

Large Language Models (LLMs) e Aprendizagem Adaptativa no Ensino Superior de Computação: uma revisão sistemática da literatura

Denise Rísia A. Carvalho^{1,2}, Paulo Guilherme de Almeida Silva^{1,2}, Paulo Ricardo M. da Silva^{1,2}, Tânia S. de Melo Pinheiro¹, Rossana M. de Castro Andrade^{1,2}

¹Universidade Federal do Ceará (UFC) - Departamento de Computação

²Grupo de Redes de Computadores, Engenharia de Software e Sistemas (GREat)

{deniserisia,pauloguilherme,sirartorias12}@alu.ufc.br,

{rossana,tania}@ufc.br

Abstract. *This Systematic Literature Review analyzes how Large Language Models (LLMs) and adaptive learning have been applied in Computing higher education. Following Kitchenham and Charters' guidelines, 35 studies were selected from five academic databases. The results reveal six modalities of LLM use and adaptive strategies such as personalized contextual feedback, cognitive scaffolding, and knowledge tracing. Evidence points to positive contributions to academic performance, engagement, and reduced teaching workload, although their relationship with student retention and dropout remains underexplored. We conclude that the main challenge is not merely adopting LLMs, but integrating them into instructional, assessment, and ethical models.*

Resumo. *Esta Revisão Sistemática da Literatura analisa como os Modelos de Linguagem de Grande Escala (LLMs) e a aprendizagem adaptativa têm sido empregados no ensino superior de Computação. Seguindo as diretrizes de Kitchenham e Charters, foram selecionados 35 estudos em cinco bases acadêmicas. Os resultados evidenciam seis modalidades de uso dos LLMs e estratégias adaptativas como feedback contextual personalizado, andaime cognitivo (scaffolding) e rastreamento do conhecimento (knowledge tracing). As evidências apontam contribuições positivas para desempenho, engajamento e redução da carga docente, embora a relação com permanência e evasão permaneça pouco explorada. Conclui-se que o principal desafio não é apenas adotar LLMs, mas integrá-los a modelos de ensino, avaliativos e éticos.*

1. Introdução

A Educação em Computação atravessa, na segunda metade da década de 2020, um momento de profunda transformação. As rápidas mudanças tecnológicas, sociais e econômicas impõem pressões crescentes sobre as instituições de ensino superior, exigindo que os cursos da área não apenas atualizem seus conteúdos, mas repensem, de forma estrutural, seus objetivos, métodos, práticas avaliativas e ecossistemas educacionais [Valente, Machado e Barbosa, 2023]. Nesse contexto, a Sociedade Brasileira de Computação (SBC), por meio do *I Seminário Grandes Desafios da Educação em Computação no Brasil*, realizado em novembro de 2024, consolidou cinco grandes desafios estratégicos para a área no período de 2025 a 2035, estabelecendo uma agenda coletiva de pesquisa e políticas para o setor [SBC, 2025].

Entre esses desafios, dois se destacam pela sua centralidade e interdependência no contexto do ensino superior: o quarto desafio, que trata da Inteligência Artificial (IA) como Agente na Educação em Computação, e o quinto, voltado ao Desenvolvimento e Avaliação de *Soft Skills* nos Cursos da Área de Computação. Ambos convergem para uma mesma necessidade: superar modelos pedagógicos historicamente centrados na transmissão técnica de conteúdos, nos quais o professor figura como detentor do conhecimento e o aluno como receptor passivo, e avançar em direção a uma formação mais integral, adaptativa, ética e humana [Lodge, Thompson and Corrin, 2023].

A aplicação da IA no ambiente educacional tem demonstrado potencial relevante ao viabilizar trajetórias de aprendizagem ajustadas às necessidades individuais dos estudantes, fornecer acompanhamento personalizado e ampliar o acesso a conteúdos pedagógicos de qualidade [Ayeni et al., 2024]. A ampla adoção dessas tecnologias, especialmente após o lançamento do ChatGPT em novembro de 2022, inaugurou uma nova era da educação potencializada pela Inteligência Artificial Generativa (IAG), impulsionando ampla produção de pesquisas sobre suas aplicações e implicações educacionais [Zarei et al., 2024]. Os Modelos de Linguagem de Grande Escala (LLMs – *Large Language Models*), como o ChatGPT, são sistemas de IA projetados para processar e gerar textos semelhantes aos produzidos por humanos, aprendendo padrões a partir de extensos conjuntos de dados de treinamento [Xu, Chen and Miao, 2024].

Por sua capacidade de manter continuidade contextual em interações prolongadas e favorecer ambientes baseados na investigação, os LLMs apoiam princípios da aprendizagem ativa e da aprendizagem autorregulada, entendida como o processo cíclico em que os estudantes definem metas, monitoram seu progresso e refletem sobre sua aprendizagem [Zimmerman, 2002], conforme apontado por estudos recentes [Fan et al., 2025]. Entretanto, embora o potencial da IAG em contextos educacionais seja crescente, revisões recentes apontam que ainda há escassez de estudos que descrevam de forma clara como essas tecnologias são integradas, adaptadas e aplicadas em ambientes virtuais de aprendizagem, além de persistirem limitações relacionadas à dependência de ferramentas proprietárias e à pouca discussão sobre ética, privacidade e segurança de dados [Castilho et al., 2025].

Esse diagnóstico converge diretamente com as lacunas identificadas pela [SBC, 2025], a qual aponta a ausência de critérios robustos para avaliar a contribuição da IA no ensino de Computação e a carência de estruturas teóricas e empíricas capazes de sustentar avaliações sistemáticas do papel epistêmico desempenhado por sistemas de IA em ambientes educacionais. Diante desse cenário, o objetivo desta Revisão Sistemática da Literatura (RSL) é analisar como os LLMs e a aprendizagem adaptativa têm sido empregados no ensino superior de Computação.

Esta revisão dá continuidade e integra-se a uma pesquisa mais ampla, que vem investigando, de forma incremental, o uso de tecnologias educacionais, incluindo IA, no apoio ao aprendizado ativo, à autorregulação e à personalização no ensino superior. Em uma primeira etapa, desenvolveu-se uma plataforma digital de autoavaliação, com foco em *feedback* imediato e práticas formativas mediadas por tecnologia [Carvalho et al.,

2025]. Em uma segunda etapa, a investigação avançou para o uso de *Machine Learning* visando subsidiar recomendações adaptativas [Carvalho and Andrade, 2026].

O restante deste trabalho está organizado da seguinte forma: a Seção 2 apresenta o Referencial Teórico; a Seção 3 descreve a Metodologia adotada; a Seção 4 discute os Resultados; a Seção 5 aborda a Discussão. Por fim, a Seção 6 apresenta as Conclusões e Trabalhos Futuros.

2. Referencial Teórico

A formação superior em Computação no Brasil enfrenta tensões estruturais que vão além da simples atualização de conteúdos. Problemas como evasão, rigidez de matrizes curriculares, uso acrítico de tecnologias educacionais e limitação do letramento computacional persistem como obstáculos centrais [Valente, Machado e Barbosa, 2023]. Estudos baseados nos microdados do Exame Nacional de Desempenho dos Estudantes (ENADE) reforçam esse diagnóstico ao evidenciar diferenças de desempenho entre cursos, instituições e áreas de conhecimento da Computação, indicando fragilidades formativas que podem subsidiar melhorias de ensino [Rendeiro et al., 2023; Trujillo and Corrêa, 2025].

No ensino específico de programação, por exemplo, a disciplina central nos cursos da área, as dificuldades são sobretudo evidentes: estudantes enfrentam obstáculos como desenvolvimento insuficiente do raciocínio lógico, habilidades limitadas na resolução de problemas, baixa capacidade de abstração e pouca motivação, fatores que contribuem para os altos índices de reprovação e evasão [Silva et al., 2024].

Diante dos desafios apontados, tecnologias emergentes como os LLMs têm sido indicadas como recursos com potencial para mitigar parte desses problemas no ensino de Computação. Os LLMs representam um avanço significativo no campo do Processamento de Linguagem Natural (PLN), sendo compreendidos como sistemas baseados em redes neurais profundas, treinados em extensos conjuntos de dados e apoiados em arquiteturas como os *Transformers*, que aprimoram a capacidade desses modelos de processar e gerar linguagem natural de forma coerente e contextualizada [Silva and Aranha, 2025]. Esses sistemas têm sido aplicados em diversas áreas, incluindo a educação, especialmente na geração automatizada de *feedbacks* em exercícios de programação [Naveed et al., 2023].

No contexto brasileiro, o documento *Grandes Desafios da Educação em Computação no Brasil 2025–2035* já identifica a presença concreta dessas tecnologias nos cursos de Computação, mencionando explicitamente o uso de copilotos de programação como *GitHub Copilot* e *Amazon CodeWhisperer*, sistemas de correção automática de código e ferramentas de geração de exercícios com base em IAG como práticas que modificam a forma como se ensina lógica, depuração e boas práticas de desenvolvimento [SBC, 2025]. Para que esse potencial gere ganhos de ensino, contudo, é necessário articular os LLMs a abordagens didáticas centradas no estudante, entre as quais se destaca a aprendizagem adaptativa [Spriggs, Lau and Passi, 2025]. A aprendizagem adaptativa pode ser compreendida como um conjunto de abordagens

educacionais e tecnológicas que buscam ajustar o processo de ensino-aprendizagem às características individuais do estudante. Seu fundamento teórico encontra-se também na Aprendizagem Autorregulada, segundo a qual os estudantes são agentes ativos de sua própria aprendizagem, organizando ciclicamente processos de antecipação, controle do desempenho e autorreflexão [Zimmerman, 2002].

Sobre aprendizagem adaptativa com LLMs, destacam-se duas vertentes: o *knowledge tracing*, voltado à estimativa do domínio do estudante, e a personalização de conteúdo, que adapta materiais ao nível, preferências e objetivos de aprendizagem [Wang et al., 2024]. Os LLMs podem apoiar esses processos pela continuidade contextual, *feedback* personalizado e ambientes de investigação ativa [Shi et al., 2026], embora ainda existam preocupações quanto à precisão das respostas, dependência excessiva e impactos no pensamento computacional [Vadaparty et al., 2024].

3. Metodologia

Esta RSL seguiu diretrizes de Kitchenham e Charters, e apoiada na ferramenta Parsifal na definição do protocolo, organização das buscas, seleção dos estudos e extração dos dados, sendo estruturada em três etapas: planejamento, condução e apresentação dos resultados [Kitchenham and Charters, 2007; Parsifal Team, 2024].

3.1. Planejamento da Revisão

Na primeira etapa, foi elaborado o protocolo inicial da RSL, contemplando objetivo, questões de pesquisa, estratégia de busca, critérios de seleção, avaliação da qualidade e formulário de extração de dados, com estrutura orientada pelo critério PICOC [Petticrew and Roberts, 2008], conforme apresentado na Tabela 1.

Tabela 1. Estrutura PICOC utilizada para definição do protocolo da RSL

Population	Estudantes do ensino superior na área de Computação
Intervention	Uso de Large Language Models (LLMs), IA generativa, aprendizagem adaptativa
Comparison	Ensino tradicional, avaliações convencionais, ambientes de aprendizagem não adaptativos
Outcome	Melhoria no desempenho acadêmico, redução de erros, diminuição de lacunas de aprendizagem, aumento do engajamento
Context	Ensino superior, no contexto de disciplinas de Computação

3.1.1 Questões de Pesquisa

A partir da delimitação de tema, protocolo e objetivo, foram estabelecidas as questões de pesquisa que orientaram a condução do trabalho. Em consequência dessa delimitação, foram definidas as seguintes questões centrais da pesquisa:

- QP1 – Como os LLMs têm sido utilizados no ensino superior de Computação?
- QP2 – Quais estratégias de aprendizagem adaptativa têm sido aplicadas em plataformas educacionais para personalizar o ensino em Computação?

- *QP3 – Quais métricas e métodos de avaliação têm sido usados para medir a eficácia de sistemas educacionais baseados em LLMs e aprendizagem adaptativa?*
- *QP4 – Quais evidências existem de que plataformas com Inteligência Artificial Generativa e aprendizagem adaptativa contribuem para o desempenho acadêmico, permanência e redução da evasão no ensino superior de Computação?*

3.1.2 String de Busca e Seleção de Fontes de Busca

As *strings* da RSL foram definidas com base nas questões de pesquisa e no protocolo PICOC, considerando os principais termos do estudo, conforme ilustrado na Figura 1.

String de Busca

```
((“Computer Science” OR “Software Engineering” OR “Computing” OR “Computer Engineering” OR “Computing education” OR “Computer Science education” OR “Software Engineering education”) AND (“Higher education” OR “University” OR “Undergraduate”)) AND (“Large Language Model” OR “Large Language Models” OR “LLM” OR “LLMs” OR “Generative AI” OR “Generative Artificial Intelligence”) AND (“Adaptive learning” OR “personalized learning” OR “adaptive education” OR “adaptive educational systems” OR “intelligent tutoring systems” OR “intelligent learning environments” OR “automated feedback”)
```

Figura 1. String de Busca

As fontes de busca foram definidas com base em estudo anterior, que destaca bases acadêmicas bem avaliadas, como *Engineering Village*, *Scopus*, *IEEE Xplore*, *ACM Digital Library* e *Springer Link* [Buchinger, Cavalcanti and Hounsell, 2014].

3.1.3 Critérios de Inclusão e Exclusão

A definição dos critérios de inclusão e exclusão orientou a triagem dos estudos, delimitando o escopo da revisão e garantindo maior relevância e especificidade aos resultados [Bartelle and Medeiros, 2025].

Tabela 2. Critérios de Inclusão (CI) e de Exclusão (CE)

Código	Critérios de Inclusão
CI1	Estudos primários que abordem o uso de LLMs, IA generativa e aprendizagem adaptativa no ensino superior em computação.
CI2	Estudos que descrevem métodos, técnicas, aplicações, algoritmos, sistemas, plataformas ou ferramentas educacionais relacionados ao tema.
Código	Critérios de Exclusão
CE1	Estudos fora do escopo da pesquisa.
CE2	Estudos que não tratem de LLMs/IA generativa nem de aprendizagem em contexto de ensino superior em computação.
CE3	Trabalhos pertencentes à literatura cinzenta (TCCs, dissertações, teses).
CE4	Trabalhos com menos de 5 páginas.
CE5	Estudos duplicados.
CE6	Trabalhos sem texto completo disponível.
CE7	Trabalhos não escritos em língua inglesa.

3.1.4 Avaliação da Qualidade e Extração de dados

Na etapa de avaliação da qualidade, foi aplicado um *checklist* composto por dez questões. Para cada questão, foram atribuídas as pontuações *Yes* (1,0 ponto), *Partially*

(0,5 ponto) ou *No* (0,0 ponto), de modo que cada estudo poderia alcançar uma pontuação máxima de 10,0 pontos. Os estudos foram avaliados conjuntamente pelos cinco pesquisadores responsáveis pela RSL. Para cada questão do *checklist*, os pesquisadores discutiram as evidências apresentadas no estudo e definiram, por consenso, uma única classificação. A pontuação final de cada artigo correspondeu à soma das notas atribuídas às dez questões de qualidade.

3.2. Condução da Revisão

A segunda etapa da RSL consistiu na condução da revisão, estruturada em quatro fases, conforme ilustrado na Figura 2. As buscas foram realizadas entre março e abril de 2026 nas bases ACM Digital Library, Scopus, IEEE Xplore, SpringerLink e Engineering Village. A string de busca principal foi adaptada à sintaxe, aos operadores e aos campos de pesquisa disponibilizados por cada base, mantendo-se os mesmos grupos conceituais definidos no protocolo da revisão. Na primeira fase, correspondente à identificação dos estudos, foram recuperados 2.039 registros. Desse total, 1.085 foram provenientes da ACM Digital Library, 90 da Scopus, 341 da IEEE Xplore, 434 da SpringerLink e 89 da Engineering Village. Na segunda fase, ocorreu a verificação dos estudos, considerando critérios como título, resumo e palavras-chave. Após essa triagem, permaneceram 489 estudos. Na terceira fase, durante a aplicação dos critérios de inclusão e exclusão, 221 estudos foram removidos, o que resultou em 268 estudos. Por fim, na quarta fase, os estudos restantes foram submetidos à avaliação de qualidade pelos pesquisadores. Nessa etapa, foram selecionados apenas os artigos que obtiveram nota igual ou superior a 8, totalizando 35 estudos, os quais foram utilizados para a análise e extração de dados.

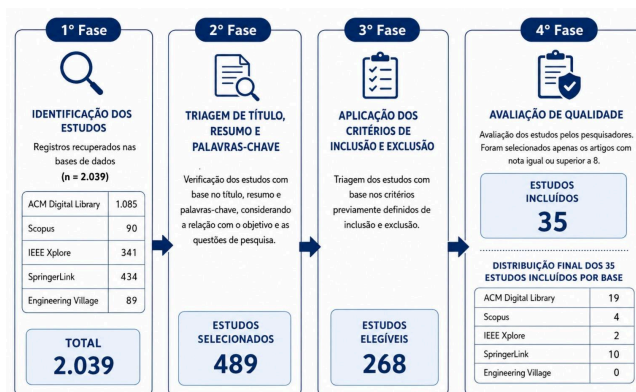


Figura 2. Fluxo de identificação, triagem e seleção dos estudos

A relação completa dos estudos incluídos na revisão, bem como o conjunto das dez questões de qualidade aplicadas, encontra-se disponível em planilha¹. Para facilitar a rastreabilidade dos resultados, cada estudo incluído recebeu um identificador único (ID), no formato A01 a A35. Esses IDs são utilizados na seção de resultados para referenciar os artigos analisados de forma padronizada.

4. Resultados

¹ <https://doi.org/10.6084/m9.figshare.32975435>

Em 2026, considerando as publicações identificadas até abril, foram registrados 11,4% dos artigos analisados. Em 2025, observou-se a maior concentração de estudos, correspondendo a 57,1% da amostra. Em 2024, foram identificadas 28,6% das publicações, enquanto 2023 apresentou a menor participação, com apenas 2,9%. Na distribuição geográfica, destacam-se os Estados Unidos (20,9%) e a Alemanha (14,9%). A Figura 3 sintetiza a evolução anual e a dispersão geográfica dos estudos.

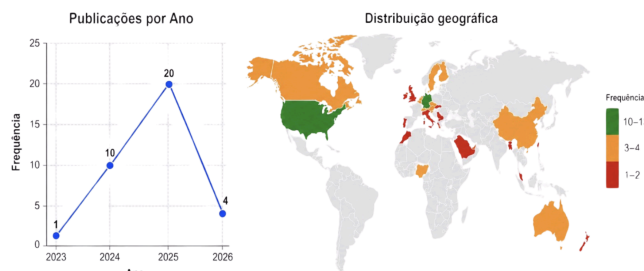


Figura 3. Distribuição dos artigos

4.1. RQP1 - Como LLMs têm sido utilizados no ensino superior de Computação?

A Figura 4 mostra a síntese das seis modalidades dos dos 35 artigos sobre o uso de LLMs no ensino superior de Computação: (i) tutoria conversacional integrada a ambientes virtuais de aprendizagem [A21], [A22], [A23]; (ii) feedback automatizado sobre submissões de código, textos e artefatos técnicos [A03], [A04], [A07], [A12]; e (iii) apoio à correção, depuração e reparo automatizado de programas [A09], [A10]. Também foram identificadas aplicações voltadas a (iv) avaliação automatizada de respostas abertas, provas matemáticas e atividades práticas [A14], [A26], [A35]; (v) geração, classificação e recomendação de exercícios personalizados [A06], [A27], [A34]; e (vi) desenvolvimento de literacia em IA, pensamento crítico e uso responsável dessas tecnologias em contextos formativos de Computação [A02], [A13], [A19].

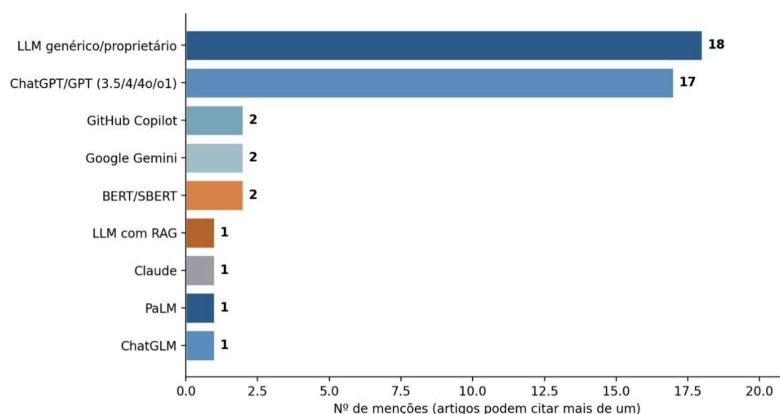


Figura 4. Modelos de linguagem reportados nos 35 estudos analisados

Além da identificação dos modelos, os estudos evidenciam que os LLMs têm sido incorporados como componentes educacionais no processo de ensino-aprendizagem, especialmente em programação, engenharia de software, bancos de dados, cibersegurança e fundamentos matemáticos da Computação. Nessas áreas, são utilizados para oferecer explicações contextualizadas, orientar a resolução de

problemas, fornecer *feedback* imediato, apoiar a avaliação formativa e adaptar intervenções às dificuldades dos estudantes [A16], [A21], [A25], [A35].

Essa tendência indica uma transição do uso genérico de chatbots para aplicações educacionais mais estruturadas, integradas a ambientes virtuais de aprendizagem, sistemas de avaliação, bases de conhecimento disciplinares, técnicas de Geração Aumentada por Recuperação (*Retrieval-Augmented Generation* — RAG) e estratégias de personalização de ensino [A22], [A23], [A29], [A34]. Assim, os LLMs têm ampliado a mediação instrutiva e o acompanhamento individualizado, embora persistam desafios relacionados à dependência excessiva, confiabilidade das respostas, integridade acadêmica e necessidade de avaliações mais robustas [A05], [A17], [A18], [A24].

4.2. RQP2 - Quais estratégias de aprendizagem adaptativa têm sido aplicadas em plataformas educacionais para personalizar o ensino em computação?

A aprendizagem adaptativa, entendida como o ajuste do conteúdo, ritmo, dificuldade ou suporte instrucional às necessidades do aprendiz [Wang et al., 2024; Spriggs, Lau and Passi, 2025], aparece de forma recorrente nos estudos analisados. Conforme sintetizado na Figura 5, foram identificadas oito estratégias principais, com destaque para o *feedback* contextual personalizado, presente em nove artigos, aplicado à correção de códigos, respostas abertas, artefatos técnicos e explicações ajustadas às dúvidas dos estudantes [A03], [A04], [A07], [A12], [A21], [A23], [A26], [A29], [A35]. Também se destaca o uso de andaime cognitivo (*scaffolding*), identificado em seis estudos, no qual o sistema conduz o estudante por meio de pistas, decomposição de problemas, perguntas orientadoras e *feedback* progressivo [A07], [A09], [A10], [A16], [A21], [A25].

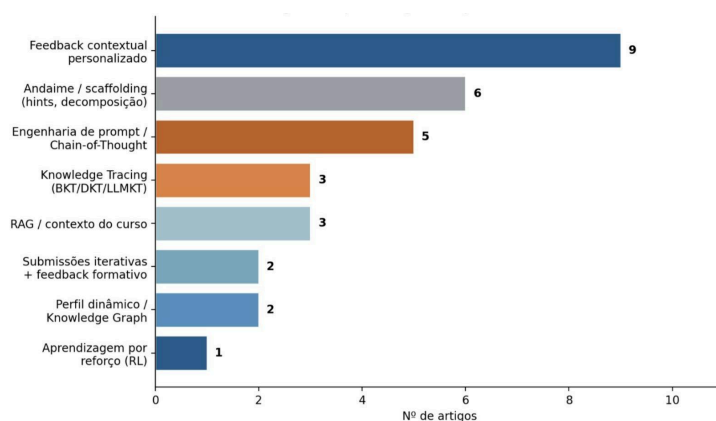


Figura 5. Estratégias de aprendizagem adaptativa identificadas nos 35 estudos

Outras estratégias incluem engenharia de prompt, utilizadas para estruturar respostas passo a passo e melhorar a qualidade das explicações [A02], [A13], [A14], [A24], [A35], além de abordagens baseadas em *Knowledge Tracing*, voltadas à estimativa do estado de conhecimento do estudante a partir de seu desempenho e histórico de interação [A16], [A34]. Também foram identificadas integrações com RAG e recuperação de contexto do curso, que tornam as respostas dos LLMs mais aderentes aos materiais institucionais [A22], [A23], [A29]. Em menor frequência, aparecem submissões iterativas com *feedback* formativo [A04], [A12], perfis dinâmicos do

estudante, *Knowledge Graphs* [A27], [A34] e aprendizagem por reforço [A34], indicando uma tendência de evolução de sistemas reativos.

4.3. RQP3 - Quais métricas e métodos de avaliação têm sido usados para medir a eficácia de sistemas educacionais baseados em LLMs e aprendizagem adaptativa?

A avaliação dos sistemas analisados combina métricas técnicas, educacionais e perceptivas. Conforme apresentado na Figura 6, os métodos mais recorrentes foram *surveys*, escalas Likert e entrevistas, identificados em 20 estudos, indicando forte interesse em compreender utilidade, usabilidade, satisfação, confiança, engajamento e aceitação das ferramentas [A03], [A05], [A06], [A07], [A14], [A17], [A18], [A21], [A22], [A23], [A24], [A25], [A29]. A análise qualitativa aparece em 14 artigos, geralmente aplicada à interpretação de respostas abertas, entrevistas, comentários dos estudantes, registros de uso e interações com os sistemas [A03], [A04], [A05], [A13], [A18], [A21], [A22], [A23], [A25], [A29].

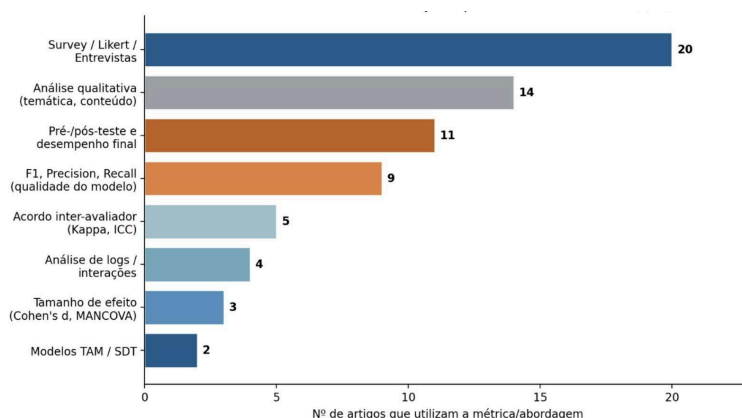


Figura 6. Frequência das métricas e métodos de avaliação

No eixo pedagógico, pré-testes, pós-testes e desempenho final foram utilizados em 11 estudos para verificar ganhos de aprendizagem, evolução conceitual e comparação entre grupos com e sem apoio de IA [A06], [A10], [A14], [A16], [A17], [A19], [A25], [A29], [A34]. No eixo técnico, métricas como *F1-score*, precisão e *recall* aparecem em 9 estudos, sobretudo em tarefas de correção automática, classificação de respostas e avaliação de artefatos educacionais [A02], [A04], [A12], [A26], [A27], [A29], [A35]. Também foram utilizados indicadores de concordância com avaliadores humanos, como Kappa e ICC, em 5 estudos [A04], [A13], [A26], [A29], [A35]. Em menor frequência, aparecem análise de *logs* e interações [A07], [A12], [A16], [A22], métricas estatísticas como tamanho de efeito, Cohen's d e MANCOVA [A06], [A17], [A19], além de modelos teóricos de aceitação e motivação, como TAM e *Self-Determination Theory* [A06], [A18].

4.4. RQP4 - Quais evidências existem de que plataformas com Inteligência Artificial Generativa e aprendizagem adaptativa contribuem para desempenho acadêmico, permanência e redução da evasão no ensino superior de computação?

A Figura 7a apresenta a direção dos efeitos reportados nos 35 estudos. A maioria dos artigos, 25 estudos ou 71% da amostra, relata evidências positivas; 8 estudos,

correspondentes a 23%, apresentam resultados mistos; e apenas 2 estudos indicam efeitos neutros, negativos ou mistos, cada um representando 3%. Esses resultados sugerem uma tendência favorável ao uso de plataformas com IAG e aprendizagem adaptativa, especialmente quando associadas a *feedback* personalizado, tutoria inteligente, apoio à programação, avaliação automatizada e personalização de atividades [A03], [A06], [A07], [A10], [A16], [A21], [A25], [A29], [A34].

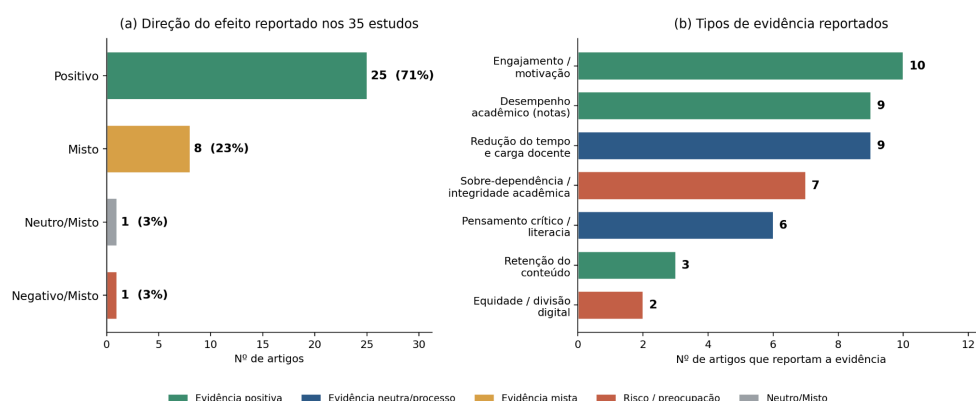


Figura 7. (a) Direção do efeito reportado nos 35 estudos e (b) tipos de evidência identificadas

Conforme a Figura 7b, as evidências concentram-se principalmente em três dimensões: engajamento e motivação, identificados em dez estudos [A05], [A06], [A18], [A21], [A22], [A23], [A25]; desempenho acadêmico, presente em nove estudos, com relatos de melhora em notas, tarefas, submissões, resolução de problemas e pré/pós-testes [A06], [A10], [A14], [A16], [A17], [A19], [A25], [A29], [A34]; e redução do tempo ou da carga docente, também observada em nove estudos, especialmente em atividades de *feedback*, correção automática e avaliação de código [A04], [A12], [A22], [A23], [A26], [A29], [A35].

Além disso, alguns estudos relatam contribuições para o pensamento crítico e literacia em IA [A13], [A18], [A19], [A24], [A25], bem como para a retenção de conteúdo [A10], [A16], [A34]. Contudo, também foram identificadas preocupações com sobre-dependência, integridade acadêmica e confiabilidade das respostas [A05], [A14], [A17], [A18], [A24], além de questões de equidade e divisão digital [A18], [A24].

5. Discussão

A análise dos estudos selecionados permite identificar três aspectos centrais sobre o uso de LLMs e aprendizagem adaptativa no ensino superior de Computação. O primeiro refere-se ao papel instrutivo dos LLMs, que apresentam maior potencial quando empregados como mediadores da aprendizagem, e não apenas como geradores automáticos de respostas. Sistemas que oferecem dicas graduais, perguntas orientadoras, *feedback* progressivo e explicações contextualizadas tendem a favorecer interações mais alinhadas aos objetivos educacionais, como em atividades de programação, depuração e resolução de problemas [Ma et al., 2025; Sölch, Dietrich and Krusche, 2025].

O segundo aspecto relaciona-se à aprendizagem adaptativa. Os estudos indicam que a personalização depende da capacidade dos sistemas de representar o estado de conhecimento do estudante, considerando seu desempenho, dificuldades recorrentes, histórico de interações e nível de autonomia. Nesse contexto, estratégias como *knowledge tracing*, classificação automática de exercícios, análise de submissões e integração com ambientes virtuais tornam-se a base para intervenções personalizadas [Scarlatos, Baker and Lan, 2025; Lombarts et al., 2025].

O terceiro ponto diz respeito à avaliação desses sistemas. As métricas utilizadas abrangem dimensões acadêmicas, técnicas e perceptivas. Embora essa diversidade permita analisar dimensões didáticas, técnicas e perceptivas, também dificulta comparações diretas entre os estudos e limita a consolidação de evidências mais robustas sobre a efetividade dessas tecnologias [Pecuchova, Benko and Drlik, 2025; Qian, Ji and Li, 2026].

5.1. Lacunas Identificadas

A primeira lacuna envolve a generalização limitada das evidências, pois grande parte dos estudos ocorre em contextos institucionais únicos, cursos introdutórios e populações semelhantes. Ainda são raras comparações entre países, instituições, níveis de ensino e diferentes condições de acesso tecnológico [Asgari et al., 2024; Haque and Hundhausen, 2025; Prather et al., 2024]. A segunda lacuna diz respeito à escassez de estudos longitudinais e experimentais mais robustos. Em muitos casos, prevalecem intervenções curtas, dados autorrelatados e ausência de grupos de controle, o que dificulta a verificação de aprendizagem sustentada, retenção de conteúdo ou transferência para novos problemas [Omeh et al., 2025; Qian, Ji and Li, 2026; Rogers et al., 2025].

A terceira lacuna refere-se à ausência de estudos que relacionem diretamente IAG e aprendizagem adaptativa a indicadores institucionais de permanência e evasão. Predominam métricas intermediárias, como satisfação, engajamento e desempenho pontual [Bassner, Lottner and Krusche, 2025; Gambo et al., 2026; Mzwri and Turcsányi-Szabo, 2025]. De forma transversal, destaca-se a necessidade de designs educacionais que reduzam a dependência da IA e promovam pensamento crítico e autoria, pois bom desempenho com apoio da ferramenta não garante aprendizagem real [Bouvier et al., 2025; Elkhodr and Gide, 2025; Rogers et al., 2025].

6. Conclusões e Trabalhos Futuros

Este trabalho apresentou uma RSL sobre o uso de LLMs e aprendizagem adaptativa no ensino superior de Computação. A investigação foi motivada pelos desafios contemporâneos da Educação em Computação, especialmente aqueles relacionados ao papel da IA como agente educacional e à necessidade de práticas instrutivas mais personalizadas, críticas e centradas no estudante, conforme destacado pela [SBC, 2025].

A partir da análise de 35 estudos, foram identificadas 6 modalidades principais de uso dos LLMs, incluindo tutoria conversacional, *feedback* automatizado, apoio à depuração e reparo de código, avaliação automatizada, recomendação de exercícios e

desenvolvimento de literacia em IA. Os resultados indicam que os LLMs tendem a apresentar maior potencial de ensino quando não são utilizados apenas como geradores de respostas, mas como mediadores do ensino capazes de oferecer explicações adaptadas, pistas graduais, *feedback* formativo e apoio à resolução de problemas.

Quanto às estratégias de aprendizagem adaptativa, observou-se predominância de *feedback* contextual personalizado, andaime cognitivo, engenharia de *prompt*, *Knowledge Tracing*, recuperação de contexto via RAG e mecanismos de personalização baseados no perfil ou histórico do estudante. Essas estratégias demonstram uma evolução dos sistemas educacionais. Em relação à avaliação, os estudos empregam métricas técnicas, pedagógicas e perceptivas, cuja diversidade permite observar diferentes dimensões do fenômeno, mas dificulta comparações diretas entre os estudos. Já as evidências encontradas sugerem contribuições positivas das plataformas com IAG e aprendizagem adaptativa para o desempenho acadêmico, o engajamento, a motivação, a redução da carga de trabalho docente e o desenvolvimento de literacia em IA. Contudo, sua relação com permanência estudantil e evasão ainda é pouco explorada, uma vez que nenhum dos estudos analisados mediu diretamente indicadores institucionais como abandono, trancamento, reprovação, continuidade ou evasão.

Este estudo contribui ao sistematizar o estado atual da literatura e ao evidenciar que o principal desafio não está apenas em incorporar LLMs ao ensino, mas em integrá-los a modelos de ensino, avaliativos e éticos capazes de promover aprendizagem efetiva, pensamento crítico e acompanhamento individualizado. Esse desafio dialoga com a agenda proposta pela SBC [SBC, 2025], reforçando a necessidade de avaliar a IA como elemento estruturante de novos ecossistemas educacionais em Computação.

Como limitações deste trabalho, destaca-se que não foram selecionados estudos conduzidos no contexto brasileiro, o que limita a compreensão sobre a apropriação dessas tecnologias em cursos de Computação no país, especialmente diante de desigualdades regionais, institucionais e de infraestrutura. Além disso, a predominância de publicações em língua inglesa pode ter restringido a identificação de produções nacionais ou regionais não indexadas nas bases consultadas. Por fim, a escassez de dados longitudinais dificulta a avaliação de efeitos sobre aprendizagem, permanência, evasão e desenvolvimento de competências.

Como trabalhos futuros, os resultados desta RSL serão utilizados como fundamentos para o desenvolvimento de uma plataforma educacional que integre LLMs, aprendizagem adaptativa e análise inteligente de interações, dando continuidade a estudos anteriores sobre autoavaliação com *feedback* imediato [Carvalho et al., 2025] e uso de *Machine Learning* para identificação de barreiras cognitivas [Carvalho and Andrade, 2026]. Na plataforma, a IA atuará na identificação de padrões, dificuldades e desempenho; os LLMs, na interpretação de respostas e na interação com os estudantes; e os recursos de IAG, na geração de *feedbacks* personalizados, explicações contextualizadas e recomendações de ensino. Com isso, espera-se apoiar o professor na detecção de lacunas formativas e na tomada de decisões de ensino. Além disso, será desenhada para uso ao longo de semestres completos, permitindo a coleta de dados longitudinais de modo a gerar evidências mais robustas e sólidas.

Declaração sobre uso de Inteligência Artificial

Este trabalho utilizou as ferramentas de Inteligência Artificial Generativa (IAG) ChatGPT 5.5 Plus e Claude 4.8 Pro. O uso dessas ferramentas concentrou-se na revisão gramatical e ortográfica, no aprimoramento da clareza, da coesão e da fluidez textual, bem como no apoio à criação de imagens e gráficos, sempre com base em instruções, informações e dados fornecidos pelas pessoas autoras. Ressalta-se que a IAG não foi utilizada para conduzir análises, interpretar dados, definir metodologia ou produzir resultados científicos. Todas as decisões teóricas, analíticas e metodológicas são de autoria das pessoas autoras, assim como a seleção, análise e discussão dos estudos.

Referências

- Asgari, M., Tsai, F.-C., Mannila, L., Strömbäck, F. and Sadique, K. M. (2024) "Students' perspectives on using digital tools in programming courses: a cross country case study between Sweden and Taiwan", *Discover Education*, v. 3, n. 1. DOI: 10.1007/s44217-024-00170-2.
- Ayeni, O. O., Al Hamad, N. M., Chisom, O. N., Osawaru, B. and Adewusi, E. O. (2024) "AI in education: a review of personalized learning and educational technology", *GSC Advanced Research and Reviews*, v. 18, n. 2, p. 261–271. DOI: 10.30574/gscarr.2024.18.2.0062.
- Bartelle, C. E. and Medeiros, F. P. A. (2025) "Inteligência artificial generativa em ambientes virtuais de aprendizagem: revisão sistemática da literatura", In: *Anais do XXXVI Simpósio Brasileiro de Informática na Educação (SBIE)*, Maceió, Sociedade Brasileira de Computação, p. 904–914. DOI: 10.5753/sbie.2025.6577.
- Bassner, P., Lottner, A. and Krusche, S. (2025) "Towards understanding the impact of context-aware AI tutors and general-purpose AI chatbots on student learning", In: *Proceedings of the 25th Koli Calling International Conference on Computing Education Research (Koli Calling '25)*, Koli, ACM. DOI: 10.1145/3769994.3770025.
- Bouvier, D. J., Cipriano, B. P., Glassey, R., Petrovska, O., Anderson, E., Birillo, A. et al. (2025) "The rest of the robots: generative AI in post-introductory computing education", In: *Proceedings of the 2025 Working Group Reports on Innovation and Technology in Computer Science Education (ITiCSE-WGR 2025)*, Nijmegen, ACM. DOI: 10.1145/3760545.3783970.
- Buchinger, D., Cavalcanti, G. A. de S. and Hounsell, M. da S. (2014) "Mecanismos de busca acadêmica: uma análise quantitativa", *Revista Brasileira de Computação Aplicada*, v. 6, n. 1, p. 108–120.
- Carvalho, D. R. A., Barroso, F. V. C., Pinheiro, T. S. de M. and Vasconcelos, P. F. C. de (2025) "Plataforma digital para autoavaliação e aprendizado de residentes em Oftalmologia", In: *Anais do IX Congresso sobre Tecnologias na Educação (Ctrl+E)*, Fortaleza, Sociedade Brasileira de Computação, p. 406–410. DOI: 10.5753/ctrl.2025.11683.
- Carvalho, D. R. A. and Andrade, R. M. de C. (2026) "Machine Learning para a identificação de barreiras cognitivas em estudantes de Engenharia de Software",

- In: Anais do VI Simpósio Brasileiro de Educação em Computação (EduComp), Sociedade Brasileira de Computação. DOI: 10.5753/educomp.2026.
- Castilho, G. U., Herrera, V. A. and Rodriguez, C. L. (2025) "Explorando o potencial da IA generativa em ambientes virtuais de aprendizagem: uma revisão sistemática", In: Anais do Simpósio Brasileiro de Informática na Educação (SBIE), Porto Alegre, Sociedade Brasileira de Computação. DOI: 10.5753/sbie.2025.11994.
- Elkhodr, M. and Gide, E. (2025) "The SAGE framework for developing critical thinking and responsible generative AI use in cybersecurity education", *Discover Education*, v. 4, n. 517. DOI: 10.1007/s44217-025-00935-3.
- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X. and Gašević, D. (2025) "Beware of metacognitive laziness: effects of generative artificial intelligence on learning motivation, processes, and performance", *British Journal of Educational Technology*, v. 56, n. 2, p. 489–530. DOI: 10.1111/bjet.13544.
- Gambo, O., Onifade, O. J., Olanrewaju-Smart, W., Abuya, O. J., Salawu, R. E., Massenon, R. and Gambo, I. (2026) "Impact of ChatGPT on learning outcomes and performance of students in computer programming courses: a mixed-method approach", *Discover Education*, v. 5, n. 164. DOI: 10.1007/s44217-026-01142-4.
- Haque, S. and Hundhausen, C. (2025) "Generative AI access, usage, and perceptions: an empirical comparison of computing students in the United States and Bangladesh", In: *Proceedings of the 2025 ACM Conference on International Computing Education Research (ICER 2025)*, v. 1, Charlottesville, ACM. DOI: 10.1145/3702652.3744231.
- Kitchenham, B. and Charters, S. (2007) *Guidelines for performing systematic literature reviews in software engineering*. Keele; Durham: Keele University; University of Durham. EBSE Technical Report, EBSE-2007-01.
- Lodge, J. M., Thompson, K. and Corrin, L. (2023) "Mapping out a research agenda for generative artificial intelligence in tertiary education", *Australasian Journal of Educational Technology*, v. 39, n. 1, p. 1–8.
- Lombarts, T., Walravens, G., Seraj, M., Ochoa, L. and van den Brand, M. (2025) "Personalization of programming education: an NLP-based bi-dimensional classification of programming exercises", In: *Proceedings of the 2025 ACM SIGPLAN International Symposium on SPLASH-E (SPLASH-E '25)*, ACM. DOI: 10.1145/3758317.3759684.
- Ma, S., Wang, J., Zhang, Y., Ma, X. and Wang, A. Y. (2025) "DBox: scaffolding algorithmic programming learning through learner-LLM co-decomposition", In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*, Yokohama, ACM. DOI: 10.1145/3706598.3713748.
- Mzwri, K. and Turcsányi-Szabo, M. (2025) "Bridging LMS and generative AI: dynamic course content integration (DCCI) for enhancing student satisfaction and engagement via the Ask ME assistant", *Journal of Computers in Education*. DOI: 10.1007/s40692-025-00367-w.

- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N. and Mian, A. (2023) "A comprehensive overview of large language models", arXiv preprint arXiv:2307.06435.
- Omeh, C. B., Ayanwale, M. A., Mnguni, L. E. and Olelewe, C. J. (2025) "Fostering programming skill and critical thinking through AI-assisted PBL integration", *Journal of New Approaches in Educational Research*, v. 14, n. 22. DOI: 10.1007/s44322-025-00041-0.
- Parsifal Team (2024) Parsifal: systematic review tool. Disponível em: <https://parsif.al/>. Acesso em: 30 mar. 2024.
- Pecuchova, J., Benko, L. and Drlik, M. (2025) "Automated grading of open-ended questions in higher education using GenAI models", *International Journal of Artificial Intelligence in Education*, v. 35, p. 3813–3846. DOI: 10.1007/s40593-025-00517-2.
- Petticrew, M. and Roberts, H. (2008) *Systematic reviews in the social sciences: a practical guide*. Malden, MA, Blackwell Publishing.
- Prather, J., Reeves, B. N., Leinonen, J., MacNeil, S., Randrianasolo, A. S., Becker, B. A., Kimmel, B., Wright, J. and Briggs, B. (2024) "The widening gap: the benefits and harms of generative AI for novice programmers", In: *Proceedings of the 2024 ACM Conference on International Computing Education Research (ICER 2024)*, v. 1, Melbourne, ACM, p. 469–486. DOI: 10.1145/3632620.3671116.
- Qian, M., Ji, H. and Li, L. (2026) "Empirical validation of a generative AI framework for personalized education assessment", *Scientific Reports*. DOI: 10.1038/s41598-026-42169-9.
- Rendeiro, P., Monteiro, E., Nicioka, D. and Santos, R. (2023) "Avaliação de desempenho em cursos de Engenharia de Computação baseada nos microdados do ENADE", In: *Anais do XXXI Workshop sobre Educação em Computação (WEI)*, João Pessoa, Sociedade Brasileira de Computação, p. 339–349. DOI: 10.5753/wei.2023.229330.
- Rogers, K., Davis, M., Maharana, M., Etheredge, P. and Chernova, S. (2025) "Playing dumb to get smart: creating and evaluating an LLM-based teachable agent within university computer science classes", In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*, Yokohama, ACM. DOI: 10.1145/3706598.3713644.
- SBC – Sociedade Brasileira de Computação (2025) *Grandes desafios da educação em computação no Brasil 2025–2035*. Porto Alegre, SBC. Disponível em: <https://books-sol.sbc.org.br/index.php/sbc/catalog/book/175>. Acesso em: 9 maio 2026.
- Scarlato, A., Baker, R. S. and Lan, A. (2025) "Exploring knowledge tracing in tutor-student dialogues using LLMs", In: *Proceedings of the 15th International Learning Analytics and Knowledge Conference (LAK '25)*, Dublin, ACM. DOI: 10.1145/3706468.3706501.
- Shi, Y., Yu, K., Dong, Y. and Chen, F. (2026) "Large language models in education: a systematic review of empirical applications, benefits, and challenges", *Computers*

- and Education: Artificial Intelligence, v. 10, p. 100529. DOI: 10.1016/j.caeai.2025.100529.
- Silva, F. G. and Aranha, E. H. S. (2025) "LLMs na educação em programação: estratégias para avaliação e feedback formativo", In: Anais do V Simpósio Brasileiro de Educação em Computação (EDUCOMP) — Workshop de Teses e Dissertações, Juiz de Fora, Sociedade Brasileira de Computação, p. 81–87. DOI: 10.5753/educomp_estendido.2025.6614.
- Silva, T. L., Vidotto, K. N. S., Tarouco, L. M. R. and Silva, P. F. (2024) "Potencialidades do uso de inteligência artificial generativa como apoio ao ensino de programação", In: Anais do XXXV Simpósio Brasileiro de Informática na Educação (SBIE), Porto Alegre, Sociedade Brasileira de Computação, p. 1942–1956. DOI: 10.5753/sbie.2024.242711.
- Sölch, M., Dietrich, F. T. J. and Krusche, S. (2025) "Direct automated feedback delivery for student submissions based on LLMs", In: Companion Proceedings of the 33rd ACM International Conference on the Foundations of Software Engineering (FSE Companion '25), Trondheim, ACM. DOI: 10.1145/3696630.3727247.
- Spriggs, K., Lau, M. C. and Passi, K. (2025) "Personalizing education through an adaptive LMS with integrated LLMs", arXiv preprint arXiv:2502.08655. DOI: 10.48550/arXiv.2502.08655.
- Trujillo, D. E. S. and Corrêa, J. M. (2025) "Análise automatizada de dados do ENADE com aplicação para os cursos de Ciência da Computação", In: Anais do XXXIII Workshop sobre Educação em Computação (WEI), Maceió, Sociedade Brasileira de Computação, p. 1043–1054. DOI: 10.5753/wei.2025.8929.
- Vadaparty, A., Zingaro, D., Smith IV, D. H., Padala, M., Alvarado, C., Gorson Benario, J. and Porter, L. (2024) "CS1-LLM: integrating LLMs into CS1 instruction", In: Proceedings of the 2024 on Innovation and Technology in Computer Science Education (ITiCSE 2024), v. 1, ACM, p. 297–303.
- Valente, J. A., Machado, L. S. and Barbosa, E. F. (2023) "Educação em computação: um campo em consolidação e os desafios emergentes para a próxima década", Revista Brasileira de Informática na Educação, v. 31, n. 1.
- Wang, S. et al. (2024) "Large language models for education: a survey and outlook", arXiv preprint arXiv:2403.18105v2. Disponível em: <https://arxiv.org/abs/2403.18105>.
- Xu, X., Chen, Y. and Miao, J. (2024) "Opportunities, challenges, and future directions of large language models, including ChatGPT in medical education: a systematic scoping review", Journal of Educational Evaluation for Health Professions, v. 21, p. 6. DOI: 10.3352/jeehp.2024.21.6.
- Zarei, M., Zarei, M., Hamzehzadeh, S., Oliyai, S. and Hosseini, M. (2024) "ChatGPT, a friend or a foe in medical education: a review of strengths, challenges, and opportunities", Shiraz E-Medical Journal, v. 25, n. 6, e145840. DOI: 10.5812/semj-145840.
- Zimmerman, B. J. (2002) "Becoming a self-regulated learner: an overview", Theory Into Practice, v. 41, n. 2, p. 64–70. DOI: 10.1207/s15430421tip4102_2.