

Reflexões sobre o Uso de LLMs no Ensino de Programação

Yasmim Silva Rosa¹, Pedro Garcia², Kattiana Constantino³, Eduardo Figueiredo¹

¹ Universidade Federal de Minas Gerais (UFMG) – Brasil

² Universidade Federal do Ouro Preto (UFOP) – Brasil

³ Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM) – Brasil

yasrrosa@ufmg.br, figueiredo@dcc.ufmg.br

pedro.garcia@ufop.edu.br

kattiana.constantino@ufvjm.edu.br

Abstract. *The objective of this article is to understand how Large Language Models (LLMs) impact programming teaching. Currently, many beginners in programming languages, in terms of creating, correcting and explaining code, have been resorting to Large Language Models due to their ease in obtaining ready-made codes and complete explanations. But do LLMs have a negative or positive impact on programming language teaching? This is what this article will discuss.*

Resumo. *O objetivo deste artigo é saber como os Modelos de Linguagem de Grande Escala (Large Language Models -LLMs) impactam no ensino de programação. Atualmente muitos iniciantes nas linguagens de programação, em termos de criação, correção e explicação de código, vêm recorrendo às Large Language Models devido à sua facilidade em obter códigos prontos e explicações completas. Mas será que as LLMs causam impacto negativo ou positivo no ensino de linguagens de programação? Isso que esse artigo irá discutir.*

1. Introdução

O mundo da programação está em constante evolução devido às mudanças na sociedade tecnológica em que vivemos (BG Lovatti, LS Vieira e K Marques et al., 2017). Por esse motivo, muitas pessoas têm se atraído pelas linguagens de programação, buscando ingressar ou migrar para essa área [Santos et al. 2019]. Com a democratização dos Large Language Models (LLMs), essas ferramentas, como o ChatGPT, tornaram-se populares devido à facilidade em fornecer códigos prontos e explicações detalhadas. As LLMs são modelos de inteligência artificial, ou mais especificamente de aprendizado de máquina [Carbonell et al. 1983], treinados em grandes volumes de dados textuais para identificar padrões de linguagem e, como consequência, serem capazes de compreender e gerar linguagem natural de forma coerente [Naveed et al., 2023].

O aprendizado de programação, em geral, demanda tempo, dedicação e prática [Renato de M. Santos e Credinê Silva de Menezes et al., 2019]. Para um iniciante na área, muitas vezes esse processo pode parecer longo e desafiador, especialmente devido à ansiedade e às propagandas de cursos que destacam salários altos e outras vantagens da carreira. Por outro lado, há também aqueles que já possuem algum conhecimento em programação, mas utilizam as LLMs como suporte para esclarecer dúvidas e aprofundar

seu aprendizado [YEPES, Igor et al., 2023]. Além disso, as LLMs podem ser ferramentas valiosas para docentes em salas de aula [YEPES, Igor et al., 2023].

No entanto, existem desafios associados ao uso das LLMs. Essas ferramentas podem, sem querer, promover plágio [JOŠT, Gregor et al., 2024], gerar respostas incorretas com convicção [Luiz C. Pereira Filho et al., 2023] e causar dependência por parte dos usuários [YEPES, Igor et al., 2023]. Portanto, esta pesquisa busca explorar os impactos das LLMs no ensino de programação, discutindo tanto os benefícios quanto as limitações dessa abordagem.

Ao final deste artigo serão apresentadas respostas para as seguintes perguntas de pesquisa:

RQ1. Qual o nível de escolaridade dos alunos em um curso de linguagem de programação usando LLMs?

RQ2. Quais as principais linguagens de programação ensinada usando LLMs?

RQ3. Quais são as principais LLMs citadas na literatura para o ensino de linguagem de programação?

RQ4. Quais os benefícios ao usarem LLMs no curso de linguagem de programação?

RQ5. Quais os impactos ao usarem LLMs no curso de linguagem de programação?

2. Trabalhos Relacionados

O estudo sobre as LLMs no ensino de programação tem sido estudado de forma crescente. Esse crescimento é devido a popularização e gratuidade, mesmo que tem algumas pagas, de LLMs na internet, para ser utilizado pelas pessoas de modo geral. A seguir um estudo relacionado as LLMs no ensino de programação:

[Natalie Kiesler et al. 2023] apresenta um estudo onde é analisado o desempenho das LLMs, utilizando especificamente o ChatGPT-3.5 e GPT-4 nas tarefas de introdução a programação. O estudo mostrou que as LLMs oferecem uma nova forma de resolução de problemas para tarefas de programação, com códigos com comentários explicativos. Por outro lado, as LLMs podem gerar respostas erradas com níveis de confiança de que está certo, devido a falta de clareza na descrição da tarefa, o que faz as LLMs oferecerem uma solução para um problema diferente.

3. Metodologia

O método utilizado para o desenvolvimento desta pesquisa foi uma revisão informal da literatura, baseado em técnicas exploradas em trabalhos anteriores [Kitchenham 2006, Souza et al. 2018, Keele et al. 2007] que nos inspirou a levantar e analisar vários artigos de pesquisadores nacionais e internacionais que publicaram artigos relacionados a este campo de pesquisa.

3.1. Pré análise

O buscador utilizado foi o Google Scholar, em que foram analisados no total 13 artigos, que foram pesquisados nas principais bibliotecas acadêmicas, artigos relacionados

ao tema e utilizadas como pesquisa as seguintes palavras chaves como: LLMs, educação, ensino, programação, benefícios e impactos. (“LLM”) AND (“ensino programação” OR “educação de programação”)

3.2. Pós análise e Coleta de dados

Depois de pesquisar os artigos, eles foram lidos e colocados em uma planilha, que continha filtros como: Base [Congresso/Biblioteca Online], Autores, Título, Ano de publicação, Linguagem de Programação, Benefícios, Malefícios, Técnicas Utilizadas e entre outros. Foram excluídos os artigos que não tinham como base [Congresso/Biblioteca Online]. A tabela 1 apresenta os artigos que discutem os impactos das LLMs no ensino da programação nas instituições brasileiras.

Tabela 1. Síntese dos resultados dos artigos que compuseram a amostra do estudo sobre o uso das LLMs no ensino de linguagem de programação.

| Referência | LLMs | Benefícios | Impactos | Ling. de Prog. | Modalidade |
|-------------------------------|---------------|-----------------|----------------------------------|----------------------|---------------|
| [Pereira Filho et al. 2023] | ChatGPT | Apoio no ensino | Códigos incorretos e Dependência | Linguagem C | Ensino Médio |
| [Yepes and Fiorin 2023] | ChatGPT | Apoio no ensino | Códigos incorretos | Linguagem C | Graduação |
| [da Silva Junior et al. 2023] | ChatGPT | Apoio no ensino | Códigos incorretos | N/D | Ensino Médio |
| [de Araujo Costa et al.] | N/D | Apoio no ensino | Dependência | N/D | Curso Técnico |
| [Jošt et al. 2024] | N/D | Apoio no ensino | Códigos incorretos e plágio | HTML/CSS/ Javascript | Graduação |
| [Geng et al. 2023] | ChatGPT | Apoio no ensino | Códigos incorretos | N/D | Graduação |
| [Leinonen et al. 2023] | GPT-3 | Apoio no ensino | Códigos incorretos | N/D | Graduação |
| [Kumar and Lan 2024] | GPT-4 | Apoio no ensino | Códigos incorretos | Java/Python | Graduação |
| [Silva et al. 2024] | ChatGPT | Apoio no ensino | Dependência | N/D | Graduação |
| [Xue et al. 2024] | ChatGPT | Apoio no ensino | Dependência | N/D | Graduação |
| [Geng et al. 2023] | ChatGPT | Apoio no ensino | Códigos incorretos | Java, Python | Curta Duração |
| [Lyu et al. 2024] | CodeTutor | Apoio no ensino | Geração de plágio | Python | Graduação |
| [Kiesler and Schiffner 2023] | ChatGPT/GPT-4 | Apoio no ensino | Respostas erradas | Python | Graduação |

4. Resultados

Esta seção apresenta os resultados de acordo com cada questão de pesquisa deste estudo. Esses resultados fornecem *insights* com base na revisão da literatura.

4.1. RQ₁ - Qual o nível de escolaridade dos alunos em um curso de linguagem de programação usando LLMs

A Figura 1 apresenta o nível de escolaridades dos alunos, o qual, segundo a revisão da literatura, constatou-se que:

Graduação: Em nove artigos analisados [Yepes and Fiorin 2023, Jošt et al. 2024, Xue et al. 2024, Lyu et al. 2024, e Kiesler and Schiffner 2023] observou-se que os testes de uso das LLMs foram realizados com alunos do ensino superior. Esses testes foram conduzidos através de atividades e exercícios envolvendo materiais como Estrutura de Dados, programação funcional, casos de testes, coleta de dados e comparação onde neste, o aluno visualizava uma explicação de um determinado código, explicava-o em voz alta para o professor e, em seguida, comparava com a explicação fornecida pelo ChatGPT para o mesmo código.

Ensino médio: Em dois artigos analisados [Pereira Filho et al. 2023, da Silva Junior et al. 2023] os testes de uso das LLMs foram realizados com alunos do ensino médio. Esses testes foram conduzidos por meio de programação básica, que envolvia a

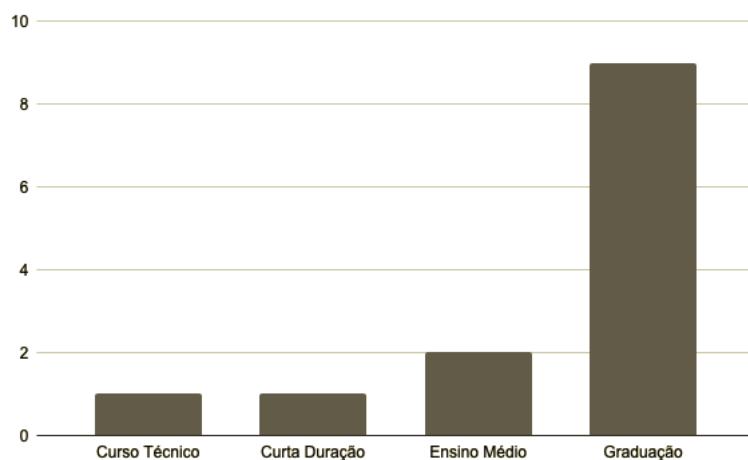


Figura 1. Nível de escolaridade dos alunos.

demonstração de habilidades de criação, correção e explicação de códigos. Além disso, foram empregadas abordagens quali-quantitativas, utilizando métodos qualitativos e quantitativos para avaliar os resultados dos testes de programação.

Curso técnico: Em um artigo [de Araujo Costa et al.] analisado, o teste do uso das LLMs foram realizados com alunos de curso técnico, no qual o teste foi aplicado por meio de testes qualitativos.

Curso de curta duração: Geng et al. [2023] testou o uso das LLMs com alunos em um curso de curta duração. O teste consistia em pesquisas anônimas com código dos exercícios, que cobriram conceitos básicos da linguagem de programação Java.

4.2. RQ₂ - Quais são as principais linguagens de programação ensinadas usando LLMs?

A Figura 2 apresenta as diferentes linguagens de programação ensinadas usando LLMs, segundo a revisão da literatura. A Linguagem C foi mencionada em 2 artigos [Pereira Filho et al. 2023, Yepes and Fiorin 2023]. Enquanto a linguagem Python foi a mais recorrente, mencionada em outros 4 artigos [Leinonen et al. 2023, Geng et al. 2023, Lyu et al. 2024, Kiesler and Schiffner 2023]. Java apareceu em 2 artigos [Leinonen et al. 2023, Geng et al. 2023], e apenas, Jošt et al. (2024) citou HTML/CSS/JavaScript. Em contraste, 6 outros artigos não especificaram as linguagens de programação [da Silva Junior et al. 2023, de Araujo Costa et al., Geng et al. 2023, Leinonen et al. 2023, Silva et al. 2024, Xue et al. 2024].

4.3. RQ₃ - Quais são as principais LLMs citadas na literatura para o ensino de linguagem de programação?

A Figura 3 apresenta os diversos tipos de LLMs encontradas nas literaturas. O ChatGPT foi abordado em 8 artigos [Pereira Filho et al. 2023]-[da Silva Junior et al. 2023, Geng et al. 2023], [Silva et al. 2024]-[Geng et al. 2023, Kiesler and Schiffner 2023]. O CodeTutor foi abordado em 1 artigo [Lyu et al. 2024]. O GPT-3 foi citado em 1 artigo [Leinonen et al. 2023], enquanto o GPT-4 foi citado em 2

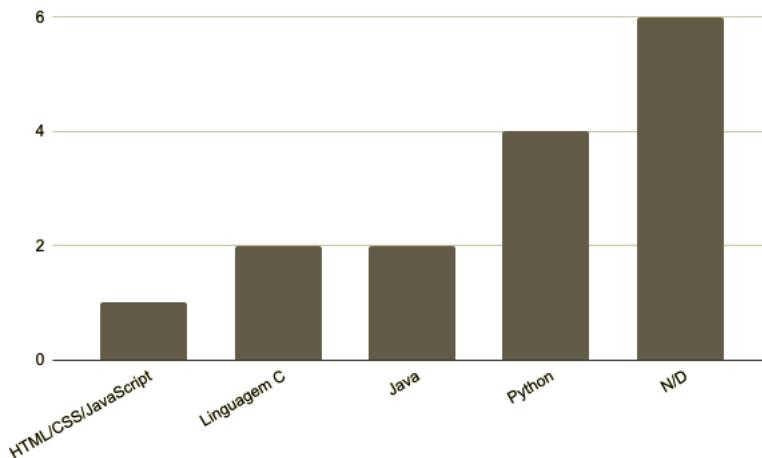


Figura 2. Linguagens de programação utilizadas.

artigos [Leinonen et al. 2023, Kiesler and Schiffner 2023]. Além de 2 artigos [de Araujo Costa et al. , Jošt et al. 2024] não especificaram qual LLM usaram.

4.4. RQ₄ - Quais os benefícios ao usarem LLMs no curso de linguagem de programação?

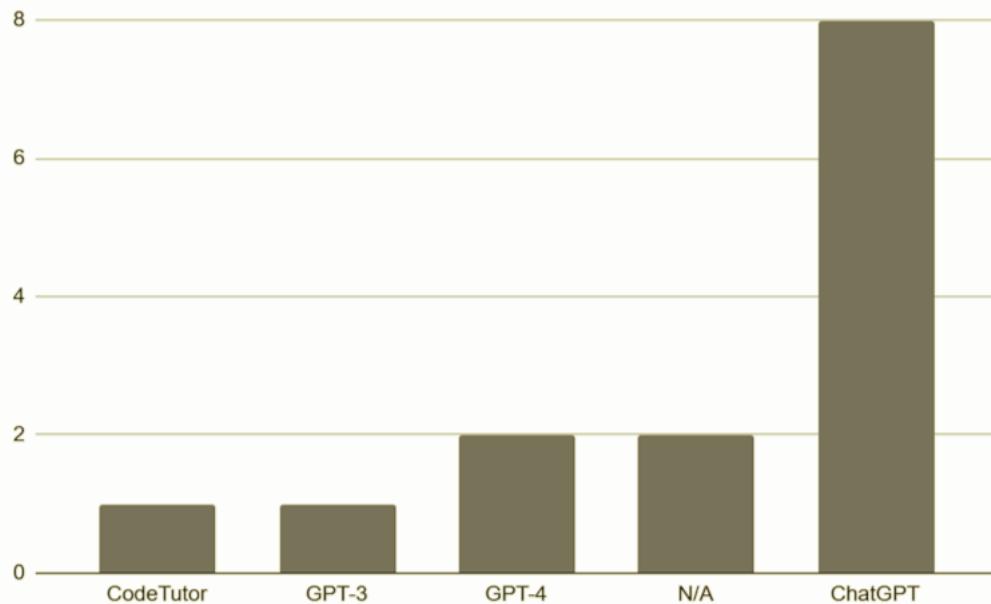
A utilização das LLMs no ensino de programação tem mostrado benefícios, conforme mostrado nos artigos acadêmicos analisados. Entre os principais benefícios, destacam-se:

Verificação da abordagem das respostas pelos alunos: Pereira Filho et al. [2023] destacou a importância de não apenas verificar a precisão das respostas, mas também o entendimento do código pelos alunos. Isso envolve entender o raciocínio e a lógica aplicada na resolução do problema, promovendo uma aprendizagem mais completa.

Verificação da precisão das respostas pelos alunos: Em três artigos [Leinonen et al. 2023, Silva et al. 2024, Xue et al. 2024] analisados foram observados que o uso das LLMs permitem uma verificação mais precisa das respostas dos alunos. Por exemplo, as LLMs ajudaram a identificar erros com maior rapidez e precisão, proporcionando *feedback* imediato e direcionado, o que facilitam o aprendizado e a correção de equívocos com mais agilidade.

Auxílio no ensino de programação: Cinco artigos [da Silva Junior et al. 2023, de Araujo Costa et al., Jošt et al. 2024, Geng et al. 2023, e Xue et al. 2024] demonstraram que as LLMs oferecem um suporte no ensino de programação. Onde atuam auxiliando tanto os professores quanto os alunos, fornecendo recursos didáticos adicionais, que melhoraram a eficiência e a eficácia do processo de ensino-aprendizagem.

Melhora no ensino de programação: Quatro artigos [Yepes and Fiorin 2023, Kumar and Lan 2024, Geng et al. 2023 e, Kiesler and Schiffner 2023] relataram uma melhoria significativa no ensino de programação com o uso das LLMs. As melhorias incluem um maior engajamento dos alunos, melhores resultados acadêmicos, uma maior retenção de conhecimento e um desenvolvimento de habilidades práticas mais sólido.

**Figura 3. LLMs utilizadas.**

4.5. RQ₅ - Quais os impactos ao usarem LLMs no curso de linguagem de programação?

Segundo a revisão da literatura, quais foram os impactos da utilização das LLMs no ensino de linguagem de programação? Entre os principais impactos, destacam-se:

Geração de códigos incorretos: Em seis artigos (Pereira Filho et al. 2023, Yepes and Fiorin 2023, da Silva Junior et al. 2023, Jošt et al. 2024, Geng et al. 2023, e Kiesler and Schiffner 2023) indicam que as LLMs podem gerar códigos incorretos e, em alguns casos, apresentá-los de maneira convincente, levando os alunos a acreditarem que estão corretos. Este fenômeno pode impactar negativamente iniciantes na programação que, devido a falta de conhecimento aprofundado, podem ter dificuldades em compreender a correção das respostas fornecidas pelos modelos.

Dependência dos alunos nas LLMs: Em quatro artigos (Pereira Filho et al. 2023, da Silva Junior et al. 2023, Silva et al. 2024, Xue et al. 2024) discutem que os alunos podem desenvolver uma dependência excessiva das LLMs, o que pode prejudicar o desenvolvimento de suas habilidades de aprendizado e resolução de problemas de programação. Além disso, a confiança total nas LLMs pode levar os alunos a não buscar esclarecer dúvidas com os professores em sala de aula.

Geração de plágio: Em dois artigos [Jošt et al. 2024, Lyu et al. 2024] discutem que as LLMs podem facilitar o plágio, pois os softwares de detecção de plágio frequentemente não conseguem identificar códigos gerados por essas LLMs. Quando esses códigos, que podem se equiparar aos códigos de outros sem a devida citação ou atribuição, podem ser utilizados para fins de avaliação e validação de avaliações acadêmicas. O que dificulta a distinção entre o trabalho de um aluno e o de uma LLM.

Respostas erradas por falta de clareza na pergunta: Kieseler and Schiffner [2023] argumentam que os erros nas respostas geradas pelas LLMs podem ser atribuídos à falta de clareza na descrição do *input* fornecido por um aluno, para uma tarefa. Esta falta de clareza, por sua vez, resulta na geração de códigos incorretos, uma vez que as LLMs podem fornecer soluções para problemas diferentes daqueles pretendidos.

Uso irresponsável das LLMs: Silva et al. [2024] É possível que as LLMs perpetuem preconceitos ou reforcem estereótipos, caso não sejam adequadamente monitoradas.

5. Discussão

Como foi discutido ao longo deste artigo, o uso das LLMs têm sido utilizadas no ensino de programação. Apesar do crescente uso, ainda há lacunas na compreensão do impacto das LLMs no ensino de programação. Tanto no desempenho dos alunos quanto nos métodos educacionais adotados pelos professores. O objetivo desse artigo é pesquisar os impactos das LLMs no ensino de programação e, para isso houve a formulação de perguntas para ajudar na pesquisa:

Os benefícios da utilização das LLMs vão desde a verificação de precisão de respostas e entendimento dos códigos pelos alunos. Além de auxiliar os alunos em seus códigos em verificação de erros com rapidez e precisão. Gerando um engajamento e habilidades práticas mais sólidas. Com os educadores, as LLMs, atuam como um auxílio dentro de sala de aula, com fornecimento de recursos didáticos adicionais

Os impactos das LLMs no ensino de programação são a geração de códigos incompletos e, em alguns casos, podem gerar códigos errados [Nunes et al. 2025], mas de maneira convincente, fazendo os alunos acreditarem que o código está correto e consequentemente atrapalhando o aprendizado. Paralelo a isso, tem a falta de clareza nas perguntas feitas pelos alunos, que consequentemente as LLMs podem fornecer soluções para um problema, bem distintos daqueles pretendidos pelos alunos.

Além de gerar dependência nos alunos, que cria uma confiança total nas LLMs, levando o mesmo a não esclarecer as dúvidas com um educador, por exemplo. A questão do plágio é um impacto bastante relevante, pois há dificuldades em distinguir o código de um aluno do código gerado por uma LLM. Outro impacto que as LLMs podem causar no ensino dos alunos em programação, se não forem usadas de maneira consciente ou sem monitoramento de um educador, é perpetuar preconceitos ou/ e reforçar estereótipos [Rodrigues et al. 2018].

6. Limitações

Durante a revisão da literatura, a *string* de busca resultou apenas em artigos que citam o ChatGPT, o que desconsidera outras LLMs que também têm aplicações no ensino de programação, levando acidentalmente à inclusão de um artigo que aborda exclusivamente o ChatGPT. Sendo que o intuito deste artigo é falar sobre as LLMs de modo geral. O intuito do artigo é saber como as LLMs impactam no ensino de programação, sem priorizar um modelo específico. Porém, como a *string* de busca só levou a artigos que citam o ChatGPT, acabou prejudicando o intuito do artigo, o que abona a diversidade e ampliação

do tema proposto. Para alinhar o artigo ao tema proposto, é preciso revisar ou criar uma nova estratégia de busca, com tese que abrange outras LLMs, garantindo um estudo amplo. Com essa revisão a finalidade do artigo é assegurada em uma análise imparcial e representativa de todas as LLMs e seus reflexos de uso no ensino de programação.

7. Considerações Finais

Essa revisão da literatura fez uma investigação sobre o impacto das LLMs no aprendizado de ensino de programação. Como resultado vimos que as LLMs podem ter impactos positivos e negativos, além de auxiliar docentes e alunos no ensino de programação, mas tem que ter responsabilidade com seu uso e cuidados com o plágio. Futuramente, pretende-se realizar pesquisas em unidades de ensino, sejam elas em nível fundamental, médio, superior ou técnico, aplicando testes e entrevistas com os alunos para observar de perto como as LLMs impactam o ensino de programação.

AGRADECIMENTOS

Esta pesquisa recebeu apoio financeiro das seguintes agências de pesquisa: CNPq (312920/2021-0, PROFIX-JD 155774/2023-9 e 157416/2024-0), CAPES, and FAPEMIG (BPD-00460-22 and APQ-01488-24).

Referências

- Carbonell, J. G., Michalski, R. S., and Mitchell, T. M. (1983). An overview of machine learning. *Machine learning*, pages 3–23.
- da Silva Junior, S. M., de Freitas, R. A. B., de Moraes, M. A. C., and Costa, D. L. V. (2023). Chatgpt no auxílio da aprendizagem de programação: Um estudo de caso. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 1375–1384. SBC.
- de Araujo Costa, D. B., Knupp, D. R., de Souza Alves, C., Júnior, W. d. S. A., Correa, M. E. L., and Cupti, M. d. C. G. Inteligência artificial na educação.
- Geng, C., Yihan, Z., Pientka, B., and Si, X. (2023). Can chatgpt pass an introductory level functional language programming course? *arXiv preprint arXiv:2305.02230*.
- Jošt, G., Taneski, V., and Karakatič, S. (2024). The impact of large language models on programming education and student learning outcomes. *Applied Sciences*, 14(10):4115.
- Keele, S. et al. (2007). Guidelines for performing systematic literature reviews in software engineering. Technical report, Technical report, ver. 2.3 ebse technical report. ebse.
- Kiesler, N. and Schiffner, D. (2023). Large language models in introductory programming education: Chatgpt’s performance and implications for assessments. *arXiv preprint arXiv:2308.08572*.
- Kitchenham, B. (2006). Evidence-based software engineering and systematic literature reviews. In *Product-Focused Software Process Improvement: 7th International Conference, PROFES 2006, Amsterdam, The Netherlands, June 12-14, 2006. Proceedings* 7, pages 3–3. Springer.

- Kumar, N. A. and Lan, A. (2024). Using large language models for student-code guided test case generation in computer science education. *arXiv preprint arXiv:2402.07081*.
- Leinonen, J., Denny, P., MacNeil, S., Sarsa, S., Bernstein, S., Kim, J., Tran, A., and Hellas, A. (2023). Comparing code explanations created by students and large language models. In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, pages 124–130.
- Lyu, W., Wang, Y., Chung, T., Sun, Y., and Zhang, Y. (2024). Evaluating the effectiveness of llms in introductory computer science education: A semester-long field study. In *Proceedings of the Eleventh ACM Conference on Learning@ Scale*, pages 63–74.
- Nunes, R., Figueiredo, E., Soares, L., Nadi, S., Ferreira, F., and Santos, G. (2025). Evaluating the effectiveness of llms in fixing maintainability issues in real-world projects. In *IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*.
- Pereira Filho, L. C., de Souza, T. d. P., and de Paula, L. B. (2023). Análise das respostas do chatgpt em relação ao conteúdo de programação para iniciantes. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 1738–1748. SBC.
- Rodrigues, P., Souza, M., and Figueiredo, E. (2018). Games and gamification in software engineering education: A survey with educators. In *IEEE Frontiers in Education Conference (FIE)*, pages 1–9.
- Santos, A., Souza, M., Dayrell, M., and Figueiredo, E. (2019). A systematic mapping study on game elements and serious games for learning programming. *Communications in Computer and Information Science (CCIS)*, 1022.
- Silva, C. A. G. d., Ramos, F. N., de Moraes, R. V., and Santos, E. L. d. (2024). Chatgpt: Challenges and benefits in software programming for higher education. *Sustainability*, 16(3):1245.
- Souza, M., Veado, L., Moreira, R. T., Figueiredo, E., and Costa, H. (2018). A systematic mapping study on game-related methods for software engineering education. *Information and Software Technology*, 95:201–218.
- Xue, Y., Chen, H., Bai, G. R., Tairas, R., and Huang, Y. (2024). Does chatgpt help with introductory programming? an experiment of students using chatgpt in cs1. In *Proceedings of the 46th International Conference on Software Engineering: Software Engineering Education and Training*, pages 331–341.
- Yepes, I. and Fiorin, A. (2023). Chatgpt como assistente de ensino na disciplina de estruturas de dados. *Anais do Encontro Anual de Tecnologia da Informação*, 12(1):36–36.