

### Integridade acadêmica na era do ChatGPT: Desafios éticos e as novas fronteiras da Inovação

*Academic integrity in the age of ChatGPT: Ethical challenges and new frontiers of innovation*

Recebido: 30/09/2024 | Revisado:  
24/10/2024 | Aceito: 11/12/2024 |  
Publicado: 07/08/2025

**Jéssica Caroline Rodrigues de Lima**  
ORCID: <https://orcid.org/0000-0002-0885-1876>  
Universidade Federal de Santa Catarina  
E-mail: [jessica.limadri@gmail.com](mailto:jessica.limadri@gmail.com)

**Maíra Longhinotti Felipe**  
ORCID: <https://orcid.org/0000-0001-9483-1654>  
Universidade Federal de Santa Catarina  
E-mail: [mairafelipe@gmail.com](mailto:mairafelipe@gmail.com)

**Como citar:** LIMA, J. C. R.; FELIPPE, M. L. Integridade acadêmica na era do ChatGPT: Desafios éticos e as novas fronteiras da Inovação. *Revista Brasileira da Educação Profissional e Tecnológica*, [S.l.], v. 03, n. 25, p.1-22 e17803, ago. 2025. ISSN 2447-1801. Disponível em: <Endereço eletrônico>.



This work is licensed under a [Creative Commons Attribution 4.0 Unported License](https://creativecommons.org/licenses/by/4.0/).

#### Resumo

Este artigo investiga as implicações éticas do uso de inteligência artificial generativa (GenAI), como o ChatGPT, na produção da pesquisa científica. Por meio de uma revisão narrativa da literatura, observou-se a presença recorrente de cinco categorias temáticas principais: confiabilidade das informações, vies informacional, direitos autorais e plágio, segurança de dados e desigualdade de acesso. A falta de confiabilidade nos dados constitui uma das principais razões pelas quais chatbots de IA não devem substituir o julgamento e a responsabilidade dos pesquisadores humanos. O ChatGPT pode adotar vieses e perpetuar preconceitos, além de ampliar desigualdades sociais, caso seu acesso não seja assegurado de maneira equitativa. A IA é amplamente vista como uma ferramenta, não podendo lhe ser atribuída a autoria de conteúdos, e a ausência de menção aos autores originais pode resultar em plágio. Além disso, a arquitetura do ChatGPT não garante plena proteção aos dados privados coletados. O estudo conclui que é necessária a implementação de instrumentos legais sólidos que exijam um gerenciamento seguro e transparente das informações pelas plataformas, além da formulação de diretrizes éticas para o uso de IAs generativas pelas instituições acadêmicas e a compreensão das limitações dessas ferramentas por parte dos autores, a fim de assegurar uma prática de pesquisa mais responsável.

**Palavras-chave:** ChatGPT; Inteligência Artificial; Ética na Pesquisa Científica; Plágio; Vies Algorítmico.

#### Abstract

This article investigates the ethical implications of using generative artificial intelligence (GenAI), such as ChatGPT, in the production of scientific research. Through a narrative review of the literature, the recurring presence of five main thematic categories was observed: information reliability, informational bias, copyright and plagiarism, data security, and inequality of access. The lack of reliability in data is one of the main reasons why AI chatbots should not replace the judgment and responsibility of human researchers. ChatGPT can adopt biases and perpetuate prejudices, in addition to

exacerbating social inequalities if its access is not ensured in an equitable manner. AI is widely seen as a tool and cannot be attributed authorship of content, and the absence of proper acknowledgment of original authors may result in plagiarism. Moreover, ChatGPT's architecture does not fully guarantee the protection of private data collected. The study concludes that the implementation of solid legal frameworks is necessary, requiring the secure and transparent management of information by platforms, alongside the formulation of ethical guidelines for the use of generative AIs by academic institutions and a comprehensive understanding of these tools' limitations by authors, in order to ensure more responsible research practices.

**Keywords:** *ChatGPT; Artificial Intelligence; Ethics in Scientific Research; Plagiarism; Algorithmic Bias.*

## 1 INTRODUÇÃO

As ferramentas de inteligência artificial (IA) têm impactado a sociedade em diferentes esferas, transformando as formas de trabalho e comunicação, influenciando o processo de tomada de decisões e moldando as interações com o ambiente ao redor (Eke, 2023). A IA abrange uma ampla gama de sistemas automatizados projetados para realizar ou auxiliar em tarefas cognitivas que antes eram desenvolvidas exclusivamente pela inteligência humana (Pcast, 2024).

Atualmente, a maioria dos sistemas de IA opera através do aprendizado de máquina (*machine learning*), que se baseia em inferência estatística obtida a partir do treinamento de modelos com grandes conjuntos de dados, permitindo que estes produzam textos, imagens e outros meios em respostas a perguntas relacionadas, com base nos parâmetros aprendidos (Khalifa e Albadawy (2024). As metodologias de aprendizado de máquina e seu subgrupo, o aprendizado profundo (*deep learning*), permitem que as máquinas se "desenvolvam" de maneira autônoma e melhorem de forma automática através da prática (Nic.br, 2022).

Na área da educação, a implementação de sistemas de IA começou ainda na década de 1990, com a introdução dos computadores e dos Sistemas de Tutoria Inteligente, que ofereciam feedback, dicas e orientações aos alunos durante o processo de aprendizagem. Os avanços recentes na IA incluem o desenvolvimento de ferramentas de IA generativa (GenIA), que se concentram na geração de novos conteúdos usando algoritmos avançados de aprendizado após serem pré-treinadas em grandes volumes de dados. Elas abrangem modelos de linguagem (*large language models* - LLMs), que não apenas geram e aprimoram a qualidade dos textos, mas também acessam informações de diversas fontes, habilidades que prometem transformar a experiência educacional e acelerar o progresso na pesquisa científica no enfrentamento de desafios globais (Isiaku, 2024).

Nos últimos anos, a inteligência artificial tornou-se um foco central das agendas políticas educacionais, sendo acompanhada por um intenso debate sobre as questões éticas envolvidas na sua implementação. Em 2021, a Unesco lançou a

"Recomendação sobre a Ética da Inteligência Artificial", um marco global que estabelece princípios éticos para o desenvolvimento e uso de IA por governos, empresas e instituições (Unesco, 2022). Em 2023, a Unesco publicou o "Guia para a IA Generativa na Educação e Pesquisa", com o objetivo de orientar a integração dessas tecnologias, otimizar seu potencial pedagógico e mitigar os riscos associados (Unesco, 2023). Em 2024, a conferência AAAI-24 (Association for the Advancement of Artificial Intelligence Conference), realizada em Washington, D.C., ressaltou a necessidade de incluir tópicos sobre IA nos currículos de todos os níveis de ensino e a importância de educar sobre integridade acadêmica (AAAI, 2024).

O lançamento do ChatGPT (*Chat Generative Pre-trained Transformer*) em 30 de novembro de 2022, pela OpenAI — uma empresa sem fins lucrativos sediada na Califórnia — popularizou as capacidades dos grandes modelos de linguagem, tornando essas ferramentas acessíveis a um público mais amplo. Diferentes de outros chatbots, o ChatGPT é especializado em interação conversacional por texto ou áudio, utilizando processamento de linguagem natural para garantir diálogos naturais e coerentes. Ele é capaz de memorizar informações inseridas pelo usuário em conversas anteriores e utilizá-las em consultas subsequentes, oferecendo uma assistência personalizada. Sua versão mais recente, o GPT-4, foi lançada 14 de março de 2023 (OpenAI, 2024).

Hirsh-Pasek e Blinkoff (2023) sugerem que, assim como as calculadoras se tornaram indispensáveis no ensino de matemática, o ChatGPT tem o potencial de se tornar uma ferramenta crucial para aprimorar as habilidades de pensamento crítico e comunicação de escritores, oferecendo uma vantagem significativa em relação aos humanos ao processar rapidamente grandes volumes de informações e identificar conexões entre dados de maneira mais eficiente.

No entanto, sua introdução na educação superior também traz novos riscos e pressões para a comunidade científica, uma vez que pode afetar as normas estabelecidas e os princípios de integridade acadêmica, especialmente, porque a adoção da IA por indivíduos tem avançado mais rapidamente do que a implementação de práticas e salvaguardas responsáveis (Isiaku et al. 2024).

Algumas universidades já adotaram respostas formais, diretrizes abrangentes e recursos para promover o uso eficiente, responsável e ético dos modelos de IA, como a Universidade de Stanford, o Instituto de Tecnologia de Massachusetts (MIT) e a Universidade da Califórnia nos EUA; a Universidade de Oxford e a Universidade de Cambridge no Reino Unido; a Universidade de Toronto no Canadá, e a Universidade da Tasmânia na Austrália (Torres-Salinas; Arroyo-Machado, 2023). Outras, no entanto, tem optado pela proibição completa do uso desses modelos de IA no campus e a aplicação de sanções rigorosas em alguns casos, como a Universidade de Harvard e a Universidade de Yale nos EUA; algumas instituições na China, incluindo a Universidade de Hong Kong; e várias universidades na França, incluindo a Université de Paris (Tlili et al., 2023).

Compreender os desafios éticos associados ao uso do ChatGPT é crucial para que as instituições educacionais, incluindo universidades e cursos de educação tecnológica e profissional, possam estabelecer políticas institucionais eficazes, diretrizes éticas e programas de desenvolvimento profissional adaptados, permitindo

assim que a comunidade acadêmica e demais atores relacionados a este universo possam usufruir do potencial desta ferramenta na pesquisa de maneira produtiva, transparente e responsável. Neste sentido, este estudo se propõe a investigar os aspectos éticos envolvidos no uso de inteligência artificial generativa (GenAI), como o ChatGPT, na produção de conhecimento científico e as medidas necessárias para a sua integração no processo.

## 2 MÉTODO

Este estudo investigou as questões éticas e de integridade acadêmica relacionadas ao uso de ferramentas de IA generativa (GenAI), como o ChatGPT, na pesquisa científica. Inicialmente, foi realizada uma revisão de literatura nas bases de dados Scielo e Scopus, utilizando a seguinte string de busca nos campos de título, resumo e palavras-chave: ("inteligência artificial" OR "artificial intelligence" OR IA OR AI OR "machine learning") AND ("pesquisa científica" OR "pesquisa acadêmica" OR "scientific research" OR "academic research") AND (ética OR ethics), resultando em 133 documentos. Foram considerados apenas documentos classificados como artigos, resultando em 71. Na primeira filtragem, excluíram-se os estudos provenientes das ciências médicas (Medicina, Enfermagem), cujas metodologias de pesquisa possuem contornos bastante particulares e por isso não contribuem para o entendimento geral das questões éticas relacionadas ao uso de IA na pesquisa, resultando na exclusão de 27 documentos. Dos 44 artigos restantes, a partir da leitura dos resumos, foram eliminados aqueles materiais cujos objetivos não estavam relacionados diretamente aos termos relativos ao escopo — ética, inteligência artificial e pesquisa acadêmica — reduzindo o total a 22 artigos.

Em seguida, o levantamento de dados foi ampliado para incluir outras fontes significativas sobre o tema, como os livros publicados pela Unesco, "Recomendação sobre a Ética da Inteligência Artificial" (2022) e o "Guia para a IA Generativa na Educação e Pesquisa" (2023), além de reportagens, instrumentos legais e relatórios de fóruns, como a Conferência da AAI – Associação para o Avanço da Inteligência Artificial (2024), caracterizando, assim, uma revisão de literatura narrativa. A análise de dados buscou encontrar temas recorrentes que caracterizassem as principais preocupações éticas no uso da IA na pesquisa científica destacadas pela literatura atual, sem definição prévia de categorias temáticas.

## 3 RESULTADOS

Durante a análise dos materiais, foi possível identificar a recorrência de determinados temas, o que levou à categorização dos resultados em cinco principais eixos: confiabilidade dos dados, viés de informações, direitos autorais, segurança dos dados e desigualdade no acesso.

### 3.1 CONFIABILIDADE DOS DADOS

A escrita acadêmica desempenha um papel crucial na pesquisa, servindo como um método organizado para apresentar ideias e argumentos baseados em dados e raciocínio lógico. Esse tipo de escrita permite aos autores aprofundarem-se em conceitos, a partir do gerenciamento de grandes volumes de informação e do tratamento de dados empíricos, partindo de uma teoria ou conduzindo a uma teoria ou conclusão bem fundamentada, cuja complexidade deve ser simplificada para permitir o entendimento do leitor (Khalifa; Albadawy, 2024). Desse modo, os textos acadêmicos devem atender a rigorosos padrões de precisão, evidências e coerência lógica, com cada afirmação apoiada por fontes confiáveis.

Embora o ChatGPT seja capaz de gerar textos a partir de uma ampla gama de fontes de informações, ele não possui uma compreensão conceitual profunda da tarefa que está executando, e sem uma verificação externa, não é possível garantir a precisão das respostas que ele fornece (Hamam; Fatouh, 2023). Segundo Nepal (2024), essa limitação reflete a falta de compreensão contextual e da expertise que são próprias dos seres humanos, o que pode levar a desafios em fornecer explicações sutis, conhecimento especializado em áreas específicas e uma compreensão adequada de contextos de pesquisa complexos.

Outro aspecto que pode afetar a exatidão e pertinência das respostas fornecidas pelo ChatGPT é a limitação de sua base de conhecimento, atualizada pela última vez em setembro de 2021. Isso implica na ausência de informações sobre eventos, desenvolvimentos ou descobertas ocorridas após essa data, e, portanto, impede que as evoluções recentes nos diversos tópicos sejam consideradas (Lund et al., 2023). Quando questionado sobre as fontes utilizadas, o ChatGPT explica que suas respostas não se baseiam em uma única fonte específica, mas sim em um vasto conjunto de dados textuais amplamente reconhecidos e discutidos em várias fontes, podendo em alguns casos sugerir referências adicionais sobre os temas abordados.

Nepal (2024) também aponta que os chatbots apresentam um vocabulário limitado, o que pode resultar na repetição de palavras ou frases e, ocasionalmente, em uma linguagem inconsistente ou estranha, levando o ChatGPT a gerar textos irregulares. Já Kamoun (2024) acrescenta que o ChatGPT pode falhar em produzir resultados corretos, precisos ou ideais, devido às limitações relacionadas à quantidade e diversidade de dados usados em seu treinamento. A qualidade de uma ferramenta de IA depende diretamente dos dados com os quais foi treinada, de modo que, se esses dados contiverem erros, o conteúdo pode refletir essas falhas, tornando-se suscetível de produzir textos imprecisos, sem sentido ou irrelevantes (Zhou, 2024).

Hamam e Fatouh (2023) destacam que as "alucinações" nos modelos de IA generativa ocorrem porque esses modelos não se baseiam exclusivamente em conhecimento factual, o que pode levar a afirmações com uma confiança enganosa ou à criação de imagens com artefatos que não correspondem à realidade física. Dessa forma, o problema com a veracidade nas gerações de IA vai além da precisão dos dados de treinamento, uma vez que a natureza probabilística dos modelos pode

gerar respostas que, embora pareçam plausíveis, são imprecisas, pois se baseiam apenas em padrões estatísticos.

Além disso, o problema se intensifica ao considerar que a pesquisa científica constitui um ecossistema interconectado, no qual os resultados produzidos por um grupo influenciam o trabalho de outros pesquisadores. O uso irresponsável de ferramentas de IA pode prejudicar esse ambiente ao gerar e propagar descobertas que não podem ser replicadas, carecem de quantificação da incerteza ou são excessivamente ajustadas aos dados de treinamento. A longo prazo, a utilização de dados sintéticos impróprios pode causar uma "poluição", prejudicando o progresso científico. Esses riscos aumentam quando ferramentas de IA totalmente automatizadas são implantadas sem supervisão especializada (Pcast, 2024).

Até o momento, não foi realizada uma avaliação estruturada da qualidade do conteúdo gerado pelo ChatGPT, o que implica que a redação de artigos científicos, especialmente em métodos complexos como revisões sistemáticas e meta-análises, ainda necessita da orientação e supervisão de pesquisadores humanos especializados para assegurar a precisão, coerência e credibilidade do conteúdo antes de sua utilização ou submissão para publicação (Salvagno, 2023). Para Hirsh-Pasek e Blinkoff (2023), o algoritmo das IA generativas funciona mais como um sintetizador de conhecimentos, fornecendo respostas rápidas e demonstrando um desempenho satisfatório em testes de comparação e contraste, porém pouco eficaz na elaboração e defesa de uma tese original.

Vilalta (2024) explica que as IAs generativas, em sua maioria, utilizam LLMs no modo *zero-shot*, ou seja, solicitam que o modelo produza a saída final *token* por *token*, sem fazer revisões ao longo do processo, o que seria semelhante a pedir a uma pessoa que escreva um ensaio de uma só vez, sem voltar atrás para corrigir ou revisar, e ainda assim esperar um resultado de alta qualidade. Embora os LLMs consigam desempenhar surpreendentemente bem essa missão, as limitações quanto à confiabilidade e clareza dos resultados permanecem. Para superar essas barreiras, o autor (Vilalta, 2024) sugere a integração de agentes em sistemas de IA, pois eles podem melhorar a explicabilidade sobre o funcionamento e as decisões desses modelos, além de permitir que o LLM revise e aperfeiçoe um documento várias vezes, o que contribuiria para sistemas mais robustos e transparentes.

Confiar excessivamente em uma ferramenta de IA sem verificar os resultados pode comprometer os princípios de integridade científica. Embora chatbots possam ser úteis, os pesquisadores devem reconhecer essas limitações e usar seu próprio discernimento ao analisar e utilizar as respostas do ChatGPT, além de validar as simulações de IA sempre que possível e compará-las com métodos tradicionais de verificação, como simulações numéricas e dados experimentais (Nepal, 2024).

A falta de confiabilidade nos dados constitui uma das principais razões do porque chatbots de IA generativa não devem substituir a experiência, o julgamento, a personalidade e, sobretudo, a responsabilidade dos pesquisadores humanos.

### 3.2 VIÉS DAS INFORMAÇÕES

Um dos aspectos mais críticos no uso do ChatGPT e de outros modelos de IA generativa é a possibilidade de que preconceitos sejam incorporados durante a geração de linguagem. Em virtude da falta de transparência e explicabilidade sobre o funcionamento interno desses sistemas, torna-se difícil entender por que certos conteúdos são gerados (Unesco, 2023).

Essa ausência de clareza sobre a geração dos resultados tende a limitar os usuários à lógica imposta pelos parâmetros da IA. Diferente de uma pesquisa no Google, onde diversos *links* são apresentados na página inicial, as plataformas de IA promovem um aprofundamento do tópico por meio do diálogo com a máquina, o que permite respostas mais rápidas para perguntas específicas, mas reduz significativamente a diversidade de opiniões e o debate em torno da questão pesquisada, criando uma tendência de que a primeira resposta da IA seja recebida como a “verdade” sobre o assunto (Sampaio, 2024). Como as respostas da IA podem refletir valores culturais ou comerciais específicos, dependendo das fontes de dados de treinamento utilizadas, há o risco de reprodução de conteúdos preconceituosos e perpetuação de estereótipos (Nepal, 2024).

Embora se reconheça que todas as tecnologias são produtos de contextos socioculturais, criadas por agentes morais influenciados por tensões políticas e econômicas e hábitos que fundamentam as sociedades modernas, e, portanto, carregam vieses e escolhas inerentes, Sampaio (2024) argumenta que as IAs generativas apresentam vieses em diversas fases de seu desenvolvimento. As empresas responsáveis por esses modelos afirmam que a dificuldade em tornar seus sistemas transparentes e acessíveis decorre da concorrência no mercado. No entanto, na prática, existem muitos aspectos confusos e questionáveis, especialmente no que diz respeito aos algoritmos e aos dados utilizados para o treinamento; o que faz com que esses modelos de linguagem frequentemente reproduzam vieses existentes desde o início (Sampaio, 2024).

Para reduzir os riscos da IA produzir conteúdos ofensivos ou antiéticos, os desenvolvedores implementaram "*guardrails*" em seus sistemas, que atuam como barreiras programadas, como filtros de conteúdo, algoritmos para detectar linguagem inadequada, regras de moderação e diretrizes éticas. Esses mecanismos visam manter as respostas da IA dentro de padrões seguros e aceitáveis. Contudo, a falta de regulamentações rígidas e de um monitoramento eficiente tem permitido que conteúdos tendenciosos gerados por IA generativa se espalhem pela internet, comprometendo uma das principais fontes de informação global (Unesco, 2023).

Considerando que o material gerado por GenAI pode parecer bastante convincente, a situação é particularmente preocupante para os grupos que não possuem um conhecimento prévio sólido, tornando-os mais vulneráveis a aceitar informações tendenciosas. Além disso, há um risco contínuo de que futuros modelos GPT sejam treinados com textos gerados por versões anteriores, perpetuando os mesmos vieses e erros. Esses problemas reforçam a visão de Salvagno (2023) de que, independentemente da ferramenta de IA utilizada, a presença de especialistas é

crucial na condução e redação de atividades científicas para garantir a qualidade e a integridade do trabalho.

A pesquisa de Nemorin (2024) examina como a inteligência artificial pode consolidar hierarquias de conhecimento dominadas por perspectivas ocidentais e eurocêntricas, evidenciando a predominância de marcos éticos baseados nessas visões. Ela descreve esse fenômeno como "violência epistêmica", onde o sistema reflete um conjunto restrito de normas culturais e desconsidera os conhecimentos e valores de comunidades locais ou indígenas, especialmente no Sul Global, na discussão sobre ética em IA, perpetuando desigualdades de poder e reforçando legados coloniais.

Assim, a IA estaria ligada a questões de poder, pois influencia quais sistemas e práticas são incorporados no desenvolvimento tecnológico e, conseqüentemente, como o conhecimento é produzido e disseminado. Para enfrentar essas questões, a autora (Nemorin, 2024) propõe uma análise epistemológica decolonial das práticas de IA na educação, com o objetivo de ampliar a diversidade de perspectivas e métodos de conhecimento nos sistemas de IA e promover práticas mais equitativas.

Corroborando a ideia de que as IAs tendem a se basear em indexadores que priorizam a produção acadêmica norte-americana, anglo-saxã e europeia, Susarla et al. (2023) discutem como isso pode levar a um maior número de citações para autores já estabelecidos e uma menor inclusão de novos pesquisadores, além de reduzir a presença de pesquisas inovadoras e emergentes, especialmente nas áreas de humanidades e nas pesquisas de caráter qualitativo.

Lund et al. (2023) apontam que o sistema do ChatGPT pode intensificar o "Efeito Mateus", conceito cunhado pelo sociólogo Robert K. Merton em 1968, que descreve a tendência de pesquisadores altamente citados continuarem a ser amplamente referenciados, enquanto aqueles com menos citações enfrentam dificuldades para obter reconhecimento, o que agrava as desigualdades acadêmicas. Assim, o ChatGPT pode fornecer respostas baseadas em artigos desatualizados, principalmente de autores dos EUA, devido ao grande número de citações que esses trabalhos já acumulam, sem considerar outros critérios. Como a maioria dos usuários tende a consultar apenas os primeiros resultados, ocorre uma perpetuação do ciclo em que os artigos mais citados dominam as referências.

Valeri (2023) também explora como o ChatGPT tende a gerar respostas enviesadas conforme a perspectiva das sociedades dominantes. Ela exemplifica com a questão da invenção do avião: enquanto brasileiros reconheceriam Santos Dumont como o inventor, a IA mencionaria os irmãos Wright, refletindo a visão predominante dos Estados Unidos. Isso evidencia a uniformização do conhecimento imposta pelos desenvolvedores de IA, a qual marginaliza as narrativas de minorias. Para a autora (Valeri, 2023), esta problemática poderia ser abordada com testes de falsos positivos e negativos para identificar erros e avaliar o impacto das decisões baseadas nos algoritmos usados no treinamento da IA.

Por outro lado, os pesquisadores do Núcleo de Informação e Coordenação do Ponto BR (2022) acreditam que é inviável eliminar completamente os vieses, visto que os dados utilizados no *input* são, por natureza, limitados, já que independentemente da quantidade de informações disponíveis, a seleção sempre apresentará algum nível

de viés. Com exceção dos vieses intencionalmente prejudiciais, os pesquisadores ressaltam a impossibilidade de se ter plena consciência dos próprios estereótipos, sejam eles de ordem pessoal ou cultural, a menos que sejam apontados por terceiros. Em meio a debates sobre uma “IA socialmente benéfica”, como definir seus valores, para quem ela deve ser desenvolvida e, principalmente, quem deveria determinar o que constitui esse benefício social, o núcleo destaca que, quando a cultura não aparece de forma explícita na equação, ela tende a se manifestar indiretamente por meio de vieses (Nic.br, 2022).

Para Dora Kaufman, especialista nos impactos éticos e sociais da IA, a formação de equipes multidisciplinares é fundamental para o desenvolvimento de sistemas de IA. A inclusão de diversidade cognitiva, racial e de gênero, juntamente com profissionais de diferentes áreas, aumenta a capacidade dessas equipes de identificar vieses, preconceitos e outras questões de cunho ético de forma mais eficaz (Moreira, 2023).

Desse modo, a adoção do ChatGPT também suscita preocupações éticas relacionadas aos preconceitos que pode adotar e agravar, levando a possíveis consequências como o aumento da discriminação, a ampliação das divisões digitais, a exclusão e as ameaças à diversidade cultural e social. Esse sistema ainda permanece como uma verdadeira caixa-preta, com um funcionamento opaco e imprevisível. Nesse sentido, assegurar maior transparência sobre os dados usados no treinamento dessas máquinas e entender como elas operam, juntamente com um debate contínuo na comunidade acadêmica, pode ajudar a detectar e corrigir esses vieses, promovendo resultados de pesquisa mais inclusivos e diversificados.

### 3.3 DIREITOS AUTORAIS E PLÁGIO

Os direitos autorais são os direitos exclusivos que protegem os autores de obras intelectuais — incluindo produções literárias, artísticas e científicas — assegurando-lhes controle sobre a utilização e reprodução de suas criações<sup>1</sup> (Brasil, 1988). A introdução do ChatGPT na redação científica levantou debates sobre a propriedade e proteção dos direitos autorais do conteúdo gerado, sendo questionado se o material produzido pode ser considerado uma obra protegida por direitos autorais e se a IA detém esses direitos (Salvagno, 2023). Até o momento, não há consenso nem um padrão legal unificado sobre o tema.

Alguns especialistas defendem que o conteúdo criado pelo ChatGPT pode ser considerado uma obra, pois demonstra certo grau de originalidade e criatividade, refletindo a inteligência e a autonomia da IA generativa, devendo, portanto, ser reconhecido como coautor. Em contrapartida, muitos pesquisadores argumentam que o material gerado pelo ChatGPT não deve ser classificado como uma obra, uma vez

---

<sup>1</sup> No Brasil, esses direitos são regulamentados pela Lei nº 9.610, de 19 de fevereiro de 1998, e também estão em conformidade com tratados internacionais, como a Convenção de Berna, conforme definido pelo Decreto nº 75.699, de 6 de maio de 1975 (Brasil, 1998).

que carece da criatividade e originalidade humanas, constituindo apenas uma combinação de cálculos e imitações de máquinas (Gao; Shen; Duan, 2023).

Sampaio (2024) questiona a ideia de criatividade associada ao ChatGPT, afirmando que os modelos de IA apenas reproduzem conhecimentos já existentes, reorganizando informações previamente geradas pela humanidade de forma lógica e estatística. Assim, não se trataria de criação original de textos ou novos códigos, mas da reutilização de *scripts* usados durante o seu treinamento e já disponíveis.

A lei brasileira nº 9.610/1998 (Brasil, 1998) define que para ser reconhecido como autor de uma obra, o indivíduo deve ter contribuído de forma significativa e ativa na concepção, desenvolvimento e análise do trabalho, assumindo total responsabilidade pelo conteúdo e pela integridade da pesquisa. Além disso, é essencial obter o consentimento de todos os coautores para a inclusão de seus nomes e seguir as diretrizes das instituições e publicações científicas. Diante desses critérios e considerando o estágio atual dos chatbots de IA, essas ferramentas não possuem a capacidade de consentir com a versão final de um trabalho, endossar as ideias apresentadas ou se responsabilizar pela integridade da pesquisa.

A Unesco (2023) recomenda que o uso de IA seja transparente, com a menção clara das ferramentas em qualquer trabalho acadêmico ou educacional. Ademais, os usuários devem assumir total responsabilidade pelo conteúdo, uma vez que a IA não pode ser responsabilizada, e, portanto, não deve ser listada como autora.

Embora o Escritório de Direitos Autorais dos EUA considere que obras criadas por IA com pouca intervenção humana não sejam protegidas pela legislação atual, a instituição permite que o proprietário do sistema de IA ou quem detém os seus direitos reivindique a propriedade do conteúdo gerado, uma vez que a IA é tratada como uma ferramenta, semelhante a um pintor que utiliza um pincel (Zhou, 2024).

No meio acadêmico, a maioria dos cientistas permanece cautelosa em relação ao uso de ferramentas de IA. Editores como *Nature*, *Springer*, *Science* e *Cell Press*, concordam que ferramentas de LLM (*Large Language Model*), ainda não atingiram o nível de maturidade necessário para serem consideradas coautoras de manuscritos. Eles entendem que esses modelos não atendem aos critérios de autoria, pois não podem assumir responsabilidade pelo trabalho submetido, declarar conflitos de interesse ou gerenciar questões de direitos autorais e licenças (Zhou, 2024).

Outro princípio presente nas diretrizes de submissão de todas as revistas da Springer Nature é que pesquisadores que utilizaram mecanismos de LLM devem documentar esse uso nas seções de métodos ou agradecimentos. Caso o artigo não inclua essas seções, a documentação deve ser feita na introdução ou em outra seção apropriada (Nature, 2023). O Comitê de Ética em Publicações apresenta a mesma orientação. A grande preocupação é que estudantes e pesquisadores possam apresentar textos gerados por máquinas como se fossem seus ou usar LLMs de maneira simplista, resultando em trabalhos frágeis (COPE, 2023).

A pesquisa de Ganjavi (2024) avaliou a disponibilidade e a consistência das diretrizes sobre o uso da Inteligência Artificial Generativa na publicação acadêmica. A análise, que se baseou nas orientações dos 100 maiores editores acadêmicos e jornais científicos, revelou uma considerável variabilidade e inconsistência nas

diretrizes sobre GenAI, além de conflitos entre as orientações dos jornais, mesmo quando vinculados a editores que já haviam fornecido diretrizes sobre o tema.

Diversas associações e editoras acadêmicas, como Taylor & Francis, Elsevier e Springer-Nature, estabeleceram que o uso de modelos de IA em artigos deve ser restrito a melhorias na linguagem, clareza textual e legibilidade dos documentos, com a condição de que haja supervisão, revisão e edição humanas (Rahman et al., 2023). Dessa forma, até o momento, há uma defesa da utilização da IA apenas para aprimorar um conteúdo existente, sem a formulação de novos conceitos ou conclusões científicas.

Apesar do crescente consenso sobre a impossibilidade de a IA substituir funções essenciais da autoria, ainda há uma carência de padronização nas diretrizes para seu uso. Essa falta de uniformidade resulta em diversas orientações personalizadas por periódicos e contribui para o "fenômeno da Torre de Babel", onde as regras divergentes geram confusão e inconsistência entre disciplinas (Ganjavi, 2024).

As ferramentas de IA generativa ainda não conseguem citar fontes específicas para documentar suas respostas, pois não acessam diretamente bancos de dados ou a internet em tempo real para fornecer referências precisas e atualizadas (Nature, 2023); o que representa uma limitação significativa na escrita científica, onde a citação precisa e a formatação correta são essenciais (Zhou, 2024).

Outra preocupação relevante em relação aos direitos autorais é o risco de plágio, já que o ChatGPT pode produzir conteúdo semelhante a materiais existentes, levando pesquisadores a violarem os padrões de integridade acadêmica ao não atribuírem corretamente a autoria do material (Eke, 2023). Para Salvagno (2023), mesmo quando programado para evitar a cópia direta, o ChatGPT pode cometer plágio ao reformular o trabalho de outros sem oferecer contribuições originais.

Em síntese, a perspectiva acadêmica atual considera a IA como uma ferramenta cuja autoria é atribuída a quem utiliza e controla o conteúdo gerado por ela. Assim, o autor é responsável pela autenticidade e pelo conteúdo do material. Embora o ChatGPT facilite a reescrita de conteúdo existente, a ausência de menções aos autores originais e de ideias originais poderia ocasionar a prática de plágio.

### 3.4 SEGURANÇA DE DADOS

Segurança dos dados refere-se às estratégias e práticas destinadas a proteger informações contra acessos não autorizados, uso indevido, alterações ou divulgação ilegal. No Brasil, a Lei Geral de Proteção de Dados Pessoais (LGPD), lei nº 13.709 de 14 de agosto de 2018, categoriza as informações em dados pessoais e dados pessoais sensíveis. A principal diferença entre eles é que dados pessoais consistem em informações que podem identificar um indivíduo específico, como nome, CPF, endereço e e-mail. Em contraste, dados pessoais sensíveis são informações cuja divulgação pode causar discriminação, como dados sobre saúde, orientação sexual, raça, religião e afiliação política (Brasil, 2018).

A LGPD não proíbe o uso dos dados pessoais, mas estabelece que sua utilização deve ser feita de maneira ética, transparente e legal, com a devida ciência e consentimento dos indivíduos. Além da questão da privacidade, o uso de dados envolve debates importantes sobre o direito de não ser julgado ou categorizado para fins específicos ou com base em critérios não claros (Nic.br, 2022).

No contexto do uso de IAs generativas, como ChatGPT, para a produção de textos acadêmicos, surgem possíveis riscos e ameaças à segurança desses dados, uma vez que o ChatGPT coleta e processa uma grande quantidade de informações, incluindo dados fornecidos pelos usuários durante o login e o uso da plataforma. Esses dados podem abarcar informações sensíveis ou privadas, como a identidade do usuário, resultados de pesquisa e o conteúdo de trabalhos (Gao; Shen; Duan, 2023). Além disso, a empresa OpenAI pode compartilhar essas informações com terceiros para fins comerciais, conforme descrito em sua plataforma (OpenAI, 2024).

Caso essas informações sejam acessadas, manipuladas, alteradas ou divulgadas de maneira ilegal, os direitos e interesses dos usuários podem ser comprometidos e configurar situações de plágio acadêmico, roubo ou adulteração de conhecimento ou outras atividades competitivas desleais. Portanto, os pesquisadores precisam prestar atenção ao problema da segurança dos dados e tomar algumas medidas necessárias para protegê-los (Gao; Shen; Duan, 2023).

Durante a *Publication Integrity Week 2023*, organizada pelo Comitê de Ética na Publicação, críticos expressaram preocupações de que o uso de IA em revisões poderia aumentar o risco de exposição de informações confidenciais ao domínio público. O debate destacou a distinção entre modelos locais, que não reintegram os dados utilizados aos conjuntos de treinamento, e modelos mais populares, como o ChatGPT, que demandam maior cautela em virtude de uma maior vulnerabilidade ao compartilhamento de informações pelos usuários. Além disso, enfatizou a importância de educar os usuários sobre o funcionamento desses diferentes modelos e como utilizar os resultados de maneira responsável (COPE, 2023).

O estudo "*Red Teaming AI: Evaluating the Risks and Vulnerabilities of Large Language Models*", de 2023, revelou que modelos de linguagem como o ChatGPT podem inadvertidamente divulgar informações pessoais quando solicitados a repetir palavras ou frases específicas. Esses modelos podem ser vulneráveis a ataques que exploram suas respostas para extrair dados sensíveis, mesmo que esses dados tenham sido incluídos de forma não intencional no treinamento. A pesquisa destacou que técnicas de filtragem e moderação podem não ser totalmente eficazes e enfatizou a importância do "*red teaming*" para identificar e corrigir fraquezas nos modelos, ajudando a mitigar riscos associados (Hendrycks; Zhao; Wallace, 2023).

Meneses (2023) examina como o desenvolvimento de IAs em países do Norte Global e a adesão às legislações dessas regiões resultam em uma predominância preocupante dos conteúdos desses locais sobre os de outras partes do mundo. Isso levanta questões sobre a inadequação das práticas às legislações locais de proteção de dados e sobre as frequentes alegações de uso indevido de propriedade intelectual nos processos de aprendizado de máquina.

Segundo Catarini (2023), a aplicação da Constituição Federal de 1988, juntamente com o diálogo das fontes, que abrange o Marco Civil da Internet, a Lei

Geral de Proteção de Dados (LGPD) e a Estratégia Brasileira de Inteligência Artificial (EBIA), é insuficiente para lidar com as complexidades da IA no Brasil. A autora (Catarini, 2023) aponta que, embora a LGPD ofereça uma proteção moderada, ainda faltam boas práticas e *compliance* comparáveis ao Regulamento Geral de Proteção de Dados (GDPR) europeu, o que, aliado à falta de clareza sobre suas responsabilidades, gera divergências interpretativas e insegurança jurídica. No contexto da IA, violações de direitos associadas ao uso de ferramentas como o ChatGPT, além dos impactos de *fake news*, evidenciam a urgência de se desenvolver uma regulamentação específica mais robusta que equilibre inovação e competitividade global.

Catarini (2023) faz uma analogia entre a proteção de dados e o dano ambiental, comparando vazamentos de informações pessoais a derramamentos de óleo, ambos praticamente irreversíveis, o que reforça a necessidade do princípio da prevenção aliado a uma abordagem baseada em riscos. A autora (Catarini (2023) discorre também sobre o reconhecimento da autodeterminação informativa como direito fundamental — o direito que cada pessoa tem de controlar suas próprias informações, decidindo como, quando e por quem seus dados podem ser coletados, usados e compartilhados.

Dentro do contexto da privacidade de dados, surge também o “direito ao esquecimento”, que permite a remoção de dados pessoais de bancos de dados digitais, incluindo plataformas de inteligência artificial (IA). No entanto, como essas plataformas frequentemente retêm dados por longos períodos para garantir a eficácia de suas previsões e operações, a aplicação do direito ao esquecimento, ao exigir a exclusão segura e irreversível de dados acumulados, representa um desafio. Trata-se de um processo complexo, que demanda a implementação de protocolos rigorosos de remoção pelas plataformas, bem como a consideração do impacto nos perfis e algoritmos gerados (Guimarães, 2021).

Desse modo, a segurança dos dados e a proteção da privacidade dos indivíduos são questões éticas que ganham contornos cada vez mais complexos à medida que cresce a adoção da inteligência artificial. Entre as soluções para enfrentar esses desafios, destaca-se a implementação de uma base legal sólida que assegure que as informações utilizadas no treinamento do ChatGPT sejam geridas de maneira segura e transparente. Além disso, a promoção da conscientização dos usuários sobre os riscos de uso e vazamento de dados em ferramentas de IA generativa, como o ChatGPT, pode incentivar uma maior exigência e seletividade na escolha das plataformas disponíveis no mercado, especialmente para a elaboração de conhecimento científico.

### 3.5 DESIGUALDADE NO ACESSO

A inteligência artificial, como outras inovações tecnológicas, pode gerar efeitos colaterais não intencionais, incluindo a ampliação de desigualdades sociais preexistentes, tanto dentro de um país quanto entre diferentes nações, devido à distribuição desigual do acesso à tecnologia (Pcast, 2024). Um relatório da Unesco de

2023 destaca preocupações sobre o crescente desnível no controle e uso da IA, especialmente entre o Norte e o Sul Global. Empresas situadas nos EUA, China e Europa detêm a maior parte da infraestrutura necessária para o treinamento e processamento dos enormes bancos de dados que sustentam o desenvolvimento da IA. Isso tem resultado na rápida disseminação dessas ferramentas em países tecnologicamente avançados do Norte Global, intensificando a concentração de riqueza e poder nesses territórios (Unesco, 2023).

De acordo com Almeida (2005), a exclusão socioeconômica e a exclusão digital são interdependentes, funcionando como causa e consequência uma da outra. A falta de recursos econômicos e sociais frequentemente leva à exclusão digital, enquanto a ausência de acesso à tecnologia agrava as desigualdades socioeconômicas.

O fato da criação e do controle de ferramentas de IA generativas permanecerem fora do alcance da maioria das empresas e países pode, assim, intensificar a chamada "pobreza digital" ou "pobreza de dados" nesses locais (Marwala, 2023). Meneses (2023) salienta que a exclusão digital imediata em regiões com escassez de dados pode, a longo prazo, resultar na subordinação dessas áreas aos padrões incorporados nos modelos do ChatGPT, que refletem os valores e normas das nações dominantes.

Montini (2024) vai ao encontro dessa visão, destacando que a falta de diversidade na criação de algoritmos e o viés nas aplicações de IA, moldados pelos interesses das elites tecnológicas, podem produzir soluções que não apenas desconsideram, mas também prejudicam as populações menos favorecidas, perpetuando ciclos de pobreza e exclusão.

Visando fomentar o debate sobre o papel da IA na perpetuação de desigualdades e no fortalecimento de estruturas de poder, além de buscar alternativas mais inclusivas, foi criada a plataforma DecolonizAI, desenvolvida por pesquisadores da Cátedra Oscar Sala, do Instituto de Estudos Avançados da Universidade de São Paulo. Conforme o manifesto do grupo, a colonização por meio da IA acontece não apenas por meio do conceito de inteligência, mas também através da linguagem, dos sistemas de classificação e das representações que moldam a prática científica, incluindo bancos de imagens que influenciam identidades e sentimentos de pertencimento ou exclusão no imaginário social (DecolonizAI, 2023).

O DecolonizAI (2023) enfatiza a importância de envolver representantes do Sul Global para acompanhar os avanços da IA em seus territórios, assegurando que toda a população possa acessar seus benefícios. Embora reconheça que interesses comerciais e científicos possam convergir em alguns momentos, o grupo argumenta que a pesquisa acadêmica deve priorizar o desenvolvimento humano com base no interesse público, respondendo às demandas da comunidade e de seus pesquisadores, orientando suas ações para promover a inclusão, a equidade e a justiça social, sem se restringir a agendas pré-definidas.

A implementação de IAs generativas trazem um paradoxo, pois, se por um lado tornam ferramentas mais acessíveis para aqueles que não dominam linguagens de programação, por outro, elas concentram ainda mais a tomada de decisão entre empresas que controlam essa nova forma de gerar conhecimento (Sampaio, 2024).

Outro aspecto relevante é a incerteza sobre a manutenção da gratuidade de serviços de IA generativa, como o ChatGPT e outros chatbots. Salvagno (2023) argumenta que a introdução de cobranças para uso desses serviços poderia ampliar a desigualdade não apenas entre países ricos e pobres, mas também entre pesquisadores jovens, mais familiarizados com essas novas tecnologias, e os mais experientes, favorecendo injustamente os primeiros. Boas e Boas (2024) acrescenta que o etarismo — discriminação baseada na idade, que gera desvantagens e desequilíbrios — pode ser reproduzido pelo ChatGPT em publicações científicas.

Nesse contexto, o desenvolvimento de ferramentas de IA pode ocasionar um aumento expressivo no volume de publicações de alguns pesquisadores, sem que isso necessariamente reflita um aumento real em sua expertise na área. Esse impacto na produtividade acadêmica também pode influenciar os processos de contratação de profissionais em instituições de ensino que, frequentemente, priorizam a quantidade de publicações em vez de sua qualidade (Salvagno, 2023).

No caso do ChatGPT, o sistema pode apresentar respostas baseadas em artigos desatualizados, principalmente de autores dos EUA, devido ao grande volume de citações que esses trabalhos acumulam. Isso faz com que esses artigos continuem sendo citados, mesmo quando existem publicações mais recentes e relevantes, amplificando o "efeito Mateus". Por isso, é fundamental que os pesquisadores façam revisões cuidadosas da literatura, para garantir a qualidade e o rigor acadêmico e evitar a perpetuação de desigualdades (Lund et al., 2023).

Além disso, dado que a pesquisa científica é inerentemente colaborativa — seu desenvolvimento depende da replicação e ampliação de resultados — é crucial que dados e modelos sejam amplamente acessíveis à comunidade. No entanto, a disponibilidade restrita dos modelos de IA mais avançados a alguns grupos de pesquisadores pode comprometer a capacidade de validação independente e a adaptação das pesquisas que utilizam essas tecnologias (Montini, 2024).

Para compreender a situação da inteligência artificial no cenário brasileiro é essencial primeiramente reconhecer a profundidade da desigualdade digital, que se manifesta na limitação do acesso à internet. Pesquisas recentes mostram que menos de um terço da população brasileira tem acesso constante à internet<sup>2</sup>, com pessoas em situação de vulnerabilidade frequentemente enfrentando longos períodos de desconexão (Cetic.br, 2023).

Esse cenário agrava ainda mais a situação em relação à IA. A pesquisa da Genial/Quaest de 2024 revelou que apenas 21% dos brasileiros já utilizaram alguma ferramenta de IA, sendo o seu uso concentrado entre indivíduos com ensino superior e pertencentes às classes sociais mais altas, que já estão em vantagem na adoção de tecnologias emergentes (Silvestre, 2024).

Essa desigualdade no acesso pode impactar significativamente a formação de novos pesquisadores, uma vez que aqueles em situação de vulnerabilidade têm menos oportunidades para desenvolver habilidades digitais, criando um *gap* de

---

<sup>2</sup> Segundo a pesquisa TIC Domicílios, divulgada pelo Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação em agosto de 2023, 29,4 milhões de brasileiros não têm acesso à internet (Cetic.br, 2023).

competência tecnológica que dificulta seu acesso à comunidade científica. Além disso, a concentração de conhecimento e inovação em IA em instituições com maior suporte financeiro e em campos científicos já consolidados, pode limitar a qualidade e a diversidade da pesquisa a nível global, acentuando as disparidades no reconhecimento acadêmico e nas oportunidades de publicação (Lund et al., 2023).

A maneira como a IA impactará as desigualdades sociais dependerá das decisões tomadas durante sua implementação. Se políticas públicas e práticas empresariais promoverem uma regulação consciente e iniciativas inclusivas, essas tecnologias poderão superar barreiras de exclusão e fomentar a inclusão digital, servindo como uma ponte para um futuro em que todos se beneficiem, e não apenas uma parte da população. No entanto, se a necessidade de garantir equidade na era digital for negligenciada, há o risco de aprofundamento e perpetuação das desigualdades sociais, tornando-as ainda mais difíceis de solucionar (Montini, 2024).

No campo acadêmico, é crucial assegurar acesso equitativo às plataformas e ferramentas de pesquisa baseadas em IA, além de fortalecer a capacitação digital para todos os pesquisadores, independentemente de sua localização, instituição ou situação econômica. Ademais, integrar uma perspectiva de justiça social no desenvolvimento de IA, incluindo a participação das comunidades afetadas, também se revela fundamental para que o uso dessa tecnologia seja capaz de promover mudanças positivas, reduzindo os níveis de desigualdade social e exclusão digital e aprimorando a qualidade e o impacto da pesquisa científica globalmente.

## 4 CONCLUSÃO

A inteligência artificial (IA) tem mostrado um grande potencial para transformar a produção científica e o ambiente acadêmico, especialmente com o surgimento de modelos generativos como o ChatGPT. Essas ferramentas introduzem novas maneiras de realizar pesquisas, adquirir conhecimento e compartilhar informações, utilizando uma linguagem mais simples e acessível, além de oferecer resultados de forma rápida e direta. No entanto, sua ampla adoção em diversas áreas traz consigo desafios éticos. Este estudo explorou os principais dilemas éticos discutidos na literatura, como a confiabilidade dos dados, a presença de vieses, questões de direitos autorais, segurança da informação e desigualdade de acesso a plataformas de IAs generativas no contexto da produção de conhecimento científico. Se a ética já era uma questão central para garantir boas práticas na pesquisa, a introdução de ferramentas que operam com base em padrões frequentemente opacos torna o debate sobre seus riscos, consequências e limitações ainda mais necessário.

Os modelos de ChatGPT ainda não são capazes de realizar análises mais profundas, o que os torna incapazes de substituir a expertise e o julgamento crítico dos pesquisadores. Dessa forma, o uso mais eficaz e seguro da IA generativa na produção científica, até o momento, é como uma ferramenta de assistência, em que seu poder de processamento e análise rápida de dados se une à criatividade, intuição e inteligência humana. Essa colaboração não apenas otimiza o processo — ao identificar padrões em grandes volumes de dados e executar tarefas repetitivas,

enquanto os pesquisadores interpretam e inferem conclusões a partir de quantidades muito menores de dados —, mas também possibilita uma supervisão contínua, assegurando maior confiabilidade e correção de erros.

Em domínios científicos onde precisão, explicabilidade e replicabilidade são essenciais, os problemas presentes nas saídas do ChatGTP — como a falta de clareza sobre o processo de geração dos seus resultados, restrições contextuais, imprecisões nas informações e o risco de perpetuação de vieses culturais e preconceitos, além da incapacidade de criar insights verdadeiramente inovadores — reforçam a necessidade de um uso mais cuidadoso e criterioso da ferramenta. Ademais, questões relacionadas à equidade digital evidenciam como o acesso desigual à IA pode intensificar as disparidades sociais já existentes, especialmente entre o Norte e o Sul global. Instituições e pesquisadores que não dispõem dos mesmos recursos financeiros para acessar ferramentas digitais enfrentam obstáculos adicionais para alcançar relevância no ambiente científico internacional.

A implementação dessas ferramentas não veio acompanhada de instrumentos que assegurem a proteção de dados privados ou sensíveis e a correta referência dos autores originais dos conteúdos, podendo não apenas violar direitos legalmente estabelecidos, como também ocasionar a prática de plágio.

A colaboração interdisciplinar entre pesquisadores, educadores, legisladores e desenvolvedores de IA surge como uma abordagem promissora para enfrentar os desafios deste cenário em constante transformação. Nesse contexto, torna-se essencial que as instituições educacionais, universidades e cursos de educação tecnológica e profissional, estabeleçam políticas e diretrizes éticas claras para o uso de plataformas de IA generativa, a fim de evitar abusos e assegurar o cumprimento dos princípios de integridade, privacidade e responsabilidade na pesquisa. Enquanto um consenso internacional sobre o que caracteriza má conduta acadêmica no uso da IA ainda está em desenvolvimento, alguns pontos parecem alcançar uma concordância geral: a necessidade de transparência quanto à adoção dessas ferramentas para a produção de artigos, reconhecendo-as como instrumentos e não como coautoras, além de maior atenção à citação das fontes utilizadas.

Em vez de proibir o uso do ChatGPT e de outras tecnologias de IA, que já se tornaram parte incontornável do mundo contemporâneo, é fundamental avançar nos debates para a criação e gestão de medidas que enfrentem as problemáticas atuais e os riscos que podem comprometer a construção de um futuro socialmente mais inclusivo. Em suma, é necessário que os autores compreendam claramente os limites éticos relacionados à adoção do ChatGPT e ao seu uso responsável, assegurando que o conteúdo produzido seja válido e confiável antes de integrá-lo em suas produções, as quais também irão alimentar a base de dados compartilhados por toda a comunidade científica.

Embora este estudo tenha abordado amplamente os dilemas éticos do uso da IA na produção científica, os impactos variam significativamente entre diferentes áreas do conhecimento, o que torna a generalização dos resultados uma limitação. Além disso, a ausência de investigações empíricas, como entrevistas com pesquisadores, restringe a avaliação direta do impacto dessas ferramentas na prática.

Nesse sentido, recomenda-se a condução de estudos que examinem os efeitos da IA em áreas específicas, uma vez que ciências humanas, biológicas, da saúde e exatas podem enfrentar desafios éticos particulares. Além disso, investigações que acompanhem a percepção de pesquisadores e docentes podem oferecer insights valiosos sobre a aceitação ou resistência ao uso dessas tecnologias ao longo do tempo, revelando seus impactos nas práticas de pesquisa e ensino no Brasil.

## REFERÊNCIAS

ALMEIDA, Lília Bilati de et al. O retrato da exclusão digital na sociedade brasileira. **Jistem - Journal of Information Systems and Technology Management**, v. 2, p. 55-67, 2005.

AAAI - ASSOCIATION FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE. **Thirty-Eighth AAI Conference on Artificial Intelligence; Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence**. Editado por Michael Wooldridge, Jennifer Dy, Sriraam Natarajan. Washington, DC: AAI Press, 2024.

BOAS, Paulo José Fortes Villas; BOAS, José Vitor Polachini do Valle Villas. The use of ChatGPT in scientific publishing. **Geriatrics, Gerontology and Aging**, v. 17, p. 1-4, 2023.

BRASIL. Lei nº 9.610, de 19 de fevereiro de 1998. **Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências**. Diário Oficial da República Federativa do Brasil, Brasília, DF, 20 fev. 1998.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Dispõe sobre a proteção de dados pessoais**. Diário Oficial da União: seção 1, Brasília, DF, 15 ago. 2018.

CATARINI, Paola. Por uma IA democrática, inclusiva e decolonial: “frameworks” de proteção a todos os direitos fundamentais. **Jornal da Usp**, 16 maio 2023. Disponível em: <https://jornal.usp.br/artigos/por-uma-ia-democratica-inclusiva-e-decolonial-frameworks-de-protecao-a-todos-os-direitos-fundamentais/>. Acesso em: 15 set. 2024.

CETIC.br. **Pesquisa TIC Domicílios**. 2023. São Paulo: Comitê Gestor da Internet no Brasil, 2023. Disponível em: <https://cetic.br/pt/pesquisa/domicilios/indicadores/>. Acesso em: 17 set. 2024.

COPE. **Artificial intelligence and peer review**. 2023. Disponível em: <https://publicationethics.org/publication-integrity-week-2023/ai-peer-review>. Acesso em: 25 set. 2024.

DECOLONIZAI. Manifesto da IA Decolonial. **DecolonizAI**, 31 mar. 2023. Disponível em: <https://decolonizai.com/manifesto-da-ia-decolonial/>. Acesso em: 16 set. 2024.

EKE, Daniel. ChatGPT and the rise of generative AI: threat to academic integrity?. **Journal of Responsible Technology**, v. 13, p. 100060, 2023.

HAMAM, Ahmed Ammar; FATOUH, Amr Hassan. Leveraging ChatGPT in Scientific Research: A Comprehensive Analysis. **Revista da Faculdade de Artes de Al-Wadi Al-Jadid**, v. 9, n. 18, 2023.

GANJAVI, Conner et al. Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. **BMJ**, v. 384, 2024.

GAO, Xiangmin; SHEN, Haoming; DUAN, Joan Huiqiong. A Study on ChatGPT to Alleviate the Difficulties in Writing Scientific Technology Papers. In: **2023 International Conference on the Cognitive Computing and Complex Data (ICCD)**. IEEE, 2023. p. 118-122.

GUIMARÃES, João Alexandre Silva Alves. O direito ao esquecimento: a última chance de sermos nós mesmos?. **Migalhas**, 3 set. 2021. Disponível em: <https://www.migalhas.com.br/coluna/migalhas-de-protecao-de-dados/351126/o-direito-ao-esquecimento-a-ultima-chance-de-sermos-nos-mesmos>. Acesso em: 15 set. 2024.

HENDRYCKS, Dan; ZHAO, Jeremy; WALLACE, Eric. **Red Teaming AI: Evaluating the Risks and Vulnerabilities of Large Language Models**. 2023. Disponível em: <https://arxiv.org/abs/2306.01723>. Acesso em: 12 set. 2024.

HIRSH-PASEK, Kathy; BLINKOFF, Elena. ChatGPT: Educational friend or foe?. **Brookings**, 9 jan. 2023. Disponível em: <https://www.brookings.edu/articles/chatgpt-educational-friend-or-foe/>. Acesso em: 11 set. 2024.

ISIAKU, Labaran et al. Academic Evolution in the Age of ChatGPT: An In-depth Qualitative Exploration of its Influence on Research, Learning, and Ethics in Higher Education. **Journal of University Teaching and Learning Practice**, v. 21, n. 06, 2024.

KAMOUN, Faouzi et al. Exploring Students' and Faculty's Knowledge, Attitudes, and Perceptions Towards ChatGPT: A Cross-Sectional Empirical Study. **Journal of Information Technology Education: Research**, v. 23, p. 004, 2024.

KHALIFA, Mohamed; ALBADAWY, Mona. Using artificial intelligence in academic writing and research: An essential productivity tool. **Computer Methods and Programs in Biomedicine Update**, p. 100145, 2024.

LUND, Brady et al. ChatGPT and a new academic reality: AI-written research papers and the ethics of the large language models in scholarly publishing. **Journal of the Association for Information Science and Technology**, v. 74, n. 5, p. 570-581, 2023.

MARWALA, Tshilidzi. Algorithm Bias — Synthetic Data Should Be Option of Last Resort When Training AI Systems. **ONU**, 28 jul. 2023. Disponível em: <https://unu.edu/article/algorithm-bias-synthetic-data-should-be-option-last-resort-when-training-ai-systems>. Acesso em: 25 set. 2024.

MENESES, Sônia. Apontamentos sobre o uso de AI na produção e na comunicação de pesquisas. **Revista Brasileira de História**, v. 43, n. 94, set.-dez. 2023.

MONTINI, Alessandra. Desigualdade digital: o papel da IA na redução ou ampliação das diferenças sociais. **Febraban Tech**, 12 ago. 2024. Disponível em: <https://febrabantech.febraban.org.br/especialista/alessandra-montini/desigualdade-digital-o-papel-da-ia-na-reducao-ou-ampliacao-das-diferencas-sociais>. Acesso em: 16 set. 2024.

MOREIRA, Sandra Seabra. Equipe multidisciplinar garante ética em projetos de IA. **Revista Ensino Superior**, n. 278, p. 34-36, set. 2023.

NATURE. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. **Nature**, v. 613, n. 7945, p. 612, 2023.

NEMORIN, Selena. Towards decolonising the ethics of AI in education. **Globalisation, Societies and Education**, 08 abr. 2024.

NEPAL, Tej Kumar. Exploring the Applications and Challenges of ChatGPT in Research and Academia: A Comprehensive Review. **West Science Interdisciplinary Studies**, v. 2, n. 05, p. 1043-1050, 2024.

NIC.br - Núcleo de Informação e Coordenação do Ponto BR. **Inteligência artificial e cultura: Perspectivas para a diversidade cultural na era digital**. São Paulo: Comitê Gestor da Internet no Brasil, 2023.

OPENAI. **GPT-4**. 14 mar. 2023. Disponível em: <https://openai.com/index/gpt-4-research/>. Acesso em: 21 ago. 2024.

PCAST - President's Council of Advisors on Science and Technology. **Supercharging research: harnessing artificial intelligence to meet global challenges**. Executive Office of the President, 2024.

RAHMAN, Md Mizanur et al. ChatGPT and academic research: A review and recommendations based on practical examples. **Journal of Education, Management and Development Studies**, v. 3, n. 1, p. 1-12, 2023.

SALVAGNO, Michele; TACCONE, Fabio Silvio; GERLI, Alberto Giovanni. Can artificial intelligence help for scientific writing?. **Critical Care**, v. 27, n. 1, p. 75, 2023.

SAMPAIO, Rafael Cardoso et al. ChatGPT e outras IAs transformarão a pesquisa científica: reflexões sobre seus usos. **Revista de Sociologia e Política**, v. 32, p. e008, 2024.

SILVESTRE, Paulo. Inteligência artificial pode ampliar abismo social no Brasil. **Estadão**, 29 jul. 2024. Disponível em: <https://www.estadao.com.br/brasil/macaco-eletrico/inteligencia-artificial-pode-ampliar-abismo-social-no-brasil/>. Acesso em: 17 set. 2024.

SUSARLA, Anjana; GOPAL, Ram; THATCHER, Jason; SARKER, Saonee. The Janus effect of generative AI: charting the path for responsible conduct of scholarly activities in information systems. **Information Systems Research**, v. 34, n. 2, p. 399-408, 2023.

TLILI, Abderezak et al. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. **Smart Learning Environments**, v. 10, n. 1, p. 15, 2023.

TORRES-SALINAS, Daniel; ARROYO-MACHADO, Wenceslao. **The Many Publics of Science: Using Altmetrics to Identify Common Communication Channels by Scientific Field.** arXiv, 2023. Disponível em: <https://arxiv.org/abs/2304.05157>. Acesso em: 28 set. 2024.

UNESCO. **Recomendação sobre a Ética da Inteligência Artificial.** Paris: Organização das Nações Unidas para a Educação, a Ciência e a Cultura, 2022.

UNESCO. **Guidance for generative AI in education and research.** Paris: UNESCO Publishing, 2023.

VALERI, Julia. Inteligência artificial utiliza base de dados que refletem preconceitos e desigualdades. **Jornal da USP**, 7 jul. 2023. Disponível em: <https://jornal.usp.br/atualidades/inteligencia-artificial-utiliza-base-de-dados-que-refletem-preconceitos-e-desigualdades/>. Acesso em: 16 set. 2024.

VILALTA, Lucas Paolo Sanches. É possível termos transparência de algoritmos para sistemas de inteligência artificial?. **Jornal da USP**, 05 ago. 2024. Disponível em: <https://jornal.usp.br/artigos/e-possivel-termos-transparencia-de-algoritmos-para-sistemas-de-inteligencia-artificial/>. Acesso em: 10 set. 2024.

ZHOU, Liangbin et al. ChatGPT for scientific writing —The coexistence of opportunities and challenges. **Journal of Orthopaedic Translation**, v. 44, p. A1, 2024.