

Título em português: Viés, Agência e Consciência: Uma Análise das Diretrizes da UNESCO para o Uso Crítico da Inteligência Artificial Generativa na Educação

Título em inglês: Bias, Agency, and Awareness: An Analysis of UNESCO Guidelines for the Critical Use of Generative Artificial Intelligence in Education

Resumo: Este trabalho teve como objetivo realizar uma análise documental de quatro publicações da UNESCO sobre Inteligência Artificial Generativa (IAgen), com o objetivo de identificar diretrizes, preocupações éticas e recomendações voltadas à promoção da formação crítica, do letramento digital e do letramento em IA no contexto educacional. Parte-se da hipótese de que estudantes sem formação crítica adequada tendem a aceitar passivamente conteúdos gerados por IA, confundindo reprodução algorítmica com conhecimento autêntico. Os resultados apontam quatro eixos principais nas recomendações da UNESCO: (1) primazia ética e humanista para orientar o uso da IA; (2) capacitação crítica que supera a mera instrumentalização tecnológica; (3) reconhecimento da tensão epistemológica entre automatização da resposta e construção ativa do conhecimento; e (4) necessidade de uma governança multiescalar e participativa. Conclui-se que a integração consciente da IAgen requer estratégias pedagógicas que integrem práticas reflexivas, governança responsável e alfabetização digital avançada, visando superar a superficialização cognitiva e promover a agência crítica dos estudantes frente às tecnologias digitais.

Palavras-chave: Inteligência Artificial Generativa; Letramento em IA; Pensamento crítico; Educação Superior; UNESCO.

Abstract: This study aimed to perform a documentary analysis of four UNESCO publications on Generative Artificial Intelligence (GenAI) to identify guidelines, ethical concerns, and recommendations focused on promoting critical education, digital literacy, and AI literacy within the educational context. The research is based on the hypothesis that students lacking solid critical training tend to passively accept AI-generated content, confusing algorithmic reproduction with authentic knowledge. The findings highlight four central axes in UNESCO's recommendations: (1) the ethical and humanistic primacy guiding AI use; (2) critical capacity-building surpassing mere technological instrumentalization; (3) acknowledgment of epistemological tensions between response automation and active knowledge construction; and (4) the need for multi-scale and participatory governance. The study concludes that conscious integration of GenAI requires pedagogical strategies incorporating reflective practices, responsible governance, and advanced digital literacy, aiming to counter cognitive superficiality and foster students' critical agency regarding digital technologies.

Keywords: Generative Artificial Intelligence; AI Literacy; Critical Thinking; Higher Education; UNESCO.

Área temática: Ética da informação e responsabilidade algorítmica

1. Introdução

O uso crescente de ferramentas de Inteligência Artificial Generativa (IAgen), como ChatGPT, Gemini e Copilot, tem transformado significativamente as práticas acadêmicas, especialmente entre estudantes do ensino superior. Com promessas de maior agilidade, produtividade e acesso instantâneo ao conhecimento, essas tecnologias tornaram-se parte integrante dos processos de leitura, escrita e pesquisa (Boa Sorte *et al.*, 2021; Pereira; Reis; Ulbricht; Santos, 2024). Diante desses avanços, torna-se urgente promover o letramento digital em IAgen, contemplando tanto as dimensões tecnológicas quanto as humanas (UNESCO, 2023). No entanto, esse uso intensivo nem sempre é acompanhado por uma formação crítica capaz de reconhecer os limites, os vieses e os riscos envolvidos na aplicação dessas ferramentas na produção do conhecimento.

Este trabalho se propôs a realizar a análise documental de quatro publicações da UNESCO sobre IAgen, com o objetivo de identificar diretrizes, preocupações éticas e recomendações voltadas à promoção da formação crítica, do letramento digital e do letramento em IA no contexto educacional. Parte-se da hipótese de que, na ausência de uma formação sólida em pensamento crítico e letramento digital, os estudantes tendem a aceitar os *outputs* das IAgen como verdades neutras e objetivas, o que

compromete não apenas a qualidade acadêmica dos trabalhos, mas também os princípios éticos e formativos que fundamentam a educação superior.

2. Fundamentação Teórica

O conceito de viés algorítmico refere-se às distorções sistemáticas que emergem de modelos computacionais em função dos dados com os quais foram treinados, das decisões de design de seus desenvolvedores ou de estruturas sociais preexistentes. Para Safiya Noble (2018), os algoritmos que operam motores de busca ou modelos de linguagem “não são neutros” e frequentemente reproduzem visões de mundo sexistas, racistas e classistas. Cathy O’Neil (2016) vai além, ao afirmar que os sistemas algorítmicos não apenas refletem, mas amplificam desigualdades sociais, transformando preconceitos em decisões automatizadas com aparência de precisão matemática.

Aplicadas ao campo da produção acadêmica, essas distorções podem ser reproduzidas quando estudantes se apropriam dos textos gerados por IAGen sem submetê-los a processos de validação, comparação ou crítica. Ao confiarem cegamente em plataformas de IAGen como uma “autoridade epistêmica”, os discentes abdicam do papel de sujeito do conhecimento para se tornarem consumidores passivos de respostas algorítmicas.

Além disso, é importante considerar o conceito de letramento digital, entendido aqui como a capacidade de compreender, avaliar criticamente e interagir com tecnologias digitais de forma ética e reflexiva. De modo complementar, o letramento em IA (Kalantzis; Cope, 2025) enfatiza a competência para analisar, questionar e co-criar com sistemas de inteligência artificial. A ausência dessa competência entre estudantes de graduação e pós-graduação contribui para a normalização do uso de IA como ferramenta de automatização da autoria, dissolvendo as fronteiras entre escrever e copiar, pensar e replicar.

Bender *et al.* (2021) alertam que os modelos de linguagem baseados em grandes volumes de dados são “papagaios estocásticos” (*stochastic parrots*), que combinam palavras com base em probabilidades, mas sem compreensão semântica ou responsabilidade factual. Ainda assim, esses textos simulam coerência e fluência, gerando a ilusão de profundidade argumentativa, o que pode ser particularmente sedutor para estudantes em formação.

3. Método

Este resumo expandido adota uma abordagem qualitativa e teórico-reflexiva, baseada em revisão crítica de literatura interdisciplinar, com ênfase em autores dos campos da ciência da informação, ética da inteligência artificial, educação crítica e estudos sobre tecnologia e sociedade. O objetivo não é oferecer um diagnóstico empírico, mas construir uma problematização conceitual sobre o uso acrítico de sistemas de IA generativa na produção acadêmica discente. A técnica utilizada para análise das fontes bibliográficas objeto do estudo foi a análise temática reflexiva de Braun e Clarke (2019). Esta perspectiva de análise está ancorada em uma perspectiva crítica que entende o conhecimento como uma construção social e situada, demandando um engajamento ativo e consciente dos sujeitos que produzem saber.

A escolha metodológica se justifica pela necessidade de compreender fenômenos emergentes no campo acadêmico, ainda pouco sistematizados empiricamente, mas que já apresentam implicações éticas e formativas relevantes.

4. Resultados e Discussão

As publicações da UNESCO selecionadas para análise documental foram: *Recommendation on the Ethics of Artificial Intelligence* (2022); *Guidance for Generative AI in Education and Research* (2023), *AI Competency Framework for Teachers* (2024a) e *AI and education: guidance for policy-makers* (2024b), revelaram uma padronização conceitual consistente que articula ética, pedagogia e governança tecnológica. Esta convergência estrutura-se em torno de quatro eixos fundamentais que refletem tanto as oportunidades quanto os desafios emergentes da integração da IAGen nos sistemas educacionais.

4.1. A primazia da ética humanista como bússola normativa

A *Recommendation on the Ethics of Artificial Intelligence* (2022) estabelece princípios universais de dignidade humana, justiça, inclusão e sustentabilidade como alicerces normativos para a IA educacional. O *Guidance for Generative AI in Education and Research* (2023) operacionaliza estes princípios ao identificar riscos específicos: desinformação, vieses algorítmicos, violação de privacidade e dependência tecnológica.

A resposta institucional centra-se na tríade transparência-explicabilidade-auditabilidade, assegurando que a IAgem funcione como ferramenta compreensível e validável pela comunidade educativa, evitando o efeito “caixa preta” nos processos de aprendizagem.

Este tema transcende uma visão de conformidade (*compliance*); ele posiciona a ética como o princípio orientador fundamental para toda inovação tecnológica. No *Guidance for Generative AI*, essa primazia se manifesta na exigência de “transparência”, “explicabilidade” e “responsabilização” (*accountability*) como condições não negociáveis para a adoção da IAgem, o que representa um contraponto direto à lógica da “caixa-preta” algorítmica.

2. A capacitação para a agência crítica: para além do letramento instrumental

Este tema captura a resposta pedagógica da UNESCO. A análise revela que a preocupação não é apenas com o *uso* da tecnologia, mas com a *qualidade da interação humana* com ela. O *AI Competency Framework for Teachers* (2024) vai além da dicotomia técnico-pedagógica, propondo competências integradas que incluem domínio operacional das ferramentas de IA e capacidade de análise crítica dos seus impactos sociais e pedagógicos. Os professores são conceptualizados como mediadores críticos que orientam os estudantes na navegação reflexiva pelo ecossistema da IA. Esse documento é explícito ao rejeitar um modelo de capacitação puramente instrumental. Códigos como “mentalidade centrada no ser humano” (*human-centred mindset*), “avaliação crítica de ferramentas de IA” e “pedagogia com IA” (*AI pedagogy*) indicam uma demanda por um letramento digital avançado. Este letramento capacita professores e alunos a questionarem os sistemas: como foram treinados? Quais dados utilizam? Quais vieses reproduzem?

O *Guidance for Generative AI in Education and Research*, por sua vez, reforça esta abordagem ao recomendar a integração transversal da literacia em IA nos currículos, não como adição de conteúdos tecnológicos, mas como reformulação dos processos de ensino-aprendizagem para uma cidadania crítica numa sociedade mediada pela IA.

3. Tensão Epistemológica: entre a automação da resposta e a construção do conhecimento

A análise revela preocupações centrais com as transformações epistemológicas induzidas pela IAgem, particularmente o risco de “superficialização cognitiva” quando estudantes substituem processos reflexivos pela reprodução de conteúdos automatizados. De um lado, códigos como “interações triangulares (professor-aluno-IA)” e “risco de superficialização cognitiva” apontam para o perigo de a IAgem promover uma epistemologia da resposta fácil, que se alinha à “educação bancária” de Freire. Esta substituição compromete a construção ativa do conhecimento, a autonomia intelectual e competências fundamentais como investigação, análise de fontes e verificação crítica. Por outro lado, códigos como “potencial para pensamento de ordem superior” e “resolução criativa de problemas” sugerem que, se bem utilizada, a IAgem pode automatizar tarefas de baixa ordem cognitiva, liberando espaço para a criatividade e a análise complexa.

A UNESCO propõe o conceito de “mediação reflexiva” como resposta pedagógica, utilizando a IAgem para complementar (não substituir) as capacidades cognitivas humanas, apoiando o pensamento crítico, criatividade e resolução de problemas complexos através de estratégias que integram a IA de forma consciente e intencional.

4. A arquitetura da governança: da responsabilidade individual à regulação sistêmica

O último tema construído aborda a necessidade de uma resposta estruturada e multiescalar. O *Guidance for Generative AI in Education and Research* propõe medidas regulatórias abrangentes: definição de critérios éticos para acesso a ferramentas de IA, protocolos de avaliação ética e

pedagógica para adoção institucional, sistemas de proteção de dados e privacidade, e programas de formação contínua para educadores e gestores. Esta abordagem sistêmica reconhece a necessidade de coordenação entre políticas nacionais e práticas institucionais locais, advogando por uma governança participativa que envolva decisores políticos, educadores, estudantes, famílias e comunidades numa gestão colaborativa dos desafios e oportunidades da IA educacional.

A UNESCO argumenta que a salvaguarda ética é uma responsabilidade compartilhada, que exige uma arquitetura de governança que vai desde a regulação de empresas de tecnologia até a criação de políticas institucionais claras para a aquisição, validação e uso de ferramentas de IAGen, garantindo que as decisões sejam informadas por critérios pedagógicos e não por pressões de mercado.

Conclusão

A análise dos documentos da UNESCO revela uma visão madura e nuançada sobre a integração da inteligência artificial generativa na educação, que evita tanto o tecno-otimismo acrítico quanto o tecno-pessimismo paralisante. A abordagem proposta caracteriza-se por um realismo construtivo que reconhece simultaneamente o potencial transformador da IA e a necessidade de precauções éticas e pedagógicas rigorosas.

Os quatro eixos identificados, ética e responsabilidade, competências críticas, transformações epistemológicas e governança regulatória, constituem um quadro integrado para pensar e implementar a IAGen na educação de forma sustentável e socialmente responsável. Este quadro sugere que o sucesso da integração na educação depende não apenas da sofisticação tecnológica das ferramentas disponíveis, mas sobretudo da capacidade das instituições educacionais para desenvolver culturas organizacionais, práticas pedagógicas e estruturas de governança adequadas aos desafios e oportunidades da era da IAGen.

As contribuições da UNESCO para este debate residem precisamente na articulação entre princípios universais e estratégias práticas, oferecendo um referencial que pode orientar políticas educacionais, práticas institucionais e iniciativas de formação de educadores em contextos diversos. Contudo, a efetivação desta visão requer um investimento sustentado em investigação, formação, infraestruturas e, sobretudo, numa cultura de diálogo crítico sobre o papel das tecnologias na educação e na sociedade.

Referências

- BOA SORTE, P.; FARIA, M. A. F.; dos SANTOS, A. E.; SANTOS, J. C. A.; DIAS, J. S. S. R. Inteligência artificial e escrita acadêmica: o que nos reserva o algoritmo GPT-3? **Revista EntreLinguas**, v. 7, p. 1-22, 2021. DOI: <https://doi.org/10.29051/el.v7i00.15352>.
- BRAUN, Virginia; CLARKE, Victoria. Reflecting on reflexive thematic analysis. **Qualitative research in sport, exercise and health**, v. 11, n. 4, p. 589-597, 2019.
- KALANTZIS, M. and COPE, B. Literacy in the Time of Artificial Intelligence. **Reading Research Q**, v. 60, n. e591, 2025. <https://doi.org/10.1002/rrq.591>.
- PEREIRA, Ricardo et al. Generative artificial intelligence and academic writing: an analysis of the perceptions of researchers in training. **Management Research: Journal of the Iberoamerican Academy of Management**, v. 22, n. 4, p. 429-450, 2024.
- UNESCO. **AI and education: guidance for policy-makers**. Paris: UNESCO, 2024b. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000376709>. Acesso em: 10 jun. 2025.
- UNESCO. **AI Competency framework for teachers**. Paris: UNESCO, 2024a. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000391104>. Acesso em: 10 jun. 2025.
- UNESCO. **Guidance for generative AI in education and research**. Paris: UNESCO, 2023. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000386693>. Acesso em: 9 jun. 2025.
- UNESCO. **Recommendation on the ethics of artificial intelligence**. Paris: UNESCO, 2022. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>. Acesso em: 9 jun. 2025.