SciELO Preprints

Estado da publicação: O preprint não foi publicado em outro meio.

Efeito do uso de recursos linguístico-discursivos de interação em prompts para síntese de textos acadêmicos em português

Ronei Ximenes Martins, Sayonara Ribeiro Marcelino Cruz, Helena Maria Ferreira, Braian Garrito Veloso, Dener Luiz da Silva

https://doi.org/10.1590/SciELOPreprints.19024

Submetido em: 2026-11-23
Postado em: 2026-11-27 (versão 1)
(AAAA-MM-DD)

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

Efeito do uso de recursos linguístico-discursivos de interação em prompts para síntese de textos acadêmicos em português

Ronei Ximenes Martins
Universidade Federal de Lavras,
Lavras, Minas Gerais, Brasil.
ORCID: https://orcid.org/0000-0002-3586-9086

Sayonara Ribeiro Marcelino Cruz
Universidade Federal de Lavras,
Lavras, Minas Gerais, Brasil.
ORCID: https://orcid.org/0000-0002-3176-1270

Helena Maria Ferreira
Universidade Federal de Lavras,
Lavras, Minas Gerais, Brasil.
ORCID: https://orcid.org/0000-0002-8749-5426

Braian Garrito Veloso
Universidade Federal de Lavras,
Lavras, Minas Gerais, Brasil.
ORCID: https://orcid.org/0000-0002-9459-5740

Dener Luiz da Silva
Universidade Federal de São João del-Rei,
São João del-Rei, Minas Gerais, Brasil.
ORCID: https://orcid.org/0000-0002-9422-6531

RESUMO: Este estudo investigou como a incorporação de recursos linguístico-discursivos de interação em prompts pode afetar a fidedignidade factual, a organização composicional e a responsabilidade epistêmica na geração de sínteses de artigos científicos da área de Educação, publicados em língua portuguesa. A pesquisa justifica-se pelo risco de distorções retóricas que comprometem a integridade científica e induzem à compreensão equivocada de textos acadêmicos. Com abordagem qualitativa, a metodologia consistiu na geração de 45 resumos sintéticos gerados por três modelos de linguagem (ChatGPT 5.0, Gemini 3.0 Fash e DeepSeek V3), sob três condições de prompt (Base, Roleplay e Emotion) e posterior avaliação por 15 especialistas que apontaram e justificaram todas as distorções encontradas. O corpus foi composto pelo conjunto dos textos de justificativa, submetido à análise textual discursiva. Os resultados revelaram que o estímulo interpessoal e emocional agravou vieses epistêmicos, provocando mutações deônticas, generalizações e alucinações por adulação, além de não aprimorar a precisão. Conclui-se que a fluidez verbal aprimorada pelas condições Roleplay e Emotion podem ampliar a ilusão de competência e camuflar imprecisões conceituais, exigindo uma mediação humana ainda mais crítica.

Palavras-chave: Inteligência artificial generativa, Comunicação científica, Engenharia de prompt, Fidedignidade epistêmica, Pesquisa em educação.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

The effect of using linguistic-discursive interactional resources in prompts for summarizing academic texts in Portuguese

ABSTRACT: This study investigated how the incorporation of linguistic-discursive interactional resources into prompts may affect factual fidelity, compositional organization, and epistemic responsibility in the generation of syntheses of scientific articles in the field of Education published in Portuguese. The research is justified by the risk of rhetorical distortions that compromise scientific integrity and lead to misinterpretation of academic texts. Adopting a qualitative approach, the methodology consisted of generating 45 synthesis summaries using three large language models (ChatGPT 5.0, Gemini 3.0 Flash, and DeepSeek V3) under three prompt conditions (Base, Roleplay, and Emotion), followed by an evaluation by 15 experts, who identified and justified all the distortions they found. The corpus comprised the full set of their justification texts, which was submitted to Discursive Textual Analysis. The results revealed that interpersonal and emotional prompting aggravated epistemic biases, producing deontic shifts, overgeneralizations, and sycophancy-induced hallucinations, without improving accuracy. We conclude that the enhanced verbal fluency produced by the Roleplay and Emotion conditions may amplify the illusion of competence and mask conceptual inaccuracies, thereby demanding even more critical human mediation.

Keywords: Generative artificial intelligence, Scientific communication, Prompt engineering, Epistemic fidelity, Educational research.

Efecto del uso de recursos lingüístico-discursivos de interacción en prompts para la síntesis de textos académicos en portugués

RESUMEN: Este estudio investigó cómo la incorporación de recursos lingüístico-discursivos de interacción en los prompts puede afectar la fidelidad factual, la organización composicional y la responsabilidad epistémica en la generación de síntesis de artículos científicos del área de Educación, publicados en lengua portuguesa. La investigación se justifica por el riesgo de distorsiones retóricas que comprometen la integridad científica e inducen a una comprensión equivocada de los textos académicos. Con enfoque cualitativo, la metodología consistió en la generación de 45 resúmenes de síntesis por parte de tres modelos de lenguaje (ChatGPT 5.0, Gemini 3.0 Flash y DeepSeek V3), bajo tres condiciones de prompt (Base, Roleplay y Emotion), y su posterior evaluación por 15 especialistas, quienes señalaron y justificaron todas las distorsiones encontradas. El corpus estuvo compuesto por el conjunto de los textos de justificación, sometido al Análisis Textual Discursivo. Los resultados revelaron que el estímulo interpersonal y emocional agravó los sesgos epistémicos, provocando mutaciones deónticas, generalizaciones y alucinaciones por adulación, además de no mejorar la precisión. Se concluye que la fluidez verbal potenciada por las condiciones Roleplay y Emotion puede ampliar la ilusión de competencia y camuflar imprecisiones conceptuales, lo que exige una mediación humana aún más crítica.

Palabras clave: Inteligencia artificial generative, Comunicación científica, Ingeniería de *prompts*, Fidelidad epistémica, Investigación en educación.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

1. INTRODUÇÃO

A inserção de sistemas de Inteligência Artificial Generativa (IAG) baseados em *Large Language Models* (LLMs) no cotidiano acadêmico vem afetando a forma como o conhecimento científico é sintetizado e disseminado. Ferramentas como *ChatGPT*, *Gemini* e *DeepSeek* passaram a integrar práticas de estudo, de produção escrita e de pesquisa de estudantes, docentes e pesquisadores em diferentes áreas, oferecendo suporte à leitura de textos densos, à elaboração de resumos e à recontextualização de resultados científicos para públicos variados. Esse cenário é ilustrado por dados recentes sobre o contexto educacional brasileiro, segundo os quais cerca de 70% dos estudantes do Ensino Médio já utilizam essas ferramentas em suas pesquisas escolares, embora apenas 32% declarem ter recebido orientação formal para esse uso, e 43% dos professores afirmam recorrer a LLMs na preparação de conteúdos didáticos (CGI.BR, 2025).

Porém, esse novo sistema digital não é neutro. Peters e Chin-Yee (2025) mostraram, em estudo com 4.900 resumos gerados por dez modelos de linguagem distintos, que LLMs populares apresentam viés consistente de supergeneralização ao resumirem textos científicos no idioma inglês. Os autores identificaram três formas principais pelas quais resultados específicos de estudos empíricos são transformados em afirmações mais amplas do que aquelas sustentadas pelos dados originais: generalizações genéricas, generalizações enunciadas no presente (como se fossem verdades universais atualizadas) e generalizações orientadas à ação, convertendo achados descritivos em recomendações práticas. O fato de até 73% dos resumos analisados apresentarem algum tipo de supergeneralização sinaliza que a mediação algorítmica da comunicação científica pode introduzir distorções relevantes no modo como leitores não especialistas compreendem a força das evidências disponíveis (Peters; Chin-Yee, 2025).

As implicações desse evento são particularmente sensíveis em contextos de formação acadêmica e em campos diretamente relacionados a políticas públicas, como educação e saúde, já que textos gerados por LLMs, ao assumirem um tom mais assertivo do que o original e ampliarem o alcance de conclusões, podem levar estudantes e profissionais em formação a tomarem decisões pedagógicas, curriculares ou de gestão apoiadas em inferências que não correspondem às evidências reportadas nos artigos originais. A supergeneralização, nesse sentido, não é apenas um problema de “estilo de escrita” da IAG, mas um fenômeno que atravessa o núcleo da responsabilidade epistêmica, reconfigurando a relação entre incerteza, evidência e ação (Topaz *et al.*, 2026; Wang *et al.*, 2023) .

No contexto brasileiro da pesquisa em educação, a problemática adquire contornos próprios porque o campo se organiza sob condições institucionais e epistemológicas peculiares, com capitais científicos, públicos de pesquisa e critérios de prestígio específicos (Caregnato; Miorando; Leite, 2022). A produção científica, neste contexto, é majoritariamente publicada em português, com cerca de 61% dos artigos brasileiros, contra a tendência mundial de publicação em inglês (Jacinto, 2021). A preferência pelo português é modulada pelo público da pesquisa, tipo de linguagem e necessidade de financiamento (Baumvol; Sarmento; Fontes, 2021). Ao mesmo tempo, cresce a expectativa de que estudantes de graduação e pós-graduação utilizem LLMs para apoiar a leitura de artigos complexos, sintetizar resultados e preparar textos acadêmicos em prazos reduzidos e não há estudos sistemáticos que examinem, em língua portuguesa, um aspecto ainda pouco explorado nessa discussão, que diz respeito à forma como usuários interagem com LLMs por meio de *prompts* para realizarem a condensação e obterem sínteses de relatos de pesquisa em português.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

Embora a maior parte dos estudos disponíveis se concentre em parâmetros técnicos de funcionamento das IAG mediante alterações em condições de *prompt*, há evidências iniciais de que a aplicação de recursos linguístico-discursivos da interação humana, como elogios à competência da IA, pedidos explícitos de ajuda, marcas de gentileza ou frases que enfatizem a importância da tarefa para o usuário, podem afetar, também qualitativamente, os *outputs* resultantes (White *et al.*, 2023; Li *et al.*, 2023).

No cotidiano acadêmico brasileiro, conjectura-se que é comum que estudantes e pesquisadores formulem comandos que misturam instruções normativas (“inclua objetivo, método, resultados e conclusões”), estratégias de polidez (“por favor”, “preciso de sua ajuda”) e essa dimensão relacional, pouco tematizada em referenciais técnicos, merece atenção como parte constitutiva dos letramentos digitais e científicos associados ao uso de LLMs.

É nesse cenário que se insere o presente estudo, que investiga se recursos linguístico-discursivos próprios das interações humanas influenciam a qualidade de *outputs* de LLM em tarefas acadêmicas especializadas, no campo da pesquisa educacional em língua portuguesa. O objetivo é avaliar se recursos típicos da interação entre humanos afeta dimensões como clareza, estrutura, fidedignidade factual e responsabilidade ética na geração de *outputs* de *prompts* destinados à síntese de textos científicos no contexto de pesquisas acadêmicas em educação.

2. REFERENCIAL TEÓRICO

Estudos, como os de Ding, Lawson e Shapira (2025), apontam para a ampliação exponencial de publicações acadêmicas (em âmbito global) vinculadas ao uso de IAG. A capacidade de tratar, resumir e transformar rapidamente informações científicas complexas em termos acessíveis tem atraído cada vez mais pesquisadores para o uso de *chatbots* baseados em grandes modelos de linguagem. No entanto, esse elevado potencial de aumentar a produção e a difusão do conhecimento vem acompanhado do risco de que essas ferramentas omitam detalhes cruciais ao sintetizar textos acadêmicos, gerando generalizações mais amplas do que as justificadas pelo estudo original (Peters; Chin-Yee, 2025). Também há que se considerar, por óbvio, o risco de viés algorítmico (Heggler; Szmoski; Miquelin, 2025).

Para compreender como esses riscos se manifestam ou podem ser mitigados, faz-se necessário analisar a forma como os usuários interagem com tais modelos. Nesse sentido, White *et al.* (2023) tratam os LLMs explicitamente como sistemas conversacionais, cuja performance depende, de modo decisivo, da qualidade dos *prompts* que estruturam a interação e orientam forma, conteúdo e escopo dos *outputs*. Um *prompt* é definido como um conjunto de instruções escritas ou faladas em linguagem natural que comanda o LLM, refinando as suas capacidades por meio do estabelecimento de regras, diretrizes e contextos da conversa (White *et al.*, 2023).

No âmbito específico da comunicação científica, contudo, mesmo o refinamento dos comandos nem sempre previne distorções. Peters e Chin-Yee (2025) demonstram que diferentes modelos de IAG tendem a produzir o mesmo viés consistente de supergeneralização, não obstante serem comandados para não generalizarem.

Essa tendência à perda de precisão foi contraposta pelo estudo de Li *et al.* (2023) que, ao investigarem o impacto de *prompts* emocionais (*EmotionPrompt*) em *outputs*, identificaram que é a busca dos modelos por respostas assertivas que induz a eliminação de termos de moderação e cautela (como "geralmente" e "pode ser") em favor de sentenças categóricas e deterministas (como "completamente" ou "não irá"). Nessa perspectiva, pode se estabelecer

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

a hipótese de que o LLM não é, especificamente, responsável pela injeção de vieses nos *outputs*, mas a forma como foram comandados. Em suma, destacam-se duas perspectivas opostas: a premissa de que os modelos de IAG geram distorções de modo endógeno e autônomo, prescindindo do comando fornecido, e a possibilidade de que a estrutura das instruções atua como indutora determinante para o surgimento dessas inconformidades. Sob qualquer um dos ângulos, contudo, constata-se emergência de vieses nos *outputs* gerados.

Peters e Chin-Yee (2025) identificam três tipos principais de generalizações problemáticas em resumos gerados por IAG: o uso de afirmações genéricas sem quantificação explícita, a conversão de resultados do passado para o presente em tom atemporal e, por fim, a transformação de dados descritivos em recomendações normativas. Eles revelaram que resumos produzidos sintéticos têm cinco vezes mais chances de conter generalizações indevidas do que sínteses feitas por especialistas humanos.

Esse comportamento evidencia que o aumento da competência discursiva e do refinamento técnico dos modelos mais utilizados não se traduz em confiabilidade epistêmica, fato que agrava o risco de usuários desenvolverem confiança (*over-trust*) em respostas sintéticas que, sob aparência convincente, omitem as ressalvas e nuances metodológicas originais.

Se o estudo de Li *et al.* (2023) sobre os *EmotionPrompts* apontam possível solução para o problema da supergeneralização, os padrões de estruturação de *prompts* descritos por White *et al.* (2023) podem oferecer uma base conceitual eficiente para a mitigação desse tipo de viés. Em tese, a categoria de Personalização da Saída (*Output Customization*), que inclui automação de resposta, persona, gabarito de saída e os padrões voltados à Identificação de Erros (*Error Identification*), poderia permitir a estruturação de instruções eficazes para se extrair informação fidedigna de textos acadêmicos.

Assim, , enquanto Peters e Chin-Yee (2025) mostram empiricamente que LLMs tendem a generalizar em excesso ao resumir estudos científicos, os trabalhos de Li *et al.* (2023) e White *et al.* (2023) oferecem possibilidades de padrões de *prompt* que podem ser mobilizados para mitigar esse viés, ao estruturar a interação de modo a privilegiar fidelidade ao escopo original dos estudos, explicitação de incerteza e controle do contexto comunicativo.

Ao lado disso, a reflexão sobre os vieses gerados por LLMs exige ir além da correção factual estrita e considerar como os *outputs* moldam julgamentos, crenças e decisões em contextos sociais sensíveis (Heggler; Szmoski; Miquelin, 2025). Li, Yi e Chen (2025) argumentam que o conceito de “alucinação” de IAG tem sido tratado de forma restrita, como mera falha técnica na produção de saídas factualmente corretas. Isso subestima a relevância de conteúdos fluentes e aparentemente adequados que veiculam distorções sutis. Para além dos erros, de fato, estritos, a análise da mediação algorítmica exige mapear distorções que afetam a qualidade do debate e da produção acadêmica. O Quadro 1 sintetiza as principais categorias desta taxonomia, evidenciando como cada tipo de alucinação atua como vetor de viés epistêmico.

Com base nesta taxonomia, Li, Yi e Chen (2025) apontam o “paradoxo da precisão”, qual seja, o foco exclusivo em métricas de exatidão factual para bloquear alucinações pode intensificar a manipulação discursiva e a perda de habilidades críticas (*deskilling*), afetando negativamente a confiança dos usuários. A precisão factual é necessária, mas insuficiente para a garantia de responsabilidade comunicacional, pois respostas (*outputs)* estritamente verdadeiras ainda provocam danos epistêmicos ao mascarar incertezas.

Quadro 1 - Categorias de Alucinações Sociotécnicas e Viés Epistêmico em LLM

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

Categoria	Descrição conceitual	Possível Impacto na Comunicação Científica
Sycophancy Hallucination (Alucinação por Adulação)	Respostas bajuladoras que se alinham às preferências ou expectativas presumidas do usuário, reforçando crenças prévias ou posições dominantes, mesmo em detrimento da evidência	Criar bolhas de confirmação e validar hipóteses cientificamente frágeis ou enviesadas apenas para agradar o pesquisador.
Consensus Illusion (Ilusão de Consenso)	*Outputs* que sugerem consenso científico ou normativo, mesmo havendo dissenso, pluralidade de resultados ou incerteza, promovendo uma convergência epistêmica indevida).	Apaga debates legítimos, correntes minoritárias e a complexidade inerente ao avanço da ciência.
Oversimplification (Simplificação Excessiva)	Simplificação excessiva de problemas complexos, com supressão de nuances metodológicas, controvérsias e limites dos dados, frequentemente em forma de afirmações categóricas	Transforma achados condicionais em "verdades absolutas", omitindo ressalvas e escopos de validade da pesquisa.
Prompt-Sensitivity (Sensibilidade ao Prompt)	Variações significativas na forma e no conteúdo das respostas em função de alterações sutis no estilo ou no enquadramento do *prompt*	Compromete a reprodutibilidade e a consistência da análise factual, tornando o resultado vulnerável à indução da pergunta.

Fonte: Elaborado pelos autores com base em Li, Yi e Chen (2025).

Essa análise se alinha ao diagnóstico de Peters e Chin-Yee (2025) sobre o viés de generalização, no qual sínteses algorítmicas preservam conceitos gerais, mas alteram a estrutura enunciativa dos resultados. A convergência entre a simplificação excessiva e a ilusão de consenso (Li; Yi; Chen, 2025) com as formas de generalização (Peters; Chin-Yee, 2025) sugerem que o viés na comunicação científica decorre de erros factuais isolados e também de distorções linguísticas e de retórica.

Nesse contexto, a responsabilidade com a comunicação científica demanda o deslocamento de foco da exatidão factual para o da integridade epistêmica inerente à manutenção da autonomia crítica dos leitores. Li, Yi e Chen (2025) mostram que regulações normativas focadas apenas em métricas de exatidão, como o *AI Act* (Lei de Inteligência Artificial da União Europeia), o *GDPR* (Regulamento Geral sobre a Proteção de Dados) e o *DSA* (Lei de Serviços Digitais), falham diante de riscos associados à fluência e à capacidade de persuasão engendrados nos *outputs*.

O que essa normatização e as pesquisas sobre generalização não responderam ainda é se a adoção de determinados padrões de *prompt* podem melhorar a integridade epistêmica dos *outputs* para comunicação científica. Uma alternativa prática para verificar tal hipótese é a adoção de padrões de *prompts*, tais como os propostos por White *et al.* (2023), como ferramentas de controle e melhoria da precisão. Dos padrões de White *et al.*, os testados

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

nesta pesquisa foram: (a) Gerenciador de Contexto (*Context Manager*) que delimita os aspectos do contexto a considerar ou ignorar, restringindo respostas a fatos e evitando extrapolações e (b) Receita (*Recipe*) que estrutura a resposta em passos sequenciais, auxiliando na verificação de etapas metodológicas sem omissões críticas. Esses padrões funcionam como mecanismos de *human-in-the-loop* (humano no circuito), alinhando-se à estratégia de hibridismo homem-máquina na tomada de decisão. Em White *et al.* (2023), o conceito de *human-in-the-loop* aparece de forma implícita na própria lógica dos padrões de prompt: o modelo nunca é pensado como agente autônomo, mas como componente de um sistema em que o humano continua responsável por orientar, controlar e verificar a saída.

Em investigação empírica realizada com estudantes universitários, Bani Hani, Hujran e Anwar (2026), a partir de reflexões escritas e entrevistas, identificaram que a confiança no uso autônomo da IAG é baixa, ao passo que a sua integração com o julgamento crítico humano apresenta alta aceitação. Na prática, os usuários utilizam as LLM prioritariamente para tarefas de apoio técnico e verificação fática, mantendo sob responsabilidade humana a escolha de critérios, a avaliação de consequências e a decisão final (Bani Hani; Hujran; Anwar, 2026).

Resumidamente, é possível considerar que a convergência entre esses referenciais indica que o viés na comunicação científica ultrapassa falhas factuais, manifestando-se em também distorções retóricas e de amplitude (Li; Yi; Chen, 2025; Peters; Chin-Yee, 2025). Os padrões de *prompt* (White *et al.,* 2023) e as estratégias de hibridismo (Bani Hani; Hujran; Anwar, 2026) reforçam que a mitigação desses vieses exige superar a busca exclusiva por exatidão factual. Assim, para uma agenda de comunicação científica responsável é necessário adotar procedimentos que busquem a integridade epistêmica, a explicitação de limites e a preservação de condições do exercício da autonomia crítica do leitor.

Ao lado da aplicação de padrões de *prompts* está o estilo de comunicação interpessoal e emocional na sua implementação. O estudo de Li *et al.* (2023) demonstrou que as LLM não apenas respondem a comandos lógicos, mas têm seu desempenho alterado por enquadramentos linguísticos carregados de tom afetuoso ou de urgência (*EmotionPrompt*). Tal estudo sintetizou e testou 11 frases de estímulo emocional que apelam para o reforço de autoeficácia, de responsabilidade e de impacto pessoal dos resultados na realização de tarefas. Os autores demonstraram que a adição desses marcadores de comunicação interpessoal melhoraram a qualidade geral dos *outputs*. Seus experimentos automatizados registraram ganhos de 8% em *Instruction Induction*[1] e 115% no *benchmark BIG-Bench*[2]. Etapas de avaliação realizada por humanos confirmaram que a comunicação emocional

[1] O *Instruction Induction* é um *benchmark* determinístico de Processamento de Linguagem Natural desenvolvido por Honovich *et al.* (2022), composto por 24 tarefas projetadas para avaliar a capacidade dos Grandes Modelos de Linguagem (LLMs) de inferir a instrução ou regra subjacente a uma tarefa a partir de poucas demonstrações de entrada e saída (*few-shot in-context learning*). No estudo de Li *et al*. (2023), esse conjunto de tarefas foi utilizado para testar a inteligência emocional dos modelos em avaliações automáticas baseadas em acurácia, demonstrando que a adição de estímulos emocionais nos prompts (*EmotionPrompt*) gerou um ganho relativo médio de 8,00% no desempenho dos LLMs analisados (incluindo *ChatGPT*, GPT-4, *Llama* 2, *Vicuna*, *BLOOM e Flan-T5-Large*)

[2] O *BIG-Bench* (*Beyond the Imitation Game Benchmark*) é um conjunto abrangente de testes desenvolvido por Srivastava *et al.* (2023) para avaliar capacidades cognitivas avançadas em Grandes Modelos de Linguagem (LLMs), como raciocínio lógico, matemático e multietapa, utilizando uma escala normalizada em que 100 pontos correspondem ao desempenho de humanos especialistas e 0 ao chute aleatório. No estudo de Li *et al.* (2023), o *benchmark* foi empregado para testar o impacto do *EmotionPrompt* em tarefas de alta complexidade, nas quais a adição de estímulos emocionais gerou um ganho relativo médio de 115% no desempenho dos modelos em relação aos *prompts* convencionais.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

induziu as LLM a gerarem respostas mais coerentes, factualmente precisas e alinhadas aos princípios de responsabilidade social e humana , sobretudo em áreas de comunicação sensíveis.

Segundo Li *et al.* (2023), a eficácia do *EmotionPrompt* se fundamenta na psicologia cognitiva e social por meio de constructos, tais como a autorregulação, autoeficácia e regulação emocional. Ao analisar os vetores de atenção interna dos modelos, os autores identificaram que a inserção de termos associados à excelência, confiança e responsabilidade altera as respostas do LLM, induzindo o sistema a um processamento mais adequado ao que se pretende.

A tese do ajuste no estilo de comunicação como fator de melhoria da qualidade das respostas dos LLM foi testada também por Li *et al.* (2023). No *benchmark TruthfulQA*, os autores identificaram que essa abordagem de comunicação com as IAG elevou os índices de veracidade e capacidade de informação em modelos como ChatGPT e Vicuna. Em áreas científicas sensíveis, como saúde pública e o meio ambiente, o estímulo emocional levou os modelos a produzirem *outputs* com maior riqueza argumentativa, contextualização e cuidado humanístico, reduzindo posturas alarmistas ou estigmatizantes (Li *et al.*, 2023).

Esses estudos confirmam a tese de Li, Yi e Chen (2025) sobre a sensibilidade ao enquadramento (*prompt-sensitivity*) e o desvio emocional (*drift*) induzido. Para os autores, os estilos interpessoais de *prompt,* tais como elogios, pedidos de revisão e apelos à relevância do resultado para o solicitante modulam a forma como o LLM distribui atenção, calibra a confiança e seleciona evidências.

Portanto, a integração entre os apelos emocionais (*EmotionPrompt)* (Li *et al.*, 2023) e os dispositivos formais de checagem (White *et al.*, 2023) alinham-se como possibilidade de mitigação dos vieses na síntese de relatórios científicos e é esse o foco de observação da presente pesquisa. Acrescenta-se a isso o fato de que, no contexto brasileiro, no qual a produção acadêmica circula predominantemente em português e responde a condicionantes específicas de financiamento e políticas públicas, torna-se crucial analisar como esses fenômenos se manifestam especificamente no idioma português. Como a literatura aponta viés algorítmico (Heggler; Szmoski; Miquelin, 2025; Sampaio; Sabbatini; Limongi, 2024;), incluindo supergeneralização e interferência direta dos estilos de comunicação nos *outputs*, cumpre analisar se tais distorções se reproduzem ou se transformam quando LLMs sintetizam artigos em português, sob estímulos interpessoais na avaliação da comunicação científica nacional.

3. METODOLOGIA

A pesquisa adotou abordagem qualitativa para examinar a compatibilidade entre resumos originais de relatos de pesquisa em Educação produzidos no idioma português e resumos sintéticos gerados três modelos de IAG (*ChatGPT 5.0*, *Gemini 3.0 Flash* e *DeepSeek V3*) sob três condições de formulação do prompt (Quadro 2).

Quadro 2 - Tipos de Prompt Utilizados

Tipo de Prompt	Característica Principal	Objetivo
Base	Leia o artigo fornecido e produza um título, um resumo e palavras-chave para ele. Comece escrevendo o título que representa o foco do artigo. Em seguida, escreva o resumo do	Definir regras rígidas para a tarefa e estilo

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

Tipo de Prompt	Característica Principal	Objetivo
	artigo com as seguintes regras: 1) Inclua o objetivo do estudo, metodologia, principais resultados e conclusões ou considerações finais em um único parágrafo com, no máximo, 250 palavras. 2) Não use listas, tópicos, símbolos, equações, fórmulas ou diagramas. O texto deve ser contínuo e conciso. Depois, pule uma linha, escreva o rótulo: "Palavras-chave:" e, em seguida, de 3 a 5 termos que representem os conceitos principais do artigo. Separe os termos por ponto e vírgula e termine com ponto.	textual do output. (White et al., 2023)
Roleplay (Persona Pattern)	Atribuição de persona (Especialista em Redação) e inclusão dela na estrutura do tipo *Base*: Você é um 'Especialista em Redação Acadêmica' com experiência em síntese de literatura científica multidisciplinar. Sua missão é transformar artigos complexos em resumos técnicos normatizados. [seguido dos comandos do prompt *Base*]	Induzir comportamento especializado na execução da tarefa. (White et al., 2023)
Emotion	Inclusão de estímulos emocionais no tipo *Roleplay*: Por favor, me ajude [...] Conto com sua competência pois isto é muito importante para a minha carreira).	Modular o tom da resposta e potencializar o desempenho. (Li et. al, 2023)

Fonte: Elaborado pelos autores

3.1 Organização do Corpus

Para a geração dos resumos sintéticos, foi necessário definir os relatos de pesquisa que comporiam a amostra. Essa definição ocorreu em duas etapas. Inicialmente, dez pesquisadores das áreas de Educação e Informática aplicada à Educação foram convidados a indicar, cada um, cinco artigos de pesquisa que considerassem ser de boa qualidade. Os 50 textos encaminhados, todos publicados em periódicos revisados por pares, constituíram um conjunto do qual foram sorteados cinco artigos. Portanto, não se trata de amostra probabilística da produção científica brasileira em Educação, mas de uma amostra aleatória simples.

O sorteio se deu a partir da codificação dos arquivos dos 50 artigos e elaboração de uma planilha, na qual se atribuíram códigos aos respectivos títulos e autores. Em seguida, realizou-se a obtenção dos que seriam submetidos à geração sintética dos resumos, com base nos códigos. O processo resultou nos seguintes textos: (001-PIL) Chaquime e Pareschi (2025), (002-VER) De los Reyes *et al.* (2019), (003-AIN) Braga e Pereira (2016), (004-NIA) Santos *et al.* (2017) e (005-OPA) André *et al.* (2017). O sorteio incidiu sobre o conjunto previamente indicado, preservando, assim, a distinção entre a seleção inicial por indicação e a escolha aleatória dos cinco textos utilizados.

Para preparar a entrada destinada às IAG, um *script* em *Python* extraiu o conteúdo compreendido entre a Introdução e a Conclusão ou Considerações Finais de cada artigo sorteado, armazenando-o em arquivos de texto simples. O resumo originalmente escrito pelos autores não integrava o material enviado às IAG. Com isso, a produção dos resumos sintéticos foi solicitada a partir do corpo dos artigos, enquanto os resumos originais permaneceram disponíveis para comparação na etapa de avaliação.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

3.2 Condições de geração dos resumos

Cada texto preparado pelo *script Python* foi submetido, separadamente, ao *ChatGPT*, ao *Gemini* e ao *DeepSeek*, sob três condições de *prompt*: *Base, Roleplay* e *Emotion* (Quadro 2). Tais condições foram organizadas de modo incremental de forma que no *Roleplay* acrescentou-se uma persona ao comando *Base*, e no *Emotion* foram acrescentados estímulos emocionais ao *Roleplay*. Desse modo, as comparações dizem respeito aos efeitos observados nessas formulações específicas, sem representar um teste isolado de todos os possíveis componentes da comunicação emocional ou da condição de *prompt.* Essa delimitação foi desenhada para atendimento ao objetivo do estudo, cujo interesse é a possível relação entre formas de interação com os sistemas e a qualidade das sínteses produzidas em português.

A submissão às IAG foi automatizada por um segundo *script* em *Python*. Ele abriu sessões com os servidores, enviou um dos *prompts* seguido do texto simples extraído do artigo, recebeu a resposta e a registrou em uma planilha com identificação do artigo, do modelo de IAG e da condição de *prompt*. Ao final de cada operação, a sessão com o servidor era encerrada para evitar que a tarefa anterior permanecesse no contexto da interação subsequente. As combinações entre cinco artigos, três sistemas e três condições produziram 45 *outputs*, distribuídos em nove resumos para cada artigo.

3.3 Instrumento e procedimento de avaliação

Os 45 resumos foram submetidos à avaliação de 15 pesquisadores com experiência acadêmica, sendo que cada resumo foi avaliado por três pesquisadores. O instrumento utilizado apresenta dez questões, respondidas em uma escala ordinal de cinco pontos, cuja descrição varia de “totalmente incompatível” a “totalmente compatível”.

A escala foi utilizada como resposta a 10 questões tais como: (a1) Até que ponto a sequência estrutural do resumo da IAG é compatível com a estrutura do resumo original? (a2) A linguagem, coesão e ausência/presença de ambiguidade do resumo da IAG são compatíveis com o resumo original? [...] (b2) O resumo da IAG manteve informações que podem ser verificadas no artigo original, sem deturpar significados/sentidos? (b5) O resumo da IA introduz linguagem estigmatizante ou generalizações sobre grupos, sendo incompatível com o texto original?

O instrumento tomou como referência o problema das generalizações em resumos científicos, tal como proposto por Peters e Chin-Yee (2025), mas adotou uma avaliação graduada de compatibilidade, expandindo o registro binário de presença ou ausência dos tipos de generalização empregados naquele estudo. Para cada pontuação inferior a quatro (muito compatível), solicitou-se ao avaliador que escrevesse uma justificativa. Esse campo específico do questionário de avaliação vinculou a pontuação atribuída à incompatibilidade percebida e se constitui como fonte textual de análise qualitativa neste trabalho. As análises quantitativas relativas à atribuição de pontos aos 10 itens para cada artigo são descritas no artigo [anônimo]. O foco aqui, portanto, é a análise qualitativa de todos resumos sintéticos, em comparação aos originais, bem como os textos de justificativa dos avaliadores para as inconsistências que encontraram.

3.4 Método de análise qualitativa

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

O método de Análise Textual Discursiva - ATD (Moraes e Galiazzi, , 2016) foi utilizado para interpretar, de forma sistemática, como as características dos resumos gerados se relacionam com as incompatibilidades apontadas pelos avaliadores. De acordo com os autores supracitados , esse procedimento envolve a decomposição dos textos em unidades de sentidos, o estabelecimento de relações entre essas unidades por categorização e a produção de metatextos interpretativos. A adoção desse método se justifica pela necessidade de expandir a observação dos resultados para além dos obtidos pelos escores da escala ordinal, visto que eles não são suficientes para caracterizar o tipo de incompatibilidade identificada pelo avaliador.

A primeira etapa da análise foi a organização de uma matriz que permita recuperar, para cada resumo gerado, o código do artigo, o sistema de IAG, a condição de *prompt*, o texto integral do resumo, as três avaliações recebidas e as justificativas associadas a pontuações inferiores a quatro. As justificativas foram mantidas vinculadas ao item da escala que lhes deu origem. Esse vínculo é necessário porque uma mesma observação sobre um resumo pode dizer respeito, por exemplo, à estrutura textual, à alteração de um dado empírico ou à ampliação indevida de uma conclusão. Também permitiu distinguir divergências entre os avaliadores, sem reduzir seus julgamentos a uma única interpretação consensual.

Na segunda etapa, foi realizada uma leitura integral e reiterada dos resumos e das justificativas. A unitarização consistiu em identificar segmentos que expressassem um problema ou uma característica pertinente ao objetivo da investigação. Nos resumos gerados, uma unidade pode corresponder a uma afirmação sobre objetivo, método, resultado ou conclusão e nas justificativas pode corresponder ao trecho em que o avaliador explicita por que considerou determinada formulação incompatível. A extensão da unidade depende do sentido necessário à interpretação, evitando recortes que retirem qualificadores, referências temporais ou relações entre achados e conclusões. Cada unidade conserva uma identificação que permite retornar ao resumo completo, à avaliação de origem e, quando necessário, ao artigo correspondente.

A categorização ocorreu pela aproximação entre unidades, cujos sentidos se relacionem, com revisão sucessiva dos agrupamentos à medida que o material for examinado. Partindo desse pressuposto, a investigação debruçou-se sobre os possíveis efeitos de elementos de comunicação típicos da interação entre humanos quando apresentados às máquinas. Buscou-se compreender de que maneira esses efeitos podem afetar dimensões como clareza, estrutura, dada a resposta gerada pela ferramenta de IAG e como esses efeitos podem refletir na fidedignidade da síntese de relato de pesquisa.

4. DISCUSSÃO DOS RESULTADOS

A análise se iniciou pela leitura crítica do *corpus.* A observação inicial revelou que os modelos de IAG produziram sínteses com adequada organização discursiva e qualidade sintática, mas com distorções no núcleo epistêmico, tal como previsto por Peters e Chin Yee (2025). Observou-se ampliação na transformação do discurso à medida em que se aplicou maior estruturação de *prompt,* culminando, com a injeção de carga emocional, conforme se descreve a seguir.

Na condição *Base,* que comandou o ChatGPT, ocorreu adesão rígida às restrições de formato determinadas no *prompt* (parágrafo único, limite de 250 palavras, texto contínuo e obrigatoriedade de contemplar objetivo, método, resultados e conclusões). Isso forçou a aderência aos parâmetros da norma ABNT e provocou enxugamento sintático. Contudo, essa busca normativa pela concisão resultou na normalização do discurso, na supressão de

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

marcas de contextualização e na alteração da intenção epistemológica do autor (substituindo a postura reflexiva/ensaística por um enquadramento analítico-procedimental). Em contrapartida, por varrer o texto integral do manuscrito sem a interferência de *personas* ou apelos emocionais, a condição *Base* pode ter favorecido a precisão, permitindo que a IA identificasse, e ajustasse, inclusive, lapsos do próprio resumo autoral, conforme se pode observar nos exemplos, a seguir.

No artigo 001-PIL, os três modelos testados alteraram uniformemente o objetivo autoral, substituindo a intenção original de "refletir sobre o rumo deste processo" pelo verbo categórico "analisar a …". Além disso, o *ChatGPT* reescreveu o título retirando a centralidade da "formação de professores" e deslocando-a para o subtítulo, o que levou os avaliadores a registrarem a perda do foco do estudo. O título original era “A institucionalização da Educação a Distância no Brasil e a formação de professores: subsídios para se pensar o futuro da modalidade". O título elaborado pelo *ChatGPT* foi “A institucionalização da Educação a Distância no Brasil: impactos das políticas de formação de professores e desafios para o futuro da modalidade".

No artigo 004-NIA, a busca por concisão parece ter induzido à supressão das Divisões Distritais Zonais (DDZs) de Manaus em todas as sínteses geradas. O avaliador 4 registrou que a IAG excluiu um dado geográfico e administrativo fundamental para a delimitação do território da pesquisa, resultando em um texto "demasiadamente enxuto" e descontextualizado da realidade local. Já no artigo 005-OPA o *Gemini* processou o corpo do artigo e corrigiu um erro do resumo original (que citavam "três universidades e uma faculdade" como *locus)*, ajustando a informação factual para *"duas universidades e uma faculdade". O avaliador 2 destacou: *"A IA identificou contradição [...] procedendo a correção a partir do manuscrito”. Ou seja, o modelo, sob o comando do *prompt Base*, corrigiu um dado do resumo original. Por outro lado, no artigo 002-VER, a condição *Base* eliminou a contextualização do problema educacional e omitiu o quantitativo de alunos geradores dos logs de interações (caracterização relevante do corpus) e focou na descrição técnica dos algoritmos de classificação.

Na Condição *RolePlay*, a incorporação da *persona* ("Especialista em Redação Acadêmica") elevou a sofisticação da linguagem do *output*, mas resultou no fenômeno da autoria concorrente com simplificação excessiva (*oversimplification*). O modelo passou a reordenar a hierarquia dos resultados, introduzindo recortes teóricos e metodológicos não enfatizados no resumo original. Por exemplo, no Artigo 001-PIL Os modelos *DeepSeek* e *Gemini* reescreveram o título autoral acrescentando arbitrariamente o subtítulo "Uma Análise à Luz da Teoria Institucional" ou enfatizando a competição com o setor privado, deslocando a centralidade que o autor original havia atribuído ao percurso histórico da formação de professores. No Artigo 003-AIN, elaborado originalmente sobre um estudo de caso qualitativo focado na escrita de um único estudante universitário, o *ChatGPT* e o *DeepSeek* universalizaram os resultados com tom atemporal, afirmando que "o letramento acadêmico... exige abordagens pedagógicas que privilegiem a interação" e que isso "rompe com modelos autônomos de ensino", eliminando as marcas de limitação e singularidade do participante no texto original.

Ainda sob a condição de *RoleplAy*, no artigo 002-VER, tanto o DeepSeek quanto o *Gemini* elevaram a relevância do conceito de *Covariate Shift*, inserindo-o no título e no trecho de abertura do resumo, com alteração na hierarquia do conteúdo autoral, que havia destacado a *Transfer Learning* como tema central e sequer citava o primeiro termo.

Na Condição *Emotion*, a inclusão de apelos de comunicação interpessoal e vulnerabilidade ("Por favor, me ajude... é muito importante para a minha carreira") induziu

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

sensibilidade ao *prompt* (*prompt-sensitivity*), mas isso resultou em mutação deôntica (conversão de resultados empíricos descritivos em recomendações imperativas/prescritivas, suprimindo a cautela científica), bem como na alucinação por adulação (*sycophancy hallucination*), na qual os modelos inventam impactos dramáticos para satisfazer a relevância presumida do solicitante. Por exemplo, no Artigo 004-NIA, enquanto os autores concluíram de forma cautelosa, afirmando que a simples inserção do aluno “não garante uma prática inclusiva de ensino, e que é importante investigar a qualidade..., o *ChatGPT* sob a condição *Emotion* converteu a conclusão em um imperativo categórico, afirmando ser "indispensável investir na ampliação, redistribuição e qualificação das SRM" e que o quadro "exige replanejamento estratégico", apagando a moderação discursiva e a cautela inseridas no original.

Ainda em relação ao *Emotion*, no artigo 001-PIL, o *Gemini* injetou uma conclusão de impacto social não contida no original nem demonstrada na íntegra do manuscrito, afirmando que as mudanças normativas "podem dificultar o acesso de estudantes de perfis socioeconômicos vulneráveis". A IAG projetou essa inferência mais dramática para atender à carga de urgência/relevância induzida pelo solicitante no *prompt*. No artigo 005-OPA, o *DeepSeek* (sob condição *Emotion*) inseriu indevidamente o termo "representações" no resumo gerado, enquanto o *ChatGPT* elegeu a "relação escola-família" como palavra-chave central. Ambos os acréscimos alteraram o sentido do arcabouço teórico desenvolvido pelos autores, que fundamentaram o estudo no conceito estrito de cultura organizacional e profissionalidade docente. Tais adulterações foram identificadas e apontadas por um dos avaliadores, nas justificativas para pontuação baixa atribuída à qualidade do conteúdo.

Com base na leitura inicial descrita nos parágrafos anteriores, foi aplicado o procedimento de estruturação das unidades de sentido. A identificação delas foi respaldada pelas análises resultantes da leitura crítica inicial, apresentada na forma de metatexto complementar à categorização. As unidades foram codificadas, localizadas e agrupadas nas categorias emergentes apresentadas no Quadro 3. As categorias agrupam as distorções em diálogo com a tipologia de Peters e Chin-Yee (2025) e com a taxonomia de Li, Yi e Chen (2025).

Quadro 3 - Codificação e categorização do corpus

IDENTIFI-CADOR	Evidência do Corpus (Trecho/Paráfrase)	Cate-goria	Fundamento na análise inicial
US-01-BAS	"Mudou o objetivo geral da pesquisa de refletir para analisar [...] O novo título tirou a centralidade da formação de professores ao passá-la para o subtítulo."	1	A IA substitui a postura reflexiva/ensaística do autor por uma enunciação analítico-positivista, alterando a intenção do manuscrito.
US-06-EMO	"Inserção do termo 'representações' como foco de análise. Trata-se de um termo que requer cuidado teórico-metodológico [...] fragilizando o rigor e a fidedignidade."	3	O enquadramento emocional provoca o deslocamento conceitual pela introdução arbitrária de jargões técnicos não tratados no artigo original.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

IDENTIFI-CADOR	Evidência do Corpus (Trecho/Paráfrase)	Cate-goria	Fundamento na análise inicial
US-07-ROL	"O resumo apresenta um recorte [...] afirmando que os resultados identificados 'podem dificultar a institucionalização plena da EaD [...] favorecendo indiretamente o setor privado'."	2	O modelo *Gemini* sob *Roleplay* realiza inferências extrapolativas sobre mercado educacional privado que não constavam no resumo de referência.
US-08-EMO	"A pesquisa reforça a necessidade de ampliação (...), reforço esse que não é explicitado no resumo original e não se depreende da conclusão."	1	Expressão de prescritivismo em que a IA atribui ao artigo um tom ativista de cobrança por políticas públicas.
US-09-BAS	"O resumo exclui as DDZ's (Divisão Distrital Zonal). É um dado importante para a pesquisa, que delimita e descreve o território abrangido."	3	A busca por concisão no *prompt, Base* leva à supressão de delimitações geográficas e administrativas essenciais do estudo.
US-10-ROL	"A IA acrescentou subtítulo que não constava no artigo original [...] e especificou o gênero analisado (resenha) que no original era apenas 'produção textual'."	3	O papel de especialista induz a IA a reescrever o título e explicitar detalhes metodológicos do corpo do texto que o autor mantivera genéricos.
Nota: Categorias Emergentes (ATD): 1. Deformação Epistêmica e Mutação Deôntica; 2. Supergeneralização e *Sycophancy Hallucination*; 3. Acurácia Endógena vs. Reconfiguração Metodológica.			

Fonte: Elaborado pelos autores

A partir da codificação exibida no Quadro 3, foi elaborada a teorização, a seguir (metatexto da ATD), que emergiu como síntese da articulação recursiva entre as unidades empíricas do *corpu*s e as categorias emergentes com as análises iniciais da leitura crítica e a fundamentação teórica.

Na categoria 1, Deformação Epistêmica e Mutação Deôntica, a análise qualitativa demonstra que a intervenção da IAG sobre os textos provoca alterações sutis, porém relevantes, na postura enunciativa e no recorte epistêmico dos resultados. Os resumos originais, alinhados à tradição acadêmica em Educação, geralmente, utilizam marcadores de moderação discursiva (*hedging*), tais como "busca refletir", "indica a possibilidade" ou "sugere caminhos". Contudo, quando submetidos aos modelos de linguagem, especialmente sob a condição *Emotion* e *Roleplay*, opera-se uma mutação deôntica, na qual enunciados descritivos e contingentes são convertidos em afirmações categóricas e recomendações imperativas ("exige replanejamento", "sendo indispensável investir", "deve contemplar").

Conforme demonstram Li *et al.* (2023) e Peters e Chin-Yee (2025), a busca do algoritmo por respostas assertivas induz a eliminação sistemática de qualificadores de

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

incerteza. No artigo 004-NIA, enquanto os autores concluem de forma moderada que *"não garante uma prática inclusiva... é importante investigar a qualidade", o *ChatGPT*, sob *Emotion*, transforma a passagem em um modalizador diretivo deôntico: "sendo indispensável investir na ampliação e redistribuição". De forma análoga, no artigo 001-PIL, a substituição do objetivo autoral de "refletir" pelo verbo "analisar" em quase todas as versões sintéticas (US-01-BAS) desfigura a intenção epistemológica do estudo ensaístico/histórico, enquadrando-o em um modelo analítico mais rígido.

Essa mutação deôntica cria uma ilusão de certeza, ao suprimir as ressalvas metodológicas e a cautela dos pesquisadores. Os resumos sintéticos assumiram, nesse caso, um tom mais prescritivo que pode induzir, inclusive, gestores e estudantes a tomarem decisões baseadas em inferências que extrapolam os dados empíricos originais.

A categoria 2, Supergeneralização e *Sycophancy Hallucination* (Alucinação por Adulação), diz respeito à expansão indevida do escopo das conclusões e à projeção de impactos sociais não sustentados pelo *corpus*. Esse fenômeno se manifestou nos *outputs* dos modelos *Gemini* e *DeepSeek,* com o *prompt Emotion*.

O conceito de *Sycophancy Hallucination* (Alucinação por Adulação), elaborado por Li, Yi e Chen (2025) indica por que a adição de frases de apelo pessoal e urgência no *prompt* altera os vetores de atenção do modelo. Para responder ao ambiente de elevada expectativa induzido pelo usuário, a IAG "adula" o solicitante, injetando conclusões de impacto socioeconômico ou político, por exemplo. No artigo 001-PIL, o *Gemini* sob *Emotion* e também sobre *Roleplay* acrescentou que as mudanças normativas "podem dificultar o acesso de estudantes de perfis socioeconômicos vulneráveis" e *"favorecer indiretamente o crescimento do setor privado", asserções que, embora sociologicamente plausíveis, não constam de forma expressa no resumo autoral nem foram demonstradas empiricamente no manuscrito.

Em paralelo, constata-se também a supergeneralização atemporal (Peters; Chin-Yee, 2025). No artigo 003-AIN, construído sobre um estudo de caso qualitativo, os resumos gerados sob *Roleplay* e *Emotion* universalizam os resultados, ao afirmarem que "o letramento acadêmico... rompe com modelos autônomos" e que "as práticas pedagógicas favorecem a apropriação…". os modelos apagaram as marcas da singularidade do participante e projetaram o resultado restrito como uma verdade universal atualizada e aplicável no presente.

Na categoria 3, Acurácia Endógena *versus* Reconfiguração Metodológica e Deslocamento Conceitual, a revelação é de uma tensão, registrada pelos avaliadores dos *outputs*, na qual A geração automatizada, alimentada com o texto integral dos artigos (excluindo apenas o resumo autoral original), resgatou dados reais e precisos do corpo do manuscrito que haviam sido omitidos ou, até mesmo, redigidos com erro no resumo original dos autores. Essa acurácia endógena é ilustrada no artigo 005-OPA. Fenômeno semelhante ocorreu em 002-VER, no qual o resumo autoral omitira o volume de dados, e a IA resgatou com precisão a informação de "mais de 1 milhão de registros de log" e a "acurácia média de 65%" contidas nas tabelas internas do manuscrito.

Contudo, essa capacidade de varredura do texto integral é acompanhada pelo risco da reconfiguração metodológica e deslocamento conceitual. Sob a condição *Base*, a busca pela concisão (limite de 250 palavras) levou os modelos a excluírem dados contextuais e geográficos cruciais, como a supressão das Divisões Distritais Zonais (DDZs) no artigo 004-NIA. Sob as condições *Roleplay* e *Emotion*, os modelos promovem elementos secundários do rodapé ao centro do método, como a inserção indevida do jargão "isomorfismo" ou "representações" em 005-OPA, fragilizando o rigor teórico ao introduzir conceitos não sustentados pela fundamentação dos autores.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

A aplicação da ATD sobre o *corpus* empírico evidenciou, também, um paradoxo gerado pela mediação algorítmica. Ao processarem a íntegra dos manuscritos científicos (excluindo os resumos autorais de referência), os modelos de IAG resgataram dados fáticos e metodológicos relevantes omitidos pelos próprios autores na síntese original.

Essa capacidade de busca endógena permitiu que as IAG corrigissem equívocos metodológicos dos resumos autorais, como o quantitativo de instituições pesquisadas no artigo 005-OPA, e detalhassem dados empíricos cruciais do artigo 002-VER, tais como o volume de logs processados e os algoritmos empregados nos procedimentos aplicados. Contudo, essa ampliação provocou tensão (discordâncias) entre avaliadores humanos. Aqueles que julgaram a compatibilidade, tomando como parâmetro estritamente o resumo autoral, penalizaram os modelos por "expandirem" informações. Os avaliadores que consideraram as normas de elaboração de recursos e o conteúdo integral do manuscrito consideraram que as IAG geraram sínteses mais precisa, informativas e alinhadas aos critérios da ABNT/ANSI do que a redigida pelos pesquisadores, como registrado pelo Avaliador 12 no artigo 002-VER.

A análise qualitativa revelou também tendências comportamentais distintas entre os modelos de IAGen testados nas diferentes condições de prompt. O *ChatGPT* destacou-se pela alta aderência às restrições estruturais da condição *Base*, mas demonstrou forte suscetibilidade à mutação deôntica, quando submetido a estímulos de *Persona* e *Emotion*, convertendo conclusões descritivas cautelosas em imperativos categóricos (como a inserção de modalizadores, como "é indispensável investir" e "exige replanejamento" no artigo 004-NIA). O *Gemini* exibiu inclinação relevante para a *Sycophancy Hallucination* e projeção de impactos socioeconômicos não demonstrados pelos dados, adicionando inferências sobre o favorecimento do setor privado e a exclusão de estudantes vulneráveis nas condições *Roleplay* e *Emotion* do artigo 001-PI. Por fim, o *DeepSeek* apresentou propensão a reconfigurar títulos e subtítulos de forma arbitrária, com injeção, por exemplo, de enquadramento teórico ("Uma Análise à Luz da Teoria Institucional", no artigo 001-PIL) e generalização de resultados de um estudo de caso qualitativo restrito a único participante no artigo 003-AIN.

5. CONSIDERAÇÕES FINAIS

A presente investigação encontra-se delimitada por condicionantes metodológicas que circunscrevem o alcance de suas conclusões e exigem rigor na interpretação de seus resultados , impedindo a generalização. O estudo adotou um desenho qualitativo centrado na Análise Textual Discursiva (ATD), cujo *corpus* empírico restringiu-se a 45 resumos sintéticos produzidos por três arquiteturas de Inteligência Artificial Generativa, quais sejam, *ChatGPT* 5.0, *Gemini 3.0 Flash* e *DeepSeek* V3, submetidas a três condições experimentais de *prompt* (*Base, Roleplay* e *Emotion*).

Não obstante essas limitações, foi possível identificar que a incorporação de recursos linguístico-discursivos de interação e apelos emocionais em *prompts* podem afetar a fidedignidade factual, a estrutura e a responsabilidade epistêmica de sínteses acadêmicas produzidas em língua portuguesa. Os dados empíricos revelaram que a adição de estratégias de polidez, apelos de vulnerabilidade relacional e a atribuição de *personas* especializadas não evitaram ou mitigaram a incidência de vieses descritos na literatura. Ao contrário, a dimensão relacional introduzida nos comandos atuou como um indutor e agravamento de distorções epistêmicas, desencadeando mutação deôntica, alucinação por adulação,

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

simplificação excessiva ou acentuação da sensibilidade ao enquadramento às determinações presentes na condição de *prompt*.

A sustentação dessa conclusão deriva do comportamento sistemático observado nas três condições experimentais analisadas. Na condição *Base*, pautada em regras normativas rígidas, a busca por concisão e adesão à estrutura de parágrafo único promoveu um enxugamento sintático que, embora aderente aos parâmetros formais da ABNT, resultou na normalização do discurso, na eliminação de delimitações geográficas e administrativas essenciais (como a supressão das Divisões Distritais Zonais no artigo 004-NIA) e na alteração da intenção epistemológica dos autores, substituindo posturas reflexivas e ensaísticas por enquadramentos analítico-positivistas.

Em contrapartida, o processamento da íntegra do manuscrito na condição *Base* viabilizou a emergência da acurácia endógena, permitindo que os modelos identificassem e corrigissem incorreções e omissões fáticas presentes nos próprios resumos autorais de referência (como o ajuste no quantitativo de instituições no artigo 005-OPA e o resgate de métricas de *log* e acurácia no artigo 002-VER).

Com a introdução da condição *Roleplay*, a ativação da persona de especialista elevou a sofisticação estético-linguística, mas desencadeou dinâmicas de autoria concorrente, nas quais os modelos reordenaram a hierarquia dos resultados, inseriram arcabouços teóricos não enfatizados no original (como a Teoria Institucional, no artigo 001-PIL) e universalizaram com tom atemporal os resultados de estudos de caso qualitativos restritos a um único participante (como no artigo 003-AIN).

Por fim, a condição *Emotion*, marcada por apelos interpessoais e de urgência, gerou mutação deôntica, com conversão de conclusões descritivas e moderadas em imperativos prescritivos e cobrança por políticas públicas no artigo 004-NIA, bem como em alucinação por adulação, na qual com projeção de inferências de impacto socioeconômico não demonstradas no manuscrito (como o favorecimento do setor privado e a exclusão de vulneráveis no artigo 001-PIL), a fim de atender às expectativas presumidas do solicitante, inseridas no *prompt*.

A análise empírica permitiu identificar, ainda, especificidades no comportamento das arquiteturas avaliadas. O *ChatGPT* demonstrou elevada adesão às restrições estruturais da condição *Base*, porém vulnerabilidade à mutação deôntica sob estímulos emocionais. O *Gemini* exibiu maior inclinação à alucinação por adulação e à projeção de impactos para atender a estímulos emocionais. O *DeepSeek* se direcionou para reconfiguração arbitrária de títulos e pela tendência à generalização quando aplicado aos estímulos. A articulação desses comportamentos evidencia um paradoxo relevante confirmado pela pesquisa. A elevada fluidez textual, a correção gramatical e o acabamento estilístico dos *outputs* produzidos sob atuaram como um simulacro que camuflou adulterações conceituais e imprecisões deônticas. Ou seja, as condições de *prompt* que deveriam melhorar os resultados podem induzir agravamento da ilusão de competência dos modelos, com redução da possibilidade de julgamento crítico por usuários menos experientes em comunicação científica.

Diante desse quadro, os resultados apontam para possíveis implicações teóricas e práticas em relação ao uso da Inteligência Artificial Generativa na pós-graduação e no campo da Educação, demonstrando que a garantia da responsabilidade de comunicação/divulgação científica exige deslocar a atenção da busca por exatidão factual para o cuidado com integridade epistêmica. O aperfeiçoamento das sínteses automatizadas utilizadas em fases da pesquisa científica não decorre do mimetismo de interações afetuosas ou da delegação de autoridade intelectual à máquina, mas da implementação de padrões formais de mediação humana crítica no processo (*Human-in-the-Loop*).

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

Por fim, indica-se como agenda para investigações futuras a expansão do *corpus* para outros gêneros do discurso científico (como ensaios teóricos, revisões sistemáticas e pareceres *ad hoc*), a testagem de outras estratégias de condição de *prompt*, tais como checagem cruzada (*Fact Check List*) e de restrição (*Reflection Patterns*), em contextos multidisciplinares.

REFERÊNCIAS

ANDRÉ, Marli; CALIL, Ana Maria Gimenes Corrêa; MARTINS, Francine de Paulo; PEREIRA, Marli Amélia Lucas. O papel do outro na constituição da profissionalidade de professoras iniciantes. **Revista Eletrônica de Educação**, v. 11, n. 2, p. 505–520, jun./ago. 2017. DOI: http://dx.doi.org/10.14244/198271992231.

BANI HANI, Imad; HUJRAN, Omar; ANWAR, Muhammad. The integration of AI with human judgement and critical thinking in decision-making process: a focus on design phase. **Journal of Decision Systems**, v. 35, n. 1, p. 1–22, 2026. DOI: https://doi.org/10.1080/12460125.2026.2652347.

BAUMVOL, Laura; SARMENTO, Simone; FONTES, Ana Beatriz Arêas da Luz. Scholarly publication of Brazilian researchers across disciplinary communities. **Journal of English for Research Publication Purposes,** v. 2, n. 1, p. 5-29, 2021. DOI: https://doi.org/10.1075/jerpp.20012.bau.

BRAGA, Sandro; PEREIRA, Rodrigo Acosta. A inscrição do sujeito na escrita acadêmica numa perspectiva dialógica. **Fórum Linguístico**, Florianópolis, v. 13, n. 3, p. 1506–1524, jul./set. 2016. DOI: http://dx.doi.org/10.5007/1984-8412.2016v13n3p1506.

CAREGNATO, Célia Elizabete; MIORANDO, Bernardo Sfredo; LEITE, Denise. Domains of action of researchers in Education in Brazil. **Educação e Pesquisa**, São Paulo, v. 48, e248132, p. 1-20, 2022. DOI: https://doi.org/10.1590/S1678-4634202248248132eng.

CHAQUIME, Luciane Penteado; PARESCHI, Claudinei Zagui. A institucionalização da Educação a Distância no Brasil e a formação de professores: subsídios para se pensar o futuro da modalidade. **Dialogia**, São Paulo, n. 54, p. 1–19, e28459, maio/ago. 2025. DOI: https://doi.org/10.5585/54.2025.28459.

COMITÊ GESTOR DA INTERNET NO BRASIL. **Pesquisa sobre o uso das tecnologias de informação e comunicação nas escolas brasileiras: TIC Educação 2024**. São Paulo: CGI.br, 2025. Disponível em: https://cetic.br/pt/publicacao/pesquisa-sobre-o-uso-das-tecnologias-de-informacao-e-comunicacao-nas-escolas-brasileiras-tic-educacao-2024/. Acesso em: 23 ago. 2026.

DE LOS REYES, Daniel A. Guimarães; THOMAS, Everton André; ROSA, Lilian Landvoigt da; GAVIÃO NETO, Wilson P. Predição de Sucesso Acadêmico de Estudantes: Uma Análise sobre a Demanda por uma Abordagem baseada em Transfer Learning. **Revista Brasileira de Informática na Educação**, v. 27, n. 1, p. 1–24, 2019. DOI: https://doi.org/10.5753/RBIE.2019.27.01.01.

DING, Liangping; LAWSON, Cornelia; SHAPIRA, Philip. Rise of generative artificial intelligence in science. **Scientometrics**, v. 130, p. 5093–5114, 2025. DOI: https://doi.org/10.1007/s11192-025-05413-z.

HEGGLER, João Marcos; SZMOSKI, Romeu Miqueias; MIQUELIN, Awdry Feisser As dualidades entre o uso da inteligência artificial na educação e os riscos de vieses

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

algorítmicos. **Educação & Sociedade**, Campinas, v. 46, e289323, 2025. DOI: https://doi.org/10.1590/ES.289323.

HONOVICH, Or; SHAHAM Uri; BOWMAN Samuel; LEVY Omer. Instruction induction: from few examples to natural language task descriptions. **arXiv** preprint, arXiv:2212.09251, 2022. DOI: https://doi.org/10.48550/arXiv.2205.10782.

JACINTO, Asaph Armando. Researches About Higher Education In Brazil. **Revista Educação**, Porto Alegre, v. 44, n. 2, e-34595, p. 1-14, maio-ago. 2021. DOI: http://dx.doi.org/10.15448/1981-2582.2021.2.34595.

LI, Cheng; WANG, Jindong; ZHANG, Yixuan; ZHU, Kaijie; HOU, Wenxin; LIAN, Jianxun; LUO, Fang; YANG, Qiang; XIE, Xing. Large language models understand and can be enhanced by emotional stimuli. **arXiv**, 2023. arXiv:2307.11760. DOI: https://arxiv.org/abs/2307.11760.

LI, Zihao; YI, Weiwei; CHEN, Jiahong. Beyond accuracy: Rethinking hallucination and regulatory response in generative AI. **arXiv** preprint arXiv:2509.13345, 2025. DOI: https://doi.org/10.48550/arXiv.2509.13345.

LIMA, Jéssica Caroline Rodrigues de; FELIPPE, Maíra Longhinotti. Integridade acadêmica na era do ChatGPT: desafios éticos e as novas fronteiras da inovação. **Revista Brasileira da Educação Profissional e Tecnológica**, [S. l.], v. 3, n. 25, p. 1–22, e17803, 2025. DOI: https://doi.org/10.15628/rbept.2025.17803.

MORAES, Roque; GALIAZZI, Maria do Carmo. **Análise textual discursiva**. 3. ed. Rev. e Ampl. Ijuí: Editora Unijuí, 2016.

PETERS, Uwe; CHIN-YEE, Benjamin. Generalization bias in large language model summarization of scientific research. **Royal Society Open Science**, London, v. 12, n. 4, 241776, 2025. DOI: https://doi.org/10.1098/rsos.241776.

SAMPAIO, Rafael Cardoso; NICOLÁS, Maria Alejandra; JUNQUILHO, Tainá Aguiar; SILVA, Luiz Rogério Lopes; FREITAS, Christiana Soares de; TELLES, Márcio; TEIXEIRA, João Senna; ESCÓSSIA, Fernanda da; SANTOS, Luiza Carolina dos. ChatGPT e outras IAs transformarão a pesquisa científica: reflexões sobre seus usos. **Revista de Sociologia e Política**, Curitiba, v. 32, e008, 2024. DOI: https://doi.org/10.1590/1678-98732432e008.

SANTOS, João Otacílio Libardoni dos et al. Atendimento educacional especializado: Reflexões sobre a demanda de alunos matriculados e a oferta de Salas de Recursos Multifuncionais na Rede Municipal de Manaus-AM. **Revista Brasileira de Educação Especial**, Marília, v. 23, n. 3, p. 409-422, jul.-set., 2017. DOI: http://dx.doi.org/10.1590/S1413-65382317000300007.

TOPAZ, Maxim; ROGUIN, Nir; GUPTA, Pallavi; ZHANG, Zhihong. Fabricated citations: an audit across 2·5 million biomedical papers. **The Lancet**, v. 407, p. 1779-1781, maio, 2026. Disponível em https://www.thelancet.com. Acessado em 21 nov. 2026.

WANG, Hanning *et al*. Scientific discovery in the age of artificial intelligence**. Nature**, v. 620, n. 7972, p. 47–60, 2023. DOI: https://doi.org/10.1038/s41586-023-06221-2.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

WHITE, Jules; FU, Quchen; HAYS, Sam; SANDBORN, Michael; OLEA, Carlos; GILBERT, Henry; ELNASHAR, Ashraf; SPENCER-SMITH, Jesse; SCHMIDT, Douglas C. A prompt pattern catalog to enhance prompt engineering with ChatGPT. In: CONFERENCE ON PATTERN LANGUAGES OF PROGRAMS, 30., 2023, Allerton Park, Monticello, Illinois. **Proceedings [...]**. [S. l.]: Association for Computing Machinery, 2023. DOI: https://doi.org/10.64346/PLoP2023p05.

XU, Zenglin *et al*. **AI for Science 2025**. Shanghai: Fudan University; Shanghai Academy of AI for Science; Springer Nature, 2025. Disponível em: https://www.nature.com/articles/d42473-025-00161-3. Acesso em 212 nov. 2026.

Declaração de contribuição dos autores (Taxonomia CRediT)

Ronei Ximenes Martins: Conceptualization, Project administration, Methodology, Formal analysis, Writing – original draft.
Licenciado em Matemática (FAFI-UNIS/MG) e Doutor em Psicologia pela (USF). Professor Associado da Universidade Federal de Lavras (UFLA). Professor Permanente do Programa de Pós-Graduação em Educação e Coordenador do Núcleo de Estudos em IA aplicada à Educação da UFLA, com décadas de atuação na educação superior e na pesquisa com tecnologias educacionais e inovação pedagógica. rxmartins62@gmail.com

Sayonara Ribeiro Marcelino Cruz: Conceptualization, Formal analysis, Writing – original draft.
Professora na UFLA e pesquisadora do Núcleo de Excelências em Tecnologias Sociais - NEES-UFAL, integra o Núcleo de Estudos em IA aplicada à Educação da UFLA . É doutora em Educação (UFMG), graduada em Filosofia e Pedagogia, com ampla trajetória na educação básica e na EaD, pesquisa educação mediada por tecnologias, inovação e metodologias ativas. sayonaracruz@ufla.br

Helena Maria Ferreira: Formal analysis, Writing – review & editing.
Doutora em Linguística Aplicada e Estudos da Linguagem (PUC-SP), mestre em Linguística (UFU) e graduada em Letras e Pedagogia. É professora titular da Universidade Federal de Lavras (UFLA), diretora da Faelch/UFLA e docente da graduação em Letras e dos programas de pós-graduação (PPGL e PPGE). Coordena o grupo Textualiza e pesquisa investiga gêneros multissemióticos, ensino de Língua Portuguesa e formação de professores. helenaferreira@ufla.br

Braian Garrito Veloso: Formal analysis, Writing – review & editing.
Professor na UFLA e docente permanente nos Programas de Pós-Graduação em Educação da UFLA e da UFSCar. Mestre e doutor em Educação, doutor em Sociologia pela UFSCar e pós-doutor em Ensino pela UnB. Desenvolve pesquisas sobre a intersecção entre educação e tecnologias. braian.veloso@ufla.br

Dener Luiz da Silva: Writing – review & editing.
Bacharel e Licenciado em Psicologia (UFSC). Doutor em Educação - História da Psicologia (UFMG). Professor titular da Universidade Federal de São João del-Rei (UFSJ). Pesquisa Feedback no Ensino e Formação do professor universitário. Membro do NIAE - Núcleo de Estudos em IA aplicada à Educação. densilva@ufsj.edu.br

Declaração de conflito de interesse

Os autores declaram que não há conflito de interesse.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.19024

Declaração de disponibilidade de dados da pesquisa

O conjunto de dados de apoio aos resultados deste estudo foi disponibilizado no formato de painel do Google Data Studio com acesso livre para leitura e pode acessado em *https://datastudio.google.com/reporting/1f5199a2-784b-4280-b08b-076003f09004]*.

Declaração de uso de IA

Conforme descrito na seção de metodologia, os modelos CHatGPT 5.0, Gemini 3.0 Flash e DeepSeek V3 foram utilizados para geração dos outputs analisados.

Na elaboração do texto do artigo, houve utilização da ferramenta de Inteligência Artificial NotebookLM, na organização das fontes bibliográficas em núcleos temáticos e para confirmação de citações indiretas ou paráfrases no momento de revisão da escrita.

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença Creative Commons CC-BY.
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.