

**LETRAMENTO EM IA (AI
LITERACY) NO ENSINO
BÁSICO: ESTRATÉGIAS
CURRICULARES PARA
DESENVOLVER A
COMPREENSÃO CRÍTICA
DOS ESTUDANTES SOBRE
VIESES ALGORÍTMICOS,
ALUCINAÇÕES DE IA E
ÉTICA NA INTERAÇÃO COM
ASSISTENTES VIRTUAIS**

**AI LITERACY IN BASIC EDUCATION: CURRICULAR STRATEGIES TO
DEVELOP STUDENTS' CRITICAL UNDERSTANDING OF ALGORITHMIC BIAS,
AI HALLUCINATIONS AND ETHICS IN INTERACTIONS WITH VIRTUAL
ASSISTANTS**

Ciências Humanas, Ciências Sociais Aplicadas, Linguística & Letras e

Artes

• 03/10/2026

REGISTRO DOI: [10.70773/revistatopicos/790790394](https://doi.org/10.70773/revistatopicos/790790394)

Galindo Pedro Ramos

RESUMO

A difusão de sistemas de inteligência artificial generativa e de assistentes virtuais modificou as práticas de pesquisa, produção textual, tomada de decisão e interação digital de crianças e adolescentes, criando a necessidade de incorporar o letramento em IA à educação básica. Este artigo analisa estratégias curriculares para desenvolver compreensão crítica sobre vieses algorítmicos, alucinações de IA e ética na interação com assistentes virtuais. Trata-se de estudo teórico-documental, de abordagem qualitativa, baseado em marcos internacionais recentes, legislação educacional brasileira e literatura científica sobre AI literacy em contextos K-12. A análise articula o Marco Referencial de Competências em IA para Estudantes da UNESCO, o AI Literacy Framework para educação primária e secundária da OCDE e Comissão Europeia, a Política Nacional de Educação Digital e as Normas sobre Computação na Educação Básica. Sustenta-se que letramento em IA não deve ser reduzido ao uso instrumental de ferramentas nem ao domínio de prompts, mas incluir fundamentos sobre dados e modelos, avaliação de confiabilidade, reconhecimento de vieses, proteção de dados, autoria, responsabilidade e limites da automação. Como contribuição didática, propõe-se o protocolo VERIFICA-IA e uma progressão curricular por etapas escolares, com atividades de auditoria de respostas, comparação de fontes, experimentos sobre classificação e recomendação, debates éticos e registro transparente do uso de IA. Conclui-se que a escola deve formar estudantes capazes de usar, questionar e recusar sistemas de IA de maneira informada, preservando agência humana, pensamento crítico e responsabilidade digital.

Palavras-chave: letramento em inteligência artificial; educação básica; viés algorítmico; alucinações de IA; ética digital; assistentes virtuais.

ABSTRACT

The spread of generative artificial intelligence systems and virtual assistants has changed how children and adolescents search for information, write, make decisions and interact digitally, creating the need to incorporate AI literacy into basic education. This article analyzes curricular strategies to develop critical understanding of algorithmic bias, AI hallucinations and ethics in interactions with virtual assistants. It is a qualitative theoretical-documentary study based on recent international frameworks, Brazilian educational legislation and scientific literature on AI literacy in K-12 contexts. The analysis connects UNESCO's AI Competency Framework for Students, the OECD and European Commission's AI Literacy Framework for primary and secondary education, Brazil's National Digital Education Policy and national standards for Computing in Basic Education. The article argues that AI literacy should not be reduced to instrumental tool use or prompt engineering; it should include foundations about data and models, reliability assessment, bias recognition, data protection, authorship, accountability and the limits of automation. As a didactic contribution, the VERIFICA-IA protocol and a school-stage curricular progression are proposed, combining response auditing, source comparison, classification and recommendation experiments, ethical debates and transparent documentation of AI use. It concludes that schools should prepare students to use, question and refuse AI systems in informed ways while preserving human agency, critical thinking and digital responsibility.

Keywords: artificial intelligence literacy; basic education; algorithmic bias; AI hallucinations; digital ethics; virtual assistants.

1. INTRODUÇÃO

A presença da inteligência artificial no cotidiano escolar deixou de depender da adoção institucional de plataformas especializadas. Sistemas de recomendação, mecanismos de busca, tradutores, filtros, corretores, chatbots, geradores de imagem e assistentes virtuais já integram o ecossistema digital no qual crianças e adolescentes estudam, se comunicam e tomam decisões. A popularização da inteligência artificial generativa ampliou esse contato ao permitir que estudantes conversem diretamente com sistemas capazes de produzir textos, explicações, imagens, códigos, sínteses e sugestões em linguagem natural. Essa acessibilidade modificou a pergunta educacional: não se trata mais de decidir se os estudantes terão contato com IA, mas de definir quais competências precisam desenvolver para compreender, avaliar e utilizar criticamente sistemas que já fazem parte de sua vida.

Esse deslocamento exige diferenciar uso de letramento. Saber pedir a um chatbot que resuma um texto ou resolva um exercício não significa compreender como uma resposta foi gerada, quais dados e escolhas influenciaram o sistema, em que situações o conteúdo pode estar errado ou enviesado e quem assume responsabilidade pelo uso da informação. O letramento em IA envolve conhecimento, habilidades e atitudes que permitem reconhecer sistemas de IA, compreender princípios básicos de funcionamento, avaliar resultados, tomar decisões éticas e preservar agência humana diante de processos automatizados. Em outras palavras, a escola precisa ensinar não apenas a usar IA, mas a pensar sobre IA e a viver criticamente com IA.

O debate ganhou densidade institucional nos últimos anos. A UNESCO publicou, em 2024, um marco referencial para estudantes com 12 competências distribuídas em quatro dimensões -

mentalidade centrada no ser humano, ética da IA, técnicas e aplicações de IA e design de sistemas de IA - organizadas nos níveis compreender, aplicar e criar (Miao; Shiohira; Lao, 2024). Em 2026, a OCDE e a Comissão Europeia apresentaram um framework específico para educação primária e secundária que define AI literacy como conjunto de conhecimentos, habilidades e atitudes necessários para compreender como sistemas de IA funcionam, avaliar criticamente seus resultados e utilizá-los de forma ética e criativa (OECD; European Commission, 2026). Esses documentos convergem ao rejeitar uma visão meramente operacional da tecnologia.

No Brasil, a Política Nacional de Educação Digital estabelece como estratégias da Educação Digital Escolar o desenvolvimento de competências para atuação responsável na sociedade conectada, a ética aplicada ao ambiente digital, o letramento midiático e a cidadania digital. A Lei nº 14.533/2023 também incorporou à LDB a educação digital, com foco em letramento digital, computação, programação, robótica e outras competências digitais. A Resolução CNE/CEB nº 1/2022, por sua vez, instituiu normas sobre Computação na Educação Básica como complemento à BNCC e determinou que currículos e formação docente considerem competências e habilidades específicas de computação (Brasil, 2022; Brasil, 2023). Embora esses marcos não esgotem o tema da IA generativa, oferecem base curricular para sua integração crítica.

A urgência curricular é reforçada por três problemas que se tornaram particularmente visíveis com modelos generativos: vieses algorítmicos, alucinações e relações éticas com assistentes virtuais. Sistemas podem reproduzir desigualdades presentes em dados ou em decisões de projeto; modelos de linguagem podem produzir

afirmações plausíveis, porém falsas ou sem sustentação; e interfaces conversacionais podem estimular antropomorfização, confiança excessiva, exposição indevida de dados e terceirização de decisões. Esses riscos não justificam proibição indiscriminada, mas tornam insuficiente a simples autorização de uso. É preciso desenvolver competência para detectar limitações, verificar evidências e decidir quando recorrer - ou não - a um sistema de IA.

A literatura sobre K-12 mostra que a educação em IA pode favorecer compreensão conceitual, resolução de problemas e reflexão ética quando estruturada pedagogicamente. Casal-Otero et al. (2023), em revisão de 179 documentos, identificaram forte diversidade de experiências e a necessidade de referenciais curriculares com continuidade entre etapas. Lee e Kwon (2024), em revisão sistemática de 25 estudos, observaram que atividades práticas e projetos são frequentes e podem favorecer letramento em IA e reflexão sobre impactos sociais. Liu e Zhong (2024), ao revisar 45 estudos empíricos, destacaram a predominância de experiências de curta duração e a necessidade de desenho instrucional mais robusto e longitudinal. O problema, portanto, não é apenas inserir IA no currículo, mas fazê-lo com progressão, intencionalidade e avaliação.

O objetivo deste artigo é analisar estratégias curriculares para desenvolver letramento crítico em IA no ensino básico, com ênfase em vieses algorítmicos, alucinações de IA e ética na interação com assistentes virtuais. Busca-se: a) delimitar o conceito de AI literacy para fins escolares; b) relacioná-lo ao marco curricular brasileiro; c) propor objetivos de aprendizagem progressivos para anos iniciais, anos finais e ensino médio; d) apresentar metodologias de ensino e avaliação; e) discutir formação docente e governança escolar; e f)

sistematizar um protocolo didático de verificação crítica aplicável ao uso cotidiano de IA pelos estudantes.

2. LETRAMENTO EM IA: CONCEITO, DIMENSÕES E FINALIDADES FORMATIVAS

O conceito de letramento em IA surgiu em diálogo com letramento digital, letramento de dados, pensamento computacional e educação midiática. A expressão procura responder a uma característica particular da inteligência artificial: muitos sistemas operam de maneira parcialmente invisível, transformando dados em classificações, previsões, recomendações ou conteúdos sem que o usuário conheça detalhadamente os procedimentos internos. Por isso, ser alfabetizado digitalmente não garante, por si só, compreensão suficiente sobre probabilidades, treinamento de modelos, limites de generalização, vieses, opacidade ou geração de conteúdo sintético.

Casal-Otero et al. (2023) sintetizam o campo em três eixos: aprender sobre IA, aprender como a IA funciona e aprender para a vida com IA. A primeira dimensão envolve reconhecer aplicações e conceitos; a segunda aproxima o estudante de dados, modelos, treinamento e inferência; a terceira examina consequências sociais, éticas e cidadãs. Essa arquitetura é produtiva para a educação básica porque impede duas reduções frequentes: tratar IA apenas como programação ou, no extremo oposto, abordá-la apenas como tema abstrato de ética sem qualquer compreensão técnica mínima.

O marco da UNESCO amplia essa visão ao associar conhecimento técnico a mentalidade centrada no ser humano. Nesse referencial, estudantes devem desenvolver julgamento crítico sobre soluções de

IA, compreender responsabilidades de cidadania, construir conhecimentos fundamentais para aprendizagem contínua e participar de práticas inclusivas e sustentáveis de design. A progressão compreender-aplicar-criar também é relevante para o currículo: primeiro o estudante reconhece conceitos e problemas, depois os utiliza em situações concretas e, em níveis mais avançados, participa da criação ou avaliação de soluções (Miao; Shiohira; Lao, 2024).

O framework da OCDE e Comissão Europeia, publicado em 2026, reforça que AI literacy é um conjunto de conhecimentos, habilidades e atitudes. Essa formulação é particularmente importante para a escola porque evita transformar o tema em disciplina exclusivamente informativa. Conhecer o termo 'viés algorítmico' não basta se o estudante aceita automaticamente uma recomendação enviesada; saber que chatbots podem errar é insuficiente se não se desenvolve o hábito de verificar informações. O currículo precisa produzir disposições: curiosidade, ceticismo proporcional, responsabilidade, cuidado com dados, reconhecimento de incerteza e disposição para manter seres humanos no circuito decisório.

A ascensão da IA generativa acrescentou uma camada específica ao letramento. Sistemas baseados em grandes modelos de linguagem produzem respostas por previsão de padrões linguísticos e podem gerar explicações convincentes mesmo quando não possuem base factual adequada. A fluidez da linguagem tende a aumentar a confiança do usuário porque precisão verbal é facilmente confundida com precisão epistêmica. Assim, o letramento em IA generativa exige competências de formulação de perguntas, avaliação de evidências, checagem de referências, reconhecimento

de incerteza, documentação do uso da ferramenta e preservação de autoria intelectual.

Por esse motivo, o ensino de prompt engineering, embora útil como habilidade operacional, não pode ocupar o centro do currículo. Um estudante pode aprender a elaborar comandos sofisticados e continuar vulnerável a desinformação, vieses e dependência cognitiva. A escola deve ensinar a interação com IA como um processo de investigação: formular uma demanda, interpretar a resposta, identificar o que precisa ser verificado, comparar fontes, revisar a própria compreensão e decidir o que pode ou não ser incorporado ao trabalho. O valor educacional está menos em obter respostas perfeitas e mais em tornar visível o processo crítico de validação.

Quadro 1. Dimensões curriculares do letramento em IA

Dimensão	Pergunta formativa	Competências escolares esperadas
Fundamentos	O que é IA e como produz resultados?	Reconhecer sistemas de IA; compreender dados, treinamento, inferência, probabilidade e limites de modelos em nível adequado à idade.
Avaliação crítica	A resposta é confiável e adequada?	Distinguir fluência de veracidade; verificar fatos; identificar incerteza; comparar fontes; reconhecer alucinações.
Justiça e viés	Quem pode ser favorecido ou prejudicado?	Analisar representatividade de dados, classificações, estereótipos, erros diferenciais e efeitos sobre grupos.

Ética e direitos	O uso é legítimo e responsável?	Proteger dados; respeitar autoria; reconhecer limites de automação; considerar segurança, privacidade, transparência e responsabilidade.
Agência humana	Quem decide e responde pela decisão?	Usar IA como apoio, não como autoridade; justificar escolhas; saber recusar recomendações e buscar ajuda humana.
Criação responsável	Como produzir com IA sem apagar autoria e aprendizagem?	Documentar o processo, revisar conteúdo, atribuir uso de IA, preservar voz própria e avaliar impactos do produto.

Fonte: elaboração própria com base em Miao, Shiohira e Lao (2024) e OECD/European Commission (2026).

3. METODOLOGIA

O estudo adota abordagem qualitativa e natureza teórico-documental, com síntese narrativa de literatura e documentos normativos. A estratégia foi escolhida porque o objetivo não é estimar efeito médio de uma intervenção específica, mas articular referenciais de competências, pesquisas educacionais e normas brasileiras para construir uma proposta curricular aplicável à educação básica.

Foram priorizados documentos de organismos internacionais que apresentassem marcos explícitos de competências ou recomendações para IA e educação, particularmente UNESCO, OCDE, Comissão Europeia e UNICEF. No plano nacional, foram considerados a Base Nacional Comum Curricular, as Normas sobre Computação na Educação Básica e a Política Nacional de Educação Digital. A literatura científica foi selecionada para sustentar quatro

eixos: conceituação de AI literacy; ensino de IA em K-12; vieses algorítmicos; confiabilidade e alucinações em modelos generativos; e ética de chatbots e assistentes virtuais.

A análise bibliográfica deu preferência a revisões sistemáticas, revisões de escopo e estudos de referência publicados em periódicos científicos, além de trabalhos diretamente relacionados a competências em IA. Entre as evidências recentes utilizadas encontram-se as revisões de Casal-Otero et al. (2023), Lee e Kwon (2024), Liu e Zhong (2024), Yim (2024), Park (2025), Tao et al. (2026) e estudos específicos sobre justiça algorítmica, alucinações e uso ético de chatbots.

A síntese foi organizada em três núcleos de risco e competência: viés algorítmico, alucinação/confiabilidade e ética da interação. Em seguida, esses núcleos foram transpostos para objetivos de aprendizagem progressivos em três segmentos: anos iniciais do ensino fundamental, anos finais do ensino fundamental e ensino médio. A proposta não pretende substituir currículos de redes ou a BNCC; funciona como arquitetura de integração transversal que pode ser incorporada a Computação, Língua Portuguesa, Matemática, Ciências Humanas, Ciências da Natureza, projetos interdisciplinares e educação midiática.

Como limite metodológico, este artigo não constitui revisão sistemática PRISMA e não apresenta contagem própria de estudos identificados, incluídos ou excluídos. As quantificações mencionadas pertencem às revisões publicadas e são atribuídas aos respectivos autores. As estratégias curriculares propostas possuem fundamentação teórica e documental, mas demandam validação

empírica em diferentes redes, faixas etárias e condições de infraestrutura.

4. MARCO CURRICULAR BRASILEIRO E O LUGAR DO LETRAMENTO EM IA

A incorporação do letramento em IA à educação básica brasileira pode ser realizada sem criar, necessariamente, uma disciplina isolada. A BNCC já estabelece competências gerais relacionadas a cultura digital, argumentação, pensamento científico, comunicação, responsabilidade e cidadania. A Resolução CNE/CEB nº 1/2022 acrescenta um referencial específico de Computação, enquanto a Política Nacional de Educação Digital explicita a necessidade de competências para atuação responsável em ambientes conectados, ética digital, lógica, algoritmos, letramento midiático e cidadania (Brasil, 2022; Brasil, 2023).

Esses marcos permitem compreender AI literacy como desenvolvimento curricular transversal. O tema pode aparecer na análise de fontes em História, na avaliação de argumentos em Língua Portuguesa, na discussão de correlações e probabilidades em Matemática, na modelagem e classificação em Computação, na ética de dados em Filosofia e Sociologia, no estudo de saúde e desinformação em Ciências, ou na produção multimodal em Arte. A transversalidade, entretanto, não pode significar ausência de responsabilidade. A escola precisa definir uma progressão mínima, registrar objetivos e avaliar aprendizagens específicas.

A Política Nacional de Educação Digital é especialmente relevante porque vincula educação digital à atuação responsável na sociedade conectada e inclui direitos digitais e proteção de dados das

populações vulneráveis. Essa base permite abordar IA não apenas como inovação tecnológica, mas como tema de cidadania. Quando um algoritmo de recomendação decide quais conteúdos um adolescente vê, quando um sistema de classificação atribui risco ou quando um chatbot responde a uma dúvida escolar, questões de autonomia, privacidade e poder informacional fazem parte do processo educativo.

A articulação com Computação na Educação Básica também evita que o estudante trate a IA como magia. Mesmo sem entrar em matemática avançada, é possível construir modelos mentais adequados: sistemas aprendem padrões a partir de dados; classificadores podem errar; dados incompletos produzem representações incompletas; modelos generativos estimam sequências prováveis; saídas dependem do contexto e da instrução; resultados precisam ser avaliados. Esse conhecimento conceitual funciona como proteção contra antropomorfização e confiança cega.

Um currículo nacional ou de rede deve ainda preservar o princípio da equidade. Letramento em IA não pode pressupor acesso individual contínuo a plataformas comerciais. Atividades desplugadas, análise de casos, dados impressos, simulações humanas de algoritmos, uso compartilhado de dispositivos e auditoria de respostas previamente geradas podem cumprir objetivos importantes. A escola deve evitar transformar desigualdade de acesso a ferramentas em desigualdade de aprendizagem sobre IA.

Quadro 2. Articulação entre marcos educacionais e objetivos de AI literacy

Referencial	Ênfase relevante	Possível tradução curricular
BNCC	Cultura digital, pensamento crítico, argumentação, responsabilidade e cidadania.	Avaliar informação gerada por IA; justificar decisões; produzir conteúdo com responsabilidade.
Computação na Educação Básica – Res. CNE/CEB nº 1/2022	Pensamento computacional, mundo digital e cultura digital.	Compreender dados, algoritmos, automação, modelos, representação e efeitos sociais.
Política Nacional de Educação Digital – Lei nº 14.533/2023	Atuação responsável, ética digital, letramento midiático, direitos digitais e proteção de dados.	Privacidade, autoria, segurança, transparência, verificação de conteúdo e cidadania algorítmica.
UNESCO AI CFS – 2024	Mentalidade centrada no ser humano, ética, técnicas/aplicações e design de IA.	Progressão compreender-aplicar-criar com visão sociotécnica.
OECD/EC AI Literacy Framework – 2026	Conhecimentos, habilidades e atitudes para compreender, avaliar e usar IA.	Currículo longitudinal com avaliação crítica e uso responsável.

Fonte: elaboração própria.

5. VIESES ALGORÍTMICOS: DO CONCEITO À EXPERIÊNCIA CURRICULAR

Viés algorítmico é um termo amplo que designa padrões sistemáticos de resultado capazes de produzir tratamento desigual, representação distorcida ou desempenho diferente entre grupos. O problema pode surgir em múltiplos pontos: seleção de dados,

definição de categorias, rotulagem, escolha de variáveis, objetivos de otimização, contexto de uso e interpretação humana. Por isso, atribuir todo viés 'ao algoritmo' pode ocultar decisões humanas e institucionais anteriores. Para fins escolares, é mais adequado ensinar IA como sistema sociotécnico: dados, modelos, pessoas, instituições e contextos interagem na produção de consequências.

Baker e Hawn (2022) revisam evidências de viés algorítmico em educação e demonstram que raça/etnia, gênero, nacionalidade, condição socioeconômica e deficiência estão entre categorias relevantes em sistemas de IA educacional. Os autores chamam atenção para o fato de que vieses podem ser produzidos em diferentes etapas do pipeline e defendem transição de uma visão abstrata de fairness para preocupações concretas de equidade. Esse enfoque é pedagogicamente poderoso: estudantes podem perguntar não apenas se um modelo acerta em média, mas se erra de forma semelhante para todos.

Uma atividade introdutória não precisa utilizar código. O professor pode apresentar um conjunto fictício de dados de estudantes e pedir que a turma construa uma regra de seleção para uma bolsa. Ao modificar quais variáveis estão disponíveis, torna-se possível discutir proxies, dados ausentes e critérios aparentemente neutros que produzem efeitos desiguais. Em seguida, os estudantes podem comparar diferentes regras e justificar quais informações deveriam ou não ser utilizadas. O foco não é chegar a um único algoritmo 'justo', mas compreender que toda classificação incorpora escolhas.

Nos anos iniciais, a ideia pode ser explorada por meio de exemplos de classificação. Se um sistema é treinado apenas com imagens de maçãs vermelhas, como reagirá diante de maçãs verdes? Se um

reconhecimento de voz aprende com poucos tipos de sotaque, quem pode ser entendido com mais dificuldade? A atividade ajuda a construir a noção de representatividade de dados sem exigir formalização estatística. O professor deve evitar associar identidades reais a desempenho inferior de forma estigmatizante; o objetivo é mostrar que dados incompletos limitam sistemas.

Nos anos finais, estudantes podem auditar recomendações, buscas e respostas gerativas. Podem, por exemplo, formular perguntas equivalentes alterando nomes, profissões, regiões ou perfis e observar diferenças de linguagem. Resultados não devem ser tratados como prova definitiva de discriminação, pois sistemas generativos variam e o experimento pode não ser controlado; ainda assim, a comparação funciona como ponto de partida para discutir estereótipos, amostragem, correlação e necessidade de investigação sistemática.

No ensino médio, a abordagem pode avançar para métricas de erro, confusão entre precisão global e equidade, seleção de variáveis, proxies e trade-offs. Um modelo hipotético de previsão de evasão pode ser avaliado por subgrupos: quem recebe falso positivo? quem deixa de receber apoio por falso negativo? A questão curricular deve deslocar-se da pergunta 'o algoritmo é bom?' para 'bom para qual finalidade, segundo qual métrica, para quais grupos e com quais mecanismos de contestação?'.

A educação ética sobre viés precisa ainda contemplar o papel do usuário. Assistentes virtuais podem reforçar estereótipos quando prompts pedem generalizações ou quando o usuário aceita respostas sem contestação. O estudante deve aprender a reformular perguntas, exigir evidências, identificar categorias essencialistas e

reconhecer quando a resposta atribui características a grupos sem fundamento. Essa competência articula letramento em IA, educação antidiscriminatória e letramento midiático.

Quadro 3. Estratégias didáticas para trabalhar vieses algorítmicos

Estratégia	Como funciona	Aprendizagem visada
Classificação desplugada	Turma cria regras para classificar cartões ou exemplos e depois testa casos não previstos.	Compreender dados, categorias, generalização e casos-limite.
Conjunto de dados incompleto	Professor oferece dados com grupos sub-representados e compara desempenho das regras.	Relacionar representatividade a erro e injustiça potencial.
Auditoria de linguagem	Estudantes comparam respostas gerativas a prompts equivalentes com variações controladas.	Reconhecer estereótipos, variação e necessidade de cautela metodológica.
Mapa de afetados	Para um sistema de IA, grupos identificam usuários, não usuários e terceiros afetados.	Ampliar análise ética para além de quem opera a ferramenta.
Matriz erro x impacto	Avaliar falsos positivos e falsos negativos em diferentes contextos.	Distinguir acurácia de justiça e considerar consequências do erro.

Fonte: elaboração própria.

6. ALUCINAÇÕES DE IA, INCERTEZA E CONFIABILIDADE EPISTÊMICA

A noção de alucinação em IA tornou-se central com grandes modelos de linguagem. O termo é utilizado para descrever saídas

que parecem linguisticamente plausíveis, mas contêm informações falsas, inventadas ou não sustentadas. Farquhar et al. (2024) destacam que modelos de linguagem podem produzir respostas incorretas e não fundamentadas apesar de apresentarem elevada fluência. O problema não é um simples 'bug' que o estudante deve esperar desaparecer: decorre, em parte, da própria forma probabilística de geração de linguagem e das limitações do modelo quanto a conhecimento, contexto e verificação.

No ambiente escolar, a alucinação assume formas variadas: datas incorretas, fatos inventados, obras inexistentes, citações bibliográficas falsas, atribuições equivocadas, simplificações de conceitos e respostas que misturam elementos verdadeiros e falsos. O risco aumenta porque uma resposta pode apresentar estrutura acadêmica, tom seguro e referências com aparência legítima. A competência-chave, portanto, é distinguir confiança linguística de evidência.

Ensinar alucinação apenas com advertências tende a ser pouco eficaz. O estudante precisa vivenciar situações de verificação. Uma atividade simples consiste em fornecer três respostas de chatbot sobre um tema estudado - uma correta, uma parcialmente incorreta e uma contendo referências inexistentes - e pedir que grupos construam um laudo de confiabilidade. Cada afirmação deve ser classificada como confirmada, não confirmada ou contradita, sempre com indicação da fonte utilizada. O produto final é menos importante do que o percurso de checagem.

Outra estratégia é a 'caça à citação'. O professor pode pedir que a IA sugira referências sobre um tema e, em seguida, orientar os estudantes a localizar DOI, periódico, autores e texto original.

Quando a referência não existe, o erro deixa de ser abstrato e passa a ser observável. O procedimento também ensina uma regra acadêmica importante: nunca citar uma fonte apenas porque um sistema a apresentou; é necessário verificar a existência e ler o material relevante.

A confiabilidade deve ser ensinada como gradiente, e não como dicotomia. Certas perguntas são de baixo risco e fáceis de verificar; outras envolvem saúde, segurança, direitos, avaliações ou decisões pessoais e exigem fontes qualificadas e supervisão humana. O estudante pode aprender a classificar tarefas segundo risco e verificabilidade. Uma sugestão de título para um texto não demanda a mesma cautela de uma orientação médica ou jurídica. A competência ética inclui saber reconhecer quando um assistente virtual não deve ser tratado como fonte decisória.

Também é necessário discutir incerteza. Modelos muitas vezes não comunicam adequadamente o grau de confiança de suas respostas. Pedir ao chatbot que 'admita quando não souber' pode reduzir alguns erros, mas não garante veracidade. O estudante deve compreender que instruções de prudência são úteis, porém não substituem verificação externa. Em vez de perguntar 'a IA tem certeza?', é mais produtivo perguntar 'que evidência independente sustenta esta afirmação?'.

Tao et al. (2026), em revisão de escopo sobre riscos da IA generativa em K-12, destacam preocupações relacionadas à agência intelectual e à dependência, além de recomendarem tarefas que valorizem processo, uso de pistas em vez de respostas prontas e integração a currículos de letramento crítico em IA. Esse achado reforça a ideia de que a escola não deve medir apenas se o estudante obteve a

resposta correta, mas se consegue explicar como avaliou a resposta, quais fontes consultou e por que decidiu aceitar ou rejeitar determinado conteúdo.

7. ÉTICA NA INTERAÇÃO COM ASSISTENTES VIRTUAIS

Assistentes virtuais baseados em IA alteram a forma de interação porque utilizam linguagem conversacional, memória de contexto e recursos de personalização que podem produzir sensação de diálogo com um agente social. A interface é pedagogicamente útil, mas traz risco de antropomorfização: estudantes podem atribuir ao sistema compreensão, intenção, empatia ou autoridade que ele não possui. Letramento em IA deve incluir a compreensão de que respostas conversacionais resultam de processamento computacional e não equivalem a julgamento humano, consciência ou compromisso moral.

Essa distinção é particularmente importante quando a conversa envolve aconselhamento pessoal, saúde mental, conflitos familiares, sexualidade, violência ou decisões de alto impacto. A escola precisa ensinar rotas de apoio humano: professor, responsável, serviço de orientação, profissional de saúde ou autoridade competente. Assistentes podem organizar perguntas e apresentar informações gerais, mas não devem ser tratados como substitutos automáticos de relações de cuidado ou de profissionais qualificados.

A privacidade constitui outro eixo. A Política Nacional de Educação Digital já associa inclusão digital a direitos sobre tratamento de dados e proteção de populações vulneráveis, especialmente crianças e adolescentes (Brasil, 2023). O currículo deve transformar essa diretriz em hábitos concretos: não inserir senhas, documentos,

endereço, dados de saúde, imagens de terceiros, avaliações confidenciais ou informações identificáveis em ferramentas não autorizadas. Estudantes também precisam compreender que apagar uma conversa da interface não significa necessariamente eliminar todos os registros do processamento ou treinamento do serviço.

A orientação da UNICEF sobre IA e crianças, atualizada em 2025, inclui requisitos de segurança, proteção de dados, não discriminação, transparência, explicabilidade, responsabilidade, inclusão e preparação de crianças para desenvolvimentos presentes e futuros em IA. A atualização também chama atenção para companheiros de IA e novos riscos associados ao uso por crianças. O interesse educacional desse marco está em deslocar a ética do plano abstrato para direitos concretos: privacidade, segurança, participação e não discriminação.

Autoria e integridade acadêmica também precisam ser ensinadas. A questão não se resolve com a regra binária 'pode ou não pode usar'. Diferentes tarefas possuem objetivos distintos. Se o objetivo é avaliar escrita autoral, delegar a redação integral à IA compromete a evidência de aprendizagem. Se o objetivo é comparar argumentos, usar IA para gerar uma posição a ser criticada pode ser pedagogicamente adequado. A escola deve especificar o uso permitido e exigir transparência sobre como a ferramenta participou do processo.

Uma prática recomendável é o registro de uso de IA. O estudante pode anexar ao trabalho uma breve declaração indicando ferramenta, finalidade, principais prompts, trechos aproveitados, verificações realizadas e decisões autorais. Esse registro não deve

funcionar como mecanismo policiaisco, mas como exercício metacognitivo. A pergunta passa a ser: como a IA influenciou meu raciocínio e o que fiz para manter responsabilidade sobre o produto final?

Kajiwara e Kawabata (2024), em estudo sobre currículo de uso ético de chatbots, observaram contribuição para compreensão de conceitos de LLM e uso ético em apoio à decisão, além de alertarem que estudantes mais jovens podem adotar respostas geradas com confiança excessiva. Os autores também identificaram importância de justiça, privacidade e proteção de dados na atitude dos estudantes em relação à tecnologia. Embora o estudo não autorize generalizações para todos os contextos, reforça a necessidade de combinar uso prático com reflexão ética explícita.

Quadro 4. Regras de interação ética com assistentes virtuais no ambiente escolar

Princípio	Pergunta do estudante	Comportamento esperado
Privacidade	Estou compartilhando dado que identifica a mim ou outra pessoa?	Omitir dados sensíveis e utilizar apenas ferramentas autorizadas pela escola.
Finalidade	Por que estou usando a IA nesta tarefa?	Usar a ferramenta de forma compatível com o objetivo pedagógico.
Autoria	O produto ainda expressa meu raciocínio?	Revisar, reescrever, justificar escolhas e declarar o uso de IA quando solicitado.
Confiabilidade	O que nesta resposta precisa ser verificado?	Checar fatos, referências e afirmações em fontes independentes.

Justiça	A resposta generaliza ou prejudica algum grupo?	Questionar estereótipos, linguagem discriminatória e efeitos desiguais.
Agência	Estou deixando a IA decidir algo que deveria ser decidido por pessoa responsável?	Manter decisão humana em situações de alto impacto e buscar apoio adequado.
Transparência	Consigo explicar como a IA foi usada?	Registrar processo, prompts relevantes e verificações.

Fonte: elaboração própria.

8. ESTRATÉGIAS CURRICULARES POR ETAPA DA EDUCAÇÃO BÁSICA

Uma progressão curricular adequada precisa respeitar desenvolvimento cognitivo, repertório linguístico e experiência digital. A mesma noção pode ser retomada em espiral com complexidade crescente. 'Viés', por exemplo, pode começar como problema de classificação com dados incompletos nos anos iniciais, evoluir para comparação de recomendações nos anos finais e chegar a análise de métricas, proxies e efeitos distributivos no ensino médio.

Nos anos iniciais do ensino fundamental, a prioridade deve ser construir modelos mentais básicos e hábitos de segurança. Estudantes podem distinguir pessoas de máquinas, identificar exemplos cotidianos de automação, discutir como computadores reconhecem padrões, perceber que dados limitados produzem erros, aprender que chatbots podem inventar respostas e praticar a regra de não compartilhar informações pessoais. Atividades desplugadas são especialmente adequadas: jogos de classificação,

histórias com escolhas automatizadas, comparação entre resposta do livro e resposta gerada, ou criação coletiva de perguntas para testar um assistente sob supervisão do professor.

Nos anos finais, a aprendizagem pode incorporar noções de treinamento, conjuntos de dados, classificação, recomendação, probabilidade e geração de linguagem. O estudante já pode formular hipóteses, registrar experimentos e comparar saídas. Projetos de auditoria são produtivos: investigar como um recomendador organiza conteúdos; comparar respostas de IA em diferentes formulações; detectar referências falsas; analisar representação de gênero ou raça em imagens sintéticas; ou construir um pequeno conjunto de dados e observar como decisões de rotulagem alteram resultados.

No ensino médio, AI literacy deve alcançar maior densidade sociotécnica. Conceitos como otimização, correlação, proxy, falso positivo, falso negativo, explicabilidade, propriedade intelectual, governança e impacto ambiental podem ser incorporados em projetos interdisciplinares. Estudantes podem analisar casos reais de decisões automatizadas, produzir relatórios de risco, comparar políticas de uso de IA, discutir limites de automação em áreas sensíveis e desenvolver protótipos simples com documentação ética.

A progressão também deve considerar a capacidade de verificação. Nos anos iniciais, o professor seleciona fontes de comparação; nos anos finais, estudantes aprendem a localizar fontes confiáveis e distinguir busca de evidência; no ensino médio, devem justificar a qualidade das fontes, identificar conflitos, rastrear dados primários e

reconhecer limites de generalização. Assim, a verificação deixa de ser tarefa acessória e se torna objetivo curricular autônomo.

O uso de IA em atividades de escrita também pode progredir. Crianças podem comparar duas versões de uma história e identificar incoerências; adolescentes podem pedir ao modelo argumentos favoráveis e contrários e avaliar evidências; no ensino médio, estudantes podem utilizar IA como parceiro de revisão, mantendo versões e justificando alterações. O princípio é constante: a ferramenta deve ampliar reflexão, não substituir a produção intelectual que se deseja desenvolver.

Quadro 5. Progressão curricular sugerida para AI literacy

Etapa	Conhecimentos centrais	Práticas curriculares	Evidências de aprendizagem
Anos iniciais	IA x pessoa; padrões; dados; erro; privacidade; respostas inventadas.	Jogos de classificação; histórias; comparação de fontes; regras de segurança.	Explica por que a IA pode errar; identifica dado pessoal; confere informação com fonte indicada.
Anos finais	Treinamento; classificação; recomendação; generativa; viés; verificação; autoria.	Auditorias de resposta; experimentos com prompts; caça a citações; comparação de recomendações.	Documenta verificação; reconhece viés possível; distingue evidência de opinião; declara uso de IA.
Ensino médio	Probabilidade; proxies; métricas de erro; explicabilidade; governança; impactos sociais.	Projetos de auditoria; análise de casos; debate regulatório; protótipos responsáveis.	Produz relatório de risco; justifica critérios; compara grupos; argumenta sobre

			limites de automação.
--	--	--	-----------------------

Fonte: elaboração própria.

9. METODOLOGIAS ATIVAS PARA ENSINAR IA CRITICAMENTE

A literatura recente sobre ensino de IA em K-12 indica predominância de experiências práticas, projetos e aprendizagem experiencial. Lee e Kwon (2024) encontraram atividades hands-on e project-based learning entre estratégias frequentes, enquanto Liu e Zhong (2024) destacam a necessidade de trabalho em grupo, contextos autênticos, envolvimento docente e andaimes de aprendizagem. A escolha metodológica é coerente com a natureza do letramento: compreender IA exige observar comportamento de sistemas, formular hipóteses e confrontar expectativas com resultados.

A investigação guiada é particularmente adequada para alucinações. Em vez de o professor afirmar que 'IA erra', os estudantes recebem respostas e precisam demonstrar quais afirmações são confiáveis. A tarefa mobiliza pesquisa, leitura lateral, rastreamento de fontes e argumentação. O professor pode graduar dificuldade: primeiro fontes previamente selecionadas, depois busca aberta, e por fim situações com fontes divergentes.

A aprendizagem baseada em casos favorece ética e viés. Um caso pode apresentar um sistema que seleciona candidatos para uma oportunidade, um recomendador de vídeos, um chatbot que aconselha estudante ou um detector automático de risco. Grupos recebem papéis diferentes - estudante, escola, família, desenvolvedor, regulador - e analisam benefícios, danos e

salvaguardas. A multiplicidade de perspectivas ajuda a perceber que sistemas de IA distribuem poder e riscos de modo desigual.

Atividades desplugadas continuam essenciais. Simular um classificador com cartões, representar uma árvore de decisão com colegas ou montar uma sequência de palavras prováveis permite compreender aspectos de IA sem dependência de plataforma. A vantagem adicional é a transparência: o estudante enxerga a regra que normalmente estaria escondida no sistema. A abstração pode ser retomada depois com ferramentas digitais.

Debates estruturados são úteis quando combinados com evidências. Questões como 'a escola deveria permitir IA em provas?', 'um chatbot pode aconselhar sobre escolha profissional?' ou 'uma plataforma deve personalizar conteúdo usando dados comportamentais?' exigem que estudantes mobilizem conceitos de privacidade, finalidade, viés, autonomia e responsabilidade. O objetivo não é produzir consenso obrigatório, mas aprender a justificar posições e reconhecer trade-offs.

Portfólios de interação com IA podem tornar visível o processo. O estudante registra tarefa, prompt, resposta, pontos problemáticos, fontes consultadas, revisões e reflexão final. Ao longo do semestre, o portfólio permite avaliar evolução: prompts podem melhorar, mas, principalmente, o estudante passa a identificar mais limites, verificar melhor e depender menos da autoridade aparente do sistema.

10. PROTOCOLO VERIFICA-IA: UMA ROTINA DE PENSAMENTO CRÍTICO

Como contribuição didática deste artigo, propõe-se o protocolo VERIFICA-IA, uma rotina de oito passos para orientar estudantes

diante de respostas geradas por IA. O protocolo não pretende garantir ausência de erro, nem substituir avaliação especializada. Sua função é transformar verificação crítica em hábito observável, aplicável a diferentes disciplinas e faixas etárias com adaptações de linguagem.

V - Verificar a tarefa e a resposta: identificar o que foi perguntado, qual produto a IA ofereceu e se realmente responde ao problema. E - Examinar evidências: localizar fatos, dados, citações e afirmações que precisam de suporte. R - Reconhecer incertezas e possíveis alucinações: marcar trechos duvidosos, excessivamente seguros ou não verificáveis. I - Identificar vieses: observar generalizações, ausências, estereótipos e grupos potencialmente afetados. F - Fazer triangulação: comparar com mais de uma fonte confiável, priorizando fontes primárias ou institucionais quando pertinentes. I - Informar o uso de IA: registrar como a ferramenta participou do trabalho. C - Cuidar de dados e direitos: avaliar privacidade, autoria e informações de terceiros. A - Assumir responsabilidade: decidir o que será aceito, rejeitado ou revisado, sabendo que a decisão final é humana.

A principal vantagem pedagógica do protocolo é deslocar atenção da qualidade do prompt para a qualidade do julgamento. Em uma tarefa de História, o estudante pode receber um texto fluente e ainda precisar localizar datas, personagens e interpretações. Em Ciências, deve distinguir explicações consolidadas de hipóteses. Em Matemática, pode conferir procedimento e resultado. Em Língua Portuguesa, pode analisar argumento e evidência. A mesma rotina organiza práticas transdisciplinares.

Nos anos iniciais, o protocolo pode ser reduzido a quatro perguntas: a resposta faz sentido? onde posso conferir? há alguma informação pessoal? quem é responsável pela resposta final? Nos anos finais, os oito passos podem ser usados com fichas. No ensino médio, cada passo pode se transformar em critério de relatório ou rubrica. Essa adaptação preserva continuidade curricular sem exigir complexidade precoce.

O protocolo também ajuda professores a desenhar tarefas avaliativas resistentes à terceirização. Em vez de pedir apenas um produto final, pode-se exigir evidências de cada etapa: marcação de trechos, fontes de verificação, análise de viés e justificativa de decisão. Mesmo que o estudante use IA, o objeto avaliado passa a incluir pensamento crítico sobre o uso da IA.

Quadro 6. Protocolo VERIFICA-IA

Etapas	Pergunta-guia	Registro possível
V – Verificar	A resposta realmente atende ao que foi pedido?	Resumo do que a IA afirmou.
E – Examinar evidências	Quais afirmações exigem comprovação?	Lista de fatos, dados e citações.
R – Reconhecer incerteza	O que pode estar errado, inventado ou incompleto?	Trechos marcados e nível de confiança.
I – Identificar vieses	Há generalizações, ausências ou tratamento desigual?	Grupos afetados e hipótese de viés.
F – Fazer triangulação	Que fontes independentes	Duas ou mais fontes, incluindo fonte primária quando possível.

	confirmam ou contradizem?	
I – Informar uso	Como a IA participou do trabalho?	Ferramenta, finalidade e etapas de uso.
C – Cuidar de dados	Há privacidade, autoria ou direito de terceiros envolvidos?	Decisão de remover ou anonimizar informação.
A – Assumir responsabilidade	O que será aceito, rejeitado ou revisado?	Justificativa final do estudante.

Fonte: elaboração própria.

11. AVALIAÇÃO DAS APRENDIZAGENS EM AI LITERACY

Avaliar letramento em IA requer observar mais do que conhecimento declarativo. Uma prova pode verificar se o estudante sabe definir viés ou alucinação, mas não demonstra necessariamente que reconhece esses problemas em situações reais. A avaliação deve combinar conhecimento, desempenho e atitude: explicar conceitos, aplicar procedimentos de verificação e demonstrar responsabilidade nas decisões.

Rubricas são úteis porque tornam critérios visíveis. Um relatório de auditoria de resposta pode ser avaliado pela identificação de afirmações verificáveis, qualidade das fontes, reconhecimento de incerteza, análise de viés, tratamento de privacidade e justificativa final. O estudante deve entender que encontrar um erro da IA não é 'pegar a máquina no erro', mas construir evidência sobre o que pode ou não ser confiável.

A avaliação também deve considerar processo. Tao et al. (2026) recomendam tarefas que valorizem processo sobre produto no contexto de IA generativa. Portfólios, versões sucessivas, defesa oral, diário de pesquisa e comparação entre rascunhos ajudam a observar aprendizagem e autoria. Em atividades com IA, perguntar 'como você chegou a esta resposta?' torna-se tão importante quanto perguntar 'qual é a resposta?'.

É recomendável evitar políticas de detecção automática como principal estratégia de integridade. Detectores de texto gerado por IA podem produzir falsos positivos e não resolvem a questão pedagógica central. Uma escola orientada por letramento deve preferir desenho de tarefas, acompanhamento do processo, critérios claros de uso permitido e diálogo sobre autoria. Tecnologias de detecção, quando utilizadas, não devem ser tratadas como prova infalível.

A autoavaliação pode ser incorporada. Ao final de uma atividade, estudantes podem indicar onde confiaram inicialmente na IA, o que mudou após verificação e qual regra pretendem adotar no futuro. Esse componente metacognitivo transforma erro em aprendizagem. O objetivo não é produzir medo da tecnologia, mas calibrar confiança.

Quadro 7. Rubrica sintética de avaliação do letramento crítico em IA

Critério	Inicial	Em desenvolvimento	Consolidado
Compreensão técnica	Trata a IA como fonte ou agente	Reconhece alguns conceitos, mas	Explica, em nível adequado, dados,

	sem reconhecer limites.	explica de forma parcial.	modelo, incerteza e limites.
Verificação	Aceita a resposta sem checagem.	Confere parte das afirmações ou usa fontes frágeis.	Identifica o que verificar, triangula e justifica qualidade das fontes.
Viés e justiça	Não identifica efeitos diferenciais.	Reconhece estereótipo evidente.	Analisa dados, contexto, grupos afetados e consequências.
Ética e privacidade	Compartilha dados ou usa conteúdo sem refletir sobre direitos.	Reconhece regras quando lembrado.	Protege dados, respeita autoria e justifica limites de uso.
Agência e autoria	Terceiriza decisão ou produto à IA.	Revisa parcialmente a saída.	Mantém responsabilidade, reescreve, documenta e defende suas escolhas.

Fonte: elaboração própria.

12. FORMAÇÃO DOCENTE E GOVERNANÇA ESCOLAR

Nenhum currículo de AI literacy se sustenta sem formação docente. A UNESCO reconhece esse problema ao propor, também em 2024, um marco de competências em IA para professores com 15 competências distribuídas em cinco dimensões: mentalidade centrada no ser humano, ética da IA, fundamentos e aplicações, pedagogia de IA e IA para desenvolvimento profissional (Miao; Cukurova, 2024). A formação docente precisa combinar domínio

conceitual, julgamento ético e capacidade de desenhar experiências de aprendizagem.

A escola não deve esperar que todo professor se torne especialista em ciência de dados. O nível necessário depende da prática. Professores precisam saber reconhecer limitações de sistemas, avaliar pertinência pedagógica, proteger dados, orientar verificação, definir usos permitidos e criar tarefas em que IA apoie objetivos curriculares. Especialistas em Computação podem aprofundar fundamentos técnicos, enquanto docentes de outras áreas contribuem com epistemologias disciplinares sobre evidência, interpretação e argumentação.

A governança escolar deve incluir uma política clara de IA. Essa política precisa diferenciar usos administrativos, usos docentes e usos estudantis; indicar ferramentas autorizadas; definir regras de dados e consentimento; estabelecer critérios para tarefas e avaliações; prever transparência de uso; e criar procedimento para revisão de incidentes. Políticas genéricas do tipo 'IA proibida' ou 'IA liberada' são insuficientes porque ignoram finalidade, idade, risco e contexto.

Também é recomendável constituir espaços de experimentação supervisionada. Professores podem testar ferramentas, registrar benefícios e falhas e compartilhar casos. A análise institucional deve considerar custo, acessibilidade, privacidade, dependência de fornecedor, interoperabilidade, vieses e suporte. A adoção de IA não deve ser decidida apenas pela novidade ou pela pressão de mercado.

A formação continuada precisa incluir tarefas autênticas. Um professor pode aprender mais auditando respostas de um chatbot sobre seu conteúdo disciplinar do que assistindo apenas a uma palestra geral sobre IA. Oficinas podem envolver criação de rubricas, comparação de fontes, elaboração de casos de viés, análise de termos de uso e desenho de atividades com o protocolo VERIFICA-IA. A prática formativa deve refletir o mesmo letramento crítico esperado dos estudantes.

A gestão escolar deve acompanhar equidade. Se determinadas turmas possuem maior acesso a dispositivos, se alunos pagantes usam versões mais avançadas ou se famílias dispõem de apoio técnico desigual, a IA pode ampliar diferenças. O currículo precisa assegurar que competências avaliadas possam ser desenvolvidas com recursos oferecidos pela escola.

13. EQUIDADE, INCLUSÃO E DESENVOLVIMENTO INFANTOJUVENIL

A discussão de AI literacy em educação básica precisa ser desenvolvimentalmente informada. Crianças e adolescentes não constituem apenas usuários menores de idade; encontram-se em processo de formação cognitiva, social e identitária. Sistemas conversacionais podem facilitar explicações e acessibilidade, mas também reduzir esforço produtivo, estimular dependência ou interferir na percepção de autoria. Tao et al. (2026) sintetizam riscos relacionados à agência intelectual, bem-estar e ecossistema educacional, recomendando uso com andaimes e supervisão.

Isso não significa presumir incapacidade. Pelo contrário, letramento crítico pressupõe participação ativa de estudantes na construção de

regras. Turmas podem discutir quando o uso de IA ajuda a aprender e quando apenas substitui uma habilidade que precisa ser exercitada. Estudantes podem avaliar políticas da própria escola, sugerir critérios de transparência e participar de comissões de cultura digital, respeitada a faixa etária.

A inclusão de estudantes com deficiência também precisa ser considerada. Recursos de transcrição, descrição, tradução, simplificação e comunicação aumentativa podem ampliar acesso, mas resultados devem ser avaliados quanto a precisão e adequação. Ferramentas de IA não devem substituir tecnologia assistiva validada ou apoio humano quando estes forem necessários. O princípio é utilizar automação para remover barreiras sem reduzir direitos ou autonomia.

Vieses podem afetar de forma desproporcional estudantes de grupos historicamente marginalizados. Por isso, atividades de auditoria devem ser culturalmente responsáveis e evitar reproduzir conteúdo ofensivo desnecessariamente. O professor precisa selecionar exemplos, contextualizar estereótipos e oferecer suporte quando a análise tocar experiências de discriminação. Ensinar sobre viés não deve significar expor estudantes ao dano como se ele fosse simples material didático.

A desigualdade de infraestrutura demanda alternativas. Um currículo de letramento em IA pode funcionar com atividades impressas, simulações, estudos de caso e respostas previamente geradas pelo professor. A competência avaliada deve ser compreensão crítica, não posse de assinatura premium. Esse princípio é coerente com a ideia de educação digital como direito e com a prioridade conferida pela PNED às populações vulneráveis.

14. DESAFIOS E LIMITAÇÕES PARA IMPLEMENTAÇÃO

O primeiro desafio é a velocidade de mudança tecnológica. Ferramentas, interfaces e modelos se atualizam mais rapidamente que currículos e livros didáticos. Por isso, objetivos devem se concentrar em princípios relativamente estáveis - dados, incerteza, viés, evidência, privacidade, autoria e responsabilidade - em vez de depender de nomes de plataformas. O currículo deve formar competências transferíveis.

O segundo desafio é evitar tecnocentrismo. AI literacy não pode se tornar treinamento para ferramentas comerciais. A escola tem compromisso com formação humana, cidadania e conhecimento disciplinar. A IA deve ser objeto de estudo e, quando pertinente, meio de aprendizagem; não deve redefinir todo o currículo a partir das funcionalidades disponíveis no mercado.

O terceiro desafio é a escassez de evidência longitudinal. Muitas pesquisas K-12 analisam intervenções curtas e amostras limitadas. Liu e Zhong (2024) destacam limitações de duração e desenho na literatura. Mesmo o crescimento recente de estudos sobre IA generativa ainda não permite conclusões fortes sobre efeitos de longo prazo em aprendizagem, criatividade, autorregulação ou relações sociais. Políticas escolares devem, portanto, adotar postura experimental responsável, com monitoramento e possibilidade de revisão.

O quarto desafio é avaliar riscos sem alarmismo. Alucinações, vieses e dependência são problemas reais, mas a resposta pedagógica não deve construir a tecnologia como inimiga ou autoridade inevitável. O objetivo é calibrar confiança e ampliar capacidade de decisão.

Estudantes precisam reconhecer tanto potencial quanto limites, aprendendo a usar IA quando agrega valor e a recusá-la quando compromete aprendizagem, privacidade ou segurança.

O quinto desafio é a coerência entre discurso e prática institucional. Uma escola que ensina proteção de dados, mas exige uso de plataforma sem analisar suas políticas, produz mensagem contraditória. Da mesma forma, não é coerente exigir transparência dos estudantes e utilizar sistemas automatizados de decisão ou monitoramento sem explicação. A governança escolar deve incorporar os princípios ensinados no currículo.

15. CONCLUSÃO

O letramento em IA tornou-se componente necessário da formação básica porque sistemas algorítmicos e generativos já participam de práticas cotidianas de informação, comunicação, criação e decisão. A escola não pode limitar sua resposta à proibição nem à celebração instrumental. Precisa formar estudantes capazes de compreender como sistemas de IA operam em nível adequado à idade, reconhecer limites e vieses, verificar evidências, proteger dados, preservar autoria e manter responsabilidade humana sobre decisões.

Vieses algorítmicos, alucinações e ética na interação com assistentes virtuais constituem eixos especialmente produtivos porque conectam fundamentos técnicos a problemas concretos de cidadania. O estudante que compreende por que dados podem produzir tratamento desigual, que sabe verificar uma resposta fluente e que reconhece quando uma decisão não deve ser

terceirizada desenvolve competências que ultrapassam qualquer plataforma específica.

Os marcos internacionais recentes da UNESCO e da OCDE/Comissão Europeia oferecem referências consistentes para construir progressão curricular, enquanto a BNCC, as Normas sobre Computação na Educação Básica e a Política Nacional de Educação Digital fornecem base normativa nacional para integração. O desafio passa a ser transformar esses referenciais em objetivos observáveis, metodologias adequadas à idade, avaliação processual e políticas escolares coerentes.

A proposta de progressão por etapas e o protocolo VERIFICA-IA apresentados neste artigo procuram contribuir para essa transposição. Ambos precisam ser adaptados e validados empiricamente, mas partem de um princípio central: a competência mais importante diante de uma IA não é produzir o melhor prompt, e sim exercer o melhor julgamento possível sobre o que a tecnologia oferece. Letrar em IA significa formar sujeitos capazes de usar, questionar, corrigir, contextualizar e, quando necessário, recusar a automação.

REFERÊNCIAS BIBLIOGRÁFICAS

BAKER, Ryan S.; HAWN, Aaron. Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, v. 32, n. 4, p. 1052-1092, 2022. DOI: 10.1007/s40593-021-00285-9.

BRASIL. Conselho Nacional de Educação. Resolução CNE/CEB nº 1, de 4 de outubro de 2022. Normas sobre Computação na Educação Básica – Complemento à Base Nacional Comum Curricular (BNCC). Brasília, DF: CNE, 2022.

BRASIL. Lei nº 14.533, de 11 de janeiro de 2023. Institui a Política Nacional de Educação Digital e altera as Leis nº 9.394/1996, nº 9.448/1997, nº 10.260/2001 e nº 10.753/2003. Brasília, DF: Presidência da República, 2023.

BRASIL. Ministério da Educação. Base Nacional Comum Curricular. Brasília, DF: MEC, 2018.

CASAL-OTERO, Lorena et al. AI literacy in K-12: a systematic literature review. *International Journal of STEM Education*, v. 10, art. 29, 2023. DOI: 10.1186/s40594-023-00418-7.

FARQUHAR, Sebastian et al. Detecting hallucinations in large language models using semantic entropy. *Nature*, v. 630, p. 625-630, 2024. DOI: 10.1038/s41586-024-07421-0.

IDOWU, Jamiu Adekunle. Debiasing education algorithms. *International Journal of Artificial Intelligence in Education*, v. 34, p. 1510-1540, 2024.

KAJIWARA, Yusuke; KAWABATA, Kouhei. AI literacy for ethical use of chatbot: will students accept AI ethics? *Computers and Education: Artificial Intelligence*, v. 6, 100251, 2024. DOI: 10.1016/j.caeai.2024.100251.

LEE, Sang Joon; KWON, Kyungbin. A systematic review of AI education in K-12 classrooms from 2018 to 2023: topics, strategies, and learning outcomes. *Computers and Education: Artificial Intelligence*, v. 6, 100211, 2024. DOI: 10.1016/j.caeai.2024.100211.

LIU, Xiaofan; ZHONG, Baichang. A systematic review on how educators teach AI in K-12 education. *Educational Research Review*,

v. 45, 100642, 2024. DOI: 10.1016/j.edurev.2024.100642.

MIAO, Fengchun; CUKUROVA, Mutlu. AI competency framework for teachers. Paris: UNESCO, 2024.

MIAO, Fengchun; HOLMES, Wayne. Guidance for generative AI in education and research. Paris: UNESCO, 2023.

MIAO, Fengchun; SHIOHIRA, Kelly; LAO, Natalie. AI competency framework for students. Paris: UNESCO, 2024.

OECD; EUROPEAN COMMISSION. Empowering learners for the age of AI: an AI literacy framework for primary and secondary education. Paris: OECD Publishing, 2026. DOI: 10.1787/65cd27d4-en.

PARK, Joonhyeong. A systematic literature review of generative artificial intelligence (GenAI) literacy in schools. Computers and Education: Artificial Intelligence, 100487, 2025. DOI: 10.1016/j.caeai.2025.100487.

TAO, Sisi; LAN, Min; WANG, Minjuan; LI, Hui. Potential risks of generative artificial intelligence integration into K-12 education: a scoping review. Computers and Education: Artificial Intelligence, v. 10, 100561, 2026. DOI: 10.1016/j.caeai.2026.100561.

UNICEF. Guidance on AI and children: version 3.0 – recommendations for AI policies and systems that uphold child rights. Florence: UNICEF Innocenti, 2025.

YIM, Iris Heung Yue; SU, Jiahong. Artificial intelligence (AI) learning tools in K-12 education: a scoping review. Journal of Computers in Education, v. 12, p. 93-131, 2025. DOI: 10.1007/s40692-023-00304-9.

YIM, Iris Heung Yue. A critical review of teaching and learning artificial intelligence (AI) literacy: developing an intelligence-based AI literacy framework for primary school education. *Computers and Education: Artificial Intelligence*, v. 7, 100319, 2024. DOI: 10.1016/j.caeai.2024.100319.

APÊNDICE A – SEQUÊNCIA DIDÁTICA SUGERIDA: “POSSO CONFIAR NESTA RESPOSTA?”

A sequência abaixo foi concebida para anos finais do ensino fundamental e pode ser adaptada ao ensino médio. O objetivo é integrar alucinação, viés, verificação e ética em cinco encontros. Não é necessário que todos os estudantes possuam conta em plataforma de IA; o professor pode gerar previamente as respostas e distribuí-las impressas.

Encontro 1 – Como um assistente produz respostas. O professor apresenta exemplos de perguntas e discute a diferença entre busca, banco de dados e geração de linguagem. A turma identifica sinais de antropomorfização e formula hipóteses sobre por que um sistema pode responder de modo convincente sem 'saber' da mesma forma que uma pessoa.

Encontro 2 – Caça à alucinação. Grupos recebem respostas contendo afirmações corretas, incorretas e referências inexistentes. Cada grupo aplica as três primeiras etapas do VERIFICA-IA e marca o que precisa ser confirmado. A socialização enfatiza a diferença entre fluência e evidência.

Encontro 3 – Laboratório de viés. A turma analisa um caso de classificação com dados desequilibrados e depois compara

respostas a prompts equivalentes. O professor discute representatividade, estereótipos, falso positivo e consequências do erro. Estudantes registram quais conclusões o experimento permite e quais não permite.

Encontro 4 – Ética e privacidade. Estudos de caso apresentam situações de compartilhamento de dados pessoais, aconselhamento de alto impacto, uso de IA em tarefa autoral e produção de imagem de terceiro. Grupos propõem regras e justificam cada uma com base em direitos, finalidade e risco.

Encontro 5 – Auditoria final. Cada grupo recebe uma nova resposta gerada por IA e produz um relatório completo usando o VERIFICA-IA. A avaliação considera qualidade da verificação, justificativa de fontes, análise de viés, cuidado com dados e responsabilidade autoral.

APÊNDICE B – FICHA DE AUDITORIA DE RESPOSTAS DE IA

Quadro 8. Ficha rápida para uso em sala de aula

Campo	Registro do estudante
Pergunta/tarefa original	
Síntese da resposta da IA	
Afirmações que precisam de verificação	
Possíveis erros ou alucinações	

Possíveis vieses ou ausências	
Fontes consultadas	
Dados pessoais envolvidos?	
Como a IA foi usada no trabalho?	
Decisão final: aceitar, revisar ou rejeitar? Por quê?	

Fonte: elaboração própria.