

**ENTRE O APOIO E A DEPENDÊNCIA: A INFLUÊNCIA DA INTELIGÊNCIA ARTIFICIAL  
GENERATIVA NO DESENVOLVIMENTO COGNITIVO*****BETWEEN SUPPORT AND DEPENDENCE: THE INFLUENCE OF GENERATIVE  
ARTIFICIAL INTELLIGENCE ON COGNITIVE DEVELOPMENT******ENTRE EL APOYO Y LA DEPENDENCIA: LA INFLUENCIA DE LA INTELIGENCIA  
ARTIFICIAL GENERATIVA EN EL DESARROLLO COGNITIVO***Guilherme Augusto Ferraz Rita<sup>1</sup>, Luiz Gustavo Marcondes de Campos Melo<sup>1</sup>, Maria Lívia de Lima Moreira<sup>1</sup>

e768322

<https://doi.org/10.47820/recima21.v7i6.8322>

PUBLICADO: 06/2026

**RESUMO**

O avanço da inteligência artificial generativa e dos grandes modelos de linguagem (LLMs) tem transformado o ambiente educacional, especialmente no ensino superior, ao oferecer respostas elaboradas de modo quase instantâneo. O presente estudo tem como objetivo analisar de que modo o uso da IA generativa como fonte primária de respostas afeta o desenvolvimento do pensamento crítico e a autonomia intelectual de estudantes universitários. Trata-se de uma pesquisa qualitativa, de caráter exploratório, desenvolvida por meio de revisão bibliográfica integrativa em bases como SciELO, Portal CAPES e PubMed, com busca complementar no Google Scholar. O referencial teórico articula a teoria sócio-histórica de Vygotsky, a pedagogia da autonomia de Freire, a epistemologia social de Hardwig e a sociologia das tecnologias intelectuais de Lévy, cotejadas com literatura empírica internacional recente. Os resultados indicam que, quando utilizada sem mediação pedagógica crítica, a IA generativa tende a operar como substituta cognitiva, favorecendo o *cognitive offloading* e a dependência epistêmica acrítica. Estudos experimentais evidenciam redução do engajamento cognitivo e da conectividade neural em usuários frequentes de LLMs. Conclui-se que a incorporação dessas ferramentas nas práticas educativas exige letramento crítico e mediação intencional, a fim de que a tecnologia potencialize, e não substitua, o desenvolvimento intelectual dos estudantes.

**PALAVRAS-CHAVE:** Inteligência artificial generativa. Pensamento crítico. Dependência epistêmica. Ensino superior. Mediação pedagógica.

**ABSTRACT**

*The advancement of generative artificial intelligence and large language models (LLMs) has transformed the educational environment, particularly in higher education, by instantly providing elaborate responses. This study aims to analyze how the use of generative AI as a primary source of answers affects the development of critical thinking and the intellectual autonomy of university students. This is a qualitative, exploratory study conducted through an integrative bibliographic review of databases such as SciELO, the CAPES Portal, and PubMed, with a complementary search on Google Scholar. The theoretical framework draws on Vygotsky's socio-historical theory, Freire's pedagogy of autonomy, Hardwig's social epistemology, and Lévy's sociology of intellectual technologies, contrasted with recent international empirical literature. The findings indicate that, when used without critical pedagogical mediation, generative AI tends to function as a cognitive substitute, fostering cognitive offloading and uncritical epistemic dependence. Experimental studies demonstrate reduced cognitive engagement and neural connectivity among frequent LLM users.*

<sup>1</sup> FATEC, graduandos.



*It is concluded that the incorporation of these tools into educational practices requires critical literacy and intentional mediation, so that technology enhances, rather than replaces, students' intellectual development.*

**KEYWORDS:** Generative artificial intelligence. Critical thinking. Epistemic dependence. Higher education. Pedagogical mediation.

### **RESUMEN**

*El avance de la inteligencia artificial generativa y de los grandes modelos de lenguaje (LLMs) ha transformado el entorno educativo, especialmente en la educación superior, al ofrecer respuestas elaboradas de modo casi instantáneo. El presente estudio tiene como objetivo analizar de qué modo el uso de la IA generativa como fuente primaria de respuestas afecta el desarrollo del pensamiento crítico y la autonomía intelectual de los estudiantes universitarios. Se trata de una investigación cualitativa, de carácter exploratorio, desarrollada mediante una revisión bibliográfica integrativa en bases como SciELO, Portal CAPES y PubMed, con búsqueda complementaria en Google Scholar. El marco teórico articula la teoría socio-histórica de Vygotsky, la pedagogía de la autonomía de Freire, la epistemología social de Hardwig y la sociología de las tecnologías intelectuales de Lévy, contrastadas con literatura empírica internacional reciente. Los resultados indican que, cuando se utiliza sin mediación pedagógica crítica, la IA generativa tiende a operar como sustituta cognitiva, favoreciendo el *cognitive offloading* y la dependencia epistémica acrítica. Estudios experimentales evidencian reducción del compromiso cognitivo y de la conectividad neural en usuarios frecuentes de LLMs. Se concluye que la incorporación de estas herramientas en las prácticas educativas exige alfabetización crítica y mediación intencional, a fin de que la tecnología potencie, y no sustituya, el desarrollo intelectual de los estudiantes.*

**PALABRAS CLAVE:** *Inteligencia artificial generativa. Pensamiento crítico. Dependencia epistémica. Educación superior. Mediación pedagógica.*

### **INTRODUÇÃO**

A inteligência artificial (IA) deixou de ser promessa futurista para se consolidar como infraestrutura da vida contemporânea, impactando setores produtivos, sistemas de saúde, comunicação e, de forma crescente, a educação. Essa transformação altera profundamente a maneira como o conhecimento é produzido, acessado e validado, e exige da comunidade acadêmica respostas teórica e empiricamente fundamentadas.

No ambiente educacional, a rápida difusão dessas ferramentas tem transformado práticas de estudo e produção acadêmica em ritmo que supera o da reflexão pedagógica sobre seus efeitos. O lançamento público do ChatGPT em novembro de 2022 marcou um ponto de inflexão: em pouco mais de dois anos, *chatbots* generalistas baseados em grandes modelos de linguagem passaram a ser empregados cotidianamente por estudantes do ensino superior na elaboração de trabalhos acadêmicos, na síntese de bibliografias e na obtenção de respostas para problemas conceituais. Como adverte Kenski (2007), a simples inserção de tecnologias não



garante melhorias no ensino: seus efeitos formativos dependem inteiramente do modo como são integradas às práticas pedagógicas.

Diante disso, emerge o problema central desta pesquisa: de que modo o uso da IA generativa como fonte primária de respostas afeta o desenvolvimento do pensamento crítico e a autonomia intelectual de estudantes do ensino superior?

Esse problema se intensifica diante da propensão dos LLMs a gerar informações incorretas com aparência confiável, fenômeno conhecido como *hallucination* (JI *et al.*, 2023). Tal característica cria uma assimetria epistêmica, na qual o estudante recebe respostas elaboradas sem dispor de instrumentos imediatos para avaliá-las criticamente, favorecendo a aceitação irrefletida (BENDER *et al.*, 2021).

A literatura educacional brasileira oferece elementos para iluminar essa questão. Freire (1996) defende que ensinar não é transferir conhecimento, mas criar possibilidades para que o educando construa e reconstrua o saber de forma autoral; Demo (2011) associa a formação intelectual à elaboração própria, em oposição à reprodução mecânica de informações recebidas; e Vygotsky (1984), em outro registro, lembra que as ferramentas culturais devem potencializar o pensamento, não substituí-lo. O conjunto dessas perspectivas indica que está em causa, no uso irrefletido da IA generativa, não apenas uma questão de procedimento didático, mas a integridade do processo formativo universitário.

Justifica-se este estudo, em primeiro lugar, pela urgência acadêmica do fenômeno investigado. Embora a literatura internacional sobre os efeitos cognitivos dos LLMs no ensino superior tenha crescido aceleradamente desde 2023, com publicações em periódicos de referência como *Learning and Individual Differences* (KASNECI *et al.*, 2023) e *Societies* (GERLICH, 2025), a produção brasileira sobre o tema permanece escassa, especialmente quando se trata de articular essas evidências empíricas a referenciais consolidados da teoria educacional nacional. Esta pesquisa contribui para reduzir essa lacuna ao cotejar a literatura empírica internacional recente com o quadro teórico oferecido por Vygotsky, Freire, Demo e Hardwig, oferecendo subsídios para a interpretação do fenômeno no contexto brasileiro.

Em segundo lugar, justifica-se a pesquisa por sua relevância social e formativa. Os estudantes universitários hoje em formação ingressarão, em curto prazo, em ambientes profissionais que demandarão precisamente as capacidades cognitivas — análise crítica, síntese autoral, julgamento epistêmico — cujo desenvolvimento pode estar sendo comprometido pela delegação sistemática dessas operações aos LLMs. A formação de quadros técnicos e profissionais no ensino superior brasileiro sustenta-se sobre a expectativa de que esses profissionais sejam capazes de pensamento autônomo diante de problemas complexos —



competência que não se forma na ausência de exercício cognitivo direto. Compreender os mecanismos pelos quais o uso irrefletido da IA generativa pode erodir essa capacidade é, portanto, condição para qualquer política institucional ou prática docente que pretenda integrar essas ferramentas sem prejuízo formativo.

Diante do exposto, este trabalho tem como objetivo geral analisar os efeitos do uso da IA generativa como fonte primária de respostas no desenvolvimento do pensamento crítico de estudantes do ensino superior. Como objetivos específicos, busca-se: (i) delimitar conceitualmente IA generativa, grandes modelos de linguagem, *chatbots* generalistas e tutores pedagógicos baseados em IA; (ii) caracterizar, a partir do referencial teórico mobilizado, a distinção entre o uso da IA generativa como ferramenta mediadora e como substituta cognitiva; (iii) identificar, na literatura empírica recente, as evidências sobre os efeitos cognitivos do uso intensivo de LLMs em contextos formativos; e (iv) articular essas evidências às contribuições teóricas brasileiras sobre pensamento crítico e autonomia intelectual.

Cabe registrar os limites do escopo desta investigação. Trata-se de uma revisão bibliográfica integrativa, sem coleta de dados primários junto a estudantes ou docentes, circunscrita ao uso de *chatbots* generalistas baseados em LLMs no contexto do ensino superior — ficando fora de seu escopo a análise sistemática de tutores pedagógicos baseados em IA, os mecanismos institucionais de detecção de uso indevido, as políticas universitárias específicas sobre IA e a integração dessas tecnologias na educação básica.

## 1. REFERENCIAL TEÓRICO

### **Delimitação conceitual: da IA generativa aos tutores pedagógicos**

Antes de discutir o impacto da inteligência artificial sobre os processos de aprendizagem, importa precisar o que está em causa quando se fala em “IA” no contexto educacional contemporâneo. O termo abrange um campo amplo de técnicas computacionais voltadas à execução de tarefas que tradicionalmente exigiram cognição humana, das quais interessa, a este estudo, uma classe específica: a inteligência artificial generativa. Diferentemente de sistemas discriminativos, voltados à classificação, predição ou detecção de padrões em dados existentes, sistemas generativos são treinados para produzir conteúdo novo — texto, imagens, áudio ou código — a partir das regularidades estatísticas aprendidas em grandes corpora. Sua consolidação na escala atual decorre, em larga medida, do desenvolvimento da arquitetura *Transformer* (VASWANI *et al.*, 2017), que tornou viável modelar dependências de longo alcance em sequências textuais e gerar linguagem em volume antes inalcançável.



No interior dessa categoria, os grandes modelos de linguagem (LLMs, *large language models*) constituem a subclasse especializada em texto. Trata-se de redes neurais profundas, frequentemente com centenas de bilhões de parâmetros, treinadas por meio de objetivos autossupervisionados, como a predição da próxima palavra, sobre conjuntos textuais massivos extraídos da web, de livros e de bases científicas (BROWN *et al.*, 2020). O resultado funcional dessa arquitetura é a capacidade de produzir respostas sintática e discursivamente coerentes diante de praticamente qualquer estímulo em linguagem natural. Importa, contudo, registrar uma característica conceitualmente decisiva para o restante deste trabalho: a coerência produzida por um LLM não decorre de compreensão semântica, mas da combinação probabilística de formas linguísticas observadas durante o treinamento, sem referência a significado, como sintetizam Bender *et al.* (2021). Sob essa perspectiva teórica, a fluência discursiva não deveria, portanto, ser tomada como índice de raciocínio.

Sobre o substrato técnico dos LLMs constroem-se aplicações com finalidades muito distintas, e é nesse ponto que parte da literatura tende a confundir categorias que devem ser mantidas separadas. *Chatbots* generalistas, como o ChatGPT, o Claude e o Gemini, são interfaces conversacionais de propósito amplo, otimizadas para utilidade ampla e disponibilidade imediata de respostas elaboradas. Sua arquitetura de interação não pressupõe nenhum projeto pedagógico: o desempenho avaliado pelos desenvolvedores é a qualidade percebida da resposta, não o aprendizado do usuário. Tutores pedagógicos baseados em IA, por sua vez, configuram uma categoria distinta, ainda que tecnicamente possam empregar os mesmos LLMs como base. Sistemas tutoriais inteligentes são desenhados com finalidade educativa explícita e com arquiteturas de interação que preveem mediação do processo de aprendizagem, andaimagem (*scaffolding*) e adaptação ao desempenho do aprendiz (VANLEHN, 2011). O Khanmigo, da Khan Academy, exemplifica essa categoria ao operar por meio de uma estratégia deliberadamente socrática: em vez de fornecer respostas, conduz o estudante por perguntas sucessivas até que ele próprio chegue à solução (KHAN ACADEMY, 2023).

A distinção não é apenas terminológica. Um mesmo LLM pode sustentar tanto um *chatbot* generalista quanto um tutor pedagógico; a diferença reside na finalidade do sistema e na arquitetura de interação que se constrói sobre o modelo. Esse esclarecimento importa porque o objeto deste estudo é, especificamente, o uso de *chatbots* generalistas baseados em LLMs como fonte primária de respostas por estudantes do ensino superior — não a integração curricular de tutores pedagógicos, nem a IA aplicada à educação em sentido amplo, que inclui sistemas de recomendação, análise de aprendizagem e correção automatizada, entre outros.



### **Ferramenta mediadora ou substituto cognitivo: uma distinção fundamental**

Compreender o impacto da inteligência artificial generativa sobre o aprendizado exige, antes de qualquer análise empírica, uma distinção conceitual que estrutura todo o debate: a diferença entre uma tecnologia que medeia o pensamento e uma que o substitui.

A teoria sócio-histórica de Vygotsky (1984) oferece o ponto de partida mais produtivo para essa distinção. Para o autor, o desenvolvimento cognitivo humano não ocorre de forma isolada, mas por meio da interação com o ambiente social e com as ferramentas culturais disponíveis. Essas ferramentas, sejam físicas, como instrumentos de trabalho, ou simbólicas, como a linguagem e a escrita, atuam como mediadoras entre o sujeito e o mundo, ampliando suas capacidades e possibilitando formas de pensamento que sem elas seriam inviáveis. O conceito central para essa dinâmica é a Zona de Desenvolvimento Proximal (ZDP): o espaço entre aquilo que o aprendiz já é capaz de realizar de forma autônoma e aquilo que pode realizar com o auxílio de um agente mais experiente ou de uma ferramenta mediadora adequada. A função da mediação, nesse quadro, não é realizar a tarefa pelo sujeito, mas criar as condições para que ele avance em sua própria capacidade.

Pierre Lévy (1993) aprofunda essa perspectiva ao distinguir entre tecnologias que externalizam o pensamento, tornando-o visível, manipulável e comunicável, e aquelas que simplesmente executam operações cognitivas em lugar do sujeito. A escrita externaliza a memória e o raciocínio, tornando-os objetos de reflexão; uma calculadora executa operações aritméticas sem que o usuário precise compreendê-las. Para Lévy, o valor cognitivo de uma tecnologia intelectual depende diretamente do modo como ela é incorporada às práticas do sujeito: uma mesma ferramenta pode ampliar ou empobrecer o pensamento conforme o uso que dela se faz.

Aplicada aos LLMs, essa distinção torna-se especialmente relevante. Quando um estudante utiliza o ChatGPT para organizar ideias que já foram por ele elaboradas, para obter retorno sobre um argumento que já construiu ou para explorar perspectivas que não havia considerado, a ferramenta opera como mediadora — potencializa o pensamento sem substituí-lo. Quando, ao contrário, o estudante a utiliza para gerar a resposta que deveria ser por ele produzida, para redigir o argumento que não se deu ao trabalho de construir ou para resumir um texto que não leu, a ferramenta tende a operar como substituta — executa a tarefa cognitiva que é, precisamente, o objeto da formação. É essa segunda modalidade de uso que constitui o foco da investigação aqui proposta, risco que Kasneci *et al.* (2023) identificam ao advertir que a facilidade de obtenção de respostas por meio de LLMs pode impactar negativamente o pensamento crítico e reduzir o interesse do estudante em conduzir suas próprias investigações.



### **IA educacional versus IA generativa: uma distinção necessária**

Nos últimos anos, consolidou-se uma categoria distinta de ferramentas digitais cujo design pedagógico parte de uma premissa oposta à dos LLMs de uso geral: em vez de fornecer respostas prontas, essas ferramentas estruturam a interação de modo a conduzir o estudante ao raciocínio autônomo. O Khanmigo, tutor inteligente desenvolvido pela Khan Academy e lançado em 2023, é o exemplo mais amplamente documentado dessa abordagem. Diferentemente de ferramentas como o ChatGPT, o Khanmigo não fornece respostas diretamente; em vez disso, guia o estudante para que encontre as respostas por conta própria (KHAN ACADEMY, 2023). Sua arquitetura é deliberadamente socrática: quando um estudante apresenta uma dúvida sobre um problema de álgebra, o sistema não resolve o problema, mas pergunta ao estudante qual seria o primeiro passo para resolvê-lo, conduzindo-o por meio de questões sucessivas até que ele mesmo chegue à solução. Como observou o próprio fundador da instituição, Sal Khan, é fácil criar um invólucro em torno de modelos de linguagem e chamá-lo de tutor, mas na prática esses sistemas funcionam essencialmente como ferramentas que fazem o dever de casa pelo estudante (KHAN *apud* CAUGHILL, 2023).

O Duolingo Max, lançado em março de 2023 com base no modelo GPT-4 da OpenAI, segue princípio semelhante no domínio do aprendizado de idiomas. A plataforma incorporou duas funcionalidades centrais — *Roleplay*, um parceiro de conversação baseado em IA, e *Explain My Answer*, que detalha as regras gramaticais quando o estudante comete um erro — ambas orientadas não à entrega de respostas, mas à prática situada e ao feedback contextualizado sobre erros cometidos pelo próprio aprendiz (DUOLINGO, 2023).

Isso não significa, contudo, que as ferramentas educacionais específicas estejam imunes ao risco de dependência epistêmica. O design pedagógico pode mitigar o risco, mas não eliminá-lo, pois, como advertem Kasneci *et al.* (2023, p. 7), apenas uma abordagem pedagógica com forte foco no pensamento crítico e em estratégias de verificação de informações — “*a clear pedagogical approach with a strong focus on critical thinking and strategies for fact checking*” — pode garantir o uso responsável dos LLMs nos currículos de ensino. A variável decisiva permanece sendo o modo como o estudante se relaciona com a ferramenta: se ela é usada para apoiar um processo de aprendizado que ainda pertence ao sujeito, ou se é usada para dispensar esse processo inteiramente.

### **Dependência epistêmica: conceito, mecanismos e especificidades da IA generativa**

Em uma sociedade complexa, algum grau de dependência epistêmica é não apenas inevitável, mas funcionalmente necessário: nenhum indivíduo pode dominar todos os campos do saber. O próprio Hardwig (1985, p. 335) adverte que buscar a autonomia epistêmica de forma



irrestrita resultaria apenas em crenças “*relatively uninformed, unreliable, crude, untested, and therefore irrational*” — imprecisas, não testadas e irracionais.

O problema, portanto, não é a dependência em si, mas sua degeneração em dependência acrítica, condição na qual o agente receptor não apenas confia na fonte, mas renuncia à capacidade de avaliar, questionar e verificar o que dela recebe. O próprio Hardwig (1985, p. 335) reconhece que a deferência epistêmica pode conduzir a uma postura de “*passively and uncritically accepting what we are given to believe*” — aceitar passiva e acriticamente o que nos é dado como crença. Nessa forma patológica, a dependência epistêmica não é uma estratégia cognitiva racional, mas uma abdicação da agência intelectual.

No caso específico dos LLMs, três características técnicas combinam-se para tornar esse risco estrutural. A primeira é o fenômeno das alucinações, sistematicamente documentado por Ji *et al.* (2023), que concluem:

Natural Language Generation (NLG) has improved exponentially in recent years thanks to the development of sequence-to-sequence deep learning technologies such as Transformer-based language models. This advancement has led to more fluent and coherent NLG, leading to improved development in downstream tasks such as abstractive summarization, dialogue generation, and data-to-text generation. However, it is also apparent that deep learning based generation is prone to hallucinate unintended text, which degrades the system performance and fails to meet user expectations in many real-world scenarios.

A segunda característica diz respeito à própria natureza dos modelos de linguagem. Bender *et al.* (2021, p. 616) definem o problema central ao afirmar que:

Contrary to how it may seem when we observe its output, an LM is a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning: a stochastic parrot.

A coerência sintática e a fluência discursiva dos LLMs funcionam, portanto, como simulacros de raciocínio — eles não argumentam, mas produzem sequências de texto que têm a aparência de argumentos.

A terceira característica é a opacidade epistêmica. Diferentemente de um artigo científico, que apresenta autor identificável, metodologia explícita, referências verificáveis e processo de revisão por pares, um LLM não oferece nenhuma dessas âncoras de verificação. O estudante recebe uma resposta, mas não tem acesso ao processo pelo qual ela foi gerada, às



fontes que a fundamentam (ou não), nem aos limites do conhecimento do modelo sobre aquele tema específico. Worledge, Hashimoto e Guestrin (2024, n.p.) sintetizam o problema ao afirmar que:

While large language models (LLMs) excel at synthesizing information, they do not provide reliable citation to sources, making it difficult to trace and verify the origins of the information they present. [...] Users prefer search engines for the variety of available sources, understanding of information provenance, and reliability of provided information.

### **Pensamento crítico no ensino: a perspectiva brasileira**

No contexto brasileiro, Pedro Demo (2011) oferece uma das formulações mais rigorosas dessa concepção ao associar o desenvolvimento intelectual à capacidade de pesquisa como princípio educativo. Para o autor, a aprendizagem genuína não é transmissão de conteúdo, mas reconstrução crítica do conhecimento — um processo ativo no qual o estudante não apenas recebe informações, mas as problematiza, compara perspectivas, identifica contradições e elabora sínteses próprias. Demo argumenta que predomina nas práticas educativas brasileiras a atitude do imitador, que copia, reproduz e faz prova, quando deveria impor-se a atitude de aprender pela elaboração própria. Esse diagnóstico, formulado antes da era dos LLMs, adquire uma acuidade impressionante à luz das práticas de uso de IA generativa observadas no ensino superior contemporâneo.

Freire (1996), por sua vez, articula o pensamento crítico à noção de autonomia intelectual como condição ética da formação humana. Para Freire, ensinar não é transferir conhecimento, mas criar possibilidades para que o educando construa e reconstrua o saber de forma autoral. A pedagogia da autonomia pressupõe um sujeito que pensa, que questiona, que se reconhece como agente de seu próprio processo formativo — e não como receptor passivo de respostas prontas, sejam elas fornecidas por um professor autoritário ou por um sistema de IA generativa.

### **IA generativa e educação: o estado do debate internacional**

Kasneci *et al.* (2023) identificaram, em seu amplo mapeamento das oportunidades e desafios dos modelos de linguagem para a educação, que entre os riscos centrais encontram-se as questões de dependência e de formação acrítica, e que o letramento para o uso responsável dessas ferramentas é condição indispensável para que seu potencial pedagógico se realize sem prejuízo ao desenvolvimento intelectual dos estudantes.

Esse ponto é especialmente relevante porque evidencia uma tensão que percorre todo o debate: os LLMs têm potencial genuíno como ferramentas de apoio à aprendizagem, para



auxiliar estudantes na decomposição de problemas complexos, na identificação de lacunas em seu próprio raciocínio ou na exploração inicial de temas que serão posteriormente aprofundados em fontes verificáveis. O problema não é a existência da ferramenta, mas a ausência de mediação pedagógica crítica em sua adoção. Esse diagnóstico se insere em um debate mais amplo sobre os riscos éticos da IA para a sociedade, no qual o avanço tecnológico não traz benefícios automaticamente: ele exige estruturas de governança, formação crítica e deliberação ética para que seus efeitos sejam socialmente desejáveis (FLORIDI *et al.*, 2018).

### **A tensão central: amplificação versus erosão cognitiva**

Essa tensão não é nova na história das tecnologias intelectuais. Lévy (1993) recorda que a própria escrita foi criticada por Platão, no *Fedro*, como uma ameaça à memória e ao pensamento genuíno: ao externalizar o conhecimento em texto, ela dispensaria o esforço de memorização e enfraqueceria a capacidade de raciocínio autônomo.

O caso dos LLMs, porém, apresenta uma diferença qualitativa importante em relação à escrita e a outras tecnologias intelectuais anteriores. A escrita externaliza o produto do pensamento, mas não o substitui. Os LLMs, em sua modalidade de uso mais difundida entre estudantes, externalizam o processo do pensamento: eles não apenas registram o raciocínio, mas o executam (BENDER *et al.*, 2021). Essa prática configura o que a literatura recente em psicologia cognitiva denomina *cognitive offloading* — a delegação sistemática de operações cognitivas a sistemas externos, fenômeno que pesquisas empíricas associam a reduções mensuráveis nas capacidades de pensamento crítico (GERLICH, 2025). Essa diferença implica que o uso irrefletido de LLMs não apenas transforma as formas de pensamento, como outras tecnologias fizeram, mas potencialmente dispensa o exercício do pensamento em contextos nos quais esse exercício é o objeto principal da formação.

## **2. METODOLOGIA**

O presente estudo caracteriza-se como uma pesquisa de natureza qualitativa e caráter exploratório, conduzida por meio de revisão bibliográfica integrativa com procedimento sistematizado de busca. A escolha desse desenho metodológico justifica-se pelo estágio ainda emergente do fenômeno investigado — o uso de *chatbots* generalistas baseados em LLMs como fonte primária de respostas no ensino superior — e pela natureza interdisciplinar do problema, que mobiliza contribuições da teoria educacional, da epistemologia social, da psicologia cognitiva e da literatura técnica em ciência da computação. A revisão integrativa permite articular, em um mesmo corpus analítico, obras teóricas fundantes e evidências empíricas recentes, oferecendo



um quadro interpretativo coerente sobre um fenômeno cuja consolidação ainda não dispõe de literatura estabelecida no contexto brasileiro.

### **Bases de dados e estratégia de busca**

O levantamento bibliográfico foi conduzido entre março de 2026 e maio de 2026 em quatro bases de dados estruturadas, selecionadas pela complementaridade de cobertura: SciELO e Portal de Periódicos da CAPES, para garantir a presença de produção acadêmica brasileira e ibero-americana; e PubMed, especificamente para os estudos de neurociência cognitiva que mobilizam metodologia experimental com mensuração neurofisiológica. O Google Scholar foi empregado como ferramenta de busca complementar, destinada a verificar a cobertura e a identificar eventuais estudos relevantes não recuperados nas bases estruturadas; em razão da natureza aproximada e instável de seus quantitativos, não foi computado no fluxo de contagem do protocolo de seleção. No Portal de Periódicos da CAPES, dada a amplitude da indexação, aplicaram-se os filtros de produção nacional e de periódicos revisados por pares, a fim de circunscrever os resultados à produção científica pertinente ao recorte investigado.

A estratégia de busca empregou descritores combinados por operadores booleanos, formulados em português e em inglês, conforme apresentado a seguir.

#### **Quadro 1. Strings de busca em inglês (PubMed e Google Scholar)**

IA generativa e educação: ("*large language models*" OR "*generative AI*" OR "*ChatGPT*") AND ("*higher education*" OR "*university students*")

Pensamento crítico e *cognitive offloading*: ("*large language models*" OR "*ChatGPT*") AND ("*critical thinking*" OR "*cognitive offloading*")

Dependência epistêmica: ("*large language models*" OR "*AI assistant*") AND ("*epistemic dependence*" OR "*epistemic autonomy*")

#### **Quadro 2. Strings de busca em português (SciELO, Portal CAPES, Google Scholar)**

IA generativa e educação: ("inteligência artificial generativa" OR "ChatGPT" OR "modelos de linguagem") AND ("ensino superior" OR "educação")

Pensamento crítico: ("inteligência artificial" OR "ChatGPT") AND ("pensamento crítico" OR "*autonomia intelectual*")

### **Crítérios de inclusão e exclusão**

Foram adotados os seguintes critérios de inclusão: (i) artigos científicos publicados em periódicos revisados por pares, capítulos de livros e *preprints* de plataformas reconhecidas (arXiv) cuja contribuição não estivesse ainda disponível em versão publicada — caso em que



sua condição não consolidada no sistema de revisão por pares foi explicitada e considerada na ponderação das evidências; (ii) obras de referência da teoria educacional brasileira e da epistemologia, ainda que anteriores ao recorte temporal, quando reconhecidas como fundantes do referencial teórico mobilizado; (iii) trabalhos publicados a partir de 2018, marco que coincide com a publicação do artigo seminal sobre a arquitetura *Transformer* (VASWANI *et al.*, 2017) e com a aceleração da difusão pública de LLMs; (iv) textos em português, inglês ou espanhol; (v) aderência temática direta ao problema investigado.

Como critérios de exclusão, foram descartados: (i) trabalhos centrados em aspectos exclusivamente técnicos de arquitetura ou *fine-tuning* de LLMs, sem implicação pedagógica; (ii) estudos voltados à educação básica (ensinos fundamental e médio), por escaparem ao recorte de ensino superior; (iii) materiais de divulgação não acadêmica, com exceção de fontes institucionais empregadas exclusivamente para documentação factual de plataformas (Khan Academy, Duolingo); (iv) artigos duplicados entre bases, mantida a versão de maior completude; (v) trabalhos cujo texto integral não esteve acessível.

### Processo de seleção

O processo de seleção foi conduzido em quatro fases, segundo lógica adaptada do protocolo PRISMA: identificação, triagem, elegibilidade e inclusão. Na fase de identificação, foram recuperados 1.910 registros nas quatro bases estruturadas consultadas (SciELO, 72; Portal CAPES, 1.125; PubMed, 713). Após a remoção de 362 duplicatas, sobretudo decorrentes da sobreposição entre a SciELO e o Portal CAPES, restaram 1.548 registros. A triagem por título e resumo, aplicando-se os critérios de inclusão e exclusão, resultou em 52 registros encaminhados para leitura integral. Na fase de elegibilidade, 31 trabalhos foram excluídos por não atenderem aos critérios de pertinência temática direta ao problema investigado. O corpus final analisado é composto por 21 obras, distribuídas entre referências teóricas fundantes (6), estudos empíricos sobre efeitos cognitivos dos LLMs (4) e literatura técnica e ético-regulatória sobre IA (11). A síntese do fluxo de seleção é apresentada no Quadro 3.

**Quadro 3.** Fluxo de seleção das obras (lógica adaptada do protocolo PRISMA)

Fase	Procedimento	Registros (n)
Identificação	Registros recuperados nas quatro bases estruturadas; remoção de 362 duplicatas	1.910
Triagem	Avaliação de 1.548 registros por título e resumo segundo os critérios de inclusão e exclusão	52
Elegibilidade	Leitura integral; exclusão de 31 trabalhos sem pertinência temática direta	21



Inclusão	Corpus final: teóricas fundantes (6), empíricas sobre efeitos cognitivos (4), técnica e ético-regulatória (11)	21
----------	--	----

Fonte: elaborado pelos autores (2026).

### Procedimento de análise

A análise do corpus foi conduzida de modo interpretativo e temático, organizada em torno de quatro eixos analíticos derivados do problema de pesquisa: (i) distinção conceitual entre IA generativa, LLMs, *chatbots* generalistas e tutores pedagógicos; (ii) condições em que a IA generativa opera como ferramenta mediadora *versus* substituta cognitiva; (iii) caracterização da dependência epistêmica e dos mecanismos pelos quais ela se instala no uso de LLMs; (iv) evidências empíricas sobre os efeitos cognitivos do uso de LLMs em contextos formativos.

As contribuições teóricas — em particular a teoria sócio-histórica de Vygotsky (1984), a pedagogia da autonomia de Freire (1996), a epistemologia social de Hardwig (1985) e a sociologia das tecnologias intelectuais de Lévy (1993) — foram cotejadas com a literatura empírica contemporânea, identificando-se convergências, tensões e lacunas entre os referenciais. Esse procedimento permitiu construir um argumento articulado sobre os riscos cognitivos do uso irrefletido de LLMs na formação universitária, base sobre a qual se desenvolve a discussão dos resultados na seção subsequente.

### 3. RESULTADOS E DISCUSSÃO

A revisão integrativa conduzida permitiu identificar um conjunto convergente de evidências empíricas e elaborações teóricas que, articuladas, sustentam o argumento central deste trabalho. A análise é organizada em três eixos, derivados dos objetivos específicos: (i) as evidências empíricas sobre os efeitos cognitivos do uso intensivo de LLMs por estudantes; (ii) a articulação dessas evidências ao quadro teórico mobilizado, com ênfase na distinção entre mediação e substituição cognitiva e no conceito de dependência epistêmica; e (iii) as implicações para o contexto formativo brasileiro.

#### Evidências empíricas dos efeitos cognitivos do uso intensivo de LLMs

A literatura empírica recente converge, a partir de desenhos metodológicos heterogêneos, para um diagnóstico preocupante sobre os efeitos cognitivos do uso intensivo de LLMs por estudantes. Em estudo experimental com participantes universitários distribuídos aleatoriamente entre uma condição assistida por ChatGPT e uma condição-controle, Georgiou (2025) reportou níveis significativamente inferiores de engajamento cognitivo no grupo assistido



após a realização de uma tarefa estruturada de escrita argumentativa. Os achados sugerem que a disponibilidade da assistência por IA induz o que o autor caracteriza como *cognitive offloading* — a transferência de operações cognitivas a sistemas externos —, mesmo quando o desempenho aparente nas tarefas avaliadas permanece equivalente entre os grupos.

A esse diagnóstico comportamental, Kosmyna *et al.* (2025) acrescentam uma camada neurofisiológica. Em estudo longitudinal de quatro meses com participantes distribuídos entre três grupos — usuários exclusivos de LLM, usuários de mecanismos de busca e participantes sem auxílio externo (*brain-only*) —, os autores empregaram eletroencefalografia (EEG) para mensurar conectividade cerebral durante tarefas de redação. Os resultados mostraram que o grupo *brain-only* exibiu redes neurais mais fortes e distribuídas; o grupo dos buscadores apresentou engajamento moderado; e o grupo dos usuários de LLM, a conectividade mais fraca dos três. Crucialmente, os usuários de LLM relataram o menor senso de autoria sobre os textos produzidos e demonstraram dificuldade em citar suas próprias produções — indícios comportamentais de que a delegação cognitiva ao modelo se traduziu não apenas em redução de esforço, mas em desconexão entre o estudante e o produto da própria atividade intelectual.

Em escala populacional mais ampla, Gerlich (2025) investigou, por meio de levantamento *survey* com participantes de diferentes faixas etárias, a relação entre frequência de uso de ferramentas de IA e desempenho em medidas de pensamento crítico. Os resultados indicaram correlação negativa significativa entre as duas variáveis, especialmente acentuada entre os participantes mais jovens, mediada pelo aumento de práticas de *cognitive offloading*. O estudo evidencia, em populações maiores e em contexto não-experimental, o mesmo padrão observado em condições controladas.

A confluência entre esses três trabalhos — um experimento comportamental, um estudo longitudinal com mensuração neural e um levantamento populacional —, conduzidos com desenhos metodológicos independentes, sugere uma convergência inicial de evidências de que o uso intensivo de LLMs como fonte primária de respostas se associa à redução do engajamento cognitivo e à possível erosão de capacidades de pensamento crítico. Cabe, contudo, registrar limitações importantes desses estudos. Todos foram conduzidos majoritariamente em contextos universitários anglo-saxões, sob práticas pedagógicas e expectativas formativas distintas das brasileiras; a janela temporal dos experimentos, ainda que estendida em Kosmyna *et al.* (2025), é curta diante da escala de uma trajetória formativa de graduação; e a estabilidade dessas associações em populações brasileiras permanece, no melhor conhecimento desta revisão, sem investigação empírica direta. Tais limites não invalidam o diagnóstico, mas indicam a urgência de pesquisa empírica nacional sobre o fenômeno.



### **Da evidência à teoria: substituição cognitiva e abdicação da agência intelectual**

As evidências reunidas dialogam diretamente com as distinções teóricas estabelecidas na fundamentação. A diferença vygotskiana entre ferramenta que medeia e ferramenta que executa em lugar do sujeito (VYGOTSKY, 1984), longe de ser uma especulação filosófica, é precisamente o eixo ao longo do qual os estudos empíricos discriminam efeitos cognitivos. O grupo *brain-only* de Kosmyna *et al.* (2025) é aquele cujo processo de elaboração permanece integralmente interno; o grupo de buscadores trabalha com ferramentas que externalizam o acesso à informação, mas não a operação cognitiva de síntese; o grupo dos usuários de LLM é aquele em que a ferramenta executa a operação intelectual em lugar do sujeito. A diferença de conectividade neural entre os três grupos é, nesse sentido, uma medida empírica daquilo que Lévy (1993) descreveu como o contraste entre externalizar o produto do pensamento e delegar o próprio processo.

Esse achado revaloriza a tese de Lévy de modo particular. O autor advertia que o valor cognitivo de uma tecnologia intelectual depende do modo como ela é incorporada às práticas do sujeito, e que uma mesma ferramenta pode ampliar ou empobrecer o pensamento conforme o uso. Os estudos empíricos atuais oferecem aquilo que essa tese teórica não podia oferecer em 1993: uma demonstração mensurável de que, no caso específico dos LLMs, a forma de uso predominante entre estudantes universitários é precisamente aquela que produz empobrecimento — não amplificação — das capacidades cognitivas.

A dimensão epistemológica do problema, por sua vez, é iluminada pela articulação entre Hardwig (1985) e os achados de Kosmyna *et al.* (2025) sobre a baixa percepção de autoria. Hardwig advertia que a dependência epistêmica pode degenerar em postura de aceitação passiva e acrítica daquilo que se recebe — uma abdicação da agência intelectual. O dado empírico de que usuários intensivos de LLM relatam menor senso de autoria sobre os próprios textos é, em termos hardwiguianos, exatamente isso: o estudante não se reconhece como produtor do conhecimento que sua produção registra. Não se trata apenas de um déficit cognitivo, mas de uma alteração na própria relação do sujeito com o saber, em que o lugar da elaboração própria é ocupado pela aceitação de uma elaboração externa que, por força das características técnicas dos LLMs — opacidade epistêmica, alucinações, ausência de referência ao significado (BENDER *et al.*, 2021; JI *et al.*, 2023) — não pode sequer ser objeto de verificação minimamente rigorosa por parte do estudante que a recebe.

### **Implicações para o contexto formativo brasileiro**

A leitura dessas evidências à luz da tradição educacional brasileira torna o diagnóstico ainda mais agudo. Demo (2011) sustenta que um dos problemas centrais da educação brasileira



é o predomínio da atitude do imitador — que copia, reproduz e faz prova — sobre a atitude do estudante que aprende pela elaboração própria. Formulado antes da consolidação dos LLMs, esse diagnóstico assume hoje uma dimensão técnica concreta que o autor não poderia antecipar: a delegação da elaboração a um sistema generativo dispensa, com eficiência inédita, exatamente o esforço cognitivo que Demo identifica como condição da formação. A diferença, em relação às formas tradicionais de cópia, é qualitativa. O texto produzido por um LLM é sintática e discursivamente sofisticado, próximo do que se esperaria de um estudante que efetivamente elaborou — o que torna a substituição cognitiva mais difícil de identificar pelo docente e mais facilmente naturalizada nas práticas estudantis.

Em registro complementar, Freire (1996) advertia que ensinar não é transferir conhecimento, mas criar possibilidades para que o educando construa e reconstrua o saber. Os achados empíricos sugerem que o uso irrefletido da IA generativa configura, em sentido preciso, uma transferência de respostas que dispensa o processo de construção — exatamente o contrário do que Freire descreve como ato educativo. Esse não é um problema da ferramenta em si, mas do modo de uso que tem se consolidado entre estudantes: uma prática que, em vez de mobilizar o LLM para apoiar um processo de aprendizagem que permanece do sujeito, o utiliza para dispensar esse processo inteiramente.

O diagnóstico aqui esboçado projeta-se sobre um conjunto de questões institucionais que extrapolam a relação individual entre estudante e ferramenta. A primeira diz respeito ao letramento algorítmico: compreender, ainda que em termos gerais, como um modelo de linguagem gera suas respostas — e por que essa geração não constitui referência ao significado nem garantia de veracidade (BENDER *et al.*, 2021) — é condição para o uso crítico dessas ferramentas e articula-se diretamente ao letramento para o uso responsável defendido por Kasneci *et al.* (2023). A segunda envolve a integridade acadêmica e a noção de autoria em contextos de coescrita com IA: quando parte da elaboração textual é delegada a um sistema generativo, torna-se necessário redefinir, em termos institucionais, o que se entende por trabalho autoral e quais usos são legítimos — questão que a baixa percepção de autoria relatada por usuários intensivos de LLMs (KOSMYNA *et al.*, 2025) torna ainda mais premente. A terceira refere-se às políticas universitárias de uso ético dessas tecnologias, que, para serem formativas e não apenas proibitivas, dependeriam de orientação pedagógica explícita sobre quando e como mobilizar a IA como mediadora — e não como substituta — do esforço cognitivo. Tais desdobramentos, embora situados além do recorte desta revisão, indicam direções concretas para a articulação entre o diagnóstico aqui apresentado e a prática docente e institucional.



Cabe, por fim, registrar uma lacuna específica desta revisão. A discussão pública sobre IA generativa na educação brasileira tem se desenvolvido majoritariamente em registro normativo — documentos institucionais, manifestos de associações, orientações pedagógicas —, sem o lastro de pesquisa empírica nacional que permitiria avaliar a extensão e as particularidades do fenômeno no contexto brasileiro. A ausência de estudos com rigor metodológico comparável aos trabalhos internacionais aqui mobilizados constitui, em si mesma, uma agenda de pesquisa, cuja urgência se acentua na medida em que a difusão das ferramentas avança em ritmo muito superior ao da produção de evidências sobre seus efeitos.

#### 4. CONSIDERAÇÕES FINAIS

O objetivo central deste estudo foi analisar de que modo o uso da inteligência artificial generativa como fonte primária de respostas afeta o desenvolvimento do pensamento crítico e pode conduzir à dependência epistêmica em estudantes do ensino superior. Retornando à pergunta que orientou a investigação, é possível afirmar que, quando desprovida de mediação pedagógica e de letramento crítico, essa prática tende a posicionar a tecnologia como substituta cognitiva, enfraquecendo a avaliação autônoma, a problematização e a elaboração própria — justamente as capacidades que caracterizam a formação universitária.

Os resultados da análise bibliográfica indicam que a distinção entre ferramenta mediadora e substituta é decisiva. O uso que transfere ao modelo de linguagem as operações de sintetizar e argumentar configura um *cognitive offloading* associado, conforme a literatura, à redução do pensamento crítico. Três características técnicas dos grandes modelos de linguagem — alucinações, ausência de referência ao significado e opacidade epistêmica — agravam esse quadro, criando uma assimetria na qual o estudante recebe respostas retoricamente convincentes sem dispor de meios para verificar sua procedência. Assim, a dependência epistêmica deixa de ser uma estratégia racional de divisão do trabalho cognitivo e degenera em aceitação acrítica.

Uma implicação prática desta pesquisa é que a simples incorporação institucional dos modelos de linguagem não garante benefícios formativos; é necessário que as instituições e os educadores promovam o letramento crítico para essas ferramentas, criando contextos em que a IA atue como mediadora, e não como substituta do esforço cognitivo.

O escopo deste estudo limitou-se a uma revisão bibliográfica de natureza qualitativa e exploratória, sem coleta de dados empíricos primários. Cabe ressaltar, por conseguinte, que as conclusões aqui apresentadas derivam de interpretação da literatura mobilizada, e não de verificação empírica original conduzida pelos autores, o que recomenda cautela quanto a



generalizações de caráter causal. Em pesquisas futuras, sugere-se a realização de investigações longitudinais que mensurem os efeitos do uso contínuo da IA generativa sobre a argumentação, a leitura crítica e a produção textual, bem como estudos aplicados que testem diferentes modelos de integração curricular, a fim de determinar as condições pedagógicas mais eficazes para que essas tecnologias amplifiquem, em vez de erodirem, o desenvolvimento cognitivo.

## REFERÊNCIAS

BENDER, Emily M.; GEBRU, Timnit; McMILLAN-MAJOR, Angelina; SHMITCHELL, Shmargaret. On the dangers of stochastic parrots: can language models be too big? In: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAcT), 2021, Virtual Event, Canada. **Proceedings...** New York: Association for Computing Machinery, 2021. p. 610–623. DOI: <https://doi.org/10.1145/3442188.3445922>. Disponível em: <https://dl.acm.org/doi/epdf/10.1145/3442188.3445922>. Acesso em: 25 mar. 2026.

BROWN, Tom B.; MANN, Benjamin; RYDER, Nick; SUBBIAH, Melanie; KAPLAN, Jared; DHARIWAL, Prafulla; NEELAKANTAN, Arvind; SHYAM, Pranav; SASTRY, Girish; ASKELL, Amanda; AGARWAL, Sandhini; HERBERT-VOSS, Ariel; KRUEGER, Gretchen; HENIGHAN, Tom; CHILD, Rewon; RAMESH, Aditya; ZIEGLER, Daniel M.; WU, Jeffrey; WINTER, Clemens; HESSE, Christopher; CHEN, Mark; SIGLER, Eric; LITWIN, Mateusz; GRAY, Scott; CHESSE, Benjamin; CLARK, Jack; BERNER, Christopher; McCANDLISH, Sam; RADFORD, Alec; SUTSKEVER, Ilya; AMODEI, Dario. Language models are few-shot learners. **arXiv**, 2020. DOI: <https://doi.org/10.48550/arXiv.2005.14165>. Disponível em: <https://arxiv.org/abs/2005.14165>. Acesso em: 22 maio 2026.

CAUGHILL, Patrick. **Sal Khan wants to give every student on Earth a personal AI tutor**. Freethink, 2023. Disponível em: <https://www.freethink.com/consumer-tech/khanmigo-ai-tutor>. Acesso em: 25 mar. 2026.

DEMO, Pedro. **Pesquisa**: princípio científico e educativo. 14. ed. São Paulo: Cortez, 2011.

DUOLINGO. Duolingo Max shows the future of AI education. **Globe Newswire**, 14 mar. 2023. Disponível em: <https://investors.duolingo.com/news-releases/news-release-details/duolingo-max-shows-future-ai-education>. Acesso em: 30 mar. 2026.

FLORIDI, Luciano; COWLS, Josh; BELTRAMETTI, Monica; CHATILA, Raja; CHAZERAND, Patrice; DIGNUM, Virginia; LUETGE, Christoph; MADELIN, Robert; PAGALLO, Ugo; ROSSI, Francesca; SCHAFFER, Burkhard; VALCKE, Peggy; VAYENA, Effy. AI4People: an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. **Minds and Machines**, Dordrecht, v. 28, n. 4, p. 689–707, dez. 2018. DOI: <https://doi.org/10.1007/s11023-018-9482-5>. Disponível em: <https://link.springer.com/article/10.1007/s11023-018-9482-5>. Acesso em: 30 mar. 2026.

FREIRE, Paulo. **Pedagogia da autonomia**: saberes necessários à prática educativa. São Paulo: Paz e Terra, 1996.

GEORGIU, Georgios P. ChatGPT produces more “lazy” thinkers: evidence of cognitive engagement decline. **arXiv**, 2025. DOI: <https://doi.org/10.48550/arXiv.2507.00181>. Disponível em: <https://arxiv.org/pdf/2507.00181>. Acesso em: 1 maio 2026.

GERLICH, Michael. AI tools in society: impacts on cognitive offloading and the future of critical thinking. **Societies**, Basel, v. 15, n. 1, artigo 6, jan. 2025. DOI:



<https://doi.org/10.3390/soc15010006>. Disponível em: <https://www.mdpi.com/2075-4698/15/1/6>. Acesso em: 1 abr. 2026.

HARDWIG, John. Epistemic dependence. **The Journal of Philosophy**, New York, v. 82, n. 7, p. 335–349, jul. 1985. DOI: <https://doi.org/10.2307/2026523>. Acesso em: 1 abr. 2026.

JI, Ziwei; LEE, Nayeon; FRIESKE, Rita; YU, Tiezheng; SU, Dan; XU, Yan; ISHII, Etsuko; BANG, Yejin; MADOTTO, Andrea; FUNG, Pascale. Survey of hallucination in natural language generation. **ACM Computing Surveys**, New York, v. 55, n. 12, artigo 248, mar. 2023. DOI: <https://doi.org/10.1145/3571730>. Disponível em: <https://dl.acm.org/doi/10.1145/3571730>. Acesso em: 3 abr. 2026.

KASNECI, Enkelejda; SEßLER, Kathrin; KÜCHEMANN, Stefan; BANNERT, Maria; DEMENTIEVA, Daryna; FISCHER, Frank; GASSER, Urs; GROH, Georg; GÜNNEMANN, Stephan; HÜLLERMEIER, Eyke; KRUSCHE, Stephan; KUTYNIOK, Gitta; MICHAELI, Tilman; NERDEL, Claudia; PFEFFER, Jürgen; POQUET, Oleksandra; SAILER, Michael; SCHMIDT, Albrecht; SEIDEL, Tina; STADLER, Matthias; WELLER, Jochen; KUHN, Jochen; KASNECI, Gjergji. ChatGPT for good? On opportunities and challenges of large language models for education. **Learning and Individual Differences**, Amsterdam, v. 103, artigo 102274, abr. 2023. DOI: <https://doi.org/10.1016/j.lindif.2023.102274>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1041608023000195>. Acesso em: 9 abr. 2026.

KENSKI, Vani Moreira. **Educação e tecnologias: o novo ritmo da informação**. Campinas: Papirus, 2007.

KHAN ACADEMY. **Meet Khanmigo**. 2023. Disponível em: <https://www.khanmigo.ai>. Acesso em: 13 abr. 2026.

KOSMYNA, Nataliya; HAUPTMANN, Eugene; YUAN, Ye Tong; SITU, Jessica; LIAO, Xian-Hao; BERESNITZKY, Ashly Vivian; BRAUNSTEIN, Iris; MAES, Pattie. Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task. **arXiv**, 2025. DOI: <https://doi.org/10.48550/arXiv.2506.08872>. Disponível em: <https://arxiv.org/abs/2506.08872>. Acesso em: 4 maio 2026.

LÉVY, Pierre. **As tecnologias da inteligência: o futuro do pensamento na era da informática**. Tradução: Carlos Irineu da Costa. Rio de Janeiro: Editora 34, 1993. (Coleção TRANS).

VANLEHN, Kurt. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. **Educational Psychologist**, Mahwah, v. 46, n. 4, p. 197–221, out. 2011. DOI: <https://doi.org/10.1080/00461520.2011.611369>. Acesso em: 22 maio 2026.

VASWANI, Ashish; SHAZEER, Noam; PARMAR, Niki; USZKOREIT, Jakob; JONES, Llion; GOMEZ, Aidan N.; KAISER, Łukasz; POLOSUKHIN, Illia. Attention is all you need. In: **ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS (NeurIPS)**, 31., 2017, Long Beach, CA. **Proceedings...** Red Hook: Curran Associates, 2017. p. 5998–6008. Disponível em: <https://papers.nips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>. Acesso em: 22 maio 2026.

VYGOTSKY, Lev Semenovich. **A formação social da mente: o desenvolvimento dos processos psicológicos superiores**. Organização: Michael Cole *et al.* Tradução: José Cipolla Neto, Luís Silveira Menna Barreto, Solange Castro Afeche. São Paulo: Martins Fontes, 1984.

WORLEDGE, Theodora; HASHIMOTO, Tatsunori; GUESTRIN, Carlos. The extractive-abstractive spectrum: uncovering verifiability trade-offs in LLM generations. **arXiv**, 2024. DOI: <https://doi.org/10.48550/arXiv.2411.17375>. Disponível em: <https://arxiv.org/abs/2411.17375>. Acesso em: 2 maio 2026.