



# Relato de experiência no uso de engenharia de *prompt* para apoio ao *desk review* editorial

Kamilla Pereira Silva<sup>1</sup>, Sophia Dalla Costa da Silveira<sup>2</sup>,  
Igor Moura Danieleviz e Silva<sup>3</sup>, Yasmin de Sales da Costa Medim<sup>4</sup>

<sup>1</sup>Lepidus Tecnologia, Manaus, Amazonas, Brasil

<sup>2</sup>Lepidus Tecnologia, Florianópolis, Santa Catarina, Brasil

<sup>3</sup>Lepidus Tecnologia, Campo Grande, Mato Grosso do Sul Brasil

<sup>4</sup>Lepidus Tecnologia, Manaus, Amazonas, Brasil

\*Autor correspondente: [kamilla@lepidus.com.br](mailto:kamilla@lepidus.com.br)

**Resumo:** Visa propor um conjunto de táticas de engenharia de *prompt* a serem aplicadas com inteligência artificial generativa como assistente no processo de *desk review* em periódicos científicos. Metodologicamente, classifica-se a pesquisa como exploratória, bibliográfica e descritiva. Resulta na eficiência da IA em tarefas simples ao gerar análises rápidas e claras, mas falhou em verificações quantitativas de maior precisão. Por fim conclui-se que os agentes de IA são úteis como apoio inicial, mas não substituem a revisão humana. Espera-se que este estudo fomente novos testes em práticas editoriais na comunicação científica com o uso das inteligências artificiais generativas.

**Palavras-chave:** Engenharia de *Prompt*; *Desk Review*; Inteligência Artificial; Periódicos Científicos; Comunicação Científica.

## Experience report on the use of prompt engineering to support editorial *desk review*

**Abstract:** This study aims to propose a set of prompt engineering tactics to be applied with generative artificial intelligence as an assistant in the desk review process of scientific journals. Methodologically, the research is classified as exploratory, bibliographic, and descriptive. The results show that AI was efficient in simple tasks by generating quick and clear analyses, but it failed in more precise quantitative verifications. Finally, it's concluded that AI agents are useful as initial support but do not replace human review. It is hoped that this study will encourage new tests in editorial practices within scientific communication using generative artificial intelligence.

**Keywords:** Prompt Engineering; Desk Review; Artificial Intelligence; Scientific Journals; Scientific Communication.

## Relato de experiencia en el uso de la ingeniería de *prompt* para apoyar el *desk review* editorial

**Resumen:** Este estudio tiene como objetivo proponer un conjunto de táticas de ingeniería de *prompt* para ser aplicadas con inteligencia artificial generativa como asistente en el proceso de *desk review* en revistas científicas. Metodológicamente, la investigación se clasifica como exploratoria, bibliográfica y descriptiva. Los resultados muestran que la IA fue eficiente en tareas simples al generar análisis rápidos y claros, pero falló en verificaciones cuantitativas de mayor precisión. Por último, se concluye que los agentes de IA son útiles como apoyo inicial, pero no reemplazan la revisión humana. Se espera que este estudio fomente nuevas pruebas en las prácticas editoriales en la comunicación científica con el uso de las inteligencias artificiales generativas.

**Palabras clave:** Ingeniería de *Prompts*; *Desk Review*; Inteligencia Artificial; Revistas Científicas; Comunicación Científica.



Associação Brasileira de Editores Científicos (ABEC Brasil)

<https://doi.org/10.21452/abecmeeting2025.292>

Esta obra está sob licença Creative Commons Atribuição 4.0 Internacional

## 1 Introdução

A rápida popularização de modelos de inteligência artificial (IA) generativa, como o ChatGPT, têm provocado impactos significativos nas dinâmicas de produção, revisão e disseminação científica (Jesus; Segundo *et al.* 2024). Entre os setores mais afetados, destaca-se o de periódicos científicos, cujas etapas editoriais têm sido transformadas pela incorporação dessas tecnologias. Entre os processos passíveis de moldagem com o uso dessa ferramentas, inclui-se a etapa de *desk review*, isto é, a pré-análise de artigos que, segundo a Revista Arco (2017), antecede o encaminhamento a pareceristas e envolve a verificação de aspectos formais, éticos e de escopo.

Dentre as ferramentas digitais que subsidiam a execução dessas atividades, a engenharia de *prompt* surge com potencialidade para integrá-las. Conforme observa Nascimento (2024), “a engenharia de *prompt* envolve a elaboração estratégica de instruções ou estímulos que guiam os modelos de IA a gerar respostas específicas, precisas e relevantes”. A aplicação de assistentes de inteligência artificial generativa orientados por essas técnicas, na visão deste autor, pode otimizar significativamente diversos processos antes dependentes exclusivamente da análise humana.

Este cenário ocorre devido aos Modelos de Linguagem (do inglês *Large Language Model* - LLM) pré-treinados com banco de dados que indexam informações advindas de diferentes fontes da internet (Kasneji *et al.*, 2023). Além das informações basilares previamente implementadas nos seus bancos, é possível alimentá-los com orientações e fontes de dados pelo próprio usuário.

Deste modo, esta pesquisa investiga como assistentes de IA generativa podem otimizar o *desk review* de periódicos científicos de acesso aberto, com a proposição de um *prompt* estruturado que oriente e direcione o uso ético e tecnicamente adequado desses sistemas na revisão editorial inicial. Para tanto, tem-se como objetivo geral propor um conjunto de táticas de engenharia de *prompt* a serem aplicadas com inteligência artificial generativa como assistente no processo de *desk review* em periódicos científicos. Como objetivos específicos, busca: a) delimitar os elementos e parâmetros de engenharia de *prompt* para composição de instruções lógicas aplicáveis a um periódico com critérios padronizados para a triagem inicial dos manuscritos; b) apresentar a experiência de uso da ferramenta por meio de um instrumento de coleta.

A justificativa apoia-se no crescente uso de sistemas de inteligência artificial para auxiliar atividades técnicas repetitivas que exigem tempo e alto nível de concentração. A adoção estratégica dessas tecnologias se apresenta como alternativa para ampliar a eficiência operacional e reduzir erros no processo de *desk review*. Além disso, os modelos generativos atuais oferecem recursos avançados, como o processamento direto de

arquivos e com *tokens* - palavras, caracteres, imagens ou áudio - com capacidade de lidar com volumes e dados suficientes para análises preliminares mais abrangentes.

## 2 Metodologia

A presente pesquisa, de natureza descritiva, exploratória e qualitativa, apoiada nos conceitos de Prodanov e Freitas (2013), caracteriza-se por buscar compreender a aplicação de táticas de engenharia de *prompt* em assistentes de inteligência artificial generativa voltados ao *desk review* de periódicos científicos. Para operacionalizar a etapa metodológica, a coleta de dados foi realizada a partir da seleção de um periódico científico da área de Ciências Sociais Aplicadas, escolhido pela clareza das diretrizes de submissão e pelo uso de critérios tradicionalmente empregados por outros periódicos. Ademais, a equipe de pesquisa possuía vínculo com a assistência editorial do periódico, o que possibilitou acesso às rotinas de *desk review*.

Os experimentos foram conduzidos em duas etapas distintas e complementares. Na primeira etapa, foram realizados três testes com o mesmo usuário (assistente editorial) o qual utilizou a versão gratuita do *Chat Generative Pre-trained Transformer* (GPT) versão 4o. A escolha desta versão justifica-se por sua maior capacidade de estabelecer conexões específicas entre diferentes entradas e saídas. Os três testes consistiram em análises de *desk review* conduzidas pelo mesmo usuário e seguindo um procedimento estruturado em três etapas sequenciais: i) envio do *prompt*<sup>1</sup> elaborado; ii) submissão dos documentos solicitados no próprio *prompt*; iii) comparação entre o relatório gerado pela IA e o arquivo efetivamente analisado para avaliar a precisão e adequação das respostas obtidas.

Posteriormente à realização de testes, aplicou-se com o usuário um questionário como instrumento de coleta (Apêndice 1) composto por 12 perguntas semiestruturadas, organizadas em duas dimensões analíticas: a) etapas de execução dos testes (questões relacionadas ao aspecto operacional); b) avaliação da eficiência e relevância das respostas obtidas. O questionário possibilitou captar percepções qualitativas do usuário referente à adequação do modelo de IA às exigências práticas do *desk review* para apresentar subsídios para uma avaliação dos resultados apresentados e discutidos na seção seguinte.

## 3 Resultados e discussão

Para alcance dos resultados de pesquisa, apresenta-se, a seguir, as *tags* e funções utilizadas que estruturam o *prompt* aplicado para a pré-avaliação de artigos científicos:

---

<sup>1</sup> *Prompt* disponível em: <https://doi.org/10.5281/zenodo.16994281>

### Quadro 1 - Tags utilizadas para a estruturação do *prompt* do *desk review*

TAGS E SUAS FUNÇÃO
<b>Persona:</b> definir o papel ou a identidade que o modelo de IA deve assumir (White <i>et al.</i> , 2023). Essa tag adota um tom, estilo e perspectiva específicos, o que aumenta a qualidade e a relevância da resposta para um determinado contexto.
<b>Prompt negativo:</b> especificar comportamentos ou palavras que a IA deve evitar. Incluir restrições explícitas, ou <i>negative prompts</i> , é uma maneira de restringir o comportamento do modelo e reduzir a probabilidade de erros de informações (Reynolds; Mcdonell, 2021).
<b>Problema:</b> contextualizar a tarefa como um desafio que o <i>prompt</i> visa resolver. Fornecer contexto é um princípio fundamental da engenharia de <i>prompts</i> , pois modelos de linguagem geram melhores resultados quando entendem a meta da tarefa (Brown <i>et al.</i> , 2020).
<b>Crterios:</b> detalhar as regras e condições que a resposta do modelo deve seguir. Ela transforma uma tarefa subjetiva em uma avaliação objetiva. Sua função é fazer com que os pontos importantes sejam verificados sistematicamente.
<b>Tom e estilo de linguagem:</b> comunicar as características linguísticas da resposta. Ela instrui o modelo sobre como a mensagem deve ser escrita, incluindo formalidade, objetividade e vocabulário. Permite adaptar a comunicação para públicos e contextos específicos (Brown <i>et al.</i> , 2020).
<b>Documentos:</b> referenciar a fontes de conhecimento externas. Conhecidas como <i>Retrieval-Augmented Generation</i> (RAG), indicam ao modelo que ele deve consultar arquivos específicos para executar a tarefa (Lewis <i>et al.</i> , 2020).
<b>Passos de execução:</b> dividir uma tarefa complexa em micro tarefas gerenciáveis. Essa abordagem, conhecida como <i>Chain-of-Thought</i> (CoT) ou decomposição de tarefas, melhora a capacidade do modelo de raciocinar e executar processos com múltiplos estágios de forma lógica e ordenada (Wei <i>et al.</i> , 2022).
<b>Análise:</b> formatar o estilo de saída desejada. Estruturar a saída esperada é um padrão de <i>prompt</i> que instrui o modelo a seguir um estilo específico de saída (White <i>et al.</i> , 2023).

Fonte: os autores, 2025.

Ao aplicar o *prompt* utilizando o Chat GPT e avaliar os resultados dos três testes realizados pela equipe editorial, conforme os resultados do questionário<sup>2</sup> aplicado, às análises levantadas foram examinadas de acordo com critérios de eficácia e eficiência.

Na eficácia, o usuário observou que o GPT conseguiu cumprir parte das análises esperadas, mas apresentou falhas relevantes em critérios que exigiam precisão. Houve divergências notáveis na contagem de caracteres: em uma análise, o agente apontou cerca de 44 mil caracteres, enquanto a contagem manual ficou entre 27 mil e 38 mil, dependendo do artigo. O mesmo ocorreu com a contagem de palavras dos resumos: em um caso, a IA registrou 295 palavras no resumo em português, mas a contagem correta era 312; em outro, a IA considerou o resumo longo demais quando, na verdade, estava dentro do limite.

<sup>2</sup> Disponível em: <https://doi.org/10.5281/zenodo.16994553>

Quanto ao número de figuras, o agente identificou 2, 5 ou 7 imagens em relatórios diferentes, mas as verificações humanas mostraram entre 7 e 15 figuras. Isso indica que, embora o GPT com o *prompt* aplicado seja capaz de identificar a existência de elementos como resumos, títulos e palavras-chave, ele não apresenta exatidão em tarefas quantitativas.

Em relação à eficiência, o tempo de resposta foi avaliado como adequado, pois os relatórios completos foram gerados em menos de dois minutos. Segundo o usuário, a organização em tópicos facilitou a leitura e a conferência manual. No entanto, o agente não buscou informações adicionais que poderiam ter melhorado a precisão das análises, como solicitar detalhes sobre a formatação dos arquivos ou confirmar contagens de forma cruzada. Assim, a experiência foi considerada eficiente no aspecto da rapidez e da clareza, mas incompleta em termos de validação.

Na análise da qualidade geral, o usuário apontou que confiaria no agente apenas com ressalvas. Para tarefas simples, como verificar a presença de resumos nos dois idiomas solicitados, identificar palavras-chave ou confirmar a existência de títulos, o desempenho foi satisfatório. Entretanto, para verificações que dependem de precisão absoluta, como contagem de palavras, caracteres e figuras, a experiência foi negativa. O mesmo vale para o relatório de similaridade: em dois testes a resposta foi coerente, mas em outro houve divergência acentuada em relação ao documento PDF.

O usuário sugeriu que o *prompt* seja ajustado para indicar com mais clareza onde e o que deve ser analisado no arquivo. Como conclusão, considerou que o agente pode ser usado como apoio inicial, mas não é seguro para substituir o trabalho humano em decisões editoriais.

## 4 Considerações finais

Para alcançar o objetivo de propor um conjunto de táticas de engenharia de prompt a serem aplicadas com IA generativa como assistente no processo de desk review em periódicos científicos, realizou-se o uso destas táticas em um agente baseado no ChatGPT. Desta maneira, buscou-se compreender em que medida essa tecnologia poderia contribuir para a revisão editorial inicial de manuscritos.

Os resultados mostraram que o agente de IA apresentou um desempenho satisfatório em tarefas simples de conferência. Destacou-se pela rapidez e pela clara estruturação por meio de relatórios completos com organização em tópicos. Contudo, as falhas em critérios quantitativos explanaram limites na confiabilidade dos resultados.

Do ponto de vista metodológico, o uso de *tags* estruturadas no *prompt* revelou-se como necessária para orientar a IA e obter respostas mais próximas das necessidades editoriais, o que potencializa a engenharia de *prompt* como uma prática emergente nas

atividades para estabelecer parâmetros replicáveis em diferentes processos editoriais.

Além disso, o estudo contribui para o campo editorial ao oferecer um relato de aplicação do uso de IA generativa em fluxos de *desk review*. Destaca-se que a tecnologia pode desempenhar um papel auxiliar, no entanto, sua adoção requer cautela, ajustes contínuos de *prompt* e revisão humana complementar.

Em suma, indica-se que estudos futuros realizam testagem considerando diferentes diretrizes de submissão, na tentativa de compará-los e estabelecer quais os parâmetros alteram a qualidade das triagens dos trabalhos em termos de quantidade e precisão. Sugere-se também a testagem com outros modelos de IA, parametrizações e ajustes no *prompt*, utilizando diferentes conjuntos de dados para análise de resultados.

## Contribuição dos autores

**Kamilla Pereira Silva:** Conceitualização, Curadoria de dados, Escrita – redação original, Escrita – revisão e edição, Investigação, Metodologia, Software, Validação.

**Sophia Dalla Costa da Silveira:** Conceitualização, Escrita – redação original, Escrita – revisão e edição, Investigação, Metodologia, Validação.

**Igor Moura Danieleviz e Silva:** Conceitualização, Escrita – redação original, Escrita – revisão e edição, Investigação, Metodologia, Validação.

**Yasmin de Sales da Costa Medim:** Curadoria de dados, Investigação, Software, Validação.

## Referências

BROWN, T. B. *et al.* Language models are few-shot learners. **Advances in Neural Information Processing Systems**, Cambridge, v. 33, p. 1877-1901, 2020. Disponível em: [https://proceedings.neurips.cc/paper\\_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf). Acesso em: 25 ago. 2025.

FIORELLO, L.; MEHTA, V. Accelerating editorial processes in scientific journals: Leveraging AI for rapid manuscript review. **Oral Oncology Reports**, Londres, v. 10, n. 3, p. 100511, 2024. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2772906024003571>. Acesso em: 7 ago. 2025.

JESUS, A. F.; SEGUNDO, J. E. S. *et al.* Aplicações de Inteligência Artificial Generativa em Revisões Sistemáticas da Literatura. **Tendências da Pesquisa Brasileira em Ciência da Informação**, João Pessoa, v. 17, edição especial, p. 1-27, 2024. Disponível em: <https://revistas.ancib.org/tpbci/article/view/675>. Acesso em: 7 ago. 2025.

KASNECI, E. *et al.* ChatGPT for good? On opportunities and challenges of large language models for education. **Learning and individual differences**, Londres, v.

103, p. 102274, 2023. Disponível em: [https://www.sciencedirect.com/science/article/pii/S1041608023000195?casa\\_token=NVFhd22TIEMAAAAA:rEPcFv-usSHI95pCDjCnvx21NKHfMqP1lzVlmyvM6vumkql2Uw4YyRMyk3Wk1ulp\\_SBSxRmMjxBf](https://www.sciencedirect.com/science/article/pii/S1041608023000195?casa_token=NVFhd22TIEMAAAAA:rEPcFv-usSHI95pCDjCnvx21NKHfMqP1lzVlmyvM6vumkql2Uw4YyRMyk3Wk1ulp_SBSxRmMjxBf). Acesso em: 25 ago. 2025.

LEWIS, P. *et al.* Retrieval-augmented generation for knowledge-intensive NLP tasks. **Advances in Neural Information Processing Systems**, Cambridge, v. 33, p. 9459-9474, 2020. Disponível em: <https://proceedings.neurips.cc/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf>. Acesso em: 27 ago. 2025.

NASCIMENTO, J. R. **Exploração de técnicas de engenharia de prompt para aprimorar os resultados do uso de LLM no TCMRio**. 2024. 64 f. Trabalho de Conclusão de Curso (Residência em Tecnologia da Informação) - Instituto MetrÓpole Digital da Universidade Federal do Rio Grande do Norte, Natal, 2024. Disponível em: <https://repositorio.ufrn.br/bitstreams/83ed7a5d-4651-4e51-8862-6c617dbbb2bb/download>. Acesso em: 27 ago. 2025.

PRODANOV, C. C.; FREITAS, E. C. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. 2 ed. Rio Grande do Sul: Editora Feevale, 2013. Disponível em: <https://www.feevale.br/Comum/midias/0163c988-1f5d-496f-b118-a6e009a7a2f9/E-book%20Metodologia%20do%20Trabalho%20Cientifico.pdf>. Acesso em: 7 ago. 2025.

REVISTA ARCO. **Desk Review**: utilizado por revistas científicas para agilizar o processo, método de avaliação de artigos retoma discussão sobre o produtivismo acadêmico. Revista Arco – Jornalismo Científico e Cultural, Santa Maria, 2021. Disponível em: <https://www.ufsm.br/midias/arco/post416>. Acesso em: 7 ago. 2025.

REYNOLDS, L.; MCDONELL, K. Prompt programming for large language models: beyond the few-shot paradigm. *In*: CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2021. **Anais** [...], [S.l.: [s.n.], 2021. p. 1-7. Disponível em: [https://dl.acm.org/doi/pdf/10.1145/3411763.3451760?casa\\_token=oKEQNpMSIPYAAAAA:Zc\\_sezrkAbc8zslk-kTQtzc7lubduC3-boYIMX7VgRVpBPWlufTL3F9\\_auRDk0M9waA87hsYFC8Z4f\\_Q](https://dl.acm.org/doi/pdf/10.1145/3411763.3451760?casa_token=oKEQNpMSIPYAAAAA:Zc_sezrkAbc8zslk-kTQtzc7lubduC3-boYIMX7VgRVpBPWlufTL3F9_auRDk0M9waA87hsYFC8Z4f_Q). Acesso em: 27 ago. 2025.

WEI, J. *et al.* Chain-of-thought prompting elicits reasoning in large language models. **Advances in Neural Information Processing Systems**, Cambridge, v. 35, p. 24824-24837, 2022. Disponível em: [https://proceedings.neurips.cc/paper\\_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf](https://proceedings.neurips.cc/paper_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf). Acesso em: 27 ago. 2025.

WHITE, J. *et al.* **A prompt pattern catalog to enhance prompt engineering with chatgpt**. Arxiv, 2023. Disponível em: <https://arxiv.org/abs/2302.11382>. Acesso em: 27 ago. 2025.

## APÊNDICE 1 - Protocolo de Avaliação da Eficácia e Eficiência de um Agente baseado em Prompt para Avaliação de Artigos Científicos

### Bloco A – Eficácia da Resposta

O GPT cumpriu todas as análises esperadas do prompt?

- Sim
- Não

Se não, quais critérios ficaram de fora ou foram avaliados incorretamente?

O relatório final gerado classificou corretamente a submissão como:

- Apta para avaliação
- Necessita alterações
- Reprovada

A classificação foi coerente com a sua avaliação humana?

- Sim
- Não
- Parcialmente

Comente possíveis divergências:

O GPT seguiu corretamente os passos (1, 2, 3) do protocolo estabelecido no prompt?

- Sim
- Não

Comente se houve algum salto ou execução fora da ordem:

O GPT apresentou interpretações equivocadas dos critérios?

- Sim
- Não

Se sim, exemplifique:

### Bloco B – Eficiência da Resposta

Quanto tempo (aproximado) levou para o GPT gerar o relatório completo?

- Menos de 1min
- Menos de 2 min
- Entre 2 e 5 min
- Mais de 5 min

A estrutura da resposta foi:

- Clara e objetiva
- Confusa ou redundante

Comente brevemente:

O agente solicitou informações adicionais de forma adequada?

Sim

Não

Se não, o que deveria ter perguntado?

### **Bloco C – Qualidade geral da experiência**

Você confiaria nesse agente para uma triagem inicial de artigos em sua revista?

Sim

Com ressalvas

Não

Justifique:

Que melhorias você sugeriria no prompt ou no comportamento do Agente?