Artificial Intelligence in Scientific Publishing: Implications for Authors, Peer Reviewers, and Editors

Inteligência Artificial na Publicação Científica: Implicações para Autores, Revisores e Editores

Helena Donato [1,2]

Introdução

Os modelos de linguagem de grande dimensão (*large language models*, LLMs) constituem a tecnologia subjacente à atual vaga de inteligência artificial generativa. Estes modelos são sistemas de IA treinados com volumes massivos de texto, concebidos para aprender padrões linguísticos complexos e, assim, gerar texto com aparência humana. A sua capacidade inclui produzir, resumir, traduzir, explicar conteúdos e responder a perguntas.[1]

Ferramentas como ChatGPT, Perplexity ou Gemini não são, em si mesmas, os modelos, mas aplicações que usam LLMs como infraestrutura tecnológica.[1] Ou seja, os LLMs são a tecnologia de base que permite gerar e compreender texto; os *chatbots*, como o ChatGPT, são aplicações que usam esses modelos para interagir com o utilizador.[1]

A distinção não é meramente terminológica: enquanto os LLMs representam a arquitetura e os mecanismos de aprendizagem, as aplicações concretizam usos específicos, com diferentes níveis de controlo, integração e transparência.[1] Importa, contudo, reconhecer uma limitação central. Nesta fase de desenvolvimento, a qualidade dos resultados gerados pela IA generativa não é perfeita. Persistem erros factuais, enviesamentos decorrentes dos dados de treino e fenómenos de "alucinação", nos quais o sistema gera informação plausível, mas incorreta. Ainda assim, a evolução tecnológica é rápida e observa-se uma melhoria progressiva na robustez e desempenho destes sistemas. Esta tensão entre potencial e limitação está no cerne do debate atual sobre o seu uso na investigação e publicação científica.[2]

A integração de sistemas de inteligência artificial (IA) no ecossistema editorial científico é hoje uma realidade incontornável.[2] Desde a triagem inicial de manuscritos até à deteção de plágio e de manipulação de imagens, as ferramentas algorítmicas têm vindo a transformar os fluxos de trabalho editoriais. No entanto, quando se desloca o foco para a revisão por pares, para a autoria e para a produção de conteúdo científico com recurso à IA, emerge uma questão central: poderá a IA substituir, ou sequer replicar, o juízo humano e a responsabilidade que sustentam a avaliação e a comunicação científica?

IA Generativa

A adoção acelerada de ferramentas de IA generativa, tem vindo a transformar práticas de investigação e redação científica. Perante esta realidade, editores e revistas científicas reagiram com celeridade, publicando orientações específicas para regular o uso de ferramentas como o ChatGPT e sistemas análogos no processo de publicação.[2] A emergência destas políticas não é arbitrária. Responde a riscos concretos: criação de conteúdos incorretos, plágio não intencional ou oculto, reprodução de preconceitos presentes nos dados de treino, bem como questões de direitos de autor e de privacidade. A ausência de enquadramento normativo poderia comprometer a integridade do registo científico. Observa-se, contudo, uma forte convergência entre os principais organismos editoriais e revistas de referência.

Para a maioria das editoras que seguem as Recomendações do International Committee of Medical Journal Editors, o uso de IA generativa na redação de artigos científicos é permitido, sob condições rigorosas[3]:

- Supervisão humana obrigatória: A IA deve ser sempre supervisionada e controlada pelos autores.
- Responsabilidade plena dos autores: Os autores são integralmente responsáveis pelo conteúdo publicado, independentemente das ferramentas utilizadas durante o processo de escrita.
- Uso limitado à melhoria da clareza: A IA pode ser usada para melhorar a clareza e legibilidade do texto, mas nunca para substituir o raciocínio científico, a interpretação dos dados ou o juízo crítico do autor. Isto implica que todo o conteúdo gerado ou revisto com recurso à IA deve ser cuidadosamente verificado, editado e validado. A possibilidade de erros, omissões ou "alucinações" exige uma revisão crítica sistemática por parte dos autores.
- Divulgação e transparência no momento da submissão: A transparência constitui um elemento central das políticas atuais. Quando usada, a IA deve ser declarada de forma explícita, podendo essa divulgação ocorrer: na carta de apresentação (*cover letter*); na secção de Agradecimentos; na secção de Métodos, caso a ferramenta tenha sido utilizada no desenho do estudo,

[1]Diretora Serviço Documentação e Informação Científica, Hospitais da Universidade de Coimbra, Unidade Local de Saúde de Coimbra, Coimbra, Portugal.
[2]Editora Técnica, Revista Portuguesa de Medicina Interna

https://doi.org/10.24950/rspmi.2846

na recolha ou análise de dados, à semelhança de qualquer outro *software* científico. A declaração deve incluir, sempre que aplicável, o nome da ferramenta, a versão, o fornecedor e a finalidade do uso.

Importa distinguir entre diferentes tipos de ferramentas. Corretores gramaticais ou ortográficos convencionais não requerem declaração. No entanto, o uso de LLMs para reformulação, resumo ou verificação gramatical exige divulgação, dado o seu potencial contributo substantivo para a produção textual.

Nunca é aceitável listar uma ferramenta de IA como autora. A autoria científica pressupõe responsabilidade autoral, capacidade de responder por decisões metodológicas e declaração de conflitos de interesse, atributos que não podem ser atribuídos a sistemas automatizados.[2-4]

Paralelamente às orientações sobre escrita e autoria, surgiram extensões específicas das principais reporting *guidelines* para estudos que envolvem inteligência artificial: CONSORT-AI: extensão das normas CONSORT para ensaios clínicos que envolvem intervenções baseadas em IA. SPIRIT-AI: extensão das orientações SPIRIT para protocolos de estudos com intervenções baseadas em IA. STARD-AI: adaptação das normas STARD para estudos de diagnóstico que utilizam IA. Estas iniciativas refletem o reconhecimento de que a investigação envolvendo IA apresenta especificidades metodológicas.

O Propósito da Revisão por Pares

A revisão por pares não é um exercício meramente técnico de verificação formal. Assenta em três pilares: imparcialidade, transparência e competência científica. É pedido ao revisor não apenas que identifique falhas metodológicas ou inconsistências estatísticas, mas que exerça um julgamento crítico informado pela experiência, pelo contexto e por uma compreensão aprofundada do campo. Delegar a escrita do relatório de revisão a um sistema de IA mina este propósito, pois dilui a responsabilidade intelectual e desvirtua a própria natureza da revisão por pares.[5,6]

Um dos problemas mais sérios associados ao uso de IA na revisão prende-se com a confidencialidade. Manuscritos em avaliação constituem comunicações privilegiadas. A introdução de qualquer parte do manuscrito em sistemas de IA públicos viola, por definição, o acordo de confidencialidade que rege a revisão. Muitos sistemas de IA retêm e processam a informação que lhes é fornecida, podendo usá-la para melhorar respostas futuras. Ainda que não exista intenção maliciosa, a simples possibilidade de reutilização de conteúdo confidencial constitui uma violação ética grave.[5]

Para além da confidencialidade, importa considerar as limitações técnicas. A IA é suscetível a "alucinações", produzindo informações incorretas, incompletas ou tendenciosas. Podem formular críticas genéricas, deturpar o âmbito do artigo ou apresentar raciocínios inconsistentes.

Várias organizações internacionais têm sublinhado este ponto. O International Committee of Medical Journal Editors (ICMJE),[3] na sua atualização mais recente (https://www.icmje.org/recommendations/ - Janeiro de 2026), recomenda que as revistas adotem políticas que proíbam o carregamento de manuscritos em ferramentas de IA em que a confidencialidade não possa ser garantida. De forma semelhante, o Committee on Publication Ethics (COPE https://publicationethics.org/) afirma que o uso de IA na revisão por pares não deve comprometer a confidencialidade nem substituir a supervisão humana. A World Association of Medical Editors (WAME https://wame.org/policies) desenvolveu orientações específicas para revisores, enfatizando a necessidade de preservar a integridade do processo e evitar pareceres enviesados ou de baixa qualidade.

Também alguns grupos editoriais concretizam estas orientações em políticas explícitas. O The Lancet Group proíbe o carregamento de manuscritos em ferramentas públicas de IA e considera os revisores plenamente responsáveis pelos seus pareceres. O *The BMJ* permite apenas o uso limitado de IA para apoio linguístico, com declaração obrigatória, vedando expressamente o envio de conteúdos confidenciais para plataformas públicas.

Apesar destas orientações claras de organismos como o ICMJE e de várias editoras, nem todos os revisores cumprem integralmente estas políticas. Atualmente é possível identificar alguns relatórios gerados por IA, pois tendem a exibir características reconhecíveis: linguagem vaga e excessivamente abrangente, críticas abstratas pouco suportadas nos dados concretos do manuscrito, ausência de subtileza argumentativa e um tom uniforme que carece da individualidade e da responsabilidade próprias de um revisor humano.[5-7]

Também preocupante é o enviesamento algorítmico. Estes modelos são treinados com base na literatura científica dominante e tendem a reproduzir os padrões mais comuns, podendo desvalorizar abordagens inovadoras ou menos convencionais. Num processo que deve promover o pensamento crítico e a diversidade de perspetivas, este risco é relevante.

Recomendações práticas para revisores:

- Proteger a confidencialidade: nunca inserir qualquer parte do manuscrito ou dos comentários de revisão em ferramentas públicas de IA.
- Uso limitado, se permitido: restringir a IA a apoio linguístico ou clarificação genérica de conceitos, sem delegar julgamentos científicos ou recomendações editoriais.
- Divulgação transparente: informar a revista sobre qualquer utilização de IA e o respetivo propósito.
- Assumir responsabilidade plena: o revisor, não a ferramenta, é responsável pela precisão, imparcialidade e adequação do relatório de revisão.

IA no Processo Editorial: Onde Reside o seu Valor

Seria, contudo, redutor rejeitar a IA em bloco. No âmbito editorial, existem aplicações legítimas e potencialmente benéficas, sobretudo quando se usam sistemas controlados institucionalmente ou ferramentas específicas, desenvolvidas para tarefas delimitadas. Na deteção de fraude, por exemplo, *softwares* especializados permitem identificar plágio, padrões associados a produção automatizada de texto e manipulação ou duplicação de imagens. Na triagem inicial, podem sinalizar falhas éticas evidentes, verificar a adequação à revista ou identificar omissões de documentos obrigatórios na submissão. Podem ainda apoiar a verificação de conformidade com diretrizes internacionais e auxiliar na seleção de revisores com base em palavras-chave. Nestes contextos, a IA atua como instrumento técnico sob supervisão humana e com finalidades bem definidas, não como substituto do juízo editorial ou científico. O valor acrescentado depende do enquadramento, da transparência e do controlo humano sobre o seu uso.

Convergência Normativa

Observa-se uma convergência notável entre revistas científicas quanto a princípios fundamentais:

- A IA não pode ser autora: apenas seres humanos podem cumprir critérios de autoria e assumir responsabilidade.
- Divulgação obrigatória do uso de IA: nome da ferramenta, versão, fornecedor, finalidade e âmbito.
- Responsabilidade humana inalienável: autores e revisores permanecem responsáveis por todo o conteúdo.
- Proibição do envio de manuscritos para ferramentas públicas de IA quando a confidencialidade não é garantida.
- Reconhecimento do caráter dinâmico das políticas, que deverão evoluir à medida que a tecnologia se desenvolve.

Esta convergência sugere que o debate não é sobre aceitar ou rejeitar a IA, mas sobre delimitar com clareza o seu âmbito legítimo de utilização.

Riscos, Evidência e o Papel Central do Julgamento Humano

As principais preocupações associadas ao uso da IA generativa na publicação científica são consistentes e interligadas: produção de informação incorreta, recurso a fontes não fiáveis, enviesamentos algorítmicos, homogeneização discursiva e perda de originalidade, proliferação de artigos falsos ou práticas editoriais não éticas, riscos para a integridade científica — incluindo falsificação ou fabrico de dados e questões de direitos de autor.[8]

Estes riscos não são meramente teóricos. Um estudo publicado na *Journal of Clinical Anesthesia* demonstrou que 21 *chatbots*, incluindo ChatGPT, Copilot e Gemini, identificaram corretamente menos de metade dos artigos retratados numa lista de referências e geraram uma proporção relevante de falsos positivos.[9]

Além disso, produziram respostas inconsistentes perante pedidos idênticos e recorreram a classificações vagas como "possivelmente retratado", sem utilidade científica clara. Resultados convergentes publicados na *Learned Publishing* mostraram que o ChatGPT não reconheceu retratações previamente sinalizadas e chegou a classificar alguns desses trabalhos como de elevada qualidade.[10]

Acresce que outros estudos indicam que LLMs podem incorporar informação proveniente de artigos retratados nas suas respostas, aumentando o risco de perpetuar erros.

Conclusão

A IA generativa não pode substituir autores e revisores. A convergência das políticas editoriais é inequívoca: a IA não pode ser autora; o seu uso deve ser transparente e devidamente declarado; e autores, revisores e editores permanecem plenamente responsáveis por todo o conteúdo publicado.

O desafio não reside em proibir a tecnologia, mas em enquadrá-la de forma ética e proporcional. A ciência pode beneficiar de ferramentas inteligentes; contudo, a confiança no conhecimento publicado depende de limites claros. A integridade científica continua, e continuará, a ser uma responsabilidade humana indelegável.

A inteligência artificial na publicação científica vai muito além dos *chatbots* generalistas como o ChatGPT. Existem também ferramentas especializadas orientadas especificamente para a escrita e gestão de artigos científicos, como o SciSpace, o Paperpal e outras plataformas semelhantes. Estas ferramentas ao contrário dos *chatbots* de uso geral, são desenhadas para contextos editoriais concretos, podendo aumentar a eficiência quando usadas com supervisão crítica. Assim, há claramente um "mundo para além do ChatGPT". ■

Ethical Disclosures

Conflicts of Interest: The authors have no conflicts of interest to de-clare.

Financial Support: This work has not received any contribution grant or scholarship.

Provenance and Peer Review: Commissioned; without external peer review.

Responsabilidades Éticas

Conflitos de Interesse: Os autores declaram a inexistência de conflitos de interesse.

Apoio Financeiro: Este trabalho não recebeu qualquer subsídio, bolsa ou financiamento.

Proveniência e Revisão por Pares: Solicitado; sem revisão externa por pares.

Received / Recebido: 2026/02/12
Accepted / Aceite: 2026/02/13
Published Online / Publicado Online: 2026/02/27
Published / Publicado: 2026/02/27

REFERÊNCIAS

1. Ibrahim EI, Voyer A. Qualitative research with LLM chatbots: Technological reflexivity for interpretative technology. Qualitative Res. 2025;26:133-59. doi: 10.1177/14687941251390794
2. Kim JK, Chua M, Rickard M, Lorenzo A. ChatGPT and large language model (LLM) chatbots: The current state of acceptability and a proposal for guidelines on utilization in academic medicine. J Pediatr Urol. 2023;19:598-604. doi: 10.1016/j.jpurol.2023.05.018.
3. International Committee of Medical Journal Editors. V. Use of Artificial Intelligence in Publishing [acedido 11 Fev 2026] Dsiponível em https://www.icmje.org/recommendations/browse/artificial-intelligence/
4. Flanagin A, Bibbins-Domingo K, Berkwits M, Christiansen SL. Nonhuman "Authors" and Implications for the Integrity of Scientific Publication and Medical Knowledge. JAMA. 2023;329:637-9. doi: 10.1001/jama.2023.1344.
5. Ben Salem D, Barrat JA, Ammari S, Lendoiro SL, Belaid A, Attyé A, et al. Bias, accuracy, and trust: no GenAI in peer reviewing. J Neuroradiol. 2026;53:101411. doi: 10.1016/j.neurad.2025.101411.
6. Rao VS, Kumar A, Lakkaraju H, Shah NB. Detecting LLM-generated peer reviews. PLoS One. 2025;20:e0331871. doi: 10.1371/journal.pone.0331871.
7. Naddaf M. AI tool detects LLM-generated text in research papers and peer reviews. Nature. 2025 (in press). doi: 10.1038/d41586-025-02936-6.
8. Flanagin A, Kendall-Taylor J, Bibbins-Domingo K. Guidance for Authors, Peer Reviewers, and Editors on Use of AI, Language Models, and Chatbots. JAMA. 2023;330:702-3. doi: 10.1001/jama.2023.12500.
9. Metze K, Morandin-Reis RC, de Ávila Reis MF, da Silva Fago M, Florindo JB. Misinformation, false positives and delegation of tasks - Large Language Models should not be used for the detection of retracted literature - A study of 21 Chatbots. J Clin Anesth. 2025;107:112032. doi: 10.1016/j.jclinane.2025.112032.
10. Thelwall M, Lehtisaari M, Katsirea I, Holmberg K, Zheng ET. Does ChatGPT ignore article retractions and other reliability concerns? Learn Publ. 2025 (in press). doi:10.1002/leap.2018.