uso de iA na elaboração de escritos, os fragmentos submetidos à verificação são analisados sob diferentes aspectos linguísticos. Então, a redação é confrontada com a base de dados da ferramenta, que identifica padrões típicos de uma escrita humana ou de textos produzidos por iA generativa, retornando ao usuário o resultado dessa comparação.

Observa-se, portanto, que os resultados da verificação podem variar de acordo com uma série de fatores. Por melhor que tenha sido o desenvolvimento da ferramenta e o conjunto de dados utilizado para o seu treinamento, esse tipo de aplicação possui restrições em sua capacidade de identificar se um texto foi ou não produzido por meios artificiais. Todavia, essas limitações não são claramente comunicadas e os usuários podem confiar que ela de fato seja capaz de detectar com precisão o uso de iA nas produções escritas. Essa percepção equivocada pode provocar graves consequências, como a utilização da análise feita pela ferramenta para embasar sanções decorrentes de se considerar um trabalho como inautêntico.

A análise dos resultados evidencia que, embora as plataformas de detecção de uso de iA se apoiem em modelos estatísticos e linguísticos sofisticados, ainda carecem de robustez para distinguir com precisão a origem humana ou artificial dos textos. Isso ocorre porque a modelagem de tais sistemas é construída sobre inferências probabilísticas, não sobre verificações factuais. Ou seja, a análise é feita com base em padrões previamente reconhecidos pela ferramenta e não de uma comprovação objetiva sobre a forma de construção do texto.

Dessa forma, sutis diferenças de estilo de redação, coesão ou vocabulário empregado pelo autor podem ser equivocadamente interpretadas como indícios de intervenção não humana. Apesar da aparente precisão estatística dos resultados das verificações, com atribuição de um percentual de uso de iA na elaboração dos textos, estas respostas são apenas estimativas sujeitas a erros de classificação. Em outros termos, o resultado depende mais da semelhança do texto com os padrões internalizados pelo modelo do que da efetiva participação de uma iA na sua redação. Contudo, tais margens de incerteza não são, em geral, comunicadas de forma transparente aos usuários finais (consumidores).

Constatou-se que a plataforma utilizada para verificar a contribuição de iA nos dados desta pesquisa não oferece os resultados fidedignos, pois apontou o uso desta tecnologia em trabalhos produzidos antes da sua disponibilidade ao público em geral. Sim, tal cenário pode levar seus usuários a atribuírem falsamente o uso de iA, gerando implicações éticas e jurídicas.

A compreensão adequada do funcionamento dessas ferramentas torna-se, portanto, condição indispensável para o seu uso responsável. Especialmente no contexto educacional, conforme indicam os estudos citados no referencial teórico, é essencial que docentes e discentes sejam capacitados quanto ao real alcance e às limitações das tecnologias baseadas em iA generativa, compreendendo suas possibilidades de aplicação ética no processo de ensino e aprendizagem. O domínio conceitual sobre o modo como elas processam dados, identificam padrões e produzem inferências permite reduzir o risco de interpretações equivocadas, favorecendo o uso consciente e transparente da inovação.

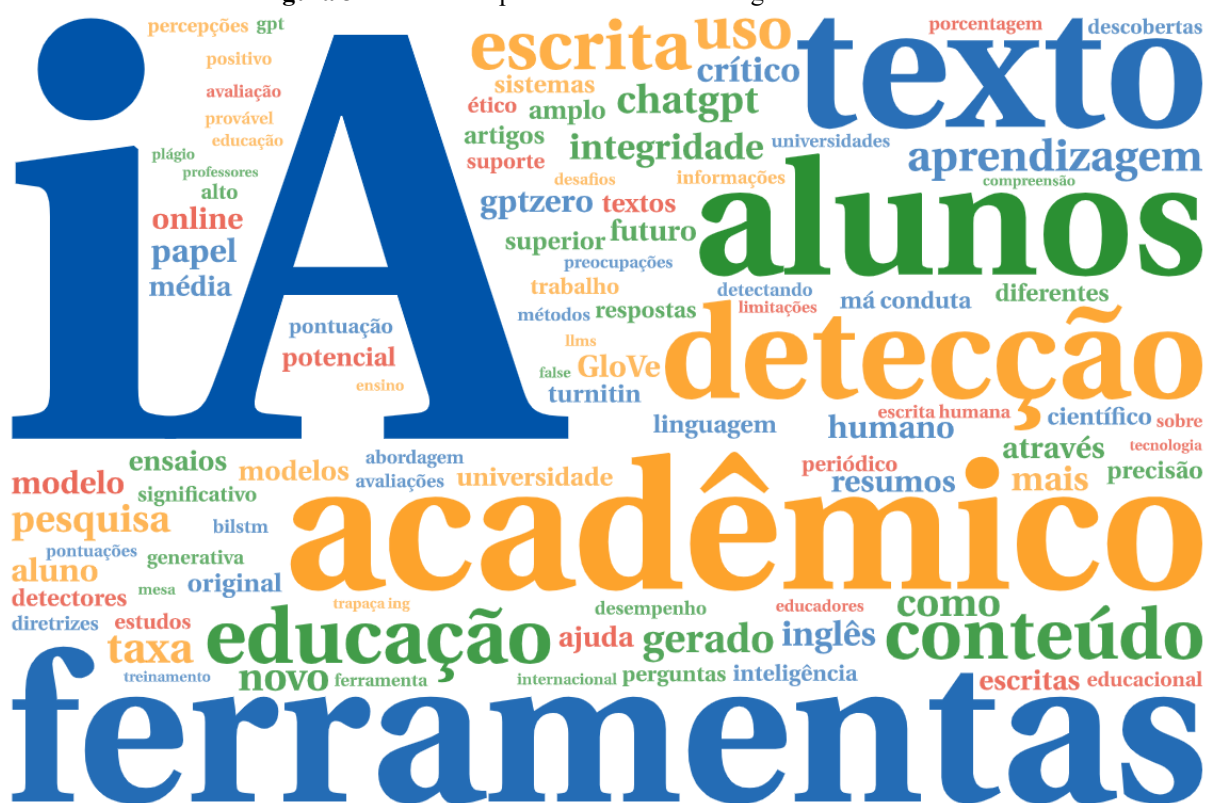
Também é importante que os usuários dessa tecnologia, especialmente quando da análise da originalidade de produções acadêmicas, tenham conhecimento prévio sobre o funcionamento destas ferramentas e compreendam, minimamente, a estrutura do modelo de negócio no qual estão sendo inseridos, conscientemente ou não. Devem, para tanto, levar esses fatores em consideração ao ponderar suas avaliações quanto à real contribuição da iA na



redação de textos acadêmicos, minimizando os efeitos da zona cinzenta da produção de trabalho não humano.

Quanto ao MSL realizado, identificou-se um conjunto de estudos que tratam de maneira crescente os desafios da integridade acadêmica diante da disseminação da iA nos ambientes educacionais, revelando três principais focos de atenção dos especialistas: **a)** preocupação com comportamento desleal e detecção de conteúdo inautêntico; **b)** impactos pedagógicos no desenvolvimento da escrita e aprendizagem; e **c)** busca por estratégias institucionais para uso ético e responsável das iAs. Na Figura 3, tem-se a representação proporcional do conjunto de palavras presentes nos estudos que compuseram o *corpus* de análise dessa etapa da pesquisa.

Figura 3 – Nuvem de palavras sobre iA e integridade acadêmica.



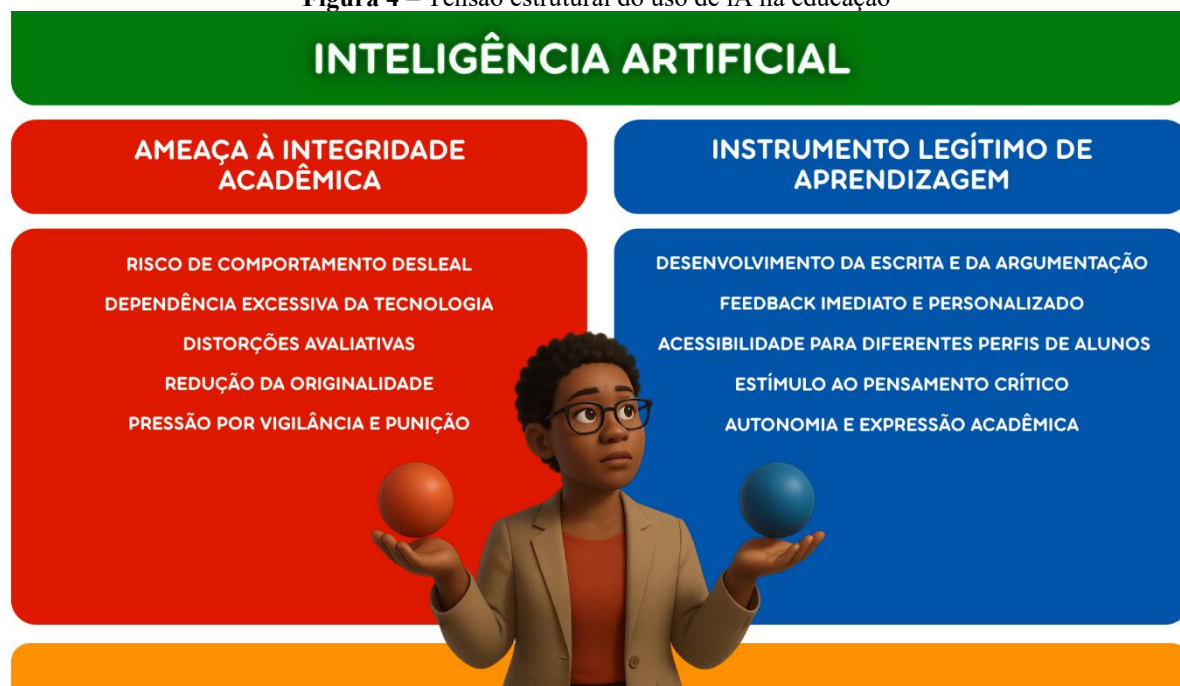
Fonte: Elaborado pelos autores

A nuvem de palavras construída a partir dos textos selecionados reforça essa distribuição temática. No infográfico produzido, termos como detecção, alunos, ferramentas, acadêmico, texto e escrita aparecem com forte destaque, indicando a centralidade do debate sobre originalidade da produção textual estudantil e o uso de métodos de aferição. Palavras como ético, integridade, crítico e papel surgem, porém com menor relevância visual, o que demonstra que a produção científica ainda se concentra mais nos mecanismos de detecção e de punição do que em abordagens formativas e de governança.

Porém, parte da literatura reconhece que ferramentas de iA também podem apoiar o desenvolvimento da escrita, a personalização da aprendizagem e o pensamento crítico, quando utilizadas com transparência e propósito educativo. Assim, os resultados convergem para uma tensão estrutural entre enxergar a iA como uma ameaça à integridade acadêmica em contraposição a encará-la como um instrumento legítimo de aprendizagem. A Figura 4 contém uma representação dessas visões conflitantes.



Figura 4 – Tensão estrutural do uso de iA na educação



Fonte: Elaborado pelos autores

Esse paradoxo pode ser observado nos estudos analisados, mostrando que as instituições de ensino têm adotado formas distintas de lidar com as complexas questões referentes à aplicação das iAs na educação. Nesse cenário, ainda repleto políticas pautadas na detecção e na vigilância convivem com iniciativas de uso pedagógico, evidenciando que a discussão ainda está em processo de maturação.

De forma ampla, considerando o consenso quanto à irreversibilidade do crescimento do uso das iAs, a literatura mapeada aponta no sentido de que a discussão sobre a aplicação dessa tecnologia na educação sob o viés da integridade acadêmica encontra-se em transição. Os estudos indicam um avanço de uma postura meramente fiscalizatória para uma abordagem sistêmica, que enfatiza as responsabilidades das pessoas envolvidas e que busca utilizar as iAs com função educativa. Nesse passo, a orientação é de que as tecnologias de iA generativa sejam efetivamente incorporadas ao processo de aprendizado, observando princípios pedagógicos claros. Nesse contexto surgem espaços para contribuições teóricas e práticas, como as propostas deste estudo.

As iAs podem ser vistas como ferramentas de mediação cognitiva, capazes de ampliar a personalização da aprendizagem, apoiar processos avaliativos e fomentar a reflexão crítica sobre a própria produção do conhecimento. No entanto, tal integração requer políticas institucionais consistentes e constante capacitação dos docentes a fim de que estes desenvolvam competências digitais éticas. Isso permitirá que tanto os educadores quanto os discentes compreendam não apenas o funcionamento das tecnologias, mas também suas implicações epistemológicas e valorativas no processo de ensino-aprendizagem.

Mais do que um avanço tecnológico, a incorporação das iAs na educação representa uma inovação humana, que exige ética, responsabilidade e intencionalidade pedagógica. Mesmo com suas imensas potencialidades e capacidades disruptivas, essa tecnologia continua sendo uma ferramenta para a realização das atividades humanas. No âmbito educacional, ela tanto pode ser aplicada em atos ilícitos ou para embasar avaliações injustas, quanto pode ser



corretamente utilizada, com exploração adequada de suas capacidades, e render bons frutos no campo educacional, com observância dos princípios de integridade acadêmica.

O verdadeiro poder transformador encontra-se, portanto, no ser humano. Cabe a todos os envolvidos no processo educacional compreender que o processo de aprendizado se desenvolve e se fortalece com as interações entre as pessoas. Assim, não basta somente a implementação de novos instrumentos, por mais avançados que sejam. É necessário que os envolvidos sejam capazes de explorar o potencial que as ferramentas tecnológicas têm para aprimorar as formas como interagem uns com os outros. É nesse espaço relacional que a tecnologia pode adquirir sentido formativo e promover uma educação orientada por valores humanos, criatividade e pensamento reflexivo.

No campo da integridade acadêmica e do *Compliance*, a situação observada suscita reflexões éticas relevantes. As plataformas de verificação de autenticidade textual, ao se apresentarem como instrumentos de preservação da integridade, assumem responsabilidade proporcional ao impacto que seus resultados podem gerar sobre a reputação e a credibilidade de autores e instituições. Assim, espera-se que atuem com transparência, comunicando de forma clara suas limitações, metodologias e margens de incerteza.

Quando esse dever de informação não é observado, o que se verifica é a transferência indevida do ônus da prova para o usuário, que passa a interpretar como conclusivos resultados meramente probabilísticos. Tal conduta contraria os princípios basilares do *Compliance* e da ética corporativa, que exigem integridade, diligência e veracidade na comunicação com os consumidores, especialmente quando os produtos ofertados têm potencial de afetar direitos da personalidade e valores fundamentais ligados à honestidade intelectual.

No caso concreto, sob uma perspectiva ética, chama a atenção o fato de que, após realizar gratuitamente a análise, a plataforma de verificação adotada neste estudo oferece, de forma destacada, uma solução para “humanizar” a redação avaliada. Entretanto, esse recurso só está disponível na versão paga, sob a promessa de fazer “o texto parecer escrito por uma pessoa” (“*make text sound human-written*”), ou seja: trata-se de um serviço de edição por meio de IA com propósito de tornar o conteúdo escrito “exclusivo”.

No que se refere à complexidade da comunicação humana e das diferentes formas de intertextualidade, nota-se que indubitavelmente, há o relevante interesse social em identificar e combater o plágio autoral e outros tipos de violação aos direitos autorais, seja no Brasil ou no mundo. Contudo, deve-se ter o cuidado de diferenciar esse tipo de ilícito das outras formas legítimas de intertextualidade, sobretudo diante de um contexto tecnológico que facilita o acesso, a interação e a reprodução de conteúdo no ambiente digital.

O zelo pela integridade intelectual não pode deixar de levar em conta a complexidade da linguagem humana e o caráter relacional que permeia toda produção textual. O elemento da intertextualidade é inerente à comunicação, sendo estruturante na criação literária, científica e acadêmica. Nem toda semelhança textual, portanto, traduz cópia ou apropriação indevida, mas pode representar, ao contrário, o diálogo natural entre diferentes vozes e referenciais teóricos. O risco de confundir esses fenômenos, ao atribuir automaticamente a textos humanos a pecha de inautenticidade com base em padrões algorítmicos, não apenas distorce a análise, como também ameaça a liberdade criativa e o desenvolvimento do pensamento científico.

Fato é, a atribuição imerecida da prática ilegal, além de provocar danos individuais, também pode prejudicar a coletividade ao inibir a produção intelectual. Neste ponto, verifica-se que a temática dessa pesquisa transcende os campos da tecnologia e da educação e alcança



o domínio jurídico, quando potencialmente afeta direitos fundamentais vinculados à honra, à imagem e à reputação dos autores.

Do ponto de vista jurídico, a prática de indevidamente conferir a alguém a acusação de “plagiador” pode gerar o dever de indenizar pelos danos morais. Nesta linha, em recurso recentemente julgado pelo STJ, foi mantida a decisão da Justiça do Distrito Federal de indenizar em R\$ 50.000,00 um autor pelos danos morais sofridos com a imputação de plágio não praticado (Brasil, 2025).

Nesse contexto, as ferramentas tecnológicas, sobretudo a iA, podem servir de apoio para a justificada preocupação quanto à integridade da produção acadêmica. Todavia, não se pode deixar de considerar a responsabilidade das empresas que as disponibilizam, e igualmente a dos usuários que se pautam exclusivamente nestas “informações” para impor a outros drásticas sanções pela suposta prática de “plágio” ou outras exposições indevidas.

Essa discussão pode ser ampliada para que sejam criados e fortalecidos marcos regulatórios que delimitem com clareza as responsabilidades dos envolvidos, abrangendo tanto as empresas fornecedoras dos serviços quanto os usuários que se valem destes resultados para avaliar produções textuais de terceiros. Dessa forma, garantindo transparência nos critérios de funcionamento e estabelecendo salvaguardas à dignidade intelectual dos autores, seria possível conjugar as responsabilidades técnicas, éticas e jurídicas para buscar um equilíbrio entre a inovação tecnológica e a proteção dos direitos fundamentais no ambiente acadêmico.

Em síntese, o estudo reforça a importância de combinar o discernimento, a prudência ética e a responsabilidade jurídica no uso das ferramentas de detecção de uso de iA na elaboração de textos, a fim de que o avanço tecnológico ocorra sem prejuízo à proteção da autoria intelectual e dos direitos dos autores.

Por fim, sugere-se que as empresas que atuam neste segmento mercadológico informem claramente seus usuários e revisem os textos das chamadas para evitar a apropriação inadvertida dos resultados fornecidos pelas plataformas, sobretudo na expectativa de se determinar o grau de contribuição da iA quando da construção de trabalhos acadêmicos.

CONCLUSÃO

No atual estágio de desenvolvimento do processo de Transformação Digital pelo qual passa a humanidade, as atenções de vários segmentos sociais e setores econômicos se voltam para o acesso ampliado às iAs e para as múltiplas possibilidades que esta inovação tecnológica pode potencializar. O emprego de modelos generativos, previamente treinados com vastos bancos de dados oriundos de produções humanas, suscita intensos debates acerca da ética e da violação de direitos autorais.

Invariavelmente, a constante e disruptiva evolução tecnológica, associada ao crescimento da população, amplia a demanda por bens e serviços com rapidez, escala e qualidade. Para atender a tais exigências crescentes, tornam-se indispensáveis recursos informacionais capazes de responder às necessidades de consumo e de distribuição quase imediatas. Nesse contexto, o uso de sistemas inteligentes revela-se inevitável em diferentes cenários socioeconômicos.

Soluções apoiadas em iA vêm favorecendo não apenas a produtividade e a qualidade, mas, também, a inovação. Cotidianamente surgem novos produtos e serviços que dependem de análise de imensos volumes de dados para serem viabilizados, ações possíveis somente com o



suporte de mecanismos automatizados de alta *performance*. Observa-se, assim, um avanço exponencial e irreversível da incorporação desses recursos tecnológicos ao contexto social.

Entretanto, apesar das amplas oportunidades que se descortinam, a utilização de tais tecnologias ainda requer o amadurecimento sob múltiplas perspectivas. Ao desenvolvimento técnico devem ser examinadas as implicações sociais, éticas e jurídicas das soluções disponíveis e disponibilizadas, assim como as interações entre os diferentes atores envolvidos.

Necessário se faz evidenciar que o uso massivo de tecnologias baseadas em iA só se tornou viável recentemente, na última década, após o desenvolvimento de equipamentos com poder computacional suficiente para processar dados em larga escala (Murer, 2025), tanto o é que em “2010, quase ninguém falava seriamente sobre iA” (Suleyman e Bhaskar, 2024).

Historicamente, o mês de novembro de 2022 ficará marcado com o lançamento ao público em geral do ChatGPT (Aranha, 2023), momento no qual suscita uma nova onda de expectativas e desafios para todos os atores que compõem esse ecossistema, demandando mais investimento, pesquisa e desenvolvimento, sobretudo, em razão da competitividade acirrada de seus *players*.

No presente estudo, constatou-se que a plataforma destinada à detecção de contribuição de iA em textos acadêmicos não apresentou *performance* satisfatória, sobretudo quando da análise de fragmentos textuais utilizados como evidências empíricas. O uso imprudente destes recursos tecnológicos, fundamentando aplicação de sanções relacionadas à ausência de originalidade, pode acarretar graves repercussões sociais e jurídicas, tanto para os indivíduos diretamente atingidos, quanto para a coletividade.

Tal assertiva se apoia no fato de a plataforma adotada no estudo ter emitido mensagens de contribuição de iA em fragmentos de trabalhos acadêmicos (evidências) elaborados em datas muito anteriores à popularização desse recurso tecnológico, explicitando falhas significativas na acurácia desta solução tecnológica, refletindo em frustração de expectativa junto aos usuários.

Como limitação da pesquisa, reconhece-se que a análise foi conduzida com base em dados coletados a partir de uma única plataforma de verificação de uso de iA, ainda que esta declare realizar cruzamentos com outras cinco ferramentas semelhantes. Tal delimitação não permite generalizar os resultados para o amplo conjunto de aplicações existentes no mercado, cada uma com lógicas de treinamento, algoritmos e bases de dados próprias.

Não obstante, essa limitação não compromete o objetivo central do estudo, que é fomentar a reflexão crítica sobre a prudência necessária ao uso dessas tecnologias e evidenciar a urgência de se conhecer suas restrições técnicas e operacionais. Mais do que oferecer uma análise comparativa entre ferramentas, a pesquisa buscou destacar o papel da consciência ética e do discernimento humano na interpretação dos resultados automatizados, reafirmando que o avanço tecnológico deve ser acompanhado de conhecimentos basilares dos usuários quanto ao funcionamento das ferramentas e de lisura na atuação das empresas que ofertam tais serviços.

Como contribuição teórica, o estudo evidencia que a confiabilidade percebida das plataformas de detecção é um fenômeno social mediado por expectativas de precisão na análise realizada por meio do algoritmo, carecendo de exame ético e pedagógico. No plano prático, propõe-se a construção de políticas institucionais de uso responsável das iAs, integrando-as de modo deliberado ao processo educativo na formação acadêmica.

Além disso, contribui-se com a recomendação para que os prestadores desses serviços assumam a obrigação de informar, de maneira inequívoca, as limitações da tecnologia disponibilizada para os seus usuários (consumidores), especialmente no que se refere à



expectativa de que tais recursos tecnológicos sejam capazes de detectar, de maneira confiável e irrefutável, eventuais contribuições de IA quando da produção de textos. Nesse sentido, tendo em vista a necessária transparência desses serviços, que têm o potencial de prejudicar a produção intelectual humana, é pertinente ainda que prestem informações claras sobre a origem dos dados utilizados na aferição de autenticidade textual.

Por fim, independentemente da implementação dos alertas sugeridos, é essencial que os usuários, especialmente integrantes da comunidade acadêmica, estejam cientes das limitações do funcionamento das plataformas de verificação de autenticidade de textos, evitando, assim, expectativas frustradas e aplicações indevidas da tecnologia.

AGRADECIMENTOS

Ao Tribunal Regional Eleitoral do Estado de Goiás (TRE-GO), à Universidade Federal do Tocantins (UFT) e ao Instituto Federal de Goiás (IFG).

Ao Programa de Pós-Graduação em Governança e Transformação Digital da UFT.

REFERÊNCIAS

- ALVES, Victor Habib Lantyer de Mello. Inteligência artificial generativa e direito autoral: investigando os limites do uso justo na era da tecnologia. *Revista de Direito, Inovação, Propriedade Intelectual e Concorrência*, Encontro Virtual, v. 10, n. 1, p. 1-19, jan./jul. 2024.
- ARANHA, Christian. *Como a tecnologia da Inteligência Artificial chegou até o ChatGPT*. MIT Technology Review Brasil, 20 abr. 2023. Disponível em: <https://mittechreview.com.br/como-a-tecnologia-da-inteligencia-artificial-chegou-ate-o-chatgpt/>. Acesso em: 20 ago. 2025.
- ARDITO, Cesare Giulio. *Generative AI detection in higher education assessments*. New Directions for Teaching and Learning, [S.l.], n. 2025, p. 11–28, 2025.
- AYANWALE, Musa Adekunle et al. *Examining artificial intelligence literacy among pre-service teachers for future classrooms*. Computers and Education Open, [S.l.], v. 6, art. 100179, 2024.
- BORDALEJO, Barbara et al. “Scarlet Cloak and the Forest Adventure”: a preliminary study of the impact of AI on commonly used writing tools. *International Journal of Educational Technology in Higher Education*, [S.l.], v. 22, art. 6, 2025.
- BRASIL. Superior Tribunal de Justiça. Recurso Especial n. 2.111.826/DF, Relator: Ministro Antonio Carlos Ferreira, Brasília, DF, j. 21 fev. 2025, DJe 25 fev. 2025.
- CONWAY, Julia M.; THOMPSON, Carrie A. *Combating the misuse of ChatGPT: A holistic approach to academic integrity*. *International Journal for Educational Integrity*, v. 21, n. 1, 2025.
- DINIZ, Debora. *Carta de uma orientadora: sobre pesquisa e escrita acadêmica*. 4. ed. Rio de Janeiro: Civilização Brasileira, 2024.
- DOMINICI, Piero. A complexidade da comunicação. A comunicação da complexidade. *MATRIZES*, São Paulo, v. 14, n. 2, p. 15–39, 2020.
- ECO, Umberto. *Como se faz uma tese*. Tradução de Gilson Cesar de Souza. 27. ed. rev. e atual. São Paulo: Perspectiva, 2019.
- FARIELLO, Serena et al. *Distinguishing human from machine: a review of advances and challenges in AI-generated text detection*. *International Journal of Interactive Multimedia and Artificial Intelligence*, [S.l.], v. 9, n. 3, p. 6–18, 2025.



GAUTAM, Naveen. Regulating the unseen hand: AI, authorship, and trust in medical science. *Annals of Medicine & Surgery*, v. 87, p. 5353–5355, 2025.

GOSSELIN, Romain-Daniel. *AI detectors are poor western blot classifiers: a study of accuracy and predictive values*. PeerJ, [S.l.], v. 13, e18988, 2025.

HAKAM, Hassan Tarek et al. Human-written vs AI-generated texts in orthopedic academic literature: comparative qualitative analysis. *JMIR Formative Research*, v. 8, p. e52164, 2024.

HAU, Mark F. *Writing with machines? Reconceptualizing student work in the age of AI*. Frontiers in Communication, [S.l.], v. 10, 2025.

HERATH, Dinuka B.; ODE, Egena; HERATH, Gayanga B. *Can AI replace humans? Comparing the capabilities of AI tools and human performance in a business management education scenario*. British Educational Research Journal, [S.l.], v. 51, p. 1073–1096, 2025.

ICAI. International Center for Academic Integrity. *The fundamental values of academic integrity*. 3. ed. Dayton: International Center for Academic Integrity, 2021. 17 p.

KAFKA, Alexander C. *Academic honesty in the ChatGPT era: Student and faculty perspectives*. The Journal of Higher Education, v. 96, n. 2, p. 120–137, 2025.

KALMANI, Vijay H. et al. *AI vs. Human Writing: Developing a Novel Method for Text Authenticity Detection in Education*. I.J. Modern Education and Computer Science, v. 17, n. 3, p. 45–58, 2025.

KIM, Soyoung et al. *Unmasking GPT: A Multidisciplinary Approach to Understanding the Use of Generative AI in Academia*. Education Sciences, v. 15, n. 5, 2025.

LEE, Kai-Fu. *Inteligência artificial: como os robôs estão mudando o mundo, a forma como amamos, nos comunicamos e vivemos*. Tradução de Marcelo Barbão. 1. ed. Rio de Janeiro: Globo Livros, 2019. 292 p. ISBN 978-65-80634-32-3.

LOPES, Fabiano Ferreira; ALMADA, Bruno Augusto Trindade. Proteção de direitos autorais na era da inteligência artificial: desafios e estratégias para o Brasil. *Revista de Direito, Inovação, Propriedade Intelectual e Concorrência*, Congresso Nacional de Brasília, v. 10, n. 2, p. 41-62, jan./jul. 2025.

LUND, Brady D. et al. *AI and academic integrity: Exploring student perceptions and implications for higher education*. Journal of Academic Ethics, v. 23, p. 1545–1565, 2025.

MALIK, Muhammad Abid; ISLAM, Amjad. *AI vs AI: How effective are Turnitin, ZeroGPT, GPTZero, and Writer AI in detecting text generated by ChatGPT, Perplexity, and Gemini?*. Journal of Applied Learning & Teaching, v. 8, n. 1, 2025.

MANFREDINI JUNIOR, Renato; VILLA-LOBOS, Eduardo Dutra; BONFÁ, Marcelo Augusto; ROCHA, Renato da Silva. *Quase sem querer*. Intérprete: Legião Urbana. In: *Dois*. Rio de Janeiro: EMI-Odeon, 1986. 1 disco sonoro, faixa 4.

MEHTA, Chirag; KHAN, Wasim A.; PATEL, Shubham. *Harnessing Generative AI in Education: Opportunities, Risks, and the Path Forward*. International Journal of Information and Education Technology, v. 15, n. 3, p. 281–287, 2025.

MILLER, Andrew S. et al. *Evaluation of the impact of large language learning models on publications in the Journal of Shoulder and Elbow Surgery*. JSES International, v. 9, p. 1803–1808, 2025.

MONDAL, Himel. *Teaching in an era of generative artificial intelligence: Ethical engagement with artificial intelligence in medical education*. Advances in Physiology Education, v. 49, p. 163–165, 2025.



- MULAUDZI, Lebohang Victoria; HAMILTON, Joleen. *Lecturer's perspective on the role of AI in personalized learning: Benefits, challenges, and ethical considerations in higher education*. Journal of Academic Ethics, v. 23, p. 1571–1591, 2025.
- MURER, Ricardo. *Fundamentos da inteligência artificial: o futuro é agora*. 1. ed. Rio de Janeiro: Alta Books, 2025.
- NASCIMENTO, Eliane Guerreiro; BRISOLARA, Valeria Silveira. A fronteira entre a intertextualidade e o plágio: análise de um caso na literatura contemporânea. In: SOUSA, Ivan Vale de (org.). *Letras, linguísticas e artes: perspectivas críticas e teóricas*. Ponta Grossa: Atena Editora, 2019. p. 35-46. (Letras, Linguísticas e Artes: Perspectivas Críticas e Teóricas; v. 3).
- PAULLET, Karen et al. *Precision check: A critical look at the reliability of AI detection tools*. Issues in Information Systems, v. 26, n. 3, p. 401–412, 2025.
- PERKINS, Mike et al. *Detection of GPT-4 generated text in higher education: combining academic judgement and software to identify generative AI tool misuse*. Journal of Academic Ethics, [S.l.], v. 22, p. 89–113, 2024.
- POPESCU-APREUTESEI, Lilia-Eliana et al. *Upholding academic integrity amidst advanced language models: evaluating BiLSTM networks with GloVe embeddings for detecting AI-generated scientific abstracts*. Computers, Materials & Continua, [S.l.], v. 84, n. 2, p. 2606–2610, 2025.
- POPESCU-APREUTESEI, Lilia-Eliana et al. *Upholding academic integrity amidst advanced language models: Evaluating BiLSTM networks with GloVe embeddings for detecting AI-generated scientific abstracts*. Computers, Materials & Continua, v. 84, n. 2, 2025.
- PRYKHODCHENKO, S. D. et al. *Software detection of Ukrainian-language texts generated by AI: methods, estimations, challenges*. Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, n. 4, p. 150–155, 2025.
- SARWANTI, Sri et al. *Are they literate on ChatGPT? University language students' perceptions, benefits and challenges in higher education learning*. Online Learning, [S.l.], v. 28, n. 3, p. 105–130, set. 2024.
- SPINELLIS, Diomidis. False authorship: an explorative case study around an AI-generated article published under my name. *Research Integrity and Peer Review*, v. 10, n. 8, 2025.
- STILLMAN, Tyler. *Maintaining academic integrity through AI-trap questions*. Journal of Educators Online, [S.l.], 2025.
- SULEYMAN, Mustafa; BHASKAR, Michael. *A próxima onda: inteligência artificial, poder e o maior dilema do século XXI*. Tradução de Alessandra Bonrruquer. 2. ed. Rio de Janeiro: Record, 2024. Tradução de: *The coming wave: technology, power, and the twenty-first century's greatest dilemma*.
- TYAGI, Anisha et al. *Evaluation of the impact of large language learning models on publications in the Journal of Shoulder and Elbow Surgery*. JSES International, v. 9, p. 1803–1808, 2025.
- ULLAH, Midrar; BIN NAEEM, Salman; KAMEL BOULOS, Maged N. *Assessing the guidelines on the use of generative artificial intelligence tools in universities: a survey of the world's top 50 universities*. Big Data and Cognitive Computing, [S.l.], v. 8, n. 12, art. 194, 2024.

