O Boletim de Conjuntura (BOCA) publica ensaios, artigos de revisão, artigos teóricos e empíricos, resenhas e vídeos relacionados às temáticas de políticas públicas.

O periódico tem como escopo a publicação de trabalhos inéditos e originais, nacionais ou internacionais que versem sobre Políticas Públicas, resultantes de pesquisas científicas e reflexões teóricas e empíricas.

Esta revista oferece acesso livre imediato ao seu conteúdo, seguindo o princípio de que disponibilizar gratuitamente o conhecimento científico ao público proporciona maior democratizacão mundial do conhecimento.



BOLETIM DE CONJUNTURA

BOCA

Ano VII | Volume 23 | Nº 68 | Boa Vista | 2025

http://www.ioles.com.br/boca

ISSN: 2675-1488

https://doi.org/10.5281/zenodo.17096444

BOLETIM DE CONJUNTURA

www.ioles.com.br/boca

AMBIVALENCIA DA INTELIGENCIA ARTIFICIAL: *FAKE NEWS* E A EDUCAÇÃO MIDIÁTICA NA EDUCAÇÃO BÁSICA

Patrick Renilton Amorim dos Santos[1]

Luiza Beth Nunes Alonso[2]

Helga Cristina Hedler[3]

Eduardo Amadeu Dutra Moresi[4]

Vicente Paulo Alves[5]

Resumo

A crescente disseminação de desinformação e fake news potencializada por modelos generativos de Inteligência Artificial representa um desafio para a educação, a gestão da informação e a coesão social. Este estudo investiga o papel da alfabetização midiática como estratégia de mitigação desses impactos, com ênfase no fortalecimento do pensamento crítico e da resiliência digital no ensino fundamental e médio. O objetivo foi mapear e analisar a produção científica relacionada à interseção entre IA, desinformação e educação midiática, identificando tendências, lacunas e oportunidades de pesquisa. Para tanto, realizou-se uma Revisão Sistemática da Literatura, seguindo o protocolo PRISMA. A busca foi conduzida nas bases Lens.org e Scopus, resultando em 4.843 artigos. Após a aplicação de critérios de elegibilidade e exclusão, foram selecionados os 20 trabalhos mais citados e aderentes ao tema. O corpus final foi submetido a análises lexicométricas no software Iramuteq, permitindo identificar padrões temáticos e relações conceituais. Os resultados revelaram quatro categorias centrais: (i) Ética e Princípios na IA Educacional, (ii) Bem-estar Digital e Impactos no Aprendizado, (iii) Governança e Regulação da AIED e (iv) Fake News e Desinformação. A análise fatorial de correspondência indicou que os debates sobre desinformação estão fortemente vinculados à dimensão tecnológica e algorítmica, enquanto as discussões éticas, embora fundamentais, permanecem pouco integradas às práticas educacionais. Conclui-se que a integração entre inovação tecnológica, ética e práticas pedagógicas é essencial para enfrentar os desafios contemporâneos, sendo a alfabetização midiática um elemento estratégico para a promoção de uma educação crítica e responsável na era digital.

Palavras-chave: Desinformação; Educação Midiática; *Fake News*; Inteligência Artificial; Letramento Digital.

Abstract

The growing dissemination of disinformation and fake news, amplified by generative Artificial Intelligence (AI) models, poses a critical challenge to education, information management, and social cohesion. This study investigates the role of media literacy as a strategy to mitigate these impacts, with an emphasis on strengthening critical thinking and digital resilience among primary and secondary students. The objective was to map and analyze the scientific literature addressing the intersection between AI, disinformation, and media literacy, identifying trends, gaps, and opportunities for research. A Systematic Literature Review was conducted following the PRISMA protocol. Searches in Lens.org and Scopus databases yielded 4,843 articles. After applying inclusion and exclusion criteria, the 20 most cited and relevant studies were selected. The final corpus was analyzed through lexicometric techniques using Iramuteq software, which enabled the identification of thematic patterns and conceptual relationships. The results revealed four central categories: (i) Ethics and Principles in AI in Education, (ii) Digital Well-being and Learning Outcomes, (iii) Governance and Regulation of AIED, and (iv) Fake News and Disinformation. Correspondence factor analysis indicated that discussions on disinformation are strongly linked to technological and algorithmic dimensions, while ethical concerns, although fundamental, remain weakly integrated into educational practices. The study concludes that integrating technological innovation, ethics, and pedagogical strategies is essential to address contemporary challenges. Media literacy emerges as a strategic element for promoting critical and responsible education in the digital age.

Keywords: Artificial Intelligence; Digital Literacy; Disinformation; Fake News; Media Literacy.

[1] Mestre em Governança, Tecnologia e Inovação pela Universidade Católica de Brasília (UCB). E-mail: patrick.rads@outlook.com

[2] Doutora em Educação pela Harvard University. E-mail: luiza.alonsoex@gmail.com

[3] Doutora em Psicologia Social, do Trabalho e das Organizações pela Universidade de Brasília (UnB). E-mail: helga@p.ucb.br

[4] Doutor em Ciência da Informação pela Universidade de Brasília (UnB). E-mail: moresi@p.ucb.br

[5] Doutor em Ciências da Religião pela Universidade Metodista de São Paulo (UMESP). E-mail: vicente@p.ucb.br

INTRODUÇÃO

A crescente capacidade dos modelos generativos de Inteligência Artificial (IA) de produzir e disseminar informações em larga escala impôs um desafio à sociedade contemporânea. Essa tecnologia tornou-se um vetor de desinformação, capaz de gerar conteúdos falsos e persuasivos com velocidade e realismo sem precedentes, ao promover avanços significativos em diferentes áreas. Esse fenômeno compromete a coesão social, a tomada de decisões em ambientes digitais saturados e nos educacionais.

No campo educacional, as tecnologias de IA influenciam o acesso e avaliação de conteúdos digitais. Por um lado, oferecem novas possibilidades de aprendizagem, mas, por outro, intensificam riscos de disseminação de informações falsas. Essa ambivalência exige que a escola vá além do uso instrumental, priorizando a formação crítica dos estudantes para interpretar e produzir informações de modo consciente e ético.

Este estudo se justifica por preencher a lacuna na literatura científica. A literatura priorizou soluções técnicas para combater a desinformação, como ferramentas de detecção automatizada, mas negligenciado abordagens pedagógicas. Há escassez de pesquisas que explorem a integração da alfabetização midiática e informacional aos currículos escolares, especialmente no ensino fundamental e médio. Do ponto de vista da sociedade, investigar esse campo é necessário para fortalecer a autonomia dos estudantes, desenvolver resiliência digital e o letramento digital de professores e alunos para torná-los cidadãos capazes de lidar com os desafios da era informacional.

Investigou-se como a educação midiática e o letramento digital podem integrar o uso crítico da inteligência artificial no combate à desinformação no ensino básico. Os objetivos envolveram identificar a produção científica sobre alfabetização midiática e uso crítico de IA no ensino fundamental e médio; mapear tendências; analisar as práticas educacionais e seu papel na mitigação da desinformação gerada por IA, e apontar lacunas e oportunidades para pesquisas futuras.

O recorte metodológico utilizou a Revisão Sistemática da Literatura, guiada pelo protocolo PRISMA (2021). As buscas foram feitas nas bases Lens.org e Scopus, a partir de descritores relacionados a inteligência artificial, *fake news*, desinformação e literacia digital e informacional. A seleção seguiu critérios de inclusão e exclusão específicos, priorizando o recorte da área de estudos educacionais e de IA. Incluíram-se marcos normativos brasileiros, como a *BNCC* e a *Lei nº 14.533/2023*, assegurando a consistência e relevância dos materiais analisados.

Além da introdução, o trabalho está estruturado em seções que apresentam o referencial teórico, o método, resultados, discussão e conclusão com contribuições, limitações e recomendações para futuras pesquisas.

REFERENCIAL TEÓRICO

A revisão de literatura explora a ambivalência da Inteligência Artificial (IA), focando sua dupla natureza: uma ferramenta transformadora na educação e um vetor de desafios sociais, como a desinformação. O texto apresenta os conceitos de forma progressiva. Inicialmente, são definidos os termos-chave: inteligência artificial, *fake news* e desinformação. Em seguida, avança para a intersecção dessas áreas, detalhando como a IA atua na disseminação e amplificação de notícias falsas. Por fim, culmina com o papel da educação midiática no desenvolvimento de habilidades críticas e na capacitação dos indivíduos para navegar no complexo ambiente digital.

Na literatura, é possível encontrar várias definições de IA, cada uma guardando os conceitos-chave da inteligência não humana programada para executar tarefas específicas. Segundo Dwivedi *et al.* (2021), descrevem a IA no contexto de interpretar e aprender de forma independente a partir de dados externos para alcançar resultados específicos por meio de adaptação flexível. De outra forma, inteligência artificial (AI) refere-se aos avanços atuais e passados que impactaram negócios e sociedade (KERR *et al*., 2025).

A definição conceitual de IA enfatiza sua capacidade adaptativa e autônoma. Aplicações práticas como o uso do GPT evidenciam o alcance e a relevância dessa tecnologia em contextos educacionais (WARDAT, *et al*., 2023). A IA pode ser usada de diferentes formas, ora favoravelmente no contexto educacional, ora gerando desinformação. A disseminação de *fake news* e estratégias de desinformação encontra terreno fértil, uma vez que os sistemas algorítmicos e modelos generativos ampliam a escala, sofisticação e potencial de convencimento desses conteúdos (ZAROCOSTAS, 2020).

A interconexão entre IA, *fake news* e desinformação não pode ser compreendida isoladamente. Ela faz parte de um fenômeno complexo que envolve dimensões tecnológicas, sociais, políticas e éticas. Isso demanda análises críticas que transcendam a descrição dos avanços tecnológicos (VELLOSO, 2025).

Além do entendimento de IA, é necessário compreender como os indivíduos processam conteúdos e como fenômenos comunicacionais como *misinformation*, *disinformation* e *fake news se configuram. Misinformation* refere-se à informação falsa ou incorreta compartilhada sem a intenção de enganar, resultado da reprodução inadvertida de conteúdo. Disinformation diz respeito à informação deliberadamente falsa ou manipulada, criada e disseminada para enganar ou causar danos. Fake *news* consistem na fabricação e difusão de conteúdos noticiosos que simulam padrões jornalísticos de credibilidade, mas veiculam informações inverídicas, podendo englobar elementos de *misinformation* e *disinformation* (BRODA; STRÖMBÄCK, 2024). Apesar das diferenças, esses conceitos compartilham a capacidade de gerar confusão informacional, comprometer a percepção da realidade e impactar negativamente a esfera pública.

A Inteligência Artificial potencializa tanto os benefícios (QUACH *et al*., 2022) quanto os riscos (WACH *et al.,* 2023) relacionados à informação. Os desafios vão além do tecnológico; eles se estendem à dimensão social e educacional, exigindo da sociedade mecanismos de preparo crítico frente à avalanche de conteúdos digitais. A educação midiática assume papel central, funcionando como um contraponto à vulnerabilidade cognitiva explorada por heurísticas e vieses (JAKESCH; HANCOCK; NAAMAN, 2022). Ela capacita indivíduos para reconhecer, interpretar e questionar criticamente as mensagens digitais. Essa perspectiva é relevante no contexto da ambivalência da inteligência artificial na disseminação de *fake news*. Enquanto ferramentas digitais podem potencializar a desinformação, a educação midiática capacita estudantes a analisar criticamente conteúdos, produzir narrativas próprias e engajar-se ativamente na sociedade. Assim, fortalecer programas de educação na educação básica é fundamental para combater a desinformação e construir uma cultura digital mais ética, inclusiva e democrática.

Ao articular o debate sobre IA e desinformação com a necessidade de desenvolver competências críticas (ALEEDAN; ALFADHLY, 2024) por meio da educação midiática, constrói-se um caminho de análise que integra os aspectos técnicos da produção e disseminação de conteúdos falsos e as estratégias pedagógicas de prevenção e enfrentamento. O caminho envolve competências desde a busca por informações até a criação de conteúdos em diversas plataformas. Em um cenário de disseminação rápida, desenvolver e aplicar práticas de letramento midiático torna-se uma habilidade indispensável.

A educação midiática é uma estratégia para enfrentar a desinformação e já consta na legislação brasileira. A Lei nº 14.533/2023 institui a Política Nacional de Educação Digital (PNED) e a Base Nacional Comum Curricular (BNCC) exigem que as escolas garantam que os alunos adquiram habilidades para interpretar, analisar e utilizar informações digitais de forma crítica e responsável (BRASIL, 2018; BRASIL, 2023). O letramento midiático e informacional é reconhecido como um direito educacional. O letramento midiático é um conceito abrangente relacionado à habilidade de acessar, analisar, avaliar e produzir conteúdo em diferentes mídias. Trata-se de um elemento fundamental para a construção de uma sociedade crítica e bem informada (ADJIN-TETTEY, 2022).

Burth *et al.* (2024) destacam que, apesar dos avanços em educação midiática na última década, ainda há uma lacuna na definição e mensuração de seu impacto. Os autores analisaram 300 estudos dos EUA e propuseram o modelo dos 6 E's, que articula diferentes dimensões da educação: resultados avaliativos, investigação crítica, expressão, aprendizagem experiencial, cidadania ativa e equidade. Esse *framework* oferece uma visão abrangente de como a educação midiática pode transformar indivíduos, comunidades e a sociedade. O estudo mostra que grande parte das iniciativas se concentra na avaliação de mudanças individuais, como desenvolvimento de habilidades técnicas e pensamento crítico para

identificar *fake news*. Entretanto, impactos mais amplos, como promoção de cidadania ativa e transformação social, ainda são pouco explorados.

O letramento midiático vai além da interpretação de notícias. Ele envolve a compreensão de como formatos influenciam a transmissão e percepção de mensagens. Plataformas como YouTube e TikTok, com conteúdos variados, moldam visões de mundo e comportamentos (BAILÉN; QUINTAS-FROUFE; GONZÁLEZ-NEIRA, 2024). É essencial capacitar jovens e adultos a analisar criticamente esses conteúdos.

Dumitru e Halpern (2023) argumentam que o pensamento crítico emerge como uma competência essencial para enfrentar os desafios do futuro do trabalho, diante do avanço da inteligência artificial e das tecnologias digitais. Os autores introduzem o conceito de *job-proof skills*, ou habilidades que permanecem relevantes em contextos de automação. Entre essas, destacam-se o pensamento crítico, a resolução de problemas, a ética, a empatia e a inteligência emocional, atributos humanos que as máquinas não conseguem replicar. O estudo mostra que ocupações que requerem pensamento crítico estão mais protegidas da substituição por IA, enquanto funções baseadas em tarefas repetitivas apresentam maior risco de automação. Além de favorecer a empregabilidade, o pensamento crítico ajuda indivíduos a se adaptarem a novas formas de trabalho e participarem ativamente na tomada de decisões, no ambiente profissional e na esfera social.

É fundamental formar cidadãos capazes de avaliar a credibilidade das fontes e identificar vieses informacionais para preservar a democracia e fortalecer a cidadania digital. Dumitru e Halpern (2023) destacam a importância da colaboração entre universidades e organizações do mercado de trabalho. Experiências europeias, demonstram que currículos flexíveis e conectados à realidade profissional promovem avanços significativos nas habilidades de pensamento crítico. A integração de casos reais, estágios e práticas interdisciplinares potencializa a aprendizagem e aproxima os estudantes das demandas do mercado. Esses achados reforçam a necessidade de incluir o pensamento crítico na educação, desde a formação básica até o ensino superior. Ao desenvolver essa competência, as instituições de ensino preparam os indivíduos para lidar com os impactos da IA no trabalho e contribuem para a construção de uma sociedade crítica, ética e preparada para os desafios da era digital.

Chao-Hsi *et al.* (2025) investigaram o letramento midiático em contextos de intensa circulação de *fake news*. Eles desenvolveram e validaram uma escala para avaliar dimensões do letramento midiático jornalístico, incluindo compreensão do processo de produção da notícia, análise crítica de informações, engajamento cívico e social e comportamentos de consumo de conteúdos midiáticos. Os resultados mostraram que indivíduos com maior letramento têm menor vulnerabilidade à desinformação, sendo mais

capazes de identificar vieses, avaliar a credibilidade das fontes e agir de maneira responsável ao compartilhar notícias.

Aleedan e Alfadhly (2024) evidenciaram que as escolas primárias promovem a educação midiática, impactando habilidades linguísticas e desenvolvimento crítico dos estudantes. A pesquisa apontou que a dimensão emocional é a mais fortalecida, contemplando ações que estimulam valores culturais, identidade nacional e atitudes éticas diante das mensagens. Outro resultado foi a relação entre educação midiática e aquisição de habilidades em língua inglesa. O uso de ferramentas como rádio escolar, vídeos e plataformas digitais mostrou-se eficaz para estimular a escuta e a fala, principalmente em atividades com sentenças curtas e pronúncia de palavras isoladas. Essa integração contribui para o aprendizado do idioma e para que os alunos desenvolvam competências comunicativas na mídia digital.

Esses resultados reforçam a importância de fortalecer programas de educação midiática nas escolas básicas para combater a desinformação e desenvolver competências digitais e linguísticas, favorecendo uma cidadania crítica e ativa na era digital. O letramento midiático integra comunicação, psicologia e sociologia, favorecendo uma compreensão das relações entre indivíduos e mídia. Essa perspectiva permite desenvolver estratégias educacionais adaptadas a diferentes públicos e contextos, tornando as práticas mais eficazes e inclusivas. Além disso, a criação de métodos para avaliar competências em letramento midiático é fundamental para identificar como distintos grupos interagem com a mídia e detectar lacunas no conhecimento. Essas avaliações orientam o ensino e a formulação de políticas públicas, acompanhando as mudanças tecnológicas e culturais.

METODOLOGIA

Está pesquisa qualitativa examina a produção científica relacionada à educação midiática e ao uso crítico da inteligência artificial na educação básica. Realizou-se uma Revisão Sistemática da Literatura, orientada pelo protocolo PRISMA 2020 (PAGE *et al*., 2021). Para aplicação do protocolo seguiram-se os passos conforme descrito em Appolinário, Hedler e Ferneda (2025). As bases utilizadas foram LES.org e SCOPUS selecionadas por sua relevância internacional e pela complementaridade entre enfoques técnicos e socioeducacionais. Para alcançar os objetivos da pesquisa, definiram-se as palavras-chave de busca (*artificial intelligence* OR AI AND *fake news* OR *Disinformation OR Social Media*).

As etapas recomendadas pelo PRISMA (2021) são: i) identificação; ii) triagem, e iii) elegibilidade e inclusão. A identificação nas bases foi realizada a partir da expressão de busca:

(title:("artificial Intelligence") OR abstract:("artificial Intelligence") OR keyword:("artificial Intelligence") OR field_of_study:("artificial Intelligence")) OR (title:("AI") OR abstract:("AI") OR keyword:("AI") OR field_of_study:("AI")) OR (title:("*Fake news*") OR abstract:("*Fake news*") OR

keyword:("*Fake news*") OR field_of_study:("*Fake news*")) OR (title:("Disinformation") OR abstract:("Disinformation") OR keyword:("Disinformation") OR field_of_study:("Disinformation")) OR (title:("social media") OR abstract:("social media") OR keyword:("social media") OR field_of_study:("social media")).

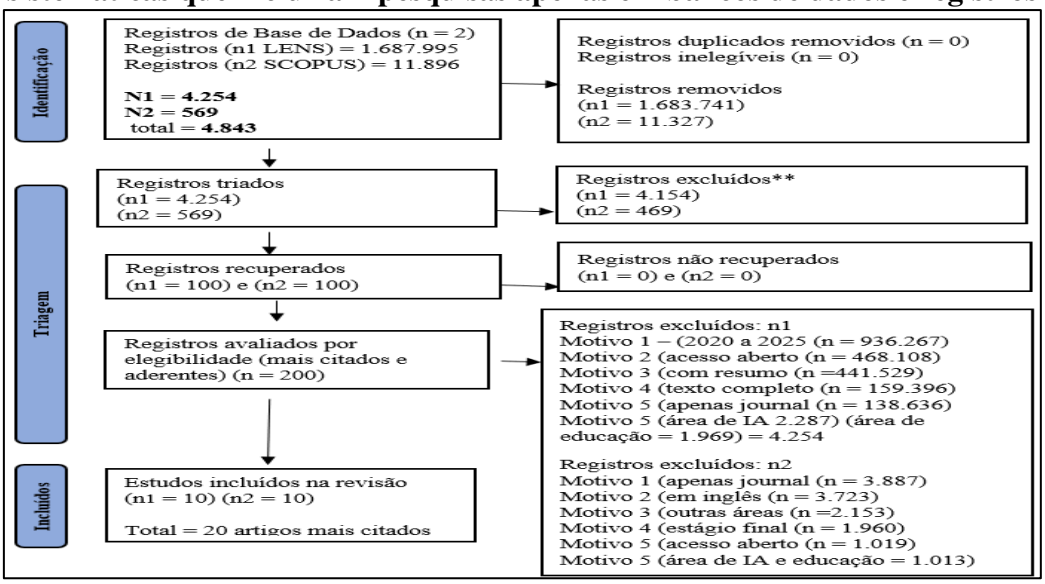
Esses termos foram aplicados aos metadados de título, resumo e palavras-chave. Retornaram nos resultados 1.687.995 publicações em inglês na base Lens e 11.896 na Scopus.

Para a triagem utilizaram-se os filtros: período (2020–2025); tipo de documento (artigos de periódicos); ter resumo e texto completo disponível. Na base Lens a busca resultou em 4.254 artigos científicos e na Scopus 569 documentos.

Para as etapas de elegibilidade e inclusão os artigos foram listados nas bases em ordem decrescente de citação, exportando-se os 100 mais citados para uma planilha formato csv. Os campos selecionados foram: título; autor(es), ano de publicação, resumo. Em seguida procedeu-se à leitura dos dados para a escolha dos 10 artigos de cada base mais aderentes à pesquisa. A amostra final foi de 20 artigos mais citados.

Os artigos foram listados em planilha por ordem de citação. Após essas etapas procedeu-se à construção do *corpus* textual, conforme orientações de Camargo e Justo (2021). Os 20 pdfs dos artigos foram convertidos em texto em formato aberto, cada artigo foi considerado na integra como uma variável, excluindo-se o nome dos autores, título do artigo, palavras-chave, resumos, subtítulos e referencias completas finais. A parametrização envolveu ainda a codificação dos 20 artigos, e exclusão de *stop words*.

Figura 1 - Diagrama PRISMA 2020 para novas revisões sistemáticas que incluíram pesquisas apenas em bancos de dados e registros



Fonte: Elaboração própria. Adaptado de PRISMA (2021).
Nota: *Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).; **If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

Para a análise dos artigos selecionados, utilizou-se a análise lexical. As categorias foram construídas com apoio do software livre Iramuteq (CAMARGO; JUSTO, 2021). Aplicaram-se as análises: i) Classificação hierárquica descendente (CHD); ii) Análise Fatorial de Correspondência; iii) Análise de similitude e, iv) Nuvem de Palavras.

RESULTADOS

Inicialmente descrevem-se os resultados da Bibliometria conforme Método PRISMA (2021) a partir das 20 pesquisas listadas na Tabela 1. A descrição dos artigos seguiu à similaridade de abordagem ou aproximação ao objeto de estudo. Em seguida foramo apresentados os resultados das análises geradas no Iramuteq: i) Classificação hierárquica descendente (CHD); ii) Análise Fatorial de Correspondência; iii) Análise de similitude e, iv) Nuvem de Palavras.

Tabela 1 - Os 20 artigos mais citados incluídos na Revisão Sistemática da Literatura

Nº	Autores	Título	Ano de Publiação	Periódico	Total de citações
Base de dados Lens.org					
1	Andy Nguyen	Ethical principles for artificial intelligence in education.	2022	Education and information technologies	1168
2	Iason Gabriel	Artificial Intelligence, Values, and Alignment	2020	Minds and Machines	1128
3	Theo Araujo	In AI we trust? Perceptions about automated decision-making by artificial intelligence	2020	AI & SOCIETY	1078
4	Kyoungwon Seo	The impact of artificial intelligence on learner-instructor interaction in *online* learning.	2021	International journal of educational technology in higher education	932
5	Esma Aïmeur	*Fake news*, disinformation and misinformation in social media: a review.	2023	Social network analysis and mining	768
6	Olga Tapalova	Artificial Intelligence in Education: AIEd for Personalised Learning Pathways	2022	Electronic Journal of e-Learning	715
7	Amy Orben	Teenagers, screens and social media: a narrative review of reviews and key studies	2020	Social psychiatry and psychiatric epidemiology	505
8	Jeffrey P. Carpenter	How and why are educators using Instagram	2020	Teaching and teacher education	452
9	Sabine Seufert	Technology-related knowledge, skills, and attitudes of pre- and in-service teachers: The current situation and emerging trends.	2020	Computers in human behavior	395
10	Daniel Schiff	Out of the laboratory and into the classroom: the future of artificial intelligence in education.	2020	AI & society	235
Base de dados SCOPUS					
1	Tlili *et al.*	What if the devil is my guardian angel: ChatGPT as a case study of using *chatbots* in education	2023	Smart Learning Environments	1.877
2	Vaccari; Chadwick	*Deepfake*s and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News	2020	Social Media and Society	975
3	Wach *et al.*	The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT	2023	Entrepreneurial Business and Economics Review	609
4	Quach *et al.*	Digital technologies: tensi1-ons in privacy and data	2022	Journal of the Academy of Marketing Science	641
5	Kozyreva; Lewandowsky; Hertwig.	Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools	2020	Psychological Science in the Public Interest	276
6	Wardat *et al.*	ChatGPT: A revolutionary tool for teaching and learning mathematics	2023	Eurasia Journal of Mathematics, Science and Technology Education	527
7	Chuah; Yu.	The future of service: The power of emotion in human-robot interaction	2021	Journal of Retailing and Consumer Services	288
8	Fagni *et al.*	TweepFake: About detecting *deepfake* tweets	2021	PLoS ONE	299
9	Yin; Zubiaga.	Towards generalisable hate speech detection: a review on obstacles and solutions	2021	PeerJ Computer Science	266
10	Jakesch; Hancock; Naaman.	Human heuristics for AI-generated language are flawed	2023	Proceedings of the National Academy of Sciences of the United States of America	313

Fonte: Elaboração própria.

Nguyen (2022) busca identificar um consenso global sobre os princípios éticos da inteligência artificial na educação (AIED). Ele realiza uma análise temática de políticas e diretrizes de organizações internacionais, como UNESCO e OCDE, para estabelecer um conjunto de princípios éticos para uma AIED confiável e ética. O estudo propõe um guia para educadores, desenvolvedores de tecnologia e formuladores de políticas na criação e implementação de sistemas de AIED.

Gabriel (2020) explora os valores que os sistemas de inteligência artificial (IA) devem alinhar. Ele apresenta diferentes visões, incluindo uma abordagem utilitarista, que defende que a IA deve criar a maior felicidade para o maior número de pessoas, e uma kantiana, que sugere que os princípios devem ser universalmente aplicáveis, como os de justiça e benevolência.

Wach *et al.* (2023) mapearam os desafios e riscos da IA generativa (IAG) em gestão e economia, usando o ChatGPT como estudo de caso. Eles fizeram isso por meio de revisão de literatura narrativa e crítica. Desenvolveram uma estrutura conceitual que organiza os aspectos negativos da IAG, suas controvérsias e ameaças, complementando uma publicação anterior sobre os positivos. Identificaram sete ameaças principais, focando no ChatGPT: a falta de regulamentação do mercado de IA; a baixa qualidade e viés algorítmico; a perda de empregos pela automação; a violação de dados pessoais e privacidade; a manipulação social e o enfraquecimento da ética; o aumento das desigualdades socioeconômicas e o tecnostresse. O artigo destaca a necessidade urgente de regulamentar o mercado para garantir concorrência justa e proteger a privacidade. Além disso, recomendam que os desenvolvedores priorizem considerações éticas e que os trabalhadores adquiram novas habilidades para se adaptarem às mudanças do mercado de trabalho.

Araujo *et al*. (2020) exploram a relação entre características pessoais e percepções de utilidade, justiça e risco das decisões automatizadas por inteligência artificial (IA). A pesquisa foi feita por meio de um experimento com cenários e uma amostra representativa da população holandesa (N=958). O estudo comparou decisões de IA e de especialistas humanos em mídia, saúde e justiça. O objetivo era compreender como fatores como conhecimento, preocupação com a privacidade *online*, autoeficácia, dados demográficos e crença em igualdade influenciam a percepção pública sobre a IA. Os resultados indicam preocupação com os riscos, com opiniões divididas sobre sua utilidade e justiça social. Embora a maioria dos participantes demonstre cautela, o estudo revela que, em certas situações, as decisões da IA foram avaliadas de forma equivalente ou superior às de especialistas humanos. Concluem que o conhecimento sobre IA e a autoeficácia *online* estão positivamente associados à percepção de utilidade e justiça. Uma maior preocupação com a privacidade *online* está associada a uma percepção crítica, de menor utilidade e justiça, e maior risco. O estudo destaca que a idade e o gênero influenciam essas

percepções, com mulheres e pessoas mais velhas vendo a IA como menos útil ou mais arriscada (ARAUJO *et al*., 2020).

Quach *et al*. (2022) analisaram as tensões das tecnologias digitais em relação à privacidade e dados nas interações entre empresas, consumidores e reguladores. A pesquisa empírica utiliza a teoria da estruturação e a literatura sobre privacidade para propor um arcabouço conceitual com três princípios e sete proposições, buscando entender as implicações das tecnologias digitais para o desempenho das empresas em um cenário de crescente preocupação com a privacidade e ramificações legais. A partir desse *framework*, propuseram uma tipologia de estratégia de dados baseada nas funções de monetização e compartilhamento, resultando na identificação de quatro tipos distintos de empresas. A pesquisa aponta que as preocupações têm abalado as relações entre empresas e consumidores, levando a intervenções regulatórias e a comportamentos de proteção dos usuários. Os achados indicam que muitos consumidores não compreendem o uso dos dados coletados e sentem-se incapazes de proteger suas informações. Concluem que a proteção eficaz da privacidade requer colaboração entre consumidores, empresas e órgãos reguladores e que, à medida que se tornam mais conscientes sobre as tecnologias digitais, adotam medidas proativas em relação a seus dados. Ressaltaram a necessidade de mais estudos sobre inovação em privacidade, especialmente para determinar se os benefícios da adoção justificam os custos.

Kozyreva; Lewandowsky; Hertwig (2020) analisam os desafios enfrentados pelos cidadãos na era da internet. Eles propõem um mapa conceitual de intervenções baseadas na ciência psicológica. O estudo descreve as diferenças entre os ambientes *online* e *offline* e identifica quatro desafios principais: arquiteturas de escolha persuasivas e manipuladoras; arquiteturas de informação assistidas por IA; informações falsas e enganosas; e ambientes de distração. A pesquisa se concentra na intervenção de *boosts*, que visam capacitar os usuários da internet, aumentando sua literacia da informação e resistência cognitiva à manipulação. Os principais achados indicaram que a internet, apesar de seus benefícios, é um ambiente repleto de arquiteturas de escolha projetadas para maximizar interesses comerciais e influenciar o comportamento dos usuários. Concluem que, embora regulamentações possam mitigar os excessos, o poder e os recursos permanecem com os agentes comerciais. A solução proposta é empoderar os usuários, por meio de ferramentas cognitivas que melhorem a capacidade de raciocínio e a resiliência à manipulação. O estudo enfatiza que a informação *online*, ao contrário da *offline*, muitas vezes carece de sinais claros de qualidade e autoria.

Aïmeur, Amri e Brassard (2023) em revisão sistemática, narrativa e crítica sobre *fake news*, *misinformation* e *disinformation,* analisaram as abordagens para sua detecção e prevenção em redes sociais *online* (OSNs). Focaram em artigos dos últimos dez anos, selecionados de Google Scholar, IEEE Xplore e Scopus, e buscaram responder a três questões: o que é *fake news* e seus conceitos relacionados,

quais os desafios para sua detecção, e quais técnicas são utilizadas para combatê-la. Os resultados demonstraram que a detecção automática de notícias falsas é um desafio complexo, pois o conteúdo se assemelha à verdade, dificultando a identificação por inteligência artificial (IA). Revelou que a propagação de informações falsas é significativamente maior que a de verdadeiras, e que a maioria dos usuários tem dificuldade em distingui-las. Destacaram que abordagens de IA ainda são insuficientes e que a colaboração multidisciplinar, incluindo psicologia, sociologia e jornalismo, é necessária para combater a *infodemia*.

Vaccari e Chadwick (2020) investigaram o impacto dos *deepfakes* - vídeos sintéticos gerados por IA - na desinformação *online*. A pesquisa integrou teorias sobre o poder da comunicação visual e o papel da incerteza para compreender a contribuição dos *deepfakes* na erosão da confiança pública. Os autores conduziram um experimento *online* com 2.005 pessoas no Reino Unido, comparando a avaliação de vídeos autênticos e manipulados. Os principais achados mostram que, em vez de serem enganadas, as pessoas se sentem incertas sobre a veracidade dos *deepfakes, resultando* em uma diminuição da confiança nas notícias das mídias sociais. O estudo conclui que a disseminação descontrolada pode contribuir para um clima de *indeterminação e cinismo*, prejudicando a cultura cívica *online*. Além disso, a maioria tem pouca capacidade de discernir, identificando-os corretamente em cerca de 50% dos casos.

Fagni *et al*. (2021) investigaram a detecção de *tweets* falsos criados por IA, conhecidos como *deepfake tweets*. Os autores introduziram um novo conjunto de dados, TweepFake, composto de *tweets* reais e falsos, para desenvolvimento de modelos de detecção. O estudo demonstrou a eficácia de uma abordagem baseada no modelo de linguagem RoBERTa, alcançando desempenho quase perfeito (FAGNI *et al*., 2021). O modelo obteve uma precisão de 99,8% na detecção, superando outros métodos, devido às diferenças linguísticas sutis entre textos gerados por IA e escritos por humanos. A alta precisão é uma contribuição significativa para a detecção de *deepfakes*, destacando a importância de *datasets* de qualidade e a necessidade de aprimorar ferramentas de combate à desinformação *online*.

Jakesch, Hancock e Naaman (2023) investigaram a capacidade humana de diferenciar textos escritos por humanos e gerados por inteligência artificial (IA). Eles realizaram seis experimentos com mais de 4.600 participantes, onde os voluntários deveriam identificar descrições pessoais em contextos profissionais, de hospitalidade e de namoro criadas por humanos ou por modelos de linguagem de última geração. O objetivo foi demonstrar a falha dos heurísticos cognitivos na detecção de textos gerados por IA levantando preocupações sobre manipulação e engano online. Os resultados mostraram que os participantes não conseguiram detectar a linguagem gerada. Uma análise computacional revelou que os julgamentos humanos são prejudicados por heurísticas intuitivas, como a associação do uso de pronomes em primeira pessoa, contrações e tópicos familiares com a autoria humana. A pesquisa demonstrou que

esses vieses tornam o julgamento previsível e manipulável, permitindo que a IA produza textos percebidos como *mais humanos*. Concluíram que, apesar das preocupações sobre o uso indevido de IA, os humanos não conseguem identificar a linguagem gerada, tornando a identificação das fontes de informação uma questão urgente.

Em uma revisão narrativa da literatura, Orben (2020) examinou a relação entre o uso de tecnologias digitais e mídias sociais e o bem-estar psicológico de adolescentes, analisando mais de 80 revisões sistemáticas e meta-análises. A autora destaca que a pesquisa existente é de baixa qualidade e focada em trabalhos transversais, impedindo uma interpretação clara das correlações. O estudo conclui que a falta de evidências robustas demonstra a necessidade de melhorias metodológicas e maior transparência na pesquisa. Os achados indicaram que a associação entre o uso de tecnologias e o bem-estar psicológico é, em média, negativa, mas muito pequena. A direção da causalidade entre as variáveis permanece incerta, pois os efeitos foram encontrados em ambas as direções. Destacaram a importância de diferenciar os tipos de uso, como o passivo (visualização de conteúdo) e o ativo (interação). Enquanto o passivo pode estar associado a um menor bem-estar devido a comparações sociais, o ativo pode ter efeitos positivos ao aumentar o suporte social. Conclui que os resultados conflitantes demonstram a necessidade de pesquisas futuras considerarem as diferenças individuais e adotar medidas mais precisas do que o *tempo de tela* para compreender os impactos.

Carpenter *et al*. (2020) investigaram como e por que educadores utilizam Instagram. Eles conduziram uma pesquisa por meio de um questionário *online*, analisando as respostas com base em *espaço de afinidade* e *empreendedorismo docente* (*teacherpreneurship*). As análises sugerem que os participantes são usuários intensivos, engajados na troca de conhecimento profissional e saberes práticos, além de oferecimento e recebimento de apoio afetivo. Uma descoberta é a tendência de educadores em misturar conteúdo pessoal e profissional, o que pode fortalecer relações *online*, mas levanta questões sobre o risco de colapso de contexto. O estudo aponta desafios como a superficialidade do conteúdo priorizado pela estética, o potencial para a criação de *câmaras de eco* e os riscos associados ao empreendedorismo docente e à exposição de informações pessoais.

Seo *et al*. (2021) investigaram a percepção de alunos e professores sobre o impacto da inteligência artificial (IA) na interação em ambientes de aprendizagem *online*. Eles usaram o *Speed Dating* com *storyboards* para analisar as opiniões de 12 estudantes e 11 instrutores sobre casos de uso na educação. O objetivo era identificar lacunas, desafios e barreiras que impedem os sistemas de IA de atingirem seu potencial, mantendo interações seguras. Os resultados indicaram que a adoção na educação *online* pode possibilitar uma interação personalizada em grande escala, mas apresenta o risco de violar limites sociais. Os participantes reconheceram os benefícios da IA na melhoria da comunicação e no fornecimento de

suporte personalizado, mas manifestaram preocupações com responsabilidade, autonomia e vigilância. Sugerem que o design futuro deve garantir a explicabilidade, a participação humana (*human-in-the-loop*) e a coleta e apresentação cuidadosa de dados para maximizar os impactos positivos e minimizar os negativos.

Seufert, Guggemos e Sailer (2021) estudaram o conhecimento, habilidades e atitudes (KSA) de professores em relação à tecnologia. Eles definem KSA, apresentam métodos de avaliação e discutem formas de aprimorá-los em formação e em serviço. A pesquisa utiliza o framework TPACK (Conhecimento Tecnológico, Pedagógico e de Conteúdo) como base teórica e explora a inteligência artificial (IA) e *estratégias de aumento* (*augmentation strategies*) como tendências emergentes na educação. As principais conclusões indicam que a tecnologia pode melhorar os processos de aprendizagem, mas sua eficácia depende do uso pelo professor. O TPACK é um modelo central para o desenvolvimento do conhecimento tecnológico do docente, embora os autores destaquem desafios em sua conceituação e na previsão de resultados práticos de aprendizagem. Além disso, ressaltam que as atitudes e crenças dos professores em relação à tecnologia são centrais na adoção e uso em sala de aula, sugerindo a importância de uma abordagem que considere a formação profissional contínua para o desenvolvimento dessas competências.

Schiff (2021) analisa o estado atual e futuro da inteligência artificial na educação (AIEd), focando em sistemas de tutoria inteligentes e agentes educacionais artificiais. O autor contextualiza o papel da AIEd historicamente, contrastando-a com tecnologias anteriores, como a educação a distância. Ele discute as possibilidades e riscos sociotécnicos por meio de duas visões de futuro: o *modelo de fábrica* e o *modelo de cidadania.* A metodologia inclui uma avaliação crítica do campo, situando o desenvolvimento da AIEd em um contexto de responsabilidade social e implementação ética. As principais conclusões indicam que, ao contrário de tecnologias como os MOOCs que carecem de recursos inteligentes, a AIEd pode superar a barreira da interação professor-aluno, simulando educadores, proporcionando diferenciação robusta para os alunos e promovendo o engajamento socioemocional. No entanto, alerta para os riscos de um *modelo de fábrica* tecnocrático, onde eficiência e automação priorizam escala em detrimento da qualidade, levando à alienação. Enfatiza a necessidade de desenvolvimento e implementação responsáveis, com a participação de pesquisadores, desenvolvedores e partes interessadas, como professores, alunos e pais, para garantir que a tecnologia melhore a educação de forma equitativa e eficaz.

Tlili *et al*. (2023) analisam o uso do ChatGPT na educação, adotando um estudo de caso qualitativo e instrumental. A pesquisa teve três etapas: análise de rede social de *tweets* para entender o discurso público, análise de conteúdo de entrevistas com usuários iniciais para explorar suas percepções e preocupações, e investigação das experiências em dez cenários educacionais. O objetivo principal foi

identificar as preocupações associadas ao uso de *chatbots*, especificamente o ChatGPT, na educação. Os resultados mostram que, embora o discurso público nas mídias sociais seja positivo e entusiástico, a formação de uma comunidade de discussão ainda é fragmentada. A análise de sentimento revelou que, embora os positivos superem os negativos, a maioria dos usuários permanece indecisa sobre o potencial da ferramenta. A investigação em cenários educacionais levantou questões e desafios, incluindo o risco de trapaça, a veracidade e honestidade do ChatGPT, preocupações com privacidade e o potencial de manipulação de informações.

Wardat *et al.* (2023) exploraram o potencial do ChatGPT no ensino de matemática. Os principais achados indicaram que ele pode atuar como recurso de apoio, oferecendo *feedback* instantâneo, gerando exemplos práticos e ajudando a resolver problemas de forma personalizada. A ferramenta se destacou por sua capacidade de adaptação ao ritmo de cada estudante, proporcionando um aprendizado mais individualizado. No entanto, apontaram a necessidade de os educadores integrarem a ferramenta de forma estratégica e ética para maximizar seus benefícios, evitando que os alunos a utilizem apenas para obter respostas prontas.

Tapalova e Zhiyenbayeva (2021) exploram o uso da IA na educação (AIEd) para criar percursos de aprendizagem personalizados. Eles propuseram uma estrutura conceitual que inclui redes sociais, *chatbots*, sistemas especialistas, mentores e agentes inteligentes, aprendizado de máquina, sistemas educacionais personalizados e ambientes de realidade virtual. A metodologia incluiu um estudo de caso com 184 estudantes de Pedagogia e Psicologia, conduzido com discussões em grupo. Os achados indicam que a IA é uma força motriz de mudança, focando nas necessidades individuais dos alunos para ajudá-los a dominar novos conhecimentos e desenvolver competências profissionais.

Chuah e Yu (2021) investigaram como robôs que expressam emoções influenciam os sentimentos dos consumidores. A pesquisa utiliza a teoria do contágio emocional e a literatura sobre interação entre humanos e robôs, empregando algoritmos de aprendizado de máquina e análise de sentimento em dados do Instagram. Os resultados indicaram que a expressão de emoções como surpresa e felicidade são essenciais para impactos positivos nos consumidores. A capacidade dos robôs de transmitir emoções é uma força significativa no serviço, influenciando a experiência do cliente. As conclusões reforçam a importância de integrar a expressão emocional no design de robôs de serviço para melhorar a aceitação da tecnologia.

Yin e Zubiaga (2021) revisaram a literatura sobre a capacidade de generalização dos modelos de detecção de discurso de ódio, que ataca ou promove ódio contra grupos ou indivíduos com base em características de identidade, como etnia, religião e orientação sexual. Apesar do crescente interesse na detecção automática como uma tarefa de processamento de linguagem natural, os modelos existentes têm

baixo desempenho em novos dados. As análises indicaram que eles generalizam mal para dados não vistos. O estudo propõe que a falta de generalização é um obstáculo significativo e sugere caminhos para superá-lo, reforçando a necessidade de uma abordagem mais robusta para melhorar a detecção de discurso de ódio *online*.

Em suma, a análise revela que a inteligência artificial (IA) e as tecnologias digitais oferecem oportunidades e desafios interligados. Estudos apontam para o potencial de ferramentas como a IA na educação (AIED) para personalizar o aprendizado e na gestão para otimizar processos, mas surgem riscos éticos, sociais e técnicos. A desinformação, a dificuldade em distinguir conteúdos falsos de verdadeiros, preocupações com privacidade e segurança dos dados, e o viés algorítmico exigem atenção. O consenso entre os autores é que, para que essas tecnologias cumpram suas promessas de forma ética e equitativa, é fundamental uma responsabilidade compartilhada. Desenvolvedores, educadores, formuladores de políticas e o público precisam atuar juntos para regulamentar o setor, capacitar os usuários com ferramentas cognitivas (literacia da informação) e priorizar considerações éticas no design e na implementação, garantindo que a inovação tecnológica sirva ao bem-estar da sociedade.

Análises Lexicais

A análise revelou uma divisão disciplinar no estudo da desinformação gerada por IA, com implicações para o avanço do campo e o desenvolvimento de soluções interdisciplinares. Os resultados foram apresentados conforme descrito no método: i) Classificação hierárquica descendente (CHD); ii) Análise Fatorial de Correspondência; iii) Análise de similitude e, iv) Nuvem de Palavras.

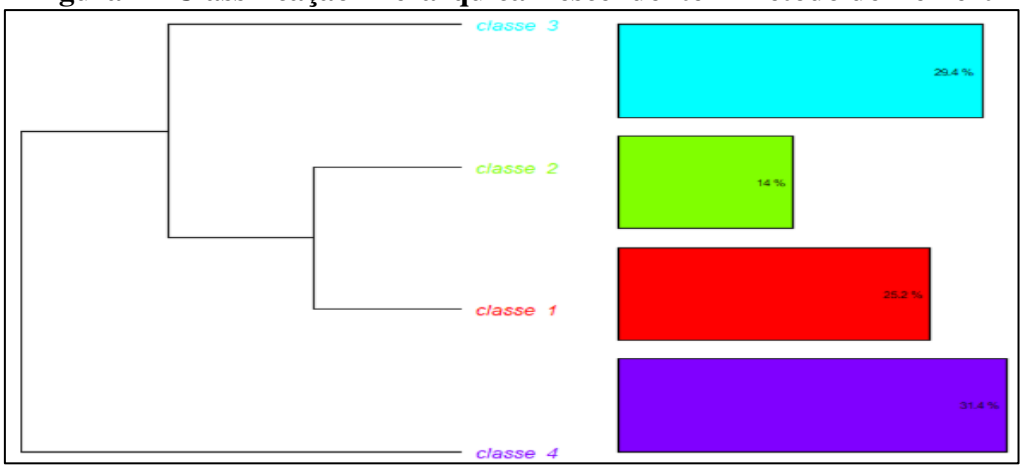
O *corpus* analisado foi composto por textos sobre inteligência artificial na educação, desinformação e educação midiática. O processamento foi feito no software Iramuteq, que organizou os dados em segmentos de texto (STs), considerando a frequência e a coocorrência das formas lexicais. O aproveitamento do *corpus* foi satisfatório, permitindo a execução completa dos métodos previstos, incluindo a Classificação Hierárquica Descendente (CHD), Análise Fatorial de Correspondência (AFC) e Análise de Similitude.

A CHD aplicada a 20 artigos sobre o problema de pesquisa visa organizar e interpretar o conteúdo textual de forma estruturada. Essa análise identifica os principais temas, conceitos e discursos na literatura. Mostra como os artigos se agrupam por semelhança de conteúdo, revelando padrões e lacunas. Auxilia na compreensão da estrutura do conhecimento existente, servindo como base para justificar a pesquisa e apontar oportunidades de investigação futura. Em suma, a CHD transforma um grande volume de texto em informações organizadas e interpretáveis, permitindo ao pesquisador enxergar quais discussões predominam, como elas se conectam e quais aspectos carecem de estudo (CAMARGO, JUSTO, 2021).

A partir do método de Reinert, foram identificadas quatro classes principais, representando agrupamentos temáticos com base na distribuição estatística das formas textuais. Para a seleção dos segmentos representativos de cada classe, foi adotado o critério relativo (média do qui-quadrado), priorizando trechos com maior densidade de termos característicos, ajustando diferenças no tamanho.

O *corpus* textual foi formado por 20 textos, 4.760 segmentos (STs) e 168.205 ocorrências (palavras, formas, vocábulos), 12.683 formas e 5.696 únicas (hápax). Foram classificados 3.668 STs, equivalendo a 77,06%. O conteúdo analisado foi categorizado em quatro classes: Classe 1, com 923 ST (25,16 %); Classe 2, com 515 ST (14,04%); Classe 3, 1.078 ST (29,39%) e Classe 4 com 1.152 ST (31,41%), conforme Figura 2.

Figura 2 - Classificação Hierárquica Descendente – Método de Reinert



Fonte: Elaboração própria.

A tabela 2 ilustra o resultado da CHD. Ela apresenta as classes e temas, segmentos típicos e score de cada uma, com dois exemplos de texto característico.

Tabela 2 - Segmentos de textos característicos por classes e temas principais

Classe	Segmento típico	Score	Tema Principal
Classe 1 Ética e Princípios na IA Educacional	The application of artificial intelligence in education has been featured as one of the most pivotal developments of the century. AIED must ensure well-informed consent from the user and maintain the confidentiality of the users' information.	42,8 34,8	Consentimento informado; Privacidade e segurança de dados; Transparência e explicabilidade; Autonomia do estudante; Mitigação de vieses e equidade; Responsabilidade e prestação de contas; Inclusão e acessibilidade
Classe 2 Bem-estar Digital e Impactos no Aprendizado	The widespread focus on screen time as the measure of digital technology use can be explained by our inability to differentiate between various forms of screen activities. The conflicting results highlight that the evidence is still too weak to promote a uniform interpretation of the correlation between time spent on digital technologies and well-being outcomes.	37,5 41,7	Tempo de tela e medição de uso; Saúde mental e bem-estar; Impactos em aprendizagem e desempenho; Uso de redes sociais; Qualidade metodológica (longitudinal vs transversal); Intervenções e literacia digital
Classe 3 Governança e Regulação da AIED	UNESCO launched global standards for AI ethics which were agreed and signed by its 193 member countries. Principle of governance and stewardship of AIED should carefully take into account multi-stakeholder perspectives and ethical considerations."	45,2 41,7	Padrões e diretrizes internacionais; Estruturas de governança e stewardship; Auditabilidade e compliance; Sustentabilidade e proporcionalidade; Políticas de acesso e infraestrutura; Capacitação de educadores; Proteção de direitos
Classe 4 Fake news e Desinformação	Fake news spreads significantly farther, faster, deeper, and more broadly than the truth, especially in politics. Automatic fake news detection remains a huge challenge primarily because the content closely resembles the truth to deceive users.	50,1 47,9	Tipologias de desinformação; Métodos automáticos de detecção (ML, redes, fact-checking); Dinâmica de propagação em redes; Limitações técnicas e éticas da detecção automática; Intervenções comportamentais; Políticas de moderação; Alfabetização midiática

Fonte: Elaboração própria.

A Tabela 2 revela quatro áreas de debate. A Classe 1) Ética e Princípios na IA Educacional, os textos enfatizam questões éticas na aplicação da IA em ambientes educacionais, destacando consentimento informado, privacidade, mitigação de vieses e inclusão. O alto score (42,8 e 34,8) indica forte representatividade desses trechos, evidenciando preocupação com o uso responsável e transparente da tecnologia.

Classe 2: Bem-estar Digital e Impactos no Aprendizado. Essa classe aborda os efeitos do uso das tecnologias digitais no bem-estar e desempenho acadêmico, focando em tempo de tela, saúde mental e qualidade das pesquisas. Os segmentos mostram evidências conflitantes, sinalizando que ainda não há consenso sobre os impactos do uso intensivo de tecnologias (scores de 37,5 e 41,7).

Classe 3: Governança e Regulação da AIED relaciona-se a estruturas normativas e políticas globais, como os padrões da UNESCO. Destaca a necessidade de governança multissetorial e mecanismos de *compliance* e auditabilidade. Os altos scores (45,2 e 41,7) indicam que esses trechos são fortemente associados às discussões sobre regulação e sustentabilidade da IA educacional.

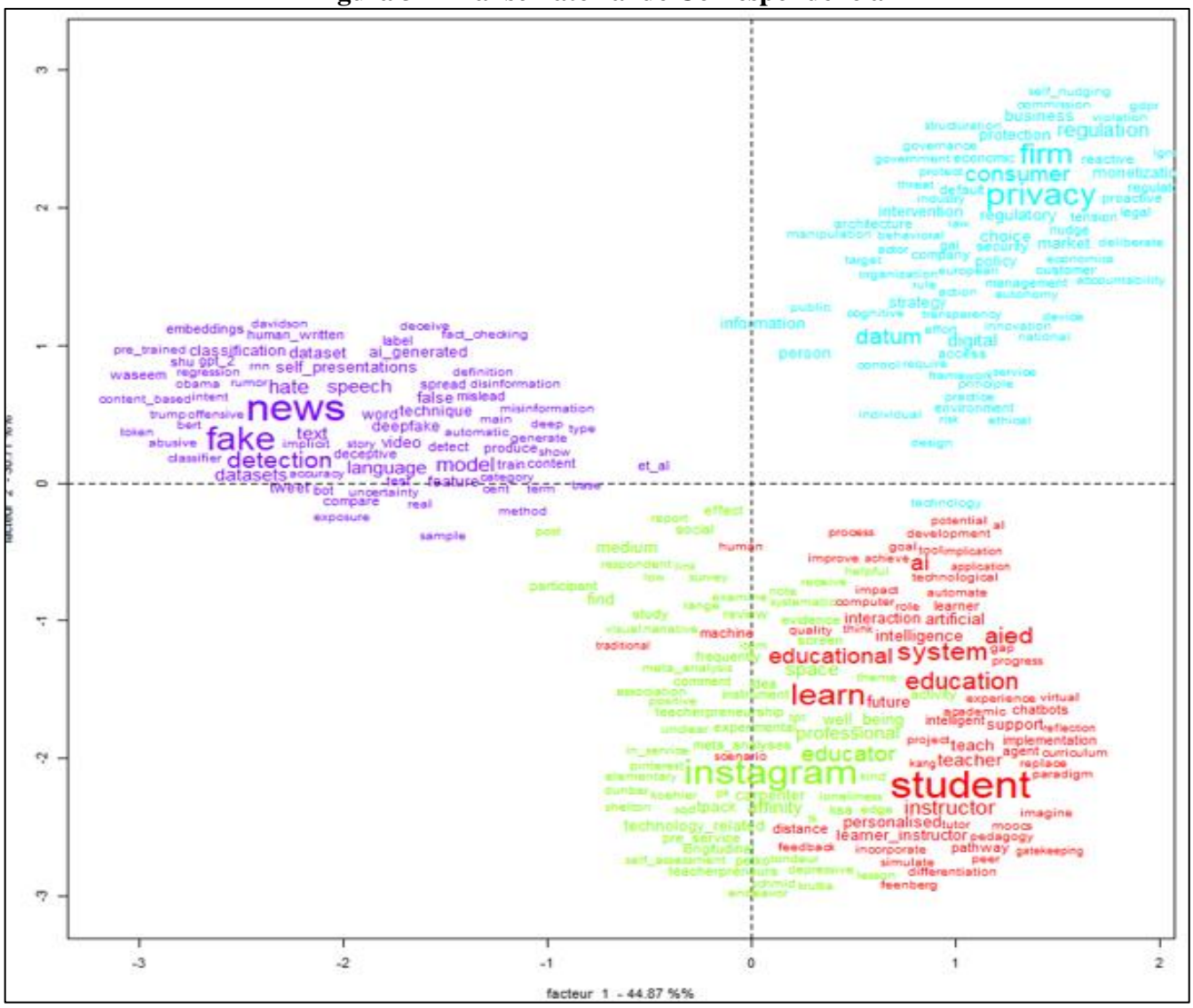
Classe 4: *Fake news* e Desinformação. Esta classe foca na propagação de desinformação, enfatizando a velocidade das *fake news* e os desafios de sua detecção automática. Inclui temas como métodos de detecção, intervenções educativas e políticas de moderação. Esta classe obteve os scores mais altos (50,1 e 47,9), sugerindo maior concentração temática e relevância no *corpus* analisado.

O destaque para a Classe 4, com os maiores scores, sugere que *fake news* e desinformação são os temas mais representados no *corpus. As* demais classes demonstram preocupações complementares, refletindo um panorama abrangente sobre os dilemas da inteligência artificial na educação.

Para apoiar a análise pelo método de Reinert, foi realizada a Análise Fatorial de Correspondência (AFC) para visualizar as relações entre as classes e identificar as proximidades entre os grupos temáticos.

Os resultados indicaram que: as Classes 3 e 4 estão próximas, evidenciando forte relação entre os discursos sobre sistemas algorítmicos, *fake news* e desinformação; a classe 1 apresenta relativa independência temática, sugerindo um núcleo específico voltado à ética e governança da inteligência artificial na educação; a classe 2 ocupa uma posição intermediária, conectando os debates sobre valores humanos às demais, funcionando como eixo entre tecnologia, ética e impactos sociais. Essa disposição visual reforça a interpretação de que os debates sobre desinformação e infodemia estão associados à dinâmica tecnológica e algorítmica, enquanto a ética e os direitos humanos aparecem como fundamentos necessários, porém menos integrados ao discurso prático.

Figura 3 - Análise Fatorial de Correspondência



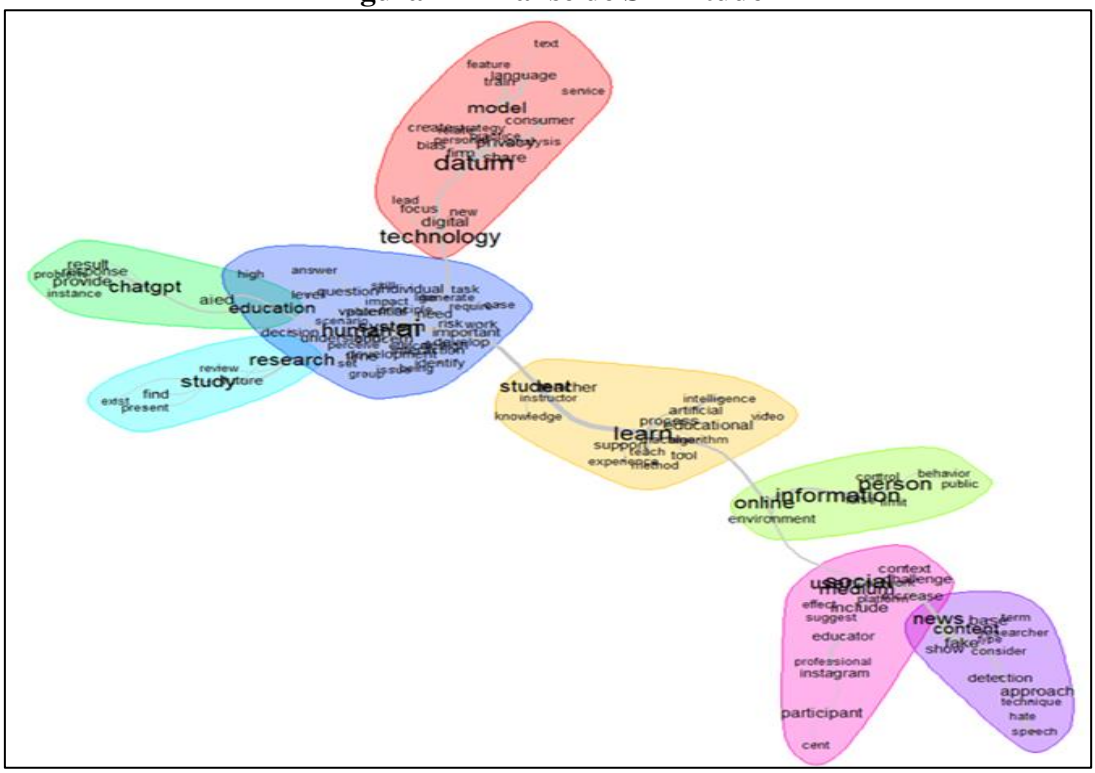
Fonte: Elaboração própria.

Na Análise de Similitude, o gráfico mostra a relação entre palavras do *corpus*, evidenciando conexões e coocorrências frequentes. Cada grupo representa um cluster temático, com termos associados. Diferente da Classificação Hierárquica Descendente (CHD), que separa o *corpus* em classes temáticas, a similitude não divide o texto. Ela mostra as conexões entre os termos, ajudando a entender como os temas se interligam.

Na Figura 4, o núcleo central (azul) indica que a educação e os estudantes são o eixo principal. Os termos *education*, *student*, *research* e *study* formam o centro. O cluster vermelho (topo) sugere foco em ferramentas e infraestrutura da IA, com palavras como *datum*, *technology* e *model*, relacionadas à tecnologia e dados. O cluster verde (esquerda) vincula o uso prático de IA em contextos educacionais e suporte acadêmico aos termos chatgpt e provide. No cluster amarelo (direita), as palavras *information*,

person e *approach*, associadas a interações digitais e uso de informações, incluindo aspectos éticos e metodológicos. O cluster roxo (inferior direito) destaca *news* e *participant*, relacionado a *fake news*, desinformação e moderação de conteúdo. A Figura 3 sugere que o *corpus* está estruturado em torno de quatro eixos principais: educação e estudantes como tema central; tecnologia e dados como base de suporte; aplicação prática da IA (ex.: ChatGPT), e, desinformação e disseminação de notícias falsas. Essas conexões reforçam a integração entre o uso de IA na educação, seus desafios técnicos e os impactos na informação e na sociedade.

Figura 4 - Análise de Similitude



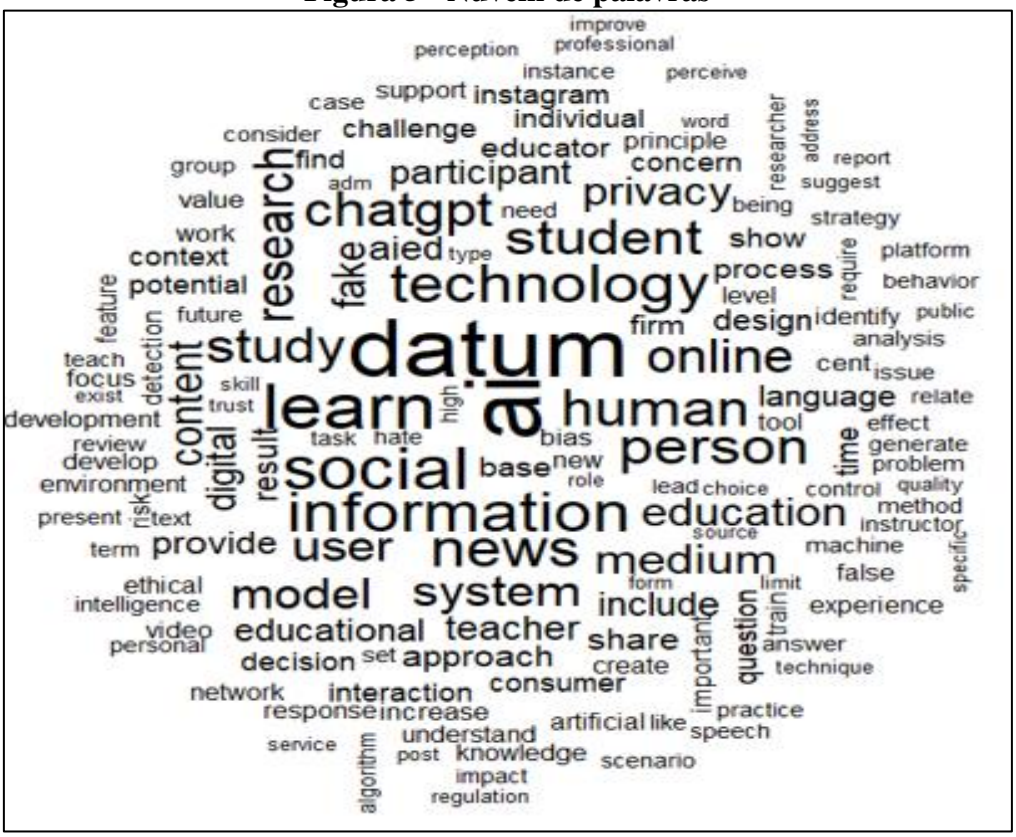
Fonte: Elaboração própria.

A nuvem de palavras (Figura 5) representa a frequência relativa dos termos no *corpus* textual analisado. Quanto maior o tamanho da palavra, maior a sua ocorrência e relevância no conjunto de textos. Esse tipo de visualização permite identificar rapidamente os conceitos centrais e recorrentes nas discussões.

A nuvem confirma os quatro temas do dendrograma e na AFC: i) Governança e privacidade – termos como *privacy*, *user*, *regulation*; ii) Educação e aprendizado com IA - centralidade de *learn*, *student*, *education*, *chatgpt*; iii) Desinformação e *fake news* - presença de *news* e *fake*; e, iv) Tecnologia e dados - destaque para *datum*, *technology*, *system*, *model*.

A nuvem confirma os quatro temas do dendrograma e na AFC: i) Governança e privacidade – termos como *privacy*, *user*, *regulation*; ii) Educação e aprendizado com IA - centralidade de *learn*, *student*, *education*, *chatgpt*; iii) Desinformação e *fake news* - presença de *news* e *fake*; e, iv) Tecnologia e dados - destaque para *datum*, *technology*, *system*, *model*.

Figura 5 - Nuvem de palavras



Fonte: Elaboração própria.

A nuvem de palavras mostra que o núcleo semântico do corpus é o debate sobre uso de dados e tecnologia na educação, interligado a aprendizado, ética, privacidade e desinformação. Ela representa visualmente os resultados do dendrograma e da AFC, reforçando a coerência das classes temáticas.

DISCUSSÃO DOS RESULTADOS

Os resultados revelam um panorama complexo sobre o papel da inteligência artificial (IA) na educação e seus desdobramentos sociais, éticos e políticos. A partir da descrição dos artigos mais citados nas bases Lens e Scopus e da análise bibliométrica pelo software Iramuteq, identificaram-se quatro eixos temáticos: (i) Ética e princípios na IA educacional, (ii) Bem-estar digital e impactos no aprendizado, (iii)

Governança e regulação da AIED e (iv) Fake news e desinformação. Esses eixos refletem a diversidade de abordagens teóricas e metodológicas e a natureza interdisciplinar do fenômeno investigado.

A Classe 1 abordou ética e princípios na IA educacional e evidenciou a centralidade de temas como privacidade, consentimento informado, mitigação de vieses e acessibilidade. Esses achados convergem com Nguyen (2022) e Schiff (2021), que defendem uma inteligência artificial na educação (AIED) pautada por princípios éticos claros para assegurar transparência e responsabilidade. Apesar dessa preocupação na literatura, a análise fatorial de correspondência (AFC) indica que ela se apresenta de forma relativamente independente em relação às demais classes, sugerindo que os debates éticos ainda não estão plenamente integrados às discussões sobre desinformação e impacto social. Essa fragmentação reforça o argumento de Wach *et al*. (2023) sobre a urgência de aproximar considerações éticas do desenho e da implementação prática das tecnologias educacionais, evitando a predominância de modelos tecnocráticos que priorizam eficiência em detrimento da equidade.

A Classe 2, relacionada ao bem-estar digital e impactos no aprendizado, destaca uma literatura com evidências conflitantes. Orben (2020) apontou a limitação metodológica das pesquisas sobre a relação entre uso de tecnologias digitais e bem-estar psicológico, reforçando que métricas simplistas, como "tempo de tela", não capturam a complexidade das interações digitais. Essa classe se posicionou como eixo intermediário entre tecnologia, ética e impactos sociais, indicando seu papel de articulação entre os debates. Essa posição reforça a necessidade de estudos que considerem variáveis contextuais, como o tipo de interação, a intencionalidade pedagógica e as características individuais dos estudantes (SEO *et al*. 2021; SEUFERT *et al*. 2021).

A Classe 3, voltada para governança e regulação da AIED, trouxe à tona documentos e padrões internacionais, como os da UNESCO, que buscam estabelecer diretrizes globais para a implementação segura e equitativa da IA na educação. O posicionamento próximo dessa classe em relação à Classe 4, evidenciado na AFC, sugere que governança e desinformação estão conectadas. Essa conexão pode ser interpretada a partir da perspectiva de Quach *et al*. (2022), que destacam a tensão entre empresas, consumidores e reguladores no uso e proteção de dados. A governança, nesse contexto, envolve disputas políticas e éticas sobre controle e justiça informacional.

A Classe 4, com os scores mais elevados, evidencia a predominância do tema fake news e desinformação no corpus analisado. Esse resultado está alinhado com pesquisas recentes sobre os desafios da detecção automática de conteúdos falsos e os impactos sociais da propagação de informações enganosas (AÏMEUR; AMRI; BRASSARD, 2023; VACCARI E CHADWICK, 2020). O foco crescente nesse tema reflete a percepção de que a desinformação, especialmente em contextos educacionais, representa um problema técnico, social e cognitivo. O estudo de Kozyreva, Lewandowsky e Hertwig

(2020) reforça essa perspectiva ao propor intervenções baseadas na literacia midiática para empoderar usuários e aumentar a resiliência cognitiva frente à manipulação.

A análise de similitude e a nuvem de palavras complementam esses achados ao mostrar que educação e estudantes são o núcleo das discussões, sustentados por três pilares: tecnologia e dados, aplicações práticas de IA como o uso do ChatGPT, e desinformação. Essa estrutura indica que, embora a educação seja o ponto de partida, as discussões se expandem para temas como privacidade, ética, regulação e cidadania digital. Esse cenário evidencia a necessidade de abordagens integradas e interdisciplinares que conectem as dimensões pedagógicas, tecnológicas e sociopolíticas.

Um aspecto relevante a discutir é a lacuna entre teoria e prática. Embora os artigos revisados proponham frameworks conceituais e recomendações éticas, os resultados da análise indicam uma distância significativa entre essas propostas e a implementação efetiva em contextos educacionais. Essa lacuna pode ser explicada pela rápida evolução tecnológica e pela dificuldade dos sistemas educacionais em acompanhar essas transformações, conforme Tlili *et al*. (2023) e Wardat *et al*. (2023). Isso reforça a necessidade de pesquisas que investiguem o potencial da IA na educação e as condições institucionais, culturais e políticas que influenciam sua adoção.

CONCLUSÃO

Este estudo contribui para a compreensão do estado da arte sobre inteligência artificial na educação, desinformação e educação midiática. Ele integra uma revisão narrativa da literatura com uma análise bibliométrica baseada em métodos lexicais. Os resultados evidenciam quatro eixos centrais que estruturam o campo: ética e princípios na AIED, bem-estar digital, governança e regulação, e fake news e desinformação.

A predominância do tema desinformação, associada à governança e regulação, indica que os desafios contemporâneos extrapolam a esfera técnica, alcançando dimensões políticas, éticas e sociais. A relativa independência dos debates éticos sugere que a construção de uma IA educacional justa e responsável ainda carece de integração prática com os sistemas tecnológicos e ambientes de aprendizagem.

Conclui-se que instituições educacionais invistam em programas de formação continuada em literacia digital e ética em IA, além de promover políticas claras sobre o uso de tecnologias emergentes, como o ChatGPT. No campo das políticas públicas, destaca-se a necessidade de regulação internacional harmonizada, capaz de equilibrar inovação e segurança, conforme as diretrizes de organismos como a UNESCO.

Este estudo possui limitações, como a análise restrita a 20 artigos e o recorte temático definido. Pesquisas futuras podem ampliar o corpus analisado, incluir bases de dados adicionais e explorar métodos mistos que integrem análises qualitativas e quantitativas. Além disso, recomenda-se investigar a implementação prática de princípios éticos em projetos educacionais que utilizem IA, para reduzir a lacuna entre teoria e prática identificada neste trabalho.

Ressaltamos que a inteligência artificial na educação tem potencial transformador, mas seu impacto dependerá da capacidade de equilibrar inovação, ética e equidade. É necessário garantir que a tecnologia sirva ao desenvolvimento humano e social. Conclui-se que a integração entre inovação tecnológica, ética e práticas pedagógicas é essencial para enfrentar os desafios contemporâneos, sendo a alfabetização midiática um elemento estratégico para a promoção de uma educação crítica e responsável na era digital

REFERÊNCIAS

ADJIN-TETTEY, D. T. “Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education”. **Cogent Arts and Humanities**, vol. 9, n. 1, 2022.

AÏMEUR, E.; AMRI, S.; BRASSARD, G. “Fake news, disinformation and misinformation in social media: a review”. **Social Network Analysis and Mining**, vol. 13, n. 30, 2023.

ALEEDAN, M. H.; ALFADHLY, A. A. “Role of primary school in promoting media education and English language skills for students”. **World Journal of English Language**, vol. 14, n. 5, 2024.

APPOLINÁRIO, D. G.; HEDLER, H. C.; FERNEDA, E. “Desafios do uso da inteligência artificial na gestão de recursos humanos: uma revisão sistemática”. **Boletim de Conjuntura (BOCA)**, vol. 21, n. 63, 2025.

ARAUJO, T. *et al*. “In AI we trust? Perceptions about automated decision-making by artificial intelligence”. **AI and Society**, vol. 35, n. 3, 2020.

BAILÉN, A. H.; QUINTAS-FROUFE, N.; GONZÁLEZ-NEIRA, A. “Audience Participation in TikTok Metadata”. **Comunicar**, n. 78, 2024.

BRASIL. **Base Nacional Comum Curricular**. Brasília: Ministério da Educação, 2018. Disponível em: <www.mec.gov.br>. Acesso em: 09/09/2025.

BRASIL. **Lei n. 14.533, de 11 de janeiro de 2023**. Brasília: Planalto, 2023. Disponível em: <www.planalto.gov.br>. Acesso em: 09/09/2025.

BRODA, E.; STRÖMBÄCK, J. “Misinformation, disinformation, and fake news: lessons from an interdisciplinary, systematic literature review”. **Annals of the International Communication Association**, vol. 48, n. 2, 2024.

BURTH, S. *et al*. "Mapping media literacy impact in the U.S.: a review of literature and call for equity". **Annals of the International Communication Association**, vol. 48, n. 4, 2024.

CAMARGO, B. C.; JUSTO, A. M. **Tutorial para uso do software Iramuteq**. Florianópolis: UFSC, 2021.

CARPENTER, J. P. *et al*. "How and why are educators using Instagram?". **Teaching and Teacher Education**, vol. 96, 2020.

CHAO-HSI, H. *et al*. "Development and exploration of news media literacy scales in Taiwan". **International Journal of Innovative Research and Scientific Studies**, vol. 8, n. 1, 2025.

CHUAH, S. H. W.; YU, J. "The future of service: The power of emotion in human-robot interaction". **Journal of Retailing and Consumer Services**, vol. 61, 2021.

DUMITRU, D.; HALPERN, D. F. "Critical thinking: creating job-proof skills for the future of work". **Journal of Intelligence**, vol. 11, n. 10, 2023.

DWIVEDI, Y. K. *et al*. "Educación, Big Data e Inteligencia Artificial". **Revista Iberoamericana de Educación a Distancia**, vol. 25, n. 1, 2022.

FAGNI, T. *et al*. "TweepFake: About detecting deepfake tweets". **Plos One**, vol. 16, n. 5, 2021.

GABRIEL, I. "Artificial intelligence, values, and alignment". **Minds and Machines**, vol. 30, 2020.

JAKESCH, M.; HANCOCK, J. T.; NAAMAN, M. "Human heuristics for AI-generated language are flawed". **Proceedings of the National Academy of Sciences of the United States of America PNAS**, vol. 120, n. 11, 2023.

KERR, D. et al. "A Review of AI and Its Impact on Management Accounting and Society". **Journal of Risk and Financial Management**, vol. 18, n. 340, 2025.

KOZYREVA, A.; LEWANDOWSKY, S.; HERTWIG, R. "Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools". **Psychological Science in the Public Interest**, vol. 21, n. 3, 2020.

NGUYEN, A. *et al*. "Ethical principles for artificial intelligence in education". **Education and Information Technologies**, vol. 28, 2022.

ORBEN, A. "Teenagers, screens and social media: a narrative review of reviews and key studies". **Social Psychiatry and Psychiatric Epidemiology**, vol. 55, 2020.

PAGE, M. J. *et al*. "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews". **The British Medical Journal**, vol. 372, n. 71, 2021.

PRISMA. "Transparent Reporting of Systematic Reviews and Meta-Analyses". **Prisma Statement** [2021]. Disponível em: <www.prisma-statement.org>. Acesso em: 15/06/2022.

QUACH, S. *et al*. "Digital technologies: tensions in privacy and data". **Journal of the Academy of Marketing Science**, vol. 50, 2022.

SEO, K. *et al*. "The impact of artificial intelligence on learner-instructor interaction in online learning". **International Journal of Educational Technology in Higher Education**, vol. 18, n. 54, 2021.

SEUFERT, S.; GUGGEMOS, J.; SAILER, M. "Technology-related knowledge, skills, and attitudes of pre- and in-service teachers: The current situation and emerging trends". **Computers in Human Behavior**, vol. 115, 2021.

TAPALOVA, O.; ZHIYENBAYEVA, N. "Artificial Intelligence in Education: AIEd for Personalised Learning Pathways". **The European Journal of Education and Technology**, n. 4, 2021.

TLILI, A. *et al*. "What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education". **Smart Learning Environments**, vol. 10, n. 15, 2023.

UNESCO. **Media and Information Literacy Curriculum for Teachers**. Paris: UNESCO, 2011. Disponível em: <www.unesdoc.unesco.org>. Acesso em: 06/03/2025.

VACCARI, C.; CHADWICK, A. "Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News". **Social Media + Society**, vol. 6, n. 1, 2020.

VELLOSO, L. C. "A vulnerabilidade do consumidor nas relações de consumo mediadas por inteligência artificial". **Revista Foco,** vol. 18, n. 6, 2025.

WACH, K. *et al*. "The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT". **Entrepreneurial Business and Economics Review,** vol. 11, n. 2, 2023.

WARDAT, Y. *et al*. "ChatGPT: a revolutionary tool for teaching and learning mathematics". **Eurasia Journal of Mathematics, Science and Technology Education**, vol. 19, n. 7, 2023.

YIN, W.; ZUBIAGA, A. "Towards generalisable hate speech detection: a review on obstacles and solutions". **arXiv**, 2021.

ZAROCOSTAS, J. "How to fight an infodemic". **The Lancet**, vol. 395, n. 10225, 2020.



BOLETIM DE CONJUNTURA (BOCA)

Ano VII | Volume 23 | Nº 68 | Boa Vista|2025

http://www.ioles.com.br/boca

Editor chefe:

Elói Martins Senhoras

Conselho Editorial

Antonio Ozai da Silva, Universidade Estadual de Maringá

Vitor Stuart Gabriel de Pieri, Universidade do Estado do Rio de Janeiro

Charles Pennaforte, Universidade Federal de Pelotas

Elói Martins Senhoras, Universidade Federal de Roraima

Julio Burdman, Universidad de Buenos Aires, Argentina

Patrícia Nasser de Carvalho, Universidade Federal de Minas Gerais

Conselho Científico

Claudete de Castro Silva Vitte, Universidade Estadual de Campinas

Fabiano de Araújo Moreira, Universidade de São Paulo

Flávia Carolina de Resende Fagundes, Universidade Feevale

Hudson do Vale de Oliveira, Instituto Federal de Roraima

Laodicéia Amorim Weersma, Universidade de Fortaleza

Marcos Antônio Fávaro Martins, Universidade Paulista

Marcos Leandro Mondardo, Universidade Federal da Grande Dourados

Reinaldo Miranda de Sá Teles, Universidade de São Paulo

Rozane Pereira Ignácio, Universidade Estadual de Roraima