



Evolução das IA Generativas na Educação: Uma Nova Perspectiva na Produção de Conhecimento em Química

DOI: 10.37702/2175-957X.COBENGE.2025.6022

Autores: RODRIGO MERO SARMENTO DA SILVA, CARLOS ANDRÉ DOS SANTOS LIMA

Resumo: O artigo discute a evolução da Inteligência Artificial Generativa no ensino de Química, destacando seu potencial para apoiar a criação de materiais didáticos e a explicação de conceitos abstratos. Apesar dos avanços, ainda existem desafios, como erros conceituais e a geração de informações imprecisas. A melhoria na formulação dos comandos para as IAs, conhecida como engenharia de prompt, mostrou ser fundamental para aprimorar a qualidade das respostas. O estudo realizou testes com diferentes modelos de linguagem e demonstrou que prompts bem estruturados aumentam a precisão, reduzem alucinações e favorecem a inclusão de referências confiáveis. Conclui-se que, embora promissora, a aplicação dessas tecnologias na educação ainda exige cuidados e pesquisas contínuas para seu uso seguro e eficaz.

Palavras-chave: IA Generativa, Química, Educação

EVOLUÇÃO DAS IA GENERATIVAS NA EDUCAÇÃO: UMA NOVA PERSPECTIVA NA PRODUÇÃO DE CONHECIMENTO EM QUÍMICA

1. INTRODUÇÃO

A Inteligência Artificial Generativa (IAG) tem desempenhado um papel cada vez mais relevante no contexto educacional, possibilitando novas formas de ensino e aprendizagem. A capacidade desses modelos de gerar textos e imagens coerentes vem sendo aplicada para auxiliar docentes na elaboração de materiais didáticos, adaptação de conteúdos para diferentes níveis de conhecimento e produção de representações visuais de conceitos abstratos (HOLMES; LUCKIN, 2016). No entanto, estudos iniciais apontaram limitações significativas desses modelos, como a presença de imprecisões conceituais, falta de referências adequadas e alucinações na geração de conteúdo (FERGUS; BOTHA; OSTOVAR, 2023).

Na educação em Química, a linguagem científica possui camadas de representação simbólica, o que impõe desafios adicionais ao uso de IAGs. Estudos recentes apontam que, apesar do avanço tecnológico, muitas ferramentas ainda apresentam dificuldades na representação correta desses aspectos (JOHNSTONE, 1993; TALANQUER, 2023). Modelos anteriores, como o ChatGPT 3.5, frequentemente cometiam erros conceituais ao descrever modelos atômicos, introduzindo termos do senso comum e negligenciando a precisão científica necessária para um ensino eficaz (LEON; VIDHANI, 2023). Girotto et al. (2024) também apontam que as IAGs apresentam desafios na precisão conceitual e na representação correta dos modelos atômicos, destacando a necessidade de aprimoramento nos prompts utilizados para interações educacionais.

Com o avanço das arquiteturas de IA e o refinamento das abordagens de engenharia de prompt, observa-se uma melhora substancial na capacidade dessas ferramentas de gerar conteúdo confiável e alinhado aos padrões acadêmicos. O presente estudo tem como objetivo avaliar o desempenho das IAGs, com ênfase no ChatGPT, na produção de conteúdo educacional em Química. Será realizado um comparativo entre respostas obtidas em estudos anteriores e abordagens emergentes, buscando identificar avanços na precisão conceitual, na redução de alucinações e na inclusão de referências embasadas.

1.1 Inteligência Artificial

A Inteligência Artificial tem suas origens na década de 1950, quando pesquisadores como Alan Turing e John McCarthy exploraram a possibilidade de criar máquinas capazes de simular a inteligência humana. Desde então, a evolução das redes neurais, do aprendizado profundo e dos modelos de processamento de linguagem natural possibilitou o desenvolvimento das Inteligências Artificiais Generativas (GOODFELLOW et al., 2014).

As IAGs ganharam notoriedade com a introdução de modelos como o GPT (Radford et al., 2019) e o DALL-E, que revolucionaram a geração de textos e imagens com base em aprendizado de padrões de dados massivos (BROWN et al., 2020).

Com o lançamento do ChatGpt 3.5 pela Openai em 30 de novembro de 2022, análises das suas capacidades vêm sendo feitas para avaliar o seu uso educacional. Girotto et al. (2024) apresenta as limitações quanto ao uso desta ferramenta para o ensino da Química. No trabalho, os autores concluem que apesar de que a maior parte das respostas continham informações corretas, haviam desvios que poderiam vir a comprometer a compreensão do conteúdo.

Desde então, avanços foram feitos em relação aos comandos que podem ser utilizados e como podem ser utilizados para possibilitar respostas mais eficientes dos modelos generativos (Zhao et al. 2021; Kong et al, 2024; Parnami et al 2022). Essas novas descobertas implicam a necessidade de reavaliação de trabalhos.

1.2 Engenharia de Prompt

A engenharia de prompt é uma estratégia essencial para aprimorar a interação com IAGs, permitindo o refinamento das respostas geradas. Segundo White et al. (2023), a qualidade da resposta de uma IA está diretamente relacionada à clareza e estrutura da instrução fornecida. Isso inclui técnicas como a definição de restrições de resposta, inclusão de exemplos e a divisão de comandos complexos em instruções menores.

Estudos apontam que ajustes em *prompts* podem reduzir significativamente a incidência de alucinações e erros conceituais em respostas geradas por IA (Girotto et al., 2024).

A engenharia e design de prompt requer entendimento sobre as limitações dos LLMs. Técnicas mais avançadas podem ser usadas para reduzir a ocorrência de alucinações e incrementar o desempenho dos modelos generativos.

Em seu livro “Prompt Engineering for Generative AI: Future-Proof Inputs for Reliable AI Outputs”, James Taylor e Mike Taylor apresentam cinco princípios fundamentais para a engenharia de prompt. Segundo os autores, o objetivo de se pensar em estratégias que sejam atemporais, e que possam ser aplicadas a diferentes modelos e que ainda assim seja possível obter respostas satisfatórias apesar das suas limitações. A saber: Direção, Formato, Exemplo, Divisão, Avaliação.

A “Direção” refere-se ao estilo no qual espera-se receber a nossa mensagem, pode-se entender como a atribuição de uma persona, a simples tarefa de atribuir uma personalidade ao modelo pode aumentar substancialmente a capacidade da IAG de realizar tarefas complexas e inclusive de simular interações humanas (Kong et al, 2024). Para a tarefa estudada neste trabalho, isto seria o equivalente a solicitar a IAG que atuasse como um experiente professor de Química.

O “Formato” auxilia a customizar a resposta do modelo generativo, essa especificação pode estar sugerida no início, com a direção dada ao modelo, entre os formatos mais populares tem-se: markdown, JSON, YAML, csv.

Fornecer “Exemplos”, pode se tornar um mecanismo capaz de auxiliar o modelo a raciocinar sobre a tarefa e produzir respostas melhores. Embora, como demonstrou Zhao et al. (2021), é preciso cautela na escolha dos exemplos uma vez que os modelos costumam apresentar variância, a depender do formato ou dos exemplos fornecidos.

Resultado satisfatórios, também podem ser obtidos quando a tarefa é “Dividida” em partes, isto possibilita ao modelo refletir sobre o comando e fornecer uma solução que melhor atende às necessidades do usuário.

Por fim a “Avaliação” da consistência dos comandos, pode ser feito solicitando o mesmo comando várias vezes, a avaliação também é uma forma de mensurar a qualidade do comando e fazer ajustes em casos de saídas com muita variabilidade.

2. METODOLOGIA

Na primeira fase, realizou-se a reformulação do prompt original, adaptando-o a um formato mais adequado a modelos autorregressivos. Essa reformulação visou otimizar as respostas fornecidas pelos modelos generativos, mitigando problemas identificados na etapa anterior, como ambiguidades e falta de precisão conceitual. O novo conjunto de prompts foi projetado para aprimorar a clareza das respostas, reduzir alucinações e garantir um alinhamento mais preciso com os conceitos químicos. Cada um dos novos prompts foi testado entre cinco e dez vezes, sendo consideradas como métricas a ocorrência de referências, a legibilidade e a proporção de respostas corretas.

A segunda fase consistiu na comparação dos resultados obtidos com aqueles encontrados no estudo de Giroto et al. (2024). Os resultados foram analisados quantitativa e qualitativamente, a fim de avaliar avanços na precisão conceitual e no alinhamento das respostas com os padrões educacionais desejados.

Por fim, consultaram-se dois especialistas na área da Química, um doutor e um engenheiro químico com o objetivo de avaliar qualitativamente as respostas geradas pelos modelos generativos.

3. ANÁLISE E DISCUSSÃO DOS RESULTADOS

Na Tabela 1, é possível observar as respostas para os prompts utilizados por Giroto et al (2024). Na Tabela 2, o prompt base usado em Giroto et al (2024) é adaptado para aprimorar a resposta fornecida pelo modelo generativo.

Tabela 1: Prompt e resultados obtidos por Giroto et al (2024).

Prompt	Resposta
Defina o modelo atômico de Thomson	[...] O modelo atômico de Thomson , proposto por Joseph John Thomson em 1897, foi um dos primeiros modelos a tentar explicar a estrutura do átomo. Ele ficou conhecido como o modelo do pudim de passas ou modelo do pudim com passas , devido à analogia utilizada para representá-lo.
Defina o modelo atômico de Rutherford	[...] O modelo atômico de Rutherford , proposto em 1911 por Ernest Rutherford, foi um marco no estudo da estrutura atômica. Baseando-se nos resultados de seu famoso experimento da folha de ouro, ele apresentou um modelo que contradizia o modelo atômico anterior, o de Thomson.

Como esperado anteriormente e constatado por Girotto et al (2024). O modelo 4o, apresenta resultado por padrão o contexto histórico para o modelo atômico, mas respeita aspectos que historicamente incorrem em erros como a má interpretação do modelo atômico atômico de Thomson (Thomson, 1904; Hon, 2013).

Tabela 2: Prompt adaptado.

Prompt Adaptado
<pre>{ "Persona": "<professor de química, paciente e experiente>", "Instruções": ["<define o modelo atômico de Thomson>", "<lista as fontes e referência para sua explicação>", "<ênfatize aspectos matemáticos>" "<a sua resposta deve conter um máximo de 500 caracteres>"], "Output format": "<Olá meu caro aluno, eis uma explicação do modelo de Thomson:>", "Exemplos": "<No artigo ...; Na revista; Segundo a Fonte..>" } "Resposta":</pre>
<pre>{ "Persona": "<professor de química, paciente e experiente>", "Instruções": ["<define o modelo atômico de Rutherford>", "<lista as fontes e referência para sua explicação>" "<ênfatize aspectos matemáticos>" "<a sua resposta deve conter um máximo de 500 caracteres>"], "Output format": "<Olá meu caro aluno, eis uma explicação do modelo de Rutherford:>", "Exemplos": "<No artigo ...; Na revista; Segundo a Fonte..>" } "Resposta":</pre>

Utilizando o prompt adaptado na Tabela 2, obtém-se os resultados com os diversos modelos de LLM's empregadas no estudo

Tabela 3: Comparativo de respostas das LLM's.

	Nova Resposta
ChatGPT 4 (Grátis)	Olá meu caro aluno, eis uma explicação do modelo de Thomson: O modelo atômico de Thomson (1904) propôs que o átomo seria como uma "pudim de passas". Ele sugeriu que o átomo era uma esfera positiva, com elétrons de carga negativa distribuídos uniformemente. A carga positiva e negativa se equilibravam, tornando o átomo neutro. Matematicamente, esse equilíbrio é representado pela soma vetorial das cargas elétricas dentro do átomo, resultando em zero. Referência: Thomson, J. J., "On the Structure of the Atom," *Philosophical Magazine* (1904).

	Olá, meu caro aluno, eis uma explicação do modelo de Rutherford: O modelo de Rutherford propõe que os átomos possuem um núcleo pequeno e denso, carregado positivamente, onde a maior parte da massa está concentrada. Os elétrons orbitam ao redor desse núcleo em órbitas circulares. A descoberta foi baseada no experimento de dispersão de partículas alfa. Matemática envolvida: a equação de Coulomb descreve a força de atração entre o núcleo e os elétrons. Fontes: "The Structure of the Atom", Rutherford et al., 1911.
ChatGPT o3-mini (Paga)	Olá, meu caro aluno! Eis uma explicação do modelo de Thomson: O modelo atômico de Thomson, conhecido como "pudim de passas", descreve o átomo com elétrons distribuídos numa esfera carregada positivamente. A distribuição dos elétrons é uniforme, e a densidade de carga pode ser descrita pela fórmula $\rho = Q/4\pi R^3$. Fontes: Thomson (1897), "Philosophical Magazine"; Zumdahl (2013), "Química Geral"; Atkins e Jones (2010), "Chemical Principles".
	Olá meu caro aluno, eis uma explicação do modelo de Rutherford: O átomo possui um núcleo central com carga positiva e quase toda a massa, enquanto os elétrons orbitam em espaço vazio. A dispersão de partículas α é descrita por $I(\theta) \propto 1/\sin^4(\theta/2)$, derivada da lei de Coulomb, $F=kq_1q_2/r^2$. Fontes: Rutherford (1911, Philosophical Magazine); Zumdahl & Zumdahl; Atkins, "Química: A Ciência Central".
Claude 3.5 Sonnet (Grátis)	Olá meu caro aluno, eis uma explicação do modelo de Thomson: O modelo atômico de Thomson (1904) propõe que o átomo é uma esfera positiva uniforme de raio R, com elétrons de carga $-e$ embebidos em equilíbrio eletrostático. A força eletrostática é dada por $F=kq_1q_2/r^2$, onde k é a constante de Coulomb. Os elétrons se organizam em anéis concêntricos para minimizar a energia potencial. Fontes principais: • Thomson, J.J. "On the Structure of the Atom", Philosophical Magazine, 1904 • Bohr, N. "The Theory of Spectra and Atomic Constitution", 1922 • Pais, A. "Niels Bohr's Times", 1991 Este modelo explicou a natureza indivisível do átomo, mas falhou em explicar o espectro atômico.

	Olá meu caro aluno, eis uma explicação do modelo de Rutherford: O modelo atômico de Rutherford (1911) é fundamentado na força de Coulomb $F=kZe^2/r^2$, onde Z é o número atômico. Propõe um núcleo denso e positivo com elétrons orbitando ao seu redor. A estabilidade orbital é dada pela igualdade entre força centrípeta mv^2/r e força eletrostática. Fontes principais: • Rutherford, E. "The Scattering of α and β Particles by Matter", 1911 • Geiger, H. & Marsden, E. "On Diffuse Reflection of α-Particles", 1909 • Wilson, W. "The Quantum Theory of Radiation", 1915 Este modelo explicou o espalhamento de partículas alfa, mas não a estabilidade eletrônica.
Deepseek - R1 (Grátis)	<Olá meu caro aluno, eis uma explicação do modelo de Thomson:> O modelo atômico de Thomson, proposto em 1897, descreve o átomo como uma esfera de carga positiva com elétrons (carga negativa) incrustados, como "passas em um pudim". Matematicamente, a neutralidade do átomo é garantida pela igualdade entre o número de elétrons e a carga positiva distribuída. Esse modelo foi baseado em experimentos com raios catódicos e cálculos de deflexão de partículas. <No artigo "On the Structure of the Atom" (1904); Na revista Philosophical Magazine; Segundo a Fonte: J.J. Thomson, "Conduction of Electricity Through Gases" (1903).>
	<Olá meu caro aluno, eis uma explicação do modelo de Rutherford:> O modelo atômico de Rutherford, proposto em 1911, descreve o átomo como um núcleo denso e positivo, onde se concentra quase toda a massa, com elétrons orbitando ao redor, em um espaço vazio. Baseado no experimento da folha de ouro, Rutherford calculou a deflexão de partículas alfa usando a lei de Coulomb e a mecânica clássica, mostrando que a força elétrica é $F=kq_1q_2/r^2$ <No artigo "The Scattering of α and β Particles" (1911); Na revista Philosophical Magazine; Segundo a Fonte: E. Rutherford, "Radioactive Substances and Their Radiations" (1913).>

Fonte: Autores

He et al. (2024) demonstraram que a escolha da formatação tem impacto significativo na acurácia dos modelos de IA. Em modelos mais sofisticados, como o GPT-4, esse impacto é reduzido, mas em modelos menores, a seleção adequada do formato pode resultar em melhorias expressivas na qualidade das respostas. Para garantir maior eficácia nos testes conduzidos neste estudo, adotou-se o formato JSON, conforme sugerido por He et al. (2024), como padrão para estruturação das interações.

Ao comparar os resultados apresentados nas Tabelas 1 e 2, é possível identificar padrões relevantes nas respostas geradas. Primeiramente, observa-se que a estruturação do prompt influenciou diretamente o teor das respostas, garantindo que os aspectos matemáticos fossem consistentemente incluídos em ambas as respostas. Essa padronização evidencia a importância da engenharia de prompt na orientação do modelo para fornecer respostas mais alinhadas ao esperado.

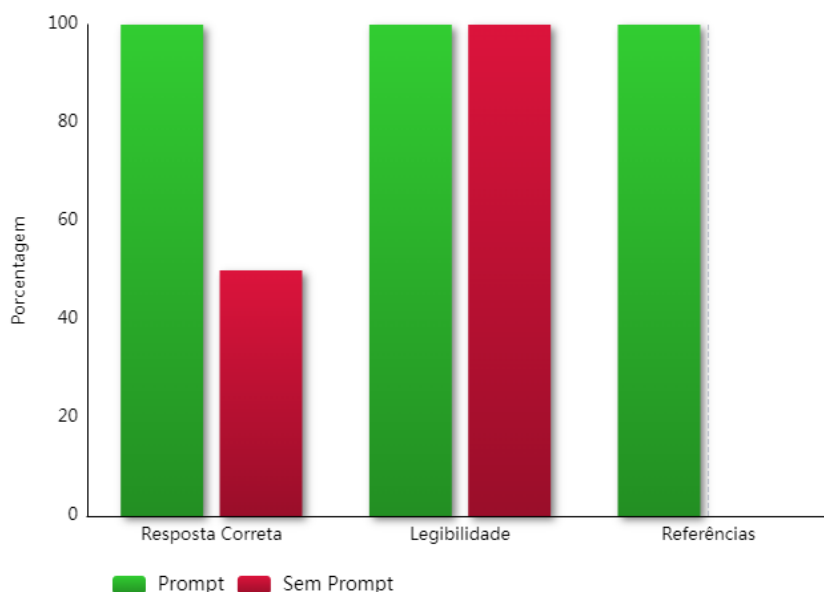
Outra vantagem identificada nessa abordagem foi o controle sobre a quantidade de caracteres na resposta, permitindo um maior refinamento das informações geradas. Adicionalmente, a inclusão explícita de referências nas respostas possibilitou não apenas a verificação da veracidade das informações pelo usuário, mas também contribuiu para reduzir significativamente o fenômeno das alucinações.

Os resultados obtidos também sugerem que, além da evolução exponencial das Large Language Models (LLMs), o uso de prompts bem estruturados desempenha um papel crucial na obtenção de respostas mais diretas, objetivas e precisas. Dessa forma, a adoção de técnicas avançadas de engenharia de prompt se mostra imprescindível para otimizar a interação com modelos de IA. Aliado a isso, a redução de alucinações se torna mais evidente quando as respostas incluem referências verificáveis, aumentando a confiabilidade dos conteúdos gerados pelo chatbot.

Este estudo avaliou o desempenho do modelo de IA da OpenAI no contexto da Química Educacional, tomando como referência os resultados obtidos por Giroto et al. (2024). A investigação demonstrou que avanços recentes na tecnologia de modelos generativos possibilitaram melhorias significativas na precisão das respostas, especialmente quando empregadas estratégias adequadas de engenharia de prompt. A estruturação das instruções revelou-se um fator determinante na orientação do conteúdo gerado, permitindo maior controle sobre a profundidade e relevância das informações apresentadas.

Conforme ilustrado na Figura 1, o uso da engenharia de prompt mostrou-se essencial para a modelagem das respostas, influenciando diretamente a precisão e a direção das informações geradas pelo modelo. A possibilidade de delimitar o escopo das respostas e minimizar a incidência de erros conceituais destaca o potencial dessa abordagem para o aprimoramento das interações com IA em contextos educacionais. O impacto positivo dessa metodologia reforça a relevância do estudo para a aplicação da IA no ensino de Química.

Figura 1: Comparativo entre Prompt e Sem Prompt



Fonte: Autores

Na etapa final, referente à consulta com especialistas, identificaram-se os seguintes pontos: observou-se o uso inadequado de analogias em livros didáticos, as quais podem induzir os alunos a erro. Não foram encontradas inconsistências matemáticas nos problemas analisados quando estas foram solicitadas; no entanto, em questões mais complexas, os especialistas identificaram falhas conceituais nos modelos utilizados. Um exemplo foi a atribuição de carga nula ao nêutron, o que denota um equívoco conceitual. Nenhum dos modelos avaliados foi capaz de responder corretamente à questão proposta, mesmo com o emprego de estratégias avançadas de engenharia de prompt.

Os resultados obtidos indicam que os avanços das IAGs (Inteligências Artificiais Generativas) têm o potencial de transformar a forma como o conhecimento é produzido e disseminado em sala de aula. O acesso gratuito a estas ferramentas amplia as possibilidades para educadores e alunos, favorecendo a inclusão digital e a democratização do aprendizado. Em comparação com metodologias tradicionais, os modelos generativos apresentaram superioridade na velocidade e eficiência na geração de materiais didáticos.

No entanto, mais pesquisas são necessárias para consolidar essas evidências e garantir a segurança e confiabilidade das respostas geradas. Este estudo preliminar representa um primeiro passo na direção do uso mais eficaz das IAGs para fins educacionais, mas é fundamental continuar investigando novas estratégias e refinamentos na interação com estes modelos. Considerando a evolução exponencial da IA, é possível que, no momento da publicação deste estudo, modelos ainda mais avançados tenham sido lançados, oferecendo resultados aprimorados e novas possibilidades para a educação.

4. CONCLUSÃO

Este trabalho consiste em uma reavaliação de Giroto et al. (2024), com o objetivo de verificar se as limitações anteriormente identificadas em modelos de Inteligência Artificial aplicados ao ensino de Química podem ser superadas por novas abordagens conversacionais ou se já foram superadas pelos avanços das próprias AGIs.

Nesta nova análise, foram utilizados outros modelos generativos, como o DeepSeek, além de novas formatações textuais e estratégias como a Chain of Thought (CoT), ambas provenientes do campo emergente da Engenharia de Prompt.

Observou-se que o uso de prompts tornou as respostas mais objetivas e reduziu a ocorrência de alucinações geradas pelos modelos. No entanto, entre as limitações identificadas, destaca-se a incapacidade de impedir que as AGIs reproduzam falácias e inconsistências perpetuadas por livros didáticos ao longo dos anos.

Conclui-se que essas ferramentas podem e devem ser incorporadas ao ambiente escolar. A aplicação dos conceitos da Engenharia de Prompt pode mitigar erros e inconsistências produzidas pelos modelos, mas não eliminá-los por completo. Por isso, é fundamental que o uso de AGIs seja acompanhado por um especialista capaz de orientar os alunos e sinalizar possíveis falhas, como alucinações.

REFERÊNCIAS

BROWN, T. et al. *Language models are few-shot learners*. Advances in Neural Information Processing Systems, v. 33, p. 1877-1901, 2020. Disponível em: <https://arxiv.org/abs/2005.14165>.

FERGUS, S.; BOTHA, M.; OSTOVAR, M. Evaluating Academic Answers Generated Using ChatGPT. *Journal of Chemical Education*, v. 100, n. 4, p. 1672-1675, 2023. Disponível em: <https://doi.org/10.1021/acs.jchemed.3c00087>.

GIROTO, G.; NASCIMENTO JUNIOR, W. J. D.; MORAIS, C. Avaliação de Tecnologias de Inteligência Artificial Generativa na Produção de Elementos Textuais e Imagéticos sobre Conceitos de Química. *Challenges 2024: Espaços e Caminhos OnLife*, p. 299-311, 2024. Disponível em: <https://www.researchgate.net/publication/381111282>.

GOODFELLOW, I. et al. *Generative Adversarial Nets*. In: Advances in Neural Information Processing Systems, 2014. Disponível em: <https://arxiv.org/abs/1406.2661>.

HOLMES, W.; LUCKIN, R. *Intelligence Unleashed: An Argument for AI in Education*. Pearson, 2016.

JOHNSTONE, A. H. The Development of Chemistry Teaching: A Changing Response to Changing Demand. *Journal of Chemical Education*, v. 70, n. 9, p. 701, 1993. Disponível em: <https://doi.org/10.1021/ed070p701>.

KONG, A.; et al. Better zero-shot reasoning with role-play prompting. *arXiv preprint*, 2024. Disponível em: <https://arxiv.org/abs/2308.07702>.

LEON, A. J.; VIDHANI, D. ChatGPT Needs a Chemistry Tutor Too. *Journal of Chemical Education*, 2023. DOI: <https://doi.org/10.1021/acs.jchemed.3c00288>.

PARNAMI, A.; LEE, M. Learning from few examples: A summary of approaches to few-shot learning. *arXiv preprint*, 2022. Disponível em: <https://arxiv.org/abs/2203.04291>.

PHOENIX, J.; TAYLOR, M. *Prompt engineering for generative AI: Future-proof inputs for reliable AI outputs at scale*. O'Reilly Media, 2024.

RADFORD, A. et al. *Language Models are Unsupervised Multitask Learners*. OpenAI Research, 2019. Disponível em: <https://openai.com/research/gpt-2>.

TALANQUER, V. Interview with the Chatbot: How Does It Reason? *Journal of Chemical Education*, v. 100, n. 8, p. 2821–2824, 2023. Disponível em: <https://doi.org/10.1021/acs.jchemed.3c00472>.

WHITE, J. et al. Prompt Engineering Strategies for Generative AI. *arXiv preprint*, 2023. Disponível em: <https://arxiv.org/abs/2303.12712>.

WENG, L. *Prompt Engineering*. Lil'Log, 2023. Disponível em: <https://lilianweng.github.io/posts/2023-03-15-prompt-engineering/>.

ZHAO, Z. et al. *Calibrate before use: Improving few-shot performance of language models*. International Conference on Machine Learning. PMLR, 2021. Disponível em: <https://arxiv.org/abs/2102.09690>.

EVOLUTION OF GENERATIVE AI IN EDUCATION: A NEW PERSPECTIVE ON KNOWLEDGE PRODUCTION IN CHEMISTRY

Abstract: *This study analyzes the impact of prompt engineering on the quality of responses generated by LLMs in the context of Chemistry education. Based on a review of previous works, especially Giroto et al. (2024), the study proposes an experimental approach divided into three stages: reproduction of the prompts previously used, reformulation to optimize model performance, and comparison of the results. The findings indicate that proper prompt structuring significantly improves response accuracy, reduces hallucinations, and ensures stricter alignment with chemical concepts. The study also highlights the importance of free access to these technologies and their potential as educational tools. Although the results are promising, the research suggests that further testing is needed to consolidate the evidence and explore new AI models in educational settings.*

Keywords: *prompt engineering, generative ai, artificial intelligence in education, chemistry education*

