Autoria Humana e Escrita Assistida por LLMs: Desafios Éticos e Jurídicos na Produção Científica

Thiago R. Lobo[1], **Josiel M. Figueiredo**[1], **Claudia A. Martins**[1]

[1]Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)
Caixa Postal 15.064 – 91.501-970 – Cuiabá – MT – Brasil

Abstract. *The increasing integration of Large Language Models (LLMs) into scientific writing has significantly reshaped research communication, enhancing productivity and reducing linguistic barriers for authors. However, these tools also introduce substantial ethical and legal challenges related to authorship, originality, and intellectual responsibility. This article examines the tensions arising from hybrid human–AI writing, particularly the risks of copyright infringement and the ambiguities surrounding creative contribution. It further analyzes technical and editorial approaches for assessing authorship and originality, highlighting the limitations of current detection methods and the need for criteria that safeguard scientific integrity. The findings indicate that transparent practices, combined with the human author's critical responsibility, are essential to delineate the role of AI and ensure compliance with ethical and legal standards in contemporary scientific production.*

Resumo. *A crescente incorporação de Large Language Models (LLMs) na produção científica tem transformado de maneira significativa os processos de escrita, revisão e comunicação acadêmica. Embora essas ferramentas ampliem a produtividade e reduzam barreiras linguísticas, elas também introduzem desafios relevantes para os conceitos de autoria, originalidade e responsabilidade intelectual. Este artigo examina as tensões éticas e jurídicas decorrentes do uso de LLMs, com ênfase nos riscos de violação de direitos autorais e nas ambiguidades associadas à escrita híbrida entre humanos e sistemas de IA. Além disso, analisa abordagens técnicas e editoriais para avaliar autoria e originalidade, discutindo as limitações dos métodos atuais de detecção e a necessidade de critérios que preservem a integridade científica. Conclui-se que a consolidação de práticas transparentes, aliada à responsabilidade crítica do autor humano, é fundamental para delimitar a contribuição da IA e assegurar conformidade com os parâmetros ético-normativos da produção científica contemporânea.*

1. Introdução

A emergência dos Large Language Models (LLMs), como o ChatGPT, desde 2022, tem provocado mudanças relevantes na produção e comunicação científica, comparáveis a transformações históricas na disseminação do conhecimento [Lin et al. 2025]. Treinados em extensos corpora textuais, esses sistemas são capazes de gerar linguagem complexa e próxima à humana, o que ampliou seu uso em atividades acadêmicas, especialmente na escrita de textos científicos.

Entre os benefícios observados estão o aumento da produtividade na redação e revisão de manuscritos [Noy and Zhang 2023] e o potencial de reduzir barreiras linguísticas

para autores não nativos de inglês, melhorando a legibilidade e a complexidade lexical de textos acadêmicos [Lin et al. 2025] [Berdejo-Espinola and Amano 2023].

Por outro lado, o uso crescente desses sistemas intensifica preocupações sobre a integridade científica. A possibilidade de geração de informações incorretas, a padronização estilística e a influência sobre processos de escrita crítica levantam questões normativas relacionadas à autoria, originalidade e responsabilidade intelectual, especialmente porque LLMs não possuem intenção comunicativa nem podem assumir responsabilidade moral ou jurídica [King 2023].

Nesse contexto, torna-se cada vez mais difícil distinguir a contribuição intelectual humana daquela assistida por IA, o que desafia critérios tradicionais de autoria e exige maior transparência no uso dessas ferramentas. Diante disso, este artigo analisa critérios éticos, jurídicos e editoriais capazes de diferenciar a contribuição humana da assistência de LLMs na escrita científica.

O trabalho está organizado da seguinte forma: i) Direitos Autorais na Produção Científica; ii) Riscos de Violação de Direitos Autorais com o Uso de LLMs; iii) Abordagens e Critérios para Avaliar Autoria e Originalidade na Escrita Científica Assistida por LLMs; iv) Conclusão.

2. Direitos Autorais na Produção Científica

A autoria científica está tradicionalmente associada à contribuição intelectual substantiva, esforço criativo e responsabilidade pelo conteúdo produzido. Diretrizes amplamente adotadas, como as do International Committee of Medical Journal Editors (ICMJE), estabelecem que apenas indivíduos capazes de assumir responsabilidade ética e intelectual podem ser considerados autores [International Committee of Medical Journal Editors 2022], o que exclui sistemas de IA, por não possuírem agência ou capacidade de endosso consciente.

No campo jurídico, o Direito Autoral protege obras resultantes do "labor intelectual" humano, incluindo artigos científicos concebidos para fins de comunicação científica [Smiraglia, 2019]. Nesse contexto, a originalidade é um critério central, sendo reconhecida quando a obra reflete escolhas criativas e a subjetividade de seu autor [Morocco-Clarke et al. 2023].

A introdução de LLMs tensiona esses parâmetros, pois ferramentas capazes de sintetizar literatura, sugerir argumentos ou estruturar textos podem influenciar partes relevantes de um manuscrito. Embora tratadas como instrumentos auxiliares, sua utilização torna menos evidente a fronteira entre suporte tecnológico e contribuição intelectual. Dessa forma, torna-se necessária a definição de critérios claros para avaliar a extensão da intervenção dessas ferramentas, considerando que a responsabilidade final pela integridade e conformidade ética do texto permanece atribuída ao pesquisador humano.

3. Riscos de Violação de Direitos Autorais com o Uso de LLMs

Os principais riscos de violação de direitos autorais decorrem tanto do processo de treinamento dos LLMs quanto da sua capacidade de reproduzir, de forma literal ou parafraseada, fragmentos de obras protegidas. Como esses modelos se apoiam em vastos corpora de dados pouco documentados, há incerteza sobre o volume de

conteúdos protegidos incluídos em seus conjuntos de treinamento [Paullada et al. 2020] [Bender et al. 2021].

Estudos recentes demonstram que LLMs podem memorizar e regurgitar trechos significativos de textos, especialmente à medida que o tamanho do modelo aumenta, o que implica risco real de redistribuição não autorizada de material protegido [Karamolegkou et al. 2023] [Carlini et al. 2022]. Além disso, a opacidade estrutural desses sistemas dificulta auditorias independentes, tornando árdua a verificação de violações potencialmente cometidas durante o treinamento ou geração de conteúdo.

Esses fatores colocam o usuário humano em posição de responsabilidade jurídica, ainda que ele não tenha controle direto sobre os dados utilizados pelo modelo. Diante da impossibilidade de atribuir responsabilidade ao sistema de IA, cabe ao autor verificar a originalidade do conteúdo produzido, revisar possíveis trechos problemáticos e assegurar conformidade com a legislação de direitos autorais.

4. Abordagens e Critérios para Avaliar Autoria e Originalidade na Escrita Científica Assistida por LLMs

A popularização de LLMs na escrita científica exige novas abordagens para avaliar autoria e originalidade em textos híbridos. O desafio consiste em identificar traços estilísticos, estruturas argumentativas ou padrões linguísticos que permitam diferenciar contribuições humanas de intervenções geradas por IA, sem penalizar injustamente textos legítimos.

Entre as abordagens técnicas, destacam-se análises linguísticas que identificam padrões recorrentes associados a modelos generativos, como o aumento abrupto de determinadas palavras de estilo [Kobak et al. 2025].

Trabalhos como [Alshareef et al. 2025], [Qorich and Ouazzani 2025], [Yin and Wang 2026] e [Yadav and MC 2025] apresentam abordagens tradicionais para a distinção entre textos humanos e sintéticos por meio de técnicas de Aprendizado de Máquina, Aprendizado Profundo e arquiteturas baseadas em Transformers, empregadas tanto para extração de características quanto para processos de fine-tuning. Apesar dos bons resultados reportados, grande parte dessas estratégias opera em contextos de classificação binária, desconsiderando cenários em que o texto não é totalmente gerado por IA, mas sim revisado ou refinado com auxílio dessas ferramentas.

Além disso, o processo de fine-tuning geralmente exige a construção de datasets específicos para o treinamento dos modelos. Na revisão sistemática conduzida por [Fraser et al. 2025], observou-se que a maioria dos datasets disponíveis encontra-se no idioma inglês. Nesse contexto, o desenvolvimento de modelos voltados para corpora em língua portuguesa, especialmente no contexto brasileiro, mostra-se particularmente desafiador.

Outra abordagem comumente usada é o watermarking, que insere sinais ocultos em textos gerados por IA com o objetivo de rastrear sua origem. Embora promissor, esse método ainda enfrenta desafios técnicos, como vulnerabilidade a ataques de paráfrase profunda, tradução reversa e substituição de sinônimos. Além disso, a escalabilidade para múltiplos idiomas ainda é um desafio. [Tudino and Qin 2024].

No âmbito editorial, cresce o entendimento de que critérios de avaliação devem se concentrar no processo criativo e na responsabilidade assumida pelo autor humano,

em vez de se apoiar exclusivamente na detecção automatizada de padrões linguísticos. A transparência no uso de IA e a clareza na descrição do papel desempenhado por ferramentas de LLM são elementos essenciais para preservar a integridade científica.

5. Conclusão

O uso crescente de Large Language Models (LLMs) na escrita científica intensifica desafios éticos e jurídicos relacionados à autoria, originalidade e responsabilidade científica. Embora essas ferramentas possam apoiar variadas atividades, elas não possuem agência ou responsabilidade, o que mantém o pesquisador humano como responsável final pela integridade, precisão e conformidade ética do conteúdo produzido.

Na Figura 1 é apresentado um modelo conceitual que organiza o problema em três dimensões: (i) fundamentos normativos da autoria científica, baseados em contribuição intelectual, originalidade e responsabilidade; (ii) riscos jurídicos associados ao uso de LLMs, incluindo possíveis violações de direitos autorais decorrentes do treinamento e da geração de conteúdo; e (iii) abordagens técnicas e editoriais para avaliar a originalidade em textos assistidos por IA, como métodos de detecção automática, análises linguísticas, watermarking e práticas editoriais baseadas em transparência.

Esse modelo evidencia que a avaliação da escrita científica assistida por IA não pode depender exclusivamente de técnicas automatizadas, as quais ainda apresentam limitações diante de modelos avançados e textos híbridos. Assim, a preservação da integridade científica requer a combinação de mecanismos técnicos, critérios editoriais e diretrizes claras de transparência quanto ao uso dessas ferramentas. Nesse contexto, a responsabilidade pela qualidade e originalidade do trabalho permanece atribuída ao autor humano.

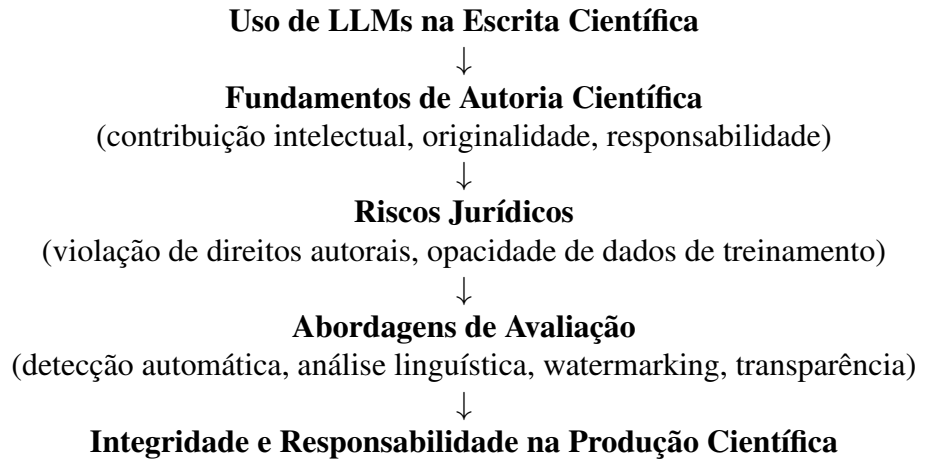

Figure 1. Modelo conceitual para análise da escrita científica assistida por LLMs.

References

Alshareef, A. M., Alsobhi, A., Khadidos, A. O., Alyoubi, K. H., Khadidos, A. O., and Ragab, M. (2025). Automated detection of chatgpt-generated text vs. human text using gannet-optimized deep learning. *Alexandria Engineering Journal*, 124:495–512.

Bender, E. M., Gebru, T., McMillan-Major, A., and Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, pages 610–623.

Berdejo-Espinola, V. and Amano, T. (2023). Ai tools can improve equity in science. *Science*, 379(6636):991.

Carlini, N., Ippolito, D., Jagielski, M., Lee, K., Tramèr, F., and Zhang, C. (2022). Quantifying memorization across neural language models. arXiv preprint arXiv:2202.07646.

Fraser, K. C., Dawkins, H., and Kiritchenko, S. (2025). Detecting ai-generated text: Factors influencing detectability with current methods. *Journal of Artificial Intelligence Research*, 82.

International Committee of Medical Journal Editors (2022). Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals.

Karamolegkou, A., Li, J., Zhou, L., and Søgaard, A. (2023). Copyright violations and large language models. arXiv preprint arXiv:2310.13771.

King, M. R. (2023). A place for large language models in scientific publishing, apart from credited authorship. *Cellular and Molecular Bioengineering*, 16:95–98.

Kobak, D., González-Márquez, R., Horvát, E.-Á., and Lause, J. (2025). Delving into llm-assisted writing in biomedical publications through excess vocabulary. *Science Advances*, 11(27).

Lin, D., Zhao, N., Tian, D., and Li, J. (2025). Chatgpt as linguistic equalizer? quantifying llm-driven lexical shifts in academic writing. arXiv preprint arXiv:2504.12317.

Morocco-Clarke, S., Sodangi, A., and Momodu, J. (2023). The implications and effects of chatgpt on academic scholarship and authorship: A death knell for original academic publications? *Information & Communications Technology Law*.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

Paullada, A., Raji, I. D., Bender, E. M., Denton, E., and Hanna, A. (2020). Data and its (dis)contents: A survey of dataset development and use in machine learning research. arXiv preprint arXiv:2012.05345.

Qorich, M. and Ouazzani, R. E. (2025). Detection of artificial intelligence-generated essays for academic assessment integrity using large language models. *Expert Systems with Applications*, 291:128405.

Tudino, G. and Qin, Y. (2024). A corpus-driven comparative analysis of ai in academic discourse: Investigating chatgpt-generated academic texts in social sciences. *Lingua*, 312:103838.

Yadav, A. and MC, S. P. (2025). Classifying ai vs. human content: Integrating bert and linguistic features for enhanced classification. *Operations Research Forum*, 6:77.

Yin, Z. and Wang, S. (2026). Span-level detection of ai-generated scientific text via contrastive learning and structural calibration. *Knowledge-Based Systems*, 334:115123.