

Uso de Agentes Conversacionais com IAGen na Personalização do Aprendizado na Educação Básica: uma Revisão Sistemática da Literatura

Angela Maria de Souza¹,
Vinicius Bertuol¹, André Augusto Bortoli², Anderson da Silva Marcolino³

¹Universidade Estadual do Oeste do Paraná (UNIOESTE)
R. Universitária, 2069 - Bloco B, Cascavel - PR, 85819-110

²Universidade Federal do Rio Grande do Sul (UFRGS)
Av. Bento Gonçalves, 9090, Porto Alegre - RS, 91540-000

³Universidade Federal do Paraná (UFPR)
R. Pioneiro, 2153 - Dallas, Palotina - PR, 85950-000

angela.souza6@unioeste.br, vinicius.bertuol@unioeste.br

andre.bortoli@inf.ufrgs.br, anderson.marcolino@ufpr.br

Abstract. *Personalization of teaching in Basic Education faces scalability challenges given student heterogeneity. Conversational agents based on Generative Artificial Intelligence (GenAI) emerge as an alternative to enable adaptive and inclusive learning paths. This Systematic Literature Review (SLR), guided by the PICOC strategy, mapped primary studies (2019-2025) across Computing and Education databases to analyze this scenario. Results show the technical hegemony of the GPT family, applied mainly to dynamic adaptation of teaching materials and automated formative feedback. However, gaps are evident, such as the incipience of hybrid architectures and the limited exploration of Retrieval-Augmented Generation (RAG) mechanisms. The mapping of these trends and gaps provides essential guidelines for developing solutions that promote equity and personalization at scale.*

Resumo. *A personalização do ensino na Educação Básica enfrenta desafios de escalabilidade frente à heterogeneidade discente. Agentes conversacionais baseados em Inteligência Artificial Generativa (IAGen) surgem como alternativa para viabilizar trilhas adaptativas e inclusivas. Esta Revisão Sistemática da Literatura (RSL), orientada pela estratégia PICOC, mapeou estudos primários (2019-2025) nas bases de Computação e Educação para analisar esse cenário. Os resultados constataam a hegemonia técnica da família GPT, aplicada majoritariamente na adaptação dinâmica de materiais didáticos e no feedback formativo automatizado. Contudo, evidenciam-se lacunas como a incipiência de arquiteturas híbridas e a escassa exploração de mecanismos de Retrieval-Augmented Generation (RAG). Conclui-se que o mapeamento dessas tendências e hiatos oferece diretrizes essenciais para o desenvolvimento de soluções que promovam equidade e personalização em larga escala.*

1. Introdução

Os avanços da Inteligência Artificial Generativa (IAGen), impulsionados por Modelos de Linguagem de Grande Escala (LLMs), têm transformado a educação ao expandir o potencial de agentes conversacionais (ACs) [Bower et al. 2024, Jauhiainen and Garagorry Guerra 2024, Kong and Yang 2024, Ravi et al. 2025]. Na Educação Básica, essas ferramentas são exploradas para otimizar o planejamento pedagógico, gerar materiais didáticos e oferecer retornos avaliativos, consolidando perspectivas para um ensino personalizado e inclusivo [Cheah et al. 2025, Uğraş et al. 2024].

Contudo, a personalização prática enfrenta desafios estruturais diante da heterogeneidade discente. Embora diretrizes como a Base Nacional Comum Curricular (BNCC) orientem a organização do ensino, a aplicação de instrumentos essenciais como o Plano Educacional Individualizado (PEI), consolidado pela nova Política Nacional de Educação Especial Inclusiva (PNEEI) [Brasil 2025], esbarra na sobrecarga docente, na escassez de recursos tecnológicos e na falta de modelos escaláveis [Kong and Yang 2024, Cheah et al. 2025]. Nesse panorama, os ACs baseados em IAGen surgem como alternativa para apoiar a adaptação dinâmica de conteúdos, mitigar a carga administrativa dos professores e prover avaliações formativas imediatas [Ravi et al. 2025, Tasdelen and Bodemer 2025, Cheah et al. 2025].

Apesar do crescente volume de publicações, a literatura apresenta-se fragmentada, carecendo de revisões sistemáticas que integrem os aspectos pedagógicos, tecnológicos e metodológicos dessa aplicação no ecossistema escolar [Bower et al. 2024, Makransky et al. 2025]. Para preencher essa lacuna, este trabalho conduz uma Revisão Sistemática da Literatura (RSL) fundamentada na estratégia PICOC, seguindo diretrizes da Computação [Kitchenham 2007, Petersen et al. 2008]. O objetivo consiste em identificar, analisar e sintetizar o estado da arte sobre ACs e personalização na Educação Básica, mapeando seus benefícios, limitações e implicações éticas, fornecendo subsídios práticos que orientem o desenho de futuras tecnologias para o contexto escolar.

Este artigo está estruturado da seguinte forma: a Seção 2 aborda a fundamentação teórica; a Seção 3 sintetiza os trabalhos relacionados; as Seções 4 e 5 detalham o protocolo e a execução da RSL; a Seção 6 discute os resultados obtidos (QP1 a QP4); a Seção 7 analisa as ameaças à validade; e a Seção 8 apresenta as considerações finais.

2. Fundamentação Teórica

A IAGen no cenário educacional, os ACs e a personalização do aprendizado na Educação Básica constituem os eixos teóricos desta RSL [Kong and Yang 2024, Makransky et al. 2025]. Impulsionada por LLMs, a IAGen expande o suporte aos processos de ensino-aprendizagem [Kong and Yang 2024, Jauhiainen and Garagorry Guerra 2024]. Adota-se aqui o termo abrangente IAGen em alinhamento à literatura recente, priorizando a capacidade gerativa e prática de criar, adaptar e personalizar conteúdos dinamicamente em detrimento da infraestrutura técnica dos modelos. Diferentemente de sistemas baseados em regras, a IAGen gera textos, respostas contextualizadas e recomendações a partir de grandes volumes de dados, potencializando o planejamento pedagógico, materiais didáticos e comentários orientadores [Blonder et al. 2024, Cheah et al. 2025] para atender à diversidade discente [Uğraş et al. 2024]. Contudo, sua aplicação evoca desafios de confiabilidade, transpa-

rência, privacidade e ética escolar [Bower et al. 2024]. Neste trabalho, os três conceitos articulam-se em níveis distintos: a IAGen designa a classe ampla de modelos gerativos, os LLMs constituem sua principal materialização técnica e os ACs correspondem à camada de interação conversacional, tipicamente construída sobre LLMs, pela qual a personalização alcança estudantes e docentes.

A incorporação de modelos de IAGen conferiu aos ACs maior capacidade de adaptação e respostas naturais [Ravi et al. 2025, Cheah et al. 2025, Makransky et al. 2025, Monib et al. 2024]. No ecossistema educativo, apoiam o esclarecimento de dúvidas, orientação de tarefas e acompanhamento do progresso [Tasdelen and Bodemer 2025], ajustando as interações ao ritmo e interesse dos alunos [Liu et al. 2025a]. Todavia, a literatura demanda avaliações sobre as abordagens metodológicas e os limites pedagógicos e éticos desses agentes na Educação Básica [Cheah et al. 2025]. Sob uma perspectiva instrucional, a opacidade inerente a esses modelos (caixas-pretas algorítmicas) impõe desafios à autonomia docente, exigindo que a mediação pedagógica atue como elemento regulador contra vieses cognitivos e alucinações técnicas que possam comprometer o aprendizado.

Sob o prisma pedagógico, a personalização do aprendizado pode ser ancorada no conceito de Zona de Desenvolvimento Proximal (ZDP) de Vygotsky [Vygotsky 1978], no qual a tecnologia atua como um “par mais capaz” ou andaime cognitivo (*scaffolding*), ajustando dinamicamente o nível de desafio às competências reais do aluno para promover uma aprendizagem significativa. Contudo, adequar conteúdos, estratégias e avaliações a essas características individuais (ritmos, estilos cognitivos e necessidades) é um desafio complexo em turmas heterogêneas e numerosas [Jauhiainen and Garagorry Guerra 2024, Makransky et al. 2025]. Nesse panorama, o PEI destaca-se como instrumento essencial para práticas inclusivas voltadas a estudantes com deficiência ou neurodivergência; contudo, sua elaboração manual e sob demanda sobrecarrega de forma severa o cotidiano docente [Brasil 2025].

Assim, os ACs baseados em IAGen surgem como potenciais facilitadores da personalização ao prover devolutivas contínuas e flexibilidade pedagógica [Ravi et al. 2025, Uğraş et al. 2024], desde que integrados criticamente [Bower et al. 2024, Makransky et al. 2025]. Sob a ótica regulatória e de governança, o tratamento de dados de menores exige estrita conformidade com a Lei Geral de Proteção de Dados Pessoais (LGPD), mitigando riscos de exposição indevida de informações sensíveis, falta de controle sobre dados de treinamento e a dependência instrumental de plataformas de propósito geral controladas por conglomerados tecnológicos transnacionais [Bower et al. 2024, Jauhiainen and Garagorry Guerra 2024].

3. Trabalhos Relacionados

As investigações sobre IA na educação variam entre abordagens puramente técnicas e aplicações de apoio pedagógico mediadas por IAGen e LLMs [Blonder et al. 2024, Ravi et al. 2025]. Estudos dedicados ao mapeamento de *chatbots* e assistentes virtuais indicam benefícios recorrentes no esclarecimento de dúvidas pontuais e no engajamento comportamental dos estudantes [Monib et al. 2024, Tasdelen and Bodemer 2025]. Todavia, essas pesquisas concentram-se majoritariamente no Ensino Superior, não tratando as especificidades cognitivas, socioemocionais e as demandas de personalização intrínsecas à Educação Básica [Cheah et al. 2025]. Esta concentração gera uma limitação, visto que

os modelos de interação projetados para adultos não consideram os estágios de desenvolvimento da infância e adolescência.

No que tange à personalização instrucional, antes da difusão da IAGen predominavam métodos adaptativos tradicionais e técnicas não generativas de aprendizado de máquina, como árvores de decisão e classificadores estáticos. Iniciativas que explorem o potencial gerativo da IAGen para mitigar a sobrecarga docente por meio da construção automatizada do PEI ou do alinhamento curricular estrito à BNCC ainda se mostram incipientes e dispersas [Uğraş et al. 2024]. Predomina na literatura uma abordagem puramente instrumental da tecnologia, que ignora as barreiras de infraestrutura e conectividade das escolas públicas brasileiras.

Ademais, as revisões disponíveis na literatura abordam o uso da IA ou de agentes conversacionais de forma genérica, sem integrar sistematicamente os eixos complementares de IAGen, personalização adaptativa e o ecossistema da Educação Básica [Bower et al. 2024], gerando a lacuna que esta RSL busca preencher ao delimitar o escopo estritamente à Educação Básica e articular os três eixos sob uma única estratégia PICOC. Essa fragmentação teórica, somada à clara predominância de pesquisas exploratórias de curto prazo e baseadas unicamente na percepção de satisfação do usuário [Makransky et al. 2025], evidencia uma lacuna na sistematização integrada do tema. Identifica-se, portanto, a necessidade de transitar de uma visão entusiasta e comercial da tecnologia para uma investigação científica, justificando a relevância desta RSL como alicerce para futuras intervenções práticas.

4. Protocolo da RSL

A condução desta RSL seguiu as diretrizes consolidadas para revisões sistemáticas na área de Computação [Kitchenham 2007, Petersen et al. 2008], visando mitigar vieses e assegurar a reprodutibilidade. O escopo foi estruturado sob a estratégia PICOC (*Population*: Educação Básica; *Intervention*: ACs baseados em IAGen; *Outcomes*: personalização do aprendizado, benefícios, limitações e desafios; *Context*: ambiente escolar). O elemento *Comparison* não foi aplicado.

Com base nesse escopo, definiram-se quatro Questões de Pesquisa (QP):

- **QP1 (Questão Principal):** Quais são as principais aplicações de ACs baseados em IAGen voltadas à personalização do aprendizado na Educação Básica?
- **QP2:** Quais abordagens tecnológicas têm sido empregadas no desenvolvimento e na implementação de ACs baseados em IAGen em contextos educacionais?
- **QP3:** Quais abordagens metodológicas de pesquisa e de avaliação são adotadas nesses estudos?
- **QP4:** Quais benefícios, limitações, desafios e implicações éticas são relatados na literatura sobre esse uso?

4.1. Estratégia de Busca e Critérios de Seleção

As buscas foram executadas em sete bases digitais: *ACM Digital Library*, *IEEE Xplore*, *ScienceDirect*, *Scopus*, *Springer Link*, *Google Scholar* e a Biblioteca Digital da Sociedade Brasileira de Computação (SBC). A *string* geral de busca estruturada foi: (“*generative artificial intelligence*” OR “*AIGen*” OR “*AIG*”) AND (“*personalization of education*” OR “*individualized learning*” OR “*personalized learning*”). Optou-se por uma *string*

ampla, sem termos exclusivos de agentes conversacionais (e.g., *chatbot*, *conversational agent*, *pedagogical agent*), a fim de não descartar estudos que descrevem tecnologias equivalentes sob outras nomenclaturas; o refinamento específico para ACs ocorreu na etapa de aplicação dos critérios **CI1** e **CE4**. A *string* foi sintaticamente adaptada aos recursos de cada base, incluindo, quando disponível, filtros de população (e.g., “K-12”, “school”); as versões completas por base constam no material complementar.

A distribuição dos registros recuperados por base digital evidencia a predominância de bases internacionais de Computação e Ciências Aplicadas no levantamento inicial. Dos 651 registros identificados, a *Scopus* concentrou o maior volume com 240 registros (TITLE-ABS-KEY combinada a K-12 ou “basic education”), seguida da *ScienceDirect* com 150 registros sob o mesmo critério. A *Springer Link* contribuiu com 119 registros a partir de buscas no título e resumo, filtradas pela área de *Computer Science*. As bases *IEEE Xplore* e *ACM Digital Library*, consultadas por *all fields* combinados a restrições no *abstract* (“school” OR “K-12”), retornaram 70 e 13 registros, respectivamente. Por fim, o *Google Scholar* e a Biblioteca Digital da Sociedade Brasileira de Computação (SBC), consultados por título e palavras-chave com termos combinados traduzidos, contribuíram com 44 e 15 registros, respectivamente, perfazendo o *corpus* bruto submetido às etapas subsequentes de triagem e seleção.

Para a seleção do *corpus*, foram estabelecidos os seguintes Critérios de Inclusão (**CI**): (**CI1**) Estudos primários que abordam explicitamente o uso de ACs baseados em IA-Gen; (**CI2**) Estudos que apresentam aplicações ou protótipos voltados à personalização do aprendizado; (**CI3**) Pesquisas publicadas em periódicos ou anais de eventos revisados por pares entre os anos de 2019 e 2025; (**CI4**) Estudos situados no contexto da Educação Básica ou contextos equivalentes de abrangência internacional; e (**CI5**) Estudos publicados nos idiomas inglês ou português.

Em contrapartida, aplicaram-se os Critérios de Exclusão (**CE**): (**CE1**) Estudos da literatura cinza ou não primários; (**CE2**) Estudos primários em formato diferente de PDF; (**CE3**) Estudos primários não disponíveis ou pagos; (**CE4**) Estudos primários que não respondem à QP1; e (**CE5**) Estudos que não foram aplicados na Educação Básica.

4.2. Seleção e Extração de Dados

O processo seguiu as recomendações do fluxo PRISMA [Rethlefsen and Page 2022], iniciando pela triagem de títulos e resumos e avançando para a leitura completa. Adicionalmente, empregou-se a técnica de *snowballing* a partir das referências dos aceitos, sem a inclusão de novos registros após a aplicação dos critérios.

A extração das evidências ocorreu de forma sistemática por meio de um formulário estruturado de apoio. O protocolo completo, o formulário de extração e a planilha de mapeamento de dados brutos dos 38 estudos primários estão acessíveis em <https://doi.org/10.6084/m9.figshare.33072413>.

5. Execução da RSL

As buscas foram conduzidas entre agosto e setembro de 2025, cobrindo o recorte temporal de 2019–2025, cujo marco inicial corresponde à consolidação dos primeiros LLMs gerativos de larga escala, para abranger a evolução recente associada aos LLMs [Jauhiainen and Garagorry Guerra 2024, Kong and Yang 2024].

5.1. Processo de Busca e Triagem (ES1 e ES2)

Foram identificados inicialmente 651 registros, agrupados conforme o escopo de indexação das bases: *Scopus & ScienceDirect* (390 registros), *IEEE Xplore & ACM Digital Library* (83 registros), *Springer Link* (119 registros) e *Google Scholar & Biblioteca Digital da SBC* (59 registros). A Figura 1 mapeia a distribuição inicial de registros brutos identificados por base, evidenciando maior concentração na Scopus e ScienceDirect. Todos os dados foram centralizados e gerenciados na ferramenta *Parsifal*¹.

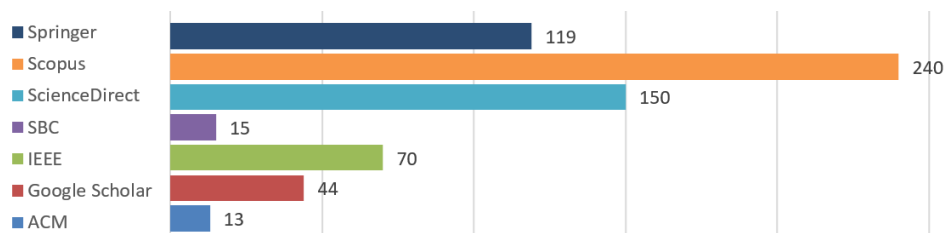


Figura 1. Distribuição dos registros identificados por base de dados

Após a deduplicação, que removeu 24 registros, restaram 627 estudos únicos submetidos à Etapa de Seleção 1 (ES1). A triagem por títulos e resumos resultou na exclusão de 500 trabalhos que não atendiam aos critérios mínimos, restando 127 artigos para leitura completa. Considerando o processo completo de seleção (ES1 e ES2), os principais motivos de exclusão envolveram a não aplicação no ecossistema da Educação Básica ($n = 316$), a falta de aderência direta à questão norteadora QP1 ($n = 135$) e manuscritos pertencentes à literatura cinza ou não primários ($n = 129$). Critérios residuais somaram 9 exclusões associadas a estudos indisponíveis para download ou pagos, totalizando 589 estudos rejeitados nas duas etapas.

A Etapa de Seleção 2 (ES2) consistiu na leitura integral dos 127 artigos pré-selecionados. Desse montante, 89 foram rejeitados por descumprirem o escopo específico de agentes conversacionais gerativos com foco em personalização, culminando em um *corpus* final de 38 estudos primários. A técnica de *backward snowballing* foi aplicada sobre as referências dos aceitos como mecanismo de validação e controle; contudo, não foram identificados novos estudos elegíveis, atestando a saturação e a abrangência da busca primária. O fluxo completo desse processo de seleção baseado nas diretrizes PRISMA está formalmente consolidado na Tabela 1.

Tabela 1. Resumo quantitativo das etapas de seleção da RSL

Etapa / Processo	Avaliados	Rejeitados	Remanescentes
Busca Inicial nas Bases	651	–	651
Remoção de Duplicatas	651	24	627
ES1 – Títulos e Resumos	627	500	127
ES2 – Leitura Completa	127	89	38
Corpus Final	–	–	38

Os 38 estudos selecionados passaram por avaliação de qualidade na plataforma *Parsifal*, com base em cinco critérios: (i) clareza dos objetivos; (ii) descrição do contexto

¹<https://parsif.al>

educacional; (iii) detalhamento da metodologia adotada; (iv) descrição da aplicação do agente conversacional; e (v) explicitação dos resultados e limitações do estudo.

No que se refere à avaliação de qualidade metodológica do *corpus* final, destacam-se os cinco estudos que obtiveram a pontuação máxima (10,0) na plataforma *Parsifal* por responderem plenamente a todos os critérios estabelecidos no protocolo. Dentre as investigações com desempenho perfeito, encontram-se os trabalhos de [Liu et al. 2025a], focado em um ambiente de aprendizagem apoiado por ChatGPT para a resolução de problemas matemáticos no ensino primário; o de [Makransky et al. 2025], voltado ao uso de IAGen para impulsionar a construção de sentido cognitivo; e o de [Tasdelen and Bodemer 2025], que investigou os impactos motivacionais e de desempenho de materiais e tarefas matemáticas personalizadas. Ainda alcançaram a melhor avaliação de qualidade o mapeamento de cocriação de ferramentas baseado em LLMs com educadores K-12 desenvolvido por [Ravi et al. 2025] e a análise do impacto de *chatbots* na autoeficácia em programação visual realizada por [Wang 2025].

Os estudos selecionados e a codificação das referências (“E” + número de identificação) que compõem o *corpus* final estão sintetizados na Tabela 2.

Tabela 2. Mapeamento Geral do *Corpus* Final Selecionado

Cód.	Referência	Cód.	Referência
E01	[Ye et al. 2025]	E20	[da Costa Silva and Ferreira 2024]
E02	[Kong and Yang 2024]	E21	[Karimy et al. 2024]
E03	[Tasdelen and Bodemer 2025]	E22	[Cheah et al. 2025]
E04	[Liu et al. 2025a]	E23	[Bower et al. 2024]
E05	[de Lima and Benini 2025]	E24	[Blonder et al. 2024]
E06	[Lee and Song 2024]	E25	[Hemalatha et al. 2024]
E07	[Liu et al. 2025b]	E26	[Jauhiainen and Garagorry Guerra 2024]
E08	[Bezirhan and von Davier 2023]	E27	[Monib et al. 2024]
E09	[Yan et al. 2025]	E28	[Almufarreh 2024]
E10	[Zha et al. 2025]	E29	[Dhungana and Paudel 2025]
E11	[Makransky et al. 2025]	E30	[Weligodapola and Kumarapperuma 2025]
E12	[Uğraş et al. 2024]	E31	[Jayaweera 2025]
E13	[Ravi et al. 2025]	E32	[Reddy et al. 2025]
E14	[Santos et al. 2025]	E33	[Oppenlaender et al. 2023]
E15	[Abdullahi et al. 2025]	E34	[Shan 2025]
E16	[dos Santos et al. 2025]	E35	[Kindenberg 2025]
E17	[Cavalcante et al. 2024]	E36	[Zhang et al. 2025]
E18	[Laranjeira and Bezerra 2025]	E37	[Wang 2025]
E19	[Wu and Liu 2025]	E38	[Abdel-Hafez and Yousef 2025]

5.2. Classificação Analítica e Temática dos Estudos

Para apoiar a síntese qualitativa indutiva [Kitchenham 2007], os estudos foram organizados em planilhas de apoio e classificados em três grupos temáticos principais (Figura 2). Essa categorização visa estabelecer o nível de proximidade das evidências com o chão da escola e as demandas práticas da docência:

- **Grupo 1 – Educação Básica (12 estudos; 31,6%):** Investiga o uso de ACs no contexto escolar regular e equivalente ao modelo K–12, com foco direto na per-

sonalização, suporte ao planejamento e adaptação de conteúdos (E01, E02, E03, E04, E05, E06, E07, E08, E09, E10, E11, E12). Constitui o núcleo empírico central de resposta à QP1 por retratar a realidade direta da sala de aula.

- **Grupo 2 – Educação Inclusiva, Linguística e Formação Docente (12 estudos; 31,6%):** Reúne pesquisas focadas em demandas pedagógicas particulares e de atendimento especializado, como suporte ao autismo, deficiência visual, bilinguismo e formação continuada de professores (E13, E14, E15, E16, E17, E18, E19, E20, E21, E22, E23, E24). Suas abordagens majoritariamente qualitativas e aplicadas fundamentam as discussões da QP2 e QP3.
- **Grupo 3 – Tecnologias Emergentes Aplicadas à Educação (14 estudos; 36,8%):** Concentra estudos de caráter transversal ao ensino formal, além de cenários exploratórios, conceituais e validações iniciais de ferramentas em ambientes puramente digitais (E25, E26, E27, E28, E29, E30, E31, E32, E33, E34, E35, E36, E37, E38). Subsidiaria as discussões de arquitetura de software e os desafios operacionais futuros (QP2 e QP4).

Essa categorização evidencia a articulação interdisciplinar entre as dimensões técnicas (LLMs, engenharia de *prompts*) e as dimensões pedagógicas (*feedback* adaptativo, alinhamento curricular).

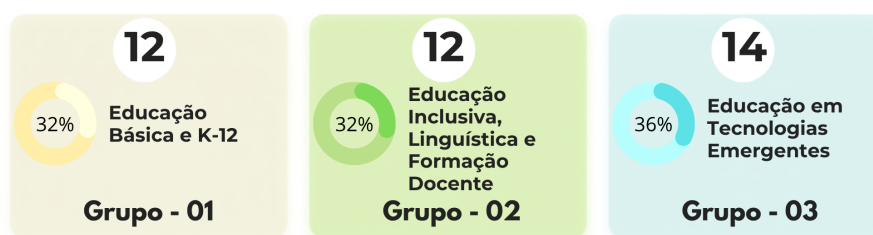


Figura 2. Distribuição dos estudos selecionados por grupo temático

6. Resultados da RSL

Os resultados obtidos a partir dos 38 estudos primários do *corpus* foram sintetizados analiticamente para responder às quatro questões de pesquisa definidas no protocolo. Salvo indicação em contrário, as categorias analíticas extraídas não são mutuamente excluentes: um mesmo estudo pode contabilizar em mais de uma categoria de uma mesma dimensão, de modo que as somas podem exceder os 38 estudos. Os percentuais reportados referem-se sempre à proporção do *corpus* ($n = 38$) que menciona cada categoria.

6.1. Aplicações Educacionais e Abordagens Tecnológicas (QP1 e QP2)

A análise combinada da QP1 e QP2 revela como as intenções pedagógicas de personalização são operacionalizadas pelas infraestruturas computacionais atuais. A Tabela 3 consolida quantitativamente as frentes de aplicação mapeadas e as respectivas escolhas de arquitetura e infraestrutura dos sistemas conversacionais.

Em resposta à QP1, a adaptação de conteúdos e materiais é relatada em 30 estudos (78,9%), concentrando-se em tutoria personalizada e geração de materiais didáticos; a oferta de *feedback* formativo aparece em 24 (63,2%), com foco em interações imediatas e correções automatizadas. Aplicações complexas de suporte a contextos específicos,

Tabela 3. Consolidação de Aplicações (QP1) e Abordagens Tecnológicas (QP2)

Eixos de Aplicação (QP1)	Qtd. (%)	Modelos e Estruturas (QP2)	Qtd. (%)
Adaptação de conteúdos e materiais	30 (78,9)	Família GPT (OpenAI) de prateleira	16 (42,1)
Oferta de <i>feedback</i> automatizado	24 (63,2)	IAGen Genérica / Agnóstica	10 (26,3)
Recomendação de trilhas e fluxos	16 (42,1)	Arquiteturas Híbridas (LLM + PLN)	8 (21,1)
Apoio ao planejamento docente	14 (36,8)	Modelos de Difusão / Multimodais	2 (5,3)
Suporte a contextos (Inclusão/PEI)	14 (36,8)	Família Gemini (Google via API)	2 (5,3)
Estratégias de Interação (QP2)		Mecanismos de Controle (QP2)	
<i>Prompt Engineering</i>	20 (52,6)	Uso explícito de RAG estruturado	6 (15,8)
<i>Fine-tuning</i> de parâmetros	11 (28,9)	Interface padrão / Uso genérico	18 (47,4)

Nota: percentuais calculados sobre o *corpus* (n = 38). Os blocos Eixos de Aplicação, Estratégias de Interação e Mecanismos de Controle contêm categorias não mutuamente excludentes; o bloco Modelos e Estruturas constitui partição exata do *corpus*.

como a automação estruturada do PEI ou a conformidade explícita à BNCC, ocorrem em 14 estudos (36,8%) e permanecem dispersas e incipientes na literatura.

6.1.1. Uso de LLMs

A utilização de LLMs constitui a principal abordagem tecnológica identificada em resposta à QP2. Dos 38 estudos analisados, 26 (68,4%) descrevem explicitamente o uso de LLMs como núcleo dos ACs educacionais: 16 adotam arquiteturas da família GPT, 2 a família Gemini e 8 arquiteturas híbridas. Os 12 restantes (31,6%) ou descrevem a tecnologia de forma inespecífica (10) ou exploram modelos multimodais de difusão (2), o que limita a transparência e a reprodutibilidade das soluções propostas.

6.1.2. Arquiteturas baseadas em GPT ou modelos equivalentes

Considerando modelos proprietários do tipo GPT ou equivalentes (famílias GPT e Gemini), 18 estudos (47,4%) explicitam sua adoção; os 20 demais (52,6%) empregam arquiteturas híbridas sem modelo proprietário identificado (8) ou descrições genéricas da tecnologia (12), o que dificulta comparações e avaliações de dependência tecnológica.

Como consequência dessa dependência massiva de ferramentas proprietárias generalistas, evidencia-se um gargalo estrutural na Informática na Educação Básica. A baixa incidência de soluções customizadas reflete as barreiras de infraestrutura local das escolas, as quais inviabilizam o processamento de modelos de código aberto (*open-source*), forçando a comunidade a delegar a mediação pedagógica aos critérios comerciais opacos das grandes corporações transnacionais [Jauhiainen and Garagorry Guerra 2024].

A hegemonia da família GPT (16 estudos; 42,1%) e a incipiência de mecanismos de *Retrieval-Augmented Generation* (RAG), técnica que ancora as respostas do modelo em bases de conteúdo verificado e presente em apenas 6 estudos (15,8%), revelam que a área ainda adota uma postura de dependência instrumental em relação a grandes monopólios tecnológicos. No cenário da Educação Básica pública brasileira, essa configuração é altamente problemática. A falta de arquiteturas híbridas locais impede que os ACs funcionem de forma customizada às realidades regionais, limitando o alinhamento estrito com as competências da BNCC [Kong and Yang 2024, Uğraş et al. 2024]. Sem o uso de

RAG estruturado para ancorar as respostas em bases de dados pedagógicas governamentais, os ACs de prateleira tendem a reproduzir conteúdos generalistas e, frequentemente, descontextualizados da matriz curricular nacional [Kong and Yang 2024].

6.2. Abordagens Metodológicas e Dimensões Críticas (QP3 e QP4)

A QP3 mapeia o delineamento das pesquisas, enquanto a QP4 investiga o balanço entre benefícios e riscos operacionais. Cruzando esses eixos metodológicos e analíticos, a Tabela 4 sistematiza os indicadores quantitativos.

Tabela 4. Delineamentos Metodológicos (QP3) e Impactos/Desafios (QP4)

Delineamento e Métodos (QP3)	Qtd. (%)	Benefícios e Desafios (QP4)	Qtd. (%)
Exploratório / Descritivo	18 (47,4)	Personalização do aprendizado	28 (73,7)
Estudo de caso	11 (28,9)	Feedback imediato percebido	25 (65,8)
Experimental / Quase-experimental	5 (13,2)	Engajamento e motivação discente	23 (60,5)
Conceitual / Teórico	4 (10,5)	Apoio ao planejamento docente	14 (36,8)
Abordagem Qualitativa	17 (44,7)	Limitações e Riscos Éticos (QP4)	
Abordagem Mista	15 (39,5)	Dependência de LLMs genéricos	21 (55,3)
Abordagem Quantitativa	6 (15,8)	Ausência de validação pedagógica	18 (47,4)
Escopo Temporal / Avaliação (QP3)		Riscos de privacidade e LGPD	18 (47,4)
Avaliações de curto prazo	23 (60,5)	Baixa explicabilidade / Opacidade	17 (44,7)
Delineamento Longitudinal	5 (13,2)	Qualidade inconsistente / Alucinações	15 (39,5)
Duração Não explicitada	10 (26,3)	Vieses algorítmicos e equidade	14 (36,8)
		Mediação docente / Uso responsável	13 (34,2)

Nota: percentuais calculados sobre o *corpus* (n = 38). As categorias da QP4 não são mutuamente excludentes; os blocos de delineamento, abordagem e escopo temporal (QP3) constituem partições exatas do *corpus*.

Os resultados da QP3 revelam que o campo de estudos encontra-se em estágio incipiente, marcado por abordagens exploratórias ou descritivas (18 estudos; 47,4%) e estudos de caso (11 estudos; 28,9%), com forte viés qualitativo (17 estudos; 44,7%) ou misto (15 estudos; 39,5%). Os instrumentos avaliativos concentram-se em questionários ou escalas de percepção (20 estudos; 52,6%) e entrevistas ou observação direta (16 estudos; 42,1%), mensurando essencialmente a aceitação tecnológica e a usabilidade. Em contraste, testes objetivos de desempenho acadêmico ocorrem em apenas 8 estudos (21,1%). Nesses casos raros de avaliação cognitiva, os autores utilizaram delineamentos de pré e pós-teste, análise de retenção de conhecimento via questionários padronizados e rubricas de avaliação aplicadas a tarefas práticas (como exercícios de programação e resolução de problemas matemáticos) para mensurar o tamanho do efeito da intervenção. Essa preferência dominante por avaliar a satisfação imediata do usuário em cenários de curto prazo (23 estudos; 60,5%) e com amostras reduzidas de até 50 participantes (16 estudos; 42,1%) configura um hiato na comprovação empírica de ganhos de aprendizagem consolidados no médio e longo prazo.

Em resposta à QP4, os benefícios mais expressivos centram-se na personalização do aprendizado (28 estudos; 73,7%) e no ganho de engajamento (23 estudos; 60,5%). Todavia, emergem limitações severas, com destaque para a ausência de validação pedagógica sistemática (18 estudos; 47,4%) e a qualidade inconsistente das respostas com risco de alucinações (15 estudos; 39,5%), que comprometem a segurança instrucional [Bower et al. 2024, Jauhiainen and Garagorry Guerra 2024].

No eixo metodológico, a predominância de avaliações de curto prazo (23 estudos; 60,5%) focadas estritamente na percepção dos participantes mascara a real eficácia pedagógica dos agentes conversacionais. O “efeito novidade” da IAGen frequentemente infla os índices de engajamento discente nos primeiros contatos, mas a ausência de delineamentos longitudinais (5 estudos; 13,2%) impede a comprovação de ganhos cognitivos consolidados a médio e longo prazo [Bower et al. 2024, Makransky et al. 2025].

Sob a perspectiva regulatória da LGPD, os dados apontam uma negligência severa: embora 18 estudos (47,4%) mencionem riscos de privacidade de menores e desconformidade com regulamentações vigentes, poucos apresentam salvaguardas técnicas para o tratamento de dados de crianças e adolescentes [Bower et al. 2024, Cheah et al. 2025, Makransky et al. 2025]. A baixa explicabilidade algorítmica (17 estudos; 44,7%) e os vieses algorítmicos com impactos na equidade (14 estudos; 36,8%) colocam o professor em uma posição de vulnerabilidade, transformando o ecossistema escolar em um campo de testes para ferramentas comerciais sem validação científica ou pedagógica prévia [Cheah et al. 2025]. Nas escolas públicas, esse cenário de governança é severamente agravado pela precariedade de conectividade e escassez de laboratórios adequados [Bower et al. 2024, Jauhiainen and Garagorry Guerra 2024], limitando a escalabilidade sustentável e segura dessas soluções.

Em síntese, o cruzamento das dimensões revela um padrão consistente: as aplicações mais frequentes (adaptação de conteúdos e *feedback*) são justamente as sustentadas pelas arquiteturas menos customizadas e pelos delineamentos menos rigorosos, sugerindo que a maturidade técnica da área avançou mais rapidamente que sua maturidade metodológica e pedagógica. A Tabela 5 consolida, em uma única visão, as principais tendências tecnológicas, metodológicas e analíticas identificadas ao longo das QP1–QP4.

Tabela 5. Síntese das principais tendências identificadas

Dimensão	Achado predominante	Qtd. (%)
Arquitetura tecnológica	Família GPT (OpenAI) de prateleira	16 (42,1)
Mecanismo de controle	Uso de interface padrão/genérica (sem RAG estruturado)	18 (47,4)
Delineamento de pesquisa	Exploratório/descritivo, seguido de estudo de caso	18 (47,4)
Abordagem de pesquisa	Qualitativa, seguida de mista	17 (44,7)
Escopo temporal	Avaliações de curto prazo	23 (60,5)
Benefício mais relatado	Personalização do aprendizado percebida	28 (73,7)
Limitação mais crítica	Dependência de LLMs genéricos	21 (55,3)

7. Ameaças à Validade do Estudo

A condução desta RSL está sujeita a ameaças mapeadas segundo as diretrizes de [Kitchenham 2007, Petersen et al. 2008]:

- **Validade de Construção:** Termos de busca podem não esgotar a literatura. Minimizou-se o risco via *strings* abrangentes com operadores booleanos no protocolo [Kitchenham 2007].
- **Validade Interna:** Vieses de seleção foram controlados por critérios de inclusão/exclusão (CI/CE), registro unificado no *Parsifal* e revisões em pares.

- **Validade Externa:** O foco é a Educação Básica, mas o *corpus* abrange K–12, inclusão e formação docente, o que expande o poder de generalização, embora exija cautela ao extrapolar os dados [Petersen et al. 2008].
- **Validade de Conclusão:** Mitigou-se o risco interpretativo com análise estruturada e em pares, quantificação explícita dos achados (frequências absolutas e percentuais sobre o *corpus* de 38 estudos) e transparência na identificação de lacunas.

8. Considerações Finais

Esta RSL identificou, analisou e sintetizou o estado da arte do uso de ACs baseados em IAGen para a personalização do aprendizado na Educação Básica, fornecendo um panorama sobre as abordagens tecnológicas, metodológicas e pedagógicas vigentes.

Os resultados indicam uma forte tendência instrumental: a tecnologia foca primordialmente em intenções imediatas e pontuais, como adaptação de conteúdos (30 estudos; 78,9%) e *feedback* automatizado (24 estudos; 63,2%). Contudo, há pouca maturidade computacional e metodológica no suporte a ecossistemas mais complexos, refletida no baixo uso de RAG (6 estudos; 15,8%), carência de validação pedagógica (18 estudos; 47,4%), predominância de avaliações baseadas em questionários e escalas de percepção (20 estudos; 52,6%) e escassez de delineamentos longitudinais (5 estudos; 13,2%). Adicionalmente, riscos éticos atrelados à LGPD e à opacidade algorítmica demandam governança estrita.

Os resultados indicam que o interesse pelos grandes modelos de linguagem fez com que essas ferramentas fossem incorporadas ao ensino antes que sua eficácia fosse devidamente avaliada. A ausência de grupos de comparação e de avaliações de longo prazo mostra que seu uso ainda ocorre, em muitos casos, com base na percepção de benefícios imediatos, sem evidências científicas robustas sobre seus efeitos na aprendizagem.

Ainda que o *corpus* internacional analisado não permita generalizar essa constatação, observa-se, no contexto brasileiro, uma lacuna no desenvolvimento de soluções computacionais integradas de forma contínua às diretrizes da BNCC e ao suporte à Educação Inclusiva [Brasil 2025]. Para responder a esse hiato, os desdobramentos futuros desta pesquisa preveem o desenvolvimento e a avaliação empírica de um AC baseado em IAGen customizado para apoiar professores na elaboração estruturada do PEI na Educação Básica. Por meio de estudos empíricos orientados a analisar a qualidade pedagógica dos planos gerados, busca-se viabilizar uma personalização ética, escalável e metodologicamente alinhada à realidade escolar pública brasileira. Como desdobramentos complementares, pretende-se aprofundar, em investigações futuras, a validação estatística de métricas pedagógicas associadas à personalização, o real impacto da personalização sobre os resultados de aprendizagem e as decisões metodológicas específicas adotadas para cada tipo de personalização identificado nesta RSL.

Declaração sobre uso de Inteligência Artificial

Declara-se que a ferramenta de IAGen Claude - Sonnet 5 foi empregada exclusivamente para fins de revisão linguística, aprimoramento de estilo, adequação de concisão textual e verificação estrutural deste manuscrito. A IA não foi utilizada para a criação de conteúdo original, análise de dados, inferências científicas ou formulação das conclusões do estudo. Os autores assumem total responsabilidade pelo conteúdo integral publicado neste trabalho.

Referências

- Abdel-Hafez, A. and Yousef, A. (2025). A dataset of media articles on chatgpt. In *Proceeding - 12th International Conference on Information Technology: Innovation Technologies, ICIT 2025*, pages 465–470. Institute of Electrical and Electronics Engineers Inc.
- Abdullahi, M. E., Abdulsalam, M. O., and Ekele, O. A. (2025). The impact of generative ai in personalized learning on students learning outcomes; teachers perception of impact of ai in the 21st century. *Babcock University Journal of Education*, 10.
- Almufarreh, A. (2024). Determinants of students’ satisfaction with ai tools in education: A pls-sem-ann approach. *Sustainability*, 16:5354.
- Bezirhan, U. and von Davier, M. (2023). Automated reading passage generation with openai’s large language model. *Computers and Education: Artificial Intelligence*, 5:100161.
- Blonder, R., Feldman-Maggor, Y., and Rap, S. (2024). Are they ready to teach? generative ai as a means to uncover pre-service science teachers’ pck and enhance their preparation program. *Journal of Science Education and Technology*.
- Bower, M., Torrington, J., Lai, J. W. M., Petocz, P., and Alfano, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative artificial intelligence? outcomes of the chatgpt teacher survey. *Education and Information Technologies*, 29:15403–15439.
- Brasil (2025). Decreto nº 12.773, de 2025. Dispõe sobre a Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva.
- Cavalcante, J. P. S., Apolinário, J. A. F., and Souza, E. P. R. (2024). Como as soluções voltadas para surdos são avaliadas? um estudo apoiado por inteligência artificial generativa. In *Anais do I Workshop de Informática na Educação Inclusiva (WIEI 2024)*, pages 85–92. Sociedade Brasileira de Computação.
- Cheah, Y. H., Lu, J., and Kim, J. (2025). Integrating generative artificial intelligence in k-12 education: Examining teachers’ preparedness, practices, and barriers. *Computers and Education: Artificial Intelligence*, 8.
- da Costa Silva, J. C. and Ferreira, H. N. M. (2024). Adaptação de questões avaliativas para estudantes neurodivergentes: Uma arquitetura computacional fundamentada em inteligência artificial generativa. In *Anais do I Workshop de Informática na Educação Inclusiva (WIEI 2024)*, pages 36–43. Sociedade Brasileira de Computação.
- de Lima, L. J. F. and Benini, F. A. V. (2025). Personalized learning with artificial intelligence: a case study in a literacy extension project. *Cuadernos de Educación y Desarrollo*, 17:e7744.
- Dhungana, D. and Paudel, S. (2025). Personalized learning paths: Automated guidance and monitoring for individual learning experiences. In *IEEE EUROCON 2025 - 21st International Conference on Smart Technologies*, pages 1–6. IEEE.
- dos Santos, D. S., Cardoso, P. C. F., and Pinto, V. H. S. C. (2025). Égua assist: Uma ferramenta de apoio ao ensino de lógica de programação para estudantes com deficiência

- visual. In *Anais do XXXIII Workshop sobre Educação em Computação (WEI 2025)*, pages 1079–1091. Sociedade Brasileira de Computação - SBC.
- Hemalatha, K., Deepika, V., M Mallika, R., and Babu, K. M. (2024). Large language model based personalized learning assistant for career-oriented skills. In *2024 2nd International Conference on Signal Processing, Communication, Power and Embedded System (SCOPEs)*, pages 1–6.
- Jauhiainen, J. S. and Garagorry Guerra, A. (2024). Generative ai and education: dynamic personalization of pupils’ school learning material with chatgpt. *Frontiers in Education*, Volume 9 - 2024.
- Jayaweera, I. (2025). Revolutionizing education: Generative ai as a catalyst for personalized learning and innovative teaching practices. In *2025 5th International Conference on Advanced Research in Computing (ICARC)*, pages 1–6. IEEE.
- Karimy, A. U., Rasuli, J., and Reddy, P. C. (2024). Ai in online educational access: Breaking barriers for female learners. In *2024 21st International Conference on Information Technology Based Higher Education and Training (ITHET)*, pages 1–9. IEEE.
- Kindenberg, B. (2025). The role of ai in historical simulation design: A tpack perspective on a french revolution simulation design experience. *Education Sciences*, 15:192.
- Kitchenham, B. (2007). Guidelines for performing systematic literature reviews in software engineering. *Technical report, Ver. 2.3 EBSE Technical Report. EBSE*.
- Kong, S.-C. and Yang, Y. (2024). A human-centered learning and teaching framework using generative artificial intelligence for self-regulated learning development through domain knowledge learning in k–12 settings. *IEEE Transactions on Learning Technologies*, 17:1562–1573.
- Laranjeira, M. L. and Bezerra, P. T. L. (2025). Gerador automático de planos de aula sobre tecnologia baseado em llms. In *Anais do XXXIII Workshop sobre Educação em Computação (WEI 2025)*, pages 515–526. Sociedade Brasileira de Computação - SBC.
- Lee, S. and Song, K.-S. (2024). Teachers’ and students’ perceptions of ai-generated concept explanations: Implications for integrating generative ai in computer science education. *Computers and Education: Artificial Intelligence*, 7:100283.
- Liu, J., Sun, D., Sun, J., Wang, J., and Yu, P. L. H. (2025a). Designing a generative ai enabled learning environment for mathematics word problem solving in primary schools: Learning performance, attitudes and interaction. *Computers and Education: Artificial Intelligence*, 9:100438.
- Liu, R., Pang, W., Chen, J., Balakrishnan, V. A., and Chin, H. L. (2025b). The application of scaffolding instruction and ai-driven diffusion models in children’s aesthetic education: A case study on teaching traditional chinese painting of the twenty-four solar terms in chinese culture. *Education and Information Technologies*, 30:9129–9160.
- Makransky, G., Shiwalia, B. M., Herlau, T., and Blurton, S. (2025). Beyond the “wow” factor: Using generative ai for increasing generative sense-making. *Educational Psychology Review*, 37:60.

- Monib, W. K., Qazi, A., Silva, L. D., and Mahmud, M. M. (2024). Exploring learners' experiences with chatgpt in personalized learning. In *2024 6th International Workshop on Artificial Intelligence and Education (WAIE)*, pages 66–70. IEEE.
- Oppenlaender, J., Silvennoinen, J., Paananen, V., and Visuri, A. (2023). Perceptions and realities of text-to-image generation. In *26th International Academic Mindtrek Conference*, pages 279–288. ACM.
- Petersen, K., Feldt, R., Mujtaba, S., and Mattsson, M. (2008). Systematic mapping studies in software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering, EASE 2008*.
- Ravi, P., Masla, J., Kakoti, G., Lin, G. C., Anderson, E., Taylor, M., Ostrowski, A. K., Breazeal, C., Klopfer, E., and Abelson, H. (2025). Co-designing large language model tools for project-based learning with k12 educators. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pages 1–25. ACM.
- Reddy, V. R. K., Tharun, N., Reddy, K. S. R., Hariharan, S., Maddala, B. P., and Baduguna, S. (2025). Student assessment using generative ai by for content specific challenge identification. In *2025 8th International Conference on Trends in Electronics and Informatics (ICOEI)*, pages 1079–1084. IEEE.
- Rethlefsen, M. L. and Page, M. J. (2022). Prisma 2020 and prisma-s: common questions on tracking records and the flow diagram.
- Santos, A., Santos, R., and Silva, C. (2025). Teachat: Promovendo o suporte educacional ao autismo com inteligência artificial conversacional. In *Anais do XXXIII Workshop sobre Educação em Computação*, pages 1370–1380, Porto Alegre, RS, Brasil. SBC.
- Shan, R. (2025). Learnrag: Implementing retrieval-augmented generation for adaptive learning systems. In *2025 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, pages 0224–0229. IEEE.
- Tasdelen, O. and Bodemer, D. (2025). Generative ai in the classroom: Effects of context-personalized learning material and tasks on motivation and performance. *International Journal of Artificial Intelligence in Education*.
- Uğraş, H., Uğraş, M., Papadakis, S., and Kalogiannakis, M. (2024). Chatgpt-supported education in primary schools: The potential of chatgpt for sustainable practices. *Sustainability*, 16:9855.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA. Edited by Michael Cole, Vera John-Steiner, Sylvia Scribner, and Ellen Souberman.
- Wang, Z. (2025). Using ai chatbots in visual programming: Effect on programming self-efficacy of upper primary school learners. *International Journal of Information and Education Technology*, 15:30–38.
- Weligodapola, M. and Kumarapperuma, C. U. (2025). Transformative impact of predictive and generative ai on education workforce in developing countries. In *2025 5th International Conference on Advanced Research in Computing*. Institute of Electrical and Electronics Engineers Inc.

- Wu, H. and Liu, W. (2025). Exploring mechanisms of effective informal genai-supported second language speaking practice: a cognitive-motivational model of achievement emotions. *Discover Computing*, 28:119.
- Yan, Y.-M., Chen, C.-Q., Hu, Y.-B., and Ye, X.-D. (2025). Llm-based collaborative programming: impact on students' computational thinking and self-efficacy. *Humanities and Social Sciences Communications*, 12:149.
- Ye, X., Zhang, W., Zhou, Y., Li, X., and Zhou, Q. (2025). Improving students' programming performance: an integrated mind mapping and generative ai chatbot learning approach. *Humanities and Social Sciences Communications*, 12:558.
- Zha, S., Liu, Y., Zheng, C., Xu, J., Yu, F., Gong, J., and Xu, Y. (2025). Mentigo: An intelligent agent for mentoring students in the creative problem solving process. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pages 1–22. ACM.
- Zhang, X., Xu, S., and Jing, Y. (2025). The mode and practice path of generative artificial intelligence empowerment classroom teaching – based on grounded theory. In *Proceedings of the 2025 International Conference on Generative Artificial Intelligence and Digital Media*, pages 58–62. ACM.