



ANACRONISMOS HISTÓRICOS NA QUÍMICA GERADOS POR CHATGPT: UM MAPEAMENTO AUTOMATIZADO

Diogo Gonzaga Monte da Costa¹

¹ PPGECEM – CCET – UFRN, diogo-monte@hotmail.com

Resumo

A crescente utilização de ferramentas de Inteligência Artificial generativa, como o ChatGPT, tem estimulado novas práticas no ensino de Ciências. No entanto, modelos de linguagem natural carecem de filtros temporais e operam sem consciência cronológica, o que pode resultar na produção de explicações anacrônicas, especialmente em contextos historicamente situados como a História da Química. Este trabalho teve como objetivo mapear anacronismos conceituais em respostas geradas pelo ChatGPT-4o a partir de um conjunto de vinte perguntas sobre eventos históricos da Química, abrangendo desde práticas alquímicas até modelos atômicos modernos. Para análise, foi elaborada uma lista de termos anacrônicos organizada por período histórico. As respostas da IA foram processadas por um script em Python capaz de detectar ocorrências desses termos. Em uma amostra de 10 perguntas, observou-se que 60% das respostas continham pelo menos um anacronismo. Os erros mais frequentes incluíram uso de termos como “átomo”, “oxigênio” e “mecânica quântica” em contextos anteriores às suas formulações científicas. Os resultados reforçam a importância da mediação crítica no uso de IA no ensino, principalmente em disciplinas com forte componente epistemológico e histórico. A metodologia proposta é replicável e pode apoiar educadores interessados em explorar IA generativa de forma contextualizada e segura.

Palavras-chave: história da química; inteligência artificial; ensino de ciências; anacronismos; modelos de linguagem; ChatGPT

Abstract

The growing use of generative Artificial Intelligence tools such as ChatGPT has introduced new possibilities in Science Education. However, large language models lack temporal filters and operate without chronological awareness, which can result in anachronistic explanations—especially when applied to historically situated content such as the History of Chemistry. This study aimed to map conceptual anachronisms in responses generated by ChatGPT-4o based on a set of twenty questions concerning major milestones in the development of Chemistry, from alchemical practices to modern atomic models. A list of anachronistic terms was constructed and classified by historical period. The AI-generated responses were analyzed using a Python script designed to detect the presence of those terms. In a sample of 10 questions, 60% of responses included at least one anachronism. Common errors involved the use of terms such as “atom,” “oxygen,” and “quantum mechanics” in pre-modern contexts. These findings

highlight the importance of critical mediation when employing generative AI in educational settings—especially in disciplines with a strong historical or epistemological component. The proposed methodology is reproducible and may assist educators interested in using generative AI tools in a historically grounded and pedagogically safe manner.

Keywords: *history of chemistry; artificial intelligence; science education; anachronism; language models; ChatGPT*

Introdução

A popularização de ferramentas baseadas em modelos de linguagem, como o *ChatGPT*, tem impulsionado o uso de Inteligência Artificial (IA) no ensino de Ciências. Esses sistemas são capazes de gerar respostas articuladas, coerentes e muitas vezes instrutivas — porém, não estão isentos de riscos epistemológicos. Um dos principais desafios refere-se à incapacidade desses modelos de operar com consciência cronológica, o que pode levar à mistura de conceitos de diferentes épocas históricas e à geração de explicações anacrônicas (Underwood *et al.*, 2025).

O uso dessas ferramentas no contexto da História da Ciência, e em especial da História da Química — marcada por mudanças conceituais profundas, controvérsias e teorias superadas —, exige atenção redobrada. Como argumenta Evangelos Pournaras (2023), modelos de IA como o *ChatGPT* podem comprometer a integridade científica de conteúdos educacionais ao produzirem alucinações e interpretações distorcidas da história do conhecimento científico, fenômeno que ele classifica como "*epistemic anachronism*".

Esse cenário é agravado pela baixa transparência e falta de replicabilidade nos processos de geração de texto por *LLMs*, como demonstrado na revisão sistemática de Yan *et al.* (2024), que identificou desafios práticos e éticos no uso de modelos de linguagem na educação. Nesse sentido, Cheung *et al.* (2024) defendem a promoção da alfabetização epistêmica como competência essencial no uso pedagógico de IA, ou seja, a habilidade crítica de compreender não apenas o conteúdo, mas os limites do conhecimento científico e tecnológico embutido nas respostas da IA.

Diante desse panorama, este trabalho teve como objetivo mapear anacronismos históricos nas respostas do *ChatGPT-4o* a perguntas sobre momentos-chave da história da Química. Por meio de uma abordagem automatizada, foram identificadas ocorrências de termos e conceitos indevidos em contextos históricos específicos, simulando uma análise que pode ser útil a professores e pesquisadores interessados em utilizar IA com criticidade no ensino.

Metodologia

Corpus de Perguntas

Foram elaboradas 20 perguntas envolvendo eventos históricos da Química, distribuídas entre os séculos XIII e XX. As perguntas abordavam desde práticas

alquímicas até o desenvolvimento de modelos atômicos e a descoberta da espectroscopia. Exemplos de perguntas:

- “Qual era a explicação do flogisto para a combustão?”
- “Como Mendeleev organizou a tabela periódica?”
- “Qual foi o papel de Bohr na estrutura do átomo?”
-

Respostas com IA

As respostas foram geradas utilizando o modelo *ChatGPT-4o* (versão junho/2024). As respostas foram armazenadas em planilha .csv, com uma linha por pergunta, contendo a resposta completa.

Lista de termos anacrônicos

Foi construída uma lista de termos anacrônicos com base em três blocos temporais:

- Antes de 1800: átomo, elétron, orbital, quântico, oxigênio, etc.
- 1800–1899: spin, teoria de orbitais moleculares, fissão, espectrometria de massa.
- 1900–1912: modelo quântico, salto quântico, níveis eletrônicos.

A lista foi associada a cada pergunta com base em seu período histórico, formando o conjunto de palavras proibidas para cada resposta.

Análise automatizada

Foi desenvolvido um *script* em *Python* que processou cada resposta e identificou a presença de termos anacrônicos. Para cada pergunta, foi contabilizado o número de ocorrências desses termos. O código também gerou um gráfico com a distribuição dos erros e percentuais de acerto/falha.

Resultados

A análise foi realizada em uma amostra de 10 das 20 perguntas inicialmente preparadas. A tabela abaixo mostra a distribuição dos anacronismos detectados:

Tabela 1 – Quantidade de anacronismos por pergunta

ID	Tema da Pergunta	Nº Acronismos
1	Transmutação dos metais	1
2	Pedra Filosofal	1
3	Teoria do flogista	2
4	Refutação do flogista	0
5	Teoria atômica de Dalton	0

6	Tabela Periódica de Mendeleev	0
9	Calor antes da Termodinâmica	1
11	Galvani e Volta	1
14	Modelo Atômico de Rutherford	1
15	Modelo Atômico de Bohr	1

Resumo estatístico:

- Total de perguntas analisadas: 10
- Respostas com anacronismos: 6
- Percentual de respostas com erro: 60%

Discussão

A análise revela que, embora o modelo *ChatGPT-4o* apresente respostas geralmente bem estruturadas, ele não possui consciência histórica. Termos como “estrutura atômica”, “oxigênio” e “mecânica quântica” foram usados em contextos anteriores à sua formulação científica real.

As respostas sobre alquimia, teoria do flogisto e modelos atômicos iniciais foram as mais propensas a conter anacronismos. Já perguntas sobre eventos pós-1850 apresentaram maior acurácia.

Esse comportamento reforça a necessidade de mediação docente ao utilizar respostas de IA generativa no ensino da História da Ciência. A clareza textual da resposta não garante sua adequação cronológica ou epistemológica.

Considerações Finais

Este trabalho demonstrou que ferramentas como o *ChatGPT-4o* ainda carecem de filtros temporais que evitem o uso de conceitos anacrônicos em contextos históricos. Apesar de serem úteis como apoio à explicação e criação de materiais, seu uso na História da Química — sem supervisão — pode reforçar visões equivocadas sobre o desenvolvimento da ciência.

A metodologia aqui proposta pode ser facilmente reproduzida por docentes e pesquisadores, oferecendo uma forma automatizada de avaliar a confiabilidade de conteúdos gerados por IA. No futuro, espera-se que modelos de linguagem possam incorporar parâmetros históricos como restrições de geração.

Referências

Cheung, K. K. C., Long, Y., Liu, Q., & Chan, H. Y. (2024). Unpacking epistemic insights of artificial intelligence (AI) in science education: A systematic review. *Science & Education*, 1-31.

OpenAI. (2024). *GPT-4o Technical Report*. <https://openai.com/research/gpt-4o>

Pournaras, E. (2023). Science in the era of ChatGPT, large language models and generative AI. *KI-Kritik/AI Critique Volume 6*, 275.

Python Software Foundation. (2023). *Python Language Reference*, version 3.11. <https://www.python.org>

Underwood, T., Nelson, L. K., & Wilkens, M. (2025). Can Language Models Represent the Past without Anachronism?. *arXiv preprint arXiv:2505.00030*.

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., ... & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55(1), 90-112.
