

Comparação de Chatbots na Formulação de Questões-Problema sobre Geração de Energia

**Bruno Henrique de Pádua Silva*. Jader Fernando Dias Breda*.
Thais de Fátima Araújo Silva*. Hélio Oliveira Ferrari***

**UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI – Instituto de Engenharia, Ciência e Tecnologia – Janaúba/MG, Brasil
e-mail: {henrique.bruno ; jader.breda; thais.araujo; helio.ferrari} @ufvjm.edu.br*

Resumo: Este estudo compara ChatGPT, Gemini e Perplexity AI na formulação de questões-problema para Trabalhos de Conclusão de Curso em Engenharia, considerando quatro tecnologias de geração de energia. As questões foram classificadas por análise de conteúdo. Os resultados indicam diferenças nas abordagens temáticas das ferramentas e evidenciam seu potencial como apoio à problematização científica, desde que utilizadas com mediação docente.

Palavras-chave: Inteligência Artificial Generativa; Educação em Engenharia; Tecnologias Educacionais; Problematização Científica; Sistemas de Geração de Energia

1. INTRODUÇÃO

A formulação de uma questão-problema constitui uma das etapas mais relevantes do processo de investigação científica, uma vez que delimita o objeto de estudo, orienta a definição dos objetivos, fundamenta a escolha dos procedimentos metodológicos e direciona a análise dos resultados (GIL, 2008; SEVERINO, 2017). No contexto dos Trabalhos de Conclusão de Curso (TCC) em cursos de Engenharia, essa etapa representa um desafio recorrente para estudantes em processo de iniciação científica, especialmente pela necessidade de transformar interesses amplos em problemas de pesquisa específicos, relevantes, exequíveis e cientificamente fundamentados. Estudos sobre metodologia científica e educação em engenharia indicam que dificuldades relacionadas à delimitação do problema de pesquisa estão frequentemente associadas à reduzida experiência investigativa, ao domínio limitado da literatura especializada e à pouca familiaridade com estratégias de problematização científica (FELDER; BRENT, 2016; GIL, 2008; SEVERINO, 2017).

Nos últimos anos, o rápido desenvolvimento dos Grandes Modelos de Linguagem (Large Language Models – LLMs) impulsionou o surgimento de ferramentas de Inteligência Artificial (IA) generativa capazes de produzir textos, sintetizar informações, responder perguntas complexas e apoiar diversas atividades acadêmicas. Sistemas como ChatGPT, Gemini e Perplexity AI passaram a ser incorporados às práticas educacionais por estudantes e docentes, oferecendo suporte à elaboração de textos, revisão bibliográfica, organização de ideias e construção inicial de propostas de pesquisa (KASNECI et al., 2023; LO et al., 2023; ZHAI, 2023). Em função de sua elevada capacidade de processamento da linguagem

natural, essas ferramentas apresentam potencial para estimular processos de ideação e auxiliar estudantes na organização inicial do pensamento científico.

Entretanto, a literatura recente demonstra que o uso de LLMs na educação deve ser compreendido tanto sob a perspectiva de suas potencialidades quanto de suas limitações. Entre os benefícios relatados destacam-se o aumento da produtividade acadêmica, o apoio à aprendizagem personalizada, a ampliação do acesso à informação, a geração de múltiplas perspectivas para um mesmo problema e o suporte ao desenvolvimento inicial de textos científicos (BAEK et al., 2023; KASNECI et al., 2023; LO, 2023). Por outro lado, diversos estudos alertam para limitações inerentes a esses modelos, como a produção de informações factualmente incorretas (hallucinations), vieses presentes nos dados de treinamento, ausência de mecanismos próprios de validação científica, inconsistências entre diferentes interações e dificuldade em distinguir evidências consolidadas de conteúdos especulativos (BENDER et al., 2021; BOMMASANI et al., 2021; OPENAI, 2023).

Essas limitações tornam particularmente relevante a discussão sobre o uso responsável da IA generativa no ensino superior. Pesquisas recentes indicam que a utilização indiscriminada dessas ferramentas pode favorecer comportamentos de dependência tecnológica, reduzir oportunidades de desenvolvimento do pensamento crítico, comprometer processos de autoria intelectual e estimular a aceitação acrítica de respostas produzidas automaticamente (KASNECI et al., 2023; UNESCO, 2023). Dessa forma, organismos internacionais e pesquisadores da área defendem que os sistemas de IA sejam empregados como instrumentos de apoio ao processo de aprendizagem, e não como substitutos da reflexão,



da investigação científica e da mediação pedagógica desenvolvida pelo professor (UNESCO, 2023; HOLMES et al., 2022).

Sob essa perspectiva, o papel do docente torna-se ainda mais relevante. A mediação do professor é fundamental para orientar a construção de problemas de pesquisa consistentes, avaliar criticamente as sugestões produzidas pelas ferramentas de IA, contextualizar conceitos científicos e estimular a autonomia intelectual dos estudantes. Em vez de substituir a orientação acadêmica, os modelos de linguagem podem atuar como recursos auxiliares para ampliar possibilidades de investigação, desde que suas respostas sejam analisadas criticamente, confrontadas com a literatura especializada e validadas por especialistas da área (KASNECI et al., 2023; HOLMES et al., 2022; UNESCO, 2023).

Apesar do crescimento expressivo das pesquisas envolvendo IA generativa na educação, observa-se que a maior parte dos estudos se concentra na avaliação da qualidade textual das respostas, no desempenho em tarefas de escrita, na resolução de exercícios ou no apoio ao ensino em diferentes disciplinas (LO, 2023; ZHAI, 2023). Permanecem relativamente escassas investigações que comparem sistematicamente diferentes LLMs quanto à sua capacidade de auxiliar especificamente a formulação de questões-problema para projetos de pesquisa, particularmente em áreas técnicas como a Engenharia Elétrica e, mais especificamente, no domínio da geração de energia. Essa lacuna dificulta a compreensão das características, tendências temáticas e limitações apresentadas por diferentes ferramentas quando utilizadas em atividades relacionadas à iniciação científica.

Diante desse contexto, este estudo tem como objetivo comparar as questões-problema produzidas por três ferramentas de Inteligência Artificial generativa — ChatGPT, Gemini e Perplexity AI — a partir de um mesmo conjunto de instruções (prompt), considerando quatro temas relacionados à geração de energia elétrica: geração fotovoltaica, geração eólica, geração termoelétrica a partir de biomassa e geração hidrelétrica. A comparação foi conduzida mediante análise qualitativa das questões geradas, organizadas em categorias temáticas previamente definidas, buscando identificar tendências de abordagem, complementaridades e limitações de cada ferramenta no apoio à elaboração de problemas de pesquisa para Trabalhos de Conclusão de Curso. Os resultados pretendem contribuir para uma utilização mais criteriosa da IA generativa no contexto da educação em Engenharia, fornecendo subsídios para docentes e estudantes compreenderem tanto o potencial quanto as limitações desses sistemas como instrumentos de apoio à formação científica.

2. MATERIAL E MÉTODOS

Esta pesquisa caracteriza-se como um estudo exploratório, descritivo e de abordagem mista (qualitativa e quantitativa), cujo objetivo consiste em comparar as questões-problema produzidas por diferentes ferramentas de Inteligência Artificial generativa a partir de um mesmo conjunto de instruções (prompt). A abordagem qualitativa foi empregada para identificar as características temáticas das questões elaboradas pelos modelos de linguagem, enquanto a abordagem quantitativa foi utilizada para contabilizar a frequência das categorias identificadas, permitindo comparar tendências de resposta entre as ferramentas analisadas.

O delineamento metodológico foi estruturado em quatro etapas principais: (i) seleção das ferramentas de IA; (ii) coleta padronizada das respostas; (iii) categorização temática das questões geradas; e (iv) análise comparativa dos resultados obtidos.

2.1 Critérios de Seleção das Ferramentas de IA

A seleção das ferramentas analisadas buscou contemplar plataformas representativas do estado da arte em modelos conversacionais baseados em Grandes Modelos de Linguagem (Large Language Models – LLMs), amplamente utilizadas no contexto acadêmico durante o ano de 2024. Foram considerados os seguintes critérios de inclusão:

- disponibilidade pública para usuários civis;
- ampla utilização por estudantes e pesquisadores;
- capacidade de compreender e produzir textos em língua portuguesa;
- suporte à geração de textos longos e elaboração de questões;
- utilização de arquitetura baseada em modelos de linguagem de grande porte;
- aplicação reconhecida em atividades de apoio à escrita acadêmica.

Com base nesses critérios foram selecionadas três ferramentas:

ChatGPT (OpenAI), escolhido por sua ampla utilização em atividades acadêmicas e por representar um dos sistemas conversacionais mais difundidos durante o período da pesquisa;

Gemini (Google), selecionado por integrar os modelos generativos desenvolvidos pelo Google e oferecer



recursos de consulta e geração textual amplamente utilizados em ambientes educacionais;

Perplexity AI, escolhido por combinar geração textual baseada em LLMs com mecanismos de recuperação de informações e apresentação explícita das referências utilizadas em suas respostas.

A escolha dessas três ferramentas não pretende estabelecer um ranking de desempenho entre todos os sistemas disponíveis no mercado, mas sim comparar três plataformas amplamente acessíveis aos estudantes brasileiros durante o período de realização da pesquisa.

2.2 Coleta das Respostas

A coleta das respostas foi realizada durante o mês de novembro de 2024, utilizando as interfaces públicas disponibilizadas por cada plataforma. Todas as interações foram conduzidas em língua portuguesa, buscando reproduzir o contexto de utilização típico de estudantes brasileiros de Engenharia. Com o objetivo de reduzir interferências metodológicas decorrentes da formulação das instruções, foi empregado exatamente o mesmo prompt em todas as ferramentas avaliadas. O prompt utilizado foi:

"Simule as respostas de um aluno que está escrevendo um projeto de conclusão de curso e está em dificuldades para encontrar questões-problema relacionadas ao tema [tipo de geração de energia]. Proponha pelo menos cinco questões-problema."

O termo entre colchetes foi substituído sucessivamente pelos seguintes temas: geração fotovoltaica; geração eólica; geração termoeletrônica a partir de biomassa; geração hidrelétrica.

Para cada tema foi realizada uma interação independente em cada ferramenta, totalizando doze consultas. Todas as respostas foram registradas integralmente imediatamente após sua geração, preservando-se a redação original produzida pelos respectivos sistemas.

Limitações da coleta

Algumas informações técnicas normalmente utilizadas em experimentos envolvendo LLMs não puderam ser controladas ou recuperadas durante a execução da pesquisa. Entre elas destacam-se:

versão exata do modelo de linguagem disponibilizado por cada plataforma no momento da consulta;

parâmetros internos de inferência (temperatura, top-p e demais parâmetros de amostragem);

limite efetivo de tokens empregado por cada sistema;

arquitetura computacional utilizada pelos fornecedores.

Essas informações não eram disponibilizadas de forma transparente pelas interfaces públicas utilizadas durante a coleta e, portanto, constituem limitações inerentes ao experimento. Da mesma forma, foi realizada apenas uma coleta para cada combinação ferramenta-tema, não sendo efetuadas múltiplas execuções para análise da variabilidade estocástica das respostas. Consequentemente, os resultados representam uma fotografia do comportamento observado durante o período de coleta e não uma caracterização estatística do desempenho dos modelos.

2.3 Critérios de Análise das Respostas

A análise das questões geradas foi conduzida mediante Análise de Conteúdo, conforme os procedimentos propostos por Bardin (2016), utilizando categorias temáticas definidas previamente a partir dos principais eixos recorrentes na literatura sobre geração de energia e pesquisa em Engenharia. Cada questão produzida pelas ferramentas foi considerada uma unidade individual de análise.

Após a leitura das respostas, cada questão foi classificada em apenas uma categoria temática, correspondente ao aspecto predominante abordado pela pergunta. Foram utilizadas seis categorias analíticas:

Impactos Ambientais – questões relacionadas aos efeitos ambientais, sustentabilidade, emissões, conservação ambiental e impactos ecológicos.

Aspectos Econômicos – questões relacionadas a custos, viabilidade financeira, investimentos, retorno econômico e mercado.

Desafios Tecnológicos – questões envolvendo eficiência, inovação tecnológica, manutenção, desenvolvimento de equipamentos e limitações técnicas.

Impactos Sociais – questões referentes aos efeitos sobre comunidades, geração de empregos, aceitação pública e desenvolvimento regional.



Regulamentações e Políticas – questões relacionadas à legislação, políticas públicas, incentivos governamentais e regulamentações do setor energético.

Integração na Matriz Energética – questões envolvendo complementaridade entre fontes de energia, estabilidade do sistema elétrico e diversificação da matriz energética.

Optou-se pela utilização de categorias mutuamente exclusivas para evitar dupla contagem de uma mesma questão. Nos casos em que uma pergunta contemplava mais de um tema, foi considerada a categoria correspondente ao assunto predominante.

A utilização dessas categorias teve como finalidade identificar tendências temáticas presentes nas respostas produzidas pelos diferentes modelos, não constituindo uma avaliação da qualidade científica ou da correção técnica das perguntas elaboradas.

2.4 Procedimentos para a Comparação

Após a categorização, todas as questões foram organizadas por ferramenta e por tema de geração de energia, originando os Quadros apresentados na Seção 3.

Em seguida, foi realizada uma contagem simples da frequência de ocorrência de cada categoria para cada chatbot, permitindo construir a Tabela 1.

A comparação quantitativa teve caráter exclusivamente descritivo, utilizando frequências absolutas das categorias identificadas.

Não foram empregados testes estatísticos inferenciais, uma vez que o objetivo da pesquisa consiste em identificar tendências temáticas das respostas produzidas pelos modelos de linguagem e não estabelecer diferenças estatisticamente significativas entre eles.

Complementarmente, realizou-se uma análise qualitativa interpretativa das questões geradas, buscando identificar características recorrentes, padrões de abordagem, complementaridades e possíveis lacunas entre as ferramentas analisadas.

Limitações metodológicas

Algumas limitações devem ser consideradas na interpretação dos resultados.

A pesquisa utilizou apenas um prompt padronizado, elaborado para simular uma situação típica enfrentada por estudantes durante a elaboração de Trabalhos de Conclusão de Curso. Diferentes formulações de prompt poderiam produzir respostas distintas.

Além disso, não foram investigados efeitos relacionados ao perfil do usuário (por exemplo, estudante iniciante, pesquisador ou especialista), tampouco diferentes estratégias de engenharia de prompts.

Outro aspecto importante refere-se à natureza dinâmica dos modelos de linguagem. Como os fornecedores atualizam continuamente seus sistemas, respostas obtidas em períodos posteriores podem diferir das apresentadas neste estudo, mesmo utilizando instruções idênticas.

Dessa forma, os resultados devem ser interpretados como representativos das ferramentas disponíveis durante o período de coleta, constituindo uma análise comparativa contextualizada ao cenário tecnológico de 2024.

3. RESULTADOS

Nesta seção, apresentamos as questões geradas por ChatGPT, Gemini e Perplexity AI sobre quatro áreas de geração de energia. As respostas são organizadas por tema, permitindo uma análise comparativa da qualidade, relevância e aplicabilidade das propostas.

Quadro 01 – Resultados obtidos por tipo de Geração

Geração Fotovoltaica

ChatGPT	Gemini	Perplexity AI
1. Quais os principais impactos ambientais da geração fotovoltaica?	1. Qual a viabilidade econômica da energia fotovoltaica em regiões urbanas densas?	1. Qual a eficiência média dos sistemas fotovoltaicos disponíveis no mercado?
2. Como o custo de implementação varia entre diferentes regiões?	2. Como a energia solar contribui para a redução de emissões de gases de efeito estufa?	2. Quais os principais custos associados à manutenção de usinas solares?
3. Quais os desafios tecnológicos	3. Quais os principais desafios para a	3. Como as regulamentações locais impactam



ChatGPT	Gemini	Perplexity AI
para aumentar a eficiência dos painéis solares?	popularização da energia fotovoltaica em países em desenvolvimento?	a instalação de sistemas fotovoltaicos?
4. Como políticas públicas podem incentivar a adoção da energia fotovoltaica no Brasil?	4. Qual o impacto das condições climáticas na eficiência dos sistemas solares?	4. Quais tecnologias recentes aumentam a durabilidade dos painéis solares?
5. Como a energia solar pode ser integrada à matriz energética brasileira sem comprometer sua estabilidade?	5. Como inovações em armazenamento podem impulsionar a adoção de sistemas fotovoltaicos?	5. Qual o impacto socioeconômico de pequenos sistemas fotovoltaicos em comunidades rurais?

Geração Eólica

ChatGPT	Gemini	Perplexity AI
1. Quais os principais impactos ambientais da construção de parques eólicos?	1. Como a energia eólica contribui para a diversificação da matriz energética brasileira?	1. Quais os custos médios para instalação de turbinas eólicas em diferentes regiões do Brasil?
2. Como a localização afeta a eficiência de turbinas eólicas?	2. Quais os impactos sociais da instalação de parques eólicos em áreas rurais?	2. Qual o impacto da energia eólica na redução da dependência de combustíveis fósseis?
3. Quais os desafios tecnológicos para reduzir os custos de manutenção?	3. Como condições climáticas extremas afetam a operação de turbinas eólicas?	3. Como a reutilização de materiais das turbinas contribui para a sustentabilidade do setor?
4. Como políticas públicas podem incentivar a expansão da	4. Quais as oportunidades de crescimento para a energia eólica	4. Quais os principais desafios enfrentados por projetos eólicos offshore no Brasil?

ChatGPT	Gemini	Perplexity AI
energia eólica no Brasil?	em mercados emergentes?	
5. Quais os desafios logísticos na instalação de parques eólicos offshore?	5. Como a tecnologia de turbinas verticais pode mudar o setor eólico?	5. Quais os avanços tecnológicos mais recentes na eficiência das turbinas eólicas?

Geração Termoeétrica em Biomassa

ChatGPT	Gemini	Perplexity AI
1. Quais os principais resíduos utilizados na geração termoeétrica a partir de biomassa?	1. Qual o potencial de expansão da biomassa na matriz energética brasileira?	1. Quais os custos operacionais médios de uma planta termoeétrica movida a biomassa?
2. Qual o impacto ambiental da queima de biomassa vs. combustíveis fósseis?	2. Quais os impactos sociais da utilização de resíduos agrícolas na geração termoeétrica?	2. Qual o papel da biomassa no cumprimento das metas de redução de emissões de carbono?
3. Quais os desafios econômicos para ampliar o uso de biomassa no Brasil?	3. Como o transporte de biomassa afeta sua viabilidade econômica?	3. Quais resíduos apresentam maior eficiência energética na geração termoeétrica?
4. Como a biomassa pode contribuir para a sustentabilidade energética em regiões rurais?	4. Quais os avanços tecnológicos que tornam a biomassa mais eficiente?	4. Como as condições climáticas afetam a disponibilidade de biomassa?
5. Quais as limitações técnicas dos sistemas de conversão termoeétrica baseados em biomassa?	5. Como as políticas ambientais regulam a expansão da biomassa no setor termoeétrico?	5. Quais os desafios logísticos no armazenamento de biomassa em larga escala?



Geração Hidrelétrica

ChatGPT	Gemini	Perplexity AI
1. Quais os impactos ambientais de grandes hidrelétricas em ecossistemas locais?	1. Qual o papel da energia hidrelétrica na matriz energética de países tropicais?	1. Quais os custos médios de construção de pequenas e grandes usinas hidrelétricas?
2. Como pequenas centrais hidrelétricas podem contribuir para a sustentabilidade energética?	2. Como os impactos ambientais podem ser mitigados na construção de hidrelétricas?	2. Como as hidrelétricas contribuem para o equilíbrio entre oferta e demanda de energia no Brasil?
3. Quais os desafios de manutenção das infraestruturas hidrelétricas?	3. Quais os desafios na operação de usinas hidrelétricas em períodos de seca extrema?	3. Quais os impactos de longo prazo das barragens no fluxo dos rios?
4. Como a variação climática afeta a capacidade de geração das hidrelétricas?	4. Como as hidrelétricas podem coexistir com outras fontes de energia renovável?	4. Quais tecnologias recentes aumentam a eficiência das turbinas hidrelétricas?
5. Quais os impactos sociais do reassentamento de comunidades em áreas de barragens?	5. Quais as vantagens das hidrelétricas reversíveis em relação às convencionais?	5. Qual o potencial de reaproveitamento de barragens existentes para geração de energia?

4. DISCUSSÃO

A categorização das questões-problema geradas pelos três modelos de linguagem permitiu identificar diferenças nas tendências temáticas das respostas produzidas por ChatGPT, Gemini e Perplexity AI. Embora todas as ferramentas tenham sido capazes de elaborar questões relacionadas aos quatro temas investigados — geração fotovoltaica, geração eólica, geração termoelétrica a partir de biomassa e geração hidrelétrica — observam-se diferenças quanto à abrangência temática, ao foco das perguntas e ao equilíbrio entre aspectos técnicos, econômicos, sociais e ambientais.

A distribuição das categorias apresentada na Tabela 1 evidencia que o ChatGPT apresentou a maior diversidade temática entre os três modelos, distribuindo suas questões de forma relativamente equilibrada entre impactos ambientais, desafios tecnológicos, aspectos econômicos, integração da matriz energética e regulamentações. Essa característica sugere maior capacidade de explorar diferentes dimensões de um mesmo problema de pesquisa, favorecendo a geração de ideias iniciais para estudantes que ainda se encontram na fase de delimitação do tema de investigação.

O Gemini apresentou comportamento distinto, concentrando maior número de questões relacionadas aos impactos sociais e aos aspectos econômicos da geração de energia. Em diversas situações, suas perguntas enfatizaram políticas públicas, desenvolvimento regional, condições socioeconômicas e efeitos das tecnologias energéticas sobre comunidades. Essa tendência aproxima-se da literatura que descreve a crescente incorporação de dimensões interdisciplinares em modelos conversacionais voltados para aplicações educacionais, particularmente quando treinados para integrar informações provenientes de diferentes domínios do conhecimento (Holmes et al., 2022; Lo, 2023).

Por sua vez, o Perplexity AI demonstrou maior concentração de questões relacionadas aos aspectos econômicos e tecnológicos. Em comparação com os demais modelos, observou-se maior frequência de perguntas envolvendo eficiência, custos operacionais, tecnologias emergentes e indicadores quantitativos. Esse comportamento é coerente com a arquitetura da plataforma, que combina modelos de linguagem com mecanismos de recuperação de informação (Retrieval-Augmented Generation – RAG), permitindo incorporar conteúdos atualizados e referências durante a construção das respostas (Lewis et al., 2020). Entretanto, verificou-se menor incidência de questões relacionadas aos impactos sociais e à integração das fontes renováveis na matriz energética, indicando menor diversidade temática nas sugestões produzidas. É importante destacar que as diferenças observadas entre os modelos não devem ser interpretadas como evidência de superioridade de uma ferramenta em relação às demais. Os resultados indicam, sobretudo, que cada sistema apresenta características próprias decorrentes de sua arquitetura computacional, de seus dados de treinamento e das estratégias utilizadas para geração das respostas. Estudos recentes ressaltam que modelos de linguagem podem produzir respostas distintas mesmo quando submetidos ao mesmo prompt, em razão da natureza probabilística do processo de geração textual e das constantes atualizações implementadas pelos seus



desenvolvedores (Bommasani et al., 2021; Kasneci et al., 2023). Dessa forma, a diversidade observada entre ChatGPT, Gemini e Perplexity AI representa uma característica inerente aos sistemas baseados em LLMs e reforça a importância de análises comparativas como a realizada neste trabalho. Do ponto de vista educacional, os resultados obtidos sugerem que as ferramentas avaliadas apresentam potencial para apoiar estudantes durante a etapa inicial de construção do problema de pesquisa. A diversidade de perguntas produzidas pode contribuir para ampliar perspectivas de investigação, estimular a identificação de diferentes abordagens para um mesmo tema e favorecer processos de reflexão inicial sobre objetos de estudo. Pesquisas recentes apontam que modelos de linguagem podem atuar como instrumentos de apoio ao desenvolvimento de ideias, organização de informações e elaboração preliminar de projetos acadêmicos, desde que utilizados de forma crítica e integrada ao processo de ensino-aprendizagem (Kasneci et al., 2023; Lo, 2023; Zhai, 2023). Entretanto, os resultados também evidenciam limitações importantes. A categorização temática demonstra que nenhuma das ferramentas contemplou de maneira uniforme todas as dimensões relevantes para a formulação de problemas científicos. Algumas categorias apresentaram baixa frequência em determinados modelos, indicando que a utilização isolada de uma única ferramenta pode conduzir o estudante a explorar apenas parte das possibilidades investigativas existentes. Além disso, este estudo não avaliou a consistência científica, a originalidade ou a viabilidade metodológica das questões propostas, restringindo-se à análise das tendências temáticas das respostas produzidas.

Outro aspecto que merece atenção refere-se aos riscos associados ao uso indiscriminado de sistemas de IA generativa na educação. Embora os modelos avaliados sejam capazes de produzir rapidamente um conjunto diversificado de perguntas, a facilidade de obtenção dessas respostas pode induzir estudantes à aceitação automática das sugestões apresentadas, reduzindo oportunidades de desenvolvimento da autonomia intelectual e do pensamento crítico. A literatura tem destacado que modelos de linguagem podem produzir informações imprecisas, omitir aspectos relevantes de determinado tema ou apresentar respostas excessivamente convincentes, mesmo quando fundamentadas em informações incorretas, fenômeno conhecido como *hallucination* (Bender et al., 2021; Kasneci et al., 2023). Consequentemente, as questões geradas pelas ferramentas não devem ser compreendidas como problemas de pesquisa prontos para utilização, mas como insumos iniciais para discussão, refinamento e validação científica.

Os resultados obtidos também sugerem que a utilização combinada das três ferramentas pode representar uma estratégia mais promissora do que a adoção exclusiva de qualquer um dos sistemas avaliados. Enquanto o ChatGPT apresentou maior diversidade temática, o Gemini enfatizou aspectos sociais e econômicos, e o Perplexity AI concentrou-se em questões técnicas e fundamentadas em informações recuperadas durante a geração das respostas. Essa complementaridade pode ampliar o repertório de possibilidades investigativas disponíveis ao estudante, permitindo que diferentes perspectivas sejam consideradas durante a construção da questão-problema. Nesse contexto, torna-se evidente que o papel do professor permanece central no processo de orientação acadêmica. A formulação de um problema de pesquisa envolve competências que extrapolam a simples geração de perguntas, incluindo análise crítica da literatura, identificação de lacunas científicas, avaliação da relevância social e técnica do tema, definição de objetivos coerentes e escolha de métodos adequados para investigação. Essas atividades dependem de julgamento científico e experiência acadêmica, competências que não são substituídas pelos modelos de linguagem atualmente disponíveis (Holmes et al., 2022; UNESCO, 2023). Assim, a utilização de ferramentas de IA deve ser compreendida como recurso complementar à orientação docente, contribuindo para ampliar possibilidades de reflexão sem substituir o processo formativo desenvolvido pelo professor.

Por fim, é importante considerar as limitações deste estudo. A comparação foi realizada utilizando um único prompt padronizado, aplicado em um único momento temporal e restrito ao contexto da geração de energia elétrica. Além disso, não foram avaliadas múltiplas execuções para mensurar a variabilidade estocástica das respostas nem diferentes perfis de usuários ou estratégias de engenharia de prompts. Pesquisas futuras poderão ampliar esse delineamento experimental por meio da utilização de protocolos sistemáticos de avaliação, múltiplas coletas, métricas quantitativas de similaridade semântica, análises de confiabilidade entre avaliadores e participação de estudantes e docentes na validação da qualidade das questões produzidas. Essas abordagens contribuirão para uma compreensão mais abrangente do potencial das ferramentas de IA generativa como apoio à iniciação científica e ao ensino de Engenharia.

CONCLUSÃO

Este estudo comparou as questões-problema produzidas por três ferramentas de Inteligência Artificial generativa — ChatGPT, Gemini e Perplexity AI — aplicadas ao contexto da elaboração de Trabalhos de Conclusão de Curso em Engenharia



Elétrica, considerando diferentes tecnologias de geração de energia elétrica. Os resultados evidenciaram que as ferramentas apresentam comportamentos distintos na formulação de questões, refletidos na distribuição das categorias temáticas identificadas. Enquanto o ChatGPT apresentou maior diversidade temática, o Gemini enfatizou aspectos socioeconômicos e o Perplexity AI concentrou-se em questões de natureza técnica e econômica. Esses achados indicam que os modelos analisados possuem características complementares e podem contribuir de formas distintas para a etapa inicial de problematização científica, sem que seja possível afirmar a superioridade de uma ferramenta em relação às demais.

Sob a perspectiva educacional, os resultados sugerem que os modelos de linguagem podem atuar como instrumentos de apoio à geração de ideias e à ampliação das possibilidades de investigação durante a construção de projetos de pesquisa. Entretanto, as diferenças observadas entre as ferramentas e a ausência de cobertura uniforme de todas as dimensões relevantes da pesquisa evidenciam que as respostas produzidas pelas IAs devem ser analisadas criticamente e confrontadas com a literatura científica. Nesse contexto, a mediação do professor permanece essencial para orientar a delimitação do problema de pesquisa, avaliar a consistência das questões propostas e promover o desenvolvimento da autonomia intelectual e do pensamento crítico dos estudantes, evitando que a Inteligência Artificial seja utilizada como substituta do processo formativo.

Como limitações, este estudo foi conduzido utilizando um único prompt padronizado, aplicado durante o cenário tecnológico de 2024, sem controle sobre parâmetros internos dos modelos nem realização de múltiplas execuções para avaliação da variabilidade das respostas. Dessa forma, os resultados representam uma análise comparativa contextualizada ao período investigado e não devem ser generalizados para versões futuras das ferramentas. Como perspectivas de pesquisa, recomenda-se ampliar o protocolo experimental por meio da utilização de diferentes estratégias de engenharia de prompts, múltiplas coletas, métricas quantitativas e participação de docentes e estudantes na avaliação das questões produzidas, contribuindo para uma compreensão mais abrangente do potencial da Inteligência Artificial generativa como ferramenta de apoio à formação científica em Engenharia.

REFERÊNCIAS

BAEK, C.; TATE, T.; DAI, S.; et al. The promises and risks of generative artificial intelligence in education. *Phi Delta Kappan*, v. 105, n. 1, p. 22–27, 2023.

BARDIN, L. *Análise de conteúdo*. Lisboa: Edições 70, 2016.

BENDER, E. M.; GEBRU, T.; MCMILLAN-MAJOR, A.; SHMITCHELL, S. On the dangers of stochastic parrots: Can language models be too big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. New York: ACM, 2021. p. 610–623.

BOMMASANI, R.; HUDSON, D. A.; ADLI, E.; et al. On the opportunities and risks of foundation models. Stanford: Stanford University, Center for Research on Foundation Models, 2021.

FELDER, R. M.; BRENT, R. *Teaching and learning STEM: A practical guide*. San Francisco: Jossey-Bass, 2016.

GIL, A. C. *Métodos e técnicas de pesquisa social*. 6. ed. São Paulo: Atlas, 2008.

HOLMES, W.; BIALIK, M.; FADEL, C. *Artificial intelligence in education: Promises and implications for teaching and learning*. Boston: Center for Curriculum Redesign, 2022.

KASNECI, E.; SESSLER, K.; KÜHNBERGER, K.; et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, v. 103, 102274, 2023.

LEWIS, P.; PEREZ, E.; POYIADZI, R.; et al. Retrieval-Augmented Generation for knowledge-intensive NLP tasks. In: *Advances in Neural Information Processing Systems (NeurIPS 2020)*. Red Hook: Curran Associates, 2020.

LO, C. K. What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, v. 13, n. 4, 410, 2023.

OPENAI. GPT-4 Technical Report. arXiv, arXiv:2303.08774, 2023.

SEVERINO, A. J. *Metodologia do trabalho científico*. 24. ed. São Paulo: Cortez, 2017.

UNESCO. *Guidance for generative AI in education and research*. Paris: UNESCO, 2023.

ZHAI, X. ChatGPT in education: Opportunities, challenges and future directions. *Educational Technology & Society*, v. 26, n. 4, p. 1–12, 2023.