

Avaliação da Aprendizagem em Programação com IA Generativa na Educação Profissional e Tecnológica

Paulo César Rodacki Gomes¹, Jomi Fred Hübner²

¹ Campus Blumenau – Instituto Federal Catarinense (IFC)
Blumenau – SC – Brasil

²Departamento de Automação e Sistemas
Universidade Federal de Santa Catarina (UFSC) – Florianópolis, SC – Brasil

paulo.gomes@ifc.edu.br, jomi.hubner@ufsc.br

Abstract. *This article presents a generative artificial intelligence (GenAI)–mediated assessment model for programming education in Vocational and Technological Education (VTE). Grounded in principles of formative assessment and self-regulated learning, the AURA^{IA} model integrates epistemological personalization, automated feedback, and teacher mediation through large language models (LLMs), learning management systems, and virtual tutoring. Anchored in students’ actual work, the AURA^{IA} prototype comprises a conceptual model and a functional implementation currently undergoing preliminary validation with graduate students and faculty members.*

Resumo. *Este artigo apresenta um modelo avaliativo mediado por Inteligência Artificial Generativa (GenAI) para o ensino de programação na Educação Profissional e Tecnológica (EPT). Fundamentado em princípios de avaliação formativa e autorregulação da aprendizagem, o modelo AURA^{IA} integra personalização epistemológica, feedback automatizado e mediação docente por meio de LLMs, sistemas de gestão de aprendizagem e tutoria virtual. A avaliação é ancorada nas produções reais dos estudantes. Já foram elaborados o modelo conceitual e seu protótipo funcional, ambos em fase de validação preliminar junto a docentes e discentes de pós-graduação.*

1. Introdução

A emergência da Inteligência Artificial Generativa (GenAI), especialmente com o lançamento público de ferramentas baseadas em LLMs como o ChatGPT em novembro de 2022, tem provocado profundas transformações em diversas esferas sociais, destacando-se seus impactos na educação [Silvestre et al. 2023].

No âmbito da Educação Profissional e Tecnológica (EPT), a GenAI apresenta desafios específicos. Por um lado, é vista como instrumento promissor para apoiar a aprendizagem e desenvolver habilidades técnicas exigidas pelo mercado de trabalho contemporâneo. Por outro, aprofundam-se dilemas éticos, pedagógicos e avaliativos que questionam paradigmas educacionais consolidados. Essas tensões requerem novas formas de compreender o papel da tecnologia na formação humana, motivando pesquisadores e educadores a revisitar os fundamentos epistemológicos e pedagógicos que orientam os processos de ensino e aprendizagem na EPT.

No campo de avaliação da aprendizagem, a disseminação da GenAI baseada em LLMs dificulta a aplicação de métodos tradicionais como provas, trabalhos de casa e listas de exercícios, pois possibilita a simulação de desempenho acadêmico sem efetiva consolidação real da aprendizagem.

No ensino de programação, os desafios agravam-se à medida que os estudantes podem automatizar soluções sem compreender conceitos subjacentes, prejudicando o desenvolvimento de competências essenciais, como raciocínio lógico, depuração e explicação do próprio código [Deus et al. 2024]. Isso evidencia a falta de oferta de modelos avaliativos que incorporem GenAI de forma pedagógica. Apesar da expansão das ferramentas de GenAI no âmbito educacional, ainda são raras as iniciativas que alinhem essa tecnologia a objetivos formativos claros e a uma mediação docente crítica [Bido et al. 2024].

Este artigo apresenta os resultados parciais de uma pesquisa em andamento cujo objetivo é desenvolver e validar um modelo avaliativo mediado por GenAI capaz de responder aos desafios emergentes da avaliação no ensino de programação na EPT. O modelo visa combinar automação e mediação pedagógica, promovendo personalização, autorregulação e *feedback* formativo a partir da análise das produções reais dos estudantes. Esse tipo de abordagem, que busca alinhar os instrumentos avaliativos aos modos próprios de pensar, resolver e expressar conhecimento de cada estudante, pode ser compreendido como personalização epistemológica. Ao mesmo tempo, busca-se preservar a centralidade do professor no processo formativo.

2. Fundamentação

As transformações trazidas pela GenAI na educação têm suscitado reflexões sobre as estruturas conceituais que sustentam as práticas de ensino e avaliação. Neste debate, destaca-se a discussão em torno dos paradigmas que orientam as concepções de aprendizagem e comunicação na educação. Griffiths et al. [2024] analisam esses impactos sob a ótica de dois paradigmas contrastantes: transmissão e coordenação. Cada um desses paradigmas combina modelos de aprendizagem com modelos de comunicação.

O paradigma da transmissão vê a educação como um fluxo linear de conhecimento [Shannon and Weaver 1949] do professor para o estudante, com a aprendizagem aferida por instrumentos padronizados, como provas e trabalhos, que avaliam a capacidade dos estudantes em reproduzir conteúdos previamente transmitidos. Já o paradigma da coordenação fundamenta-se no construtivismo e na comunicação não linear, que introduz o conceito de acoplamento estrutural como base da comunicação humana [Maturana and Varela 2011]: a aprendizagem emerge de interações dialógicas entre estudantes, professores e contexto, aproximando-se de abordagens críticas como as de Freire [2014], que enfatizam o protagonismo do estudante.

Os dois paradigmas coexistem nas instituições de ensino em um equilíbrio frágil, mas funcional. De um lado, os instrumentos organizacionais como currículos, avaliações padronizadas e planos de ensino refletem a lógica da transmissão; de outro, muitos professores adotam abordagens mais interativas e centradas no estudante, inspiradas no paradigma da coordenação. Essa convivência constituiu ao longo do tempo um *modus vivendi* que permitiu a continuidade das atividades educativas sem a necessidade de resolver plenamente a tensão entre os dois modelos. No entanto, a emergência da GenAI ameaça romper esse equilíbrio [Griffiths et al. 2024].

Apesar de o paradigma da coordenação ser valorizado na educação brasileira, a transmissão permanece essencial em disciplinas como programação, pois a exposição direta de conceitos complexos (fundamentos, algoritmos, estruturas de dados) sustenta aprendizagens iniciais. Estudos mostram que essa instrução direta facilita a compreensão de conteúdos difíceis e reduz a evasão em cursos introdutórios [Hermans and Smit 2018, Stockard et al. 2018]. Adotar o paradigma da transmissão não exclui práticas construtivistas. No ensino técnico, uma abordagem híbrida pode ser eficaz: conceitos são apresentados diretamente e depois consolidados por atividades exploratórias, favorecendo compreensão e engajamento.

Os LLMs conseguem gerar textos plausíveis e resolver problemas com elevada competência aparente, criando um simulacro da inteligência humana. Esse fato impõe um paradoxo aos instrumentos avaliativos ancorados no paradigma de transmissão devido à dificuldade de distinguir produções autênticas de respostas artificiais. Porém, em lugar de abandonar a transmissão, é preciso aprimorá-la tanto no desenho de instrumentos avaliativos quanto na articulação com abordagens de coordenação que contextualizem melhor o desempenho dos estudantes.

Embora Griffiths et al. [2024] proponham adotar o paradigma da coordenação para integrar a GenAI ao ensino de forma mais ética e eficaz, argumenta-se que o real desafio está em reconfigurar o equilíbrio entre os paradigmas. Isso envolve não apenas valorizar abordagens construtivistas, mas também adaptar o paradigma da transmissão, ainda essencial em muitos contextos. A Figura 1 ilustra esse contraste e a proposta de coexistência diante do avanço da GenAI.

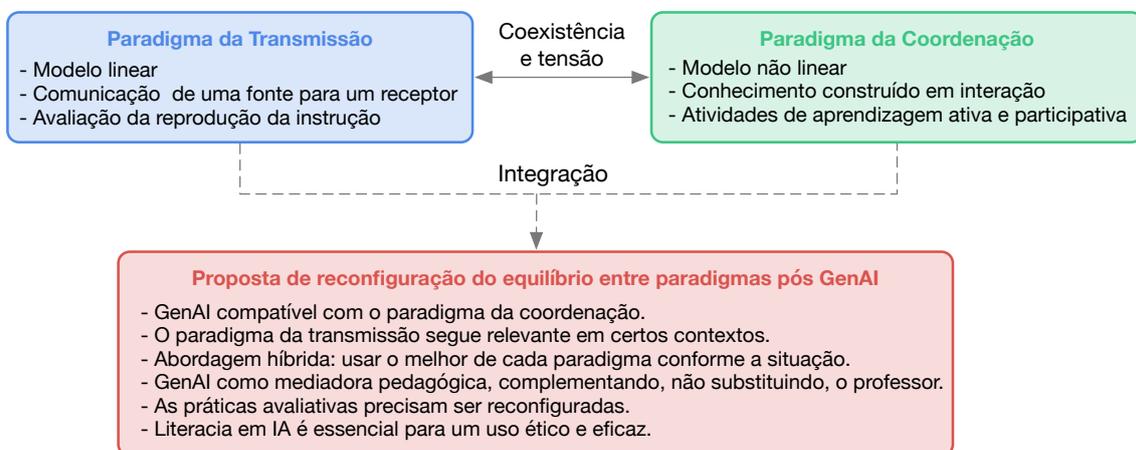


Figura 1. Reconfiguração de paradigmas pós GenAI

Com a crescente presença de GenAI nos contextos educacionais, a literacia em IA vem sendo reconhecida como uma competência fundamental para educadores, incluindo a literacia em avaliação no contexto de GenAI [Sousa and Cruz 2024]. Mais do que atualizar instrumentos avaliativos, trata-se de repensar criticamente os fundamentos que sustentam as práticas de avaliação da aprendizagem, incluindo o uso ético de ferramentas como os LLMs, a construção de critérios adequados de julgamento e a capacidade de interpretar evidências de aprendizagem em ambientes mediados por tecnologias [Moorhouse and Wong 2023], [Bower et al. 2024]. Essa reconfiguração implica não apenas domínio técnico, mas também uma mudança nos modelos pedagógicos, com in-

centivo à autonomia discente por meio de práticas como a autoavaliação e a avaliação por pares com suporte de IA [Saini et al. 2024], [Xia et al. 2024]. Além das oportunidades, esse novo cenário traz desafios importantes no que se refere à integridade acadêmica e à transparência nos processos avaliativos [Currie 2023], [Kilinç 2024].

No caso do ensino de programação de computadores, o desempenho do estudante em exercícios práticos fornece insumos valiosos sobre sua compreensão conceitual e suas estratégias de resolução de problemas [Lishinski et al. 2016]. A partir da análise dessas resoluções, é possível utilizar ferramentas de IA para gerar avaliações personalizadas, ajustadas ao histórico individual de respostas de cada estudante. Essa abordagem permite construir instrumentos avaliativos que dialogam diretamente com os desafios enfrentados por cada estudante, favorecendo o diagnóstico mais preciso de suas lacunas e avanços. Ao mesmo tempo, o uso crescente de ferramentas de GenAI pelos estudantes levanta desafios específicos para o ensino de programação, tornando-se difícil assegurar a autoria das soluções entregues nos exercícios, o que compromete a eficácia das avaliações tradicionais baseadas apenas no produto final. Nesse contexto, avaliações personalizadas que se originam a partir das próprias produções anteriores dos estudantes podem oferecer uma estratégia mais robusta. Essas discussões fornecem os fundamentos para o modelo de avaliação personalizada mediada por IA que será apresentado na próxima seção, com foco no ensino de programação de computadores.

3. O modelo AURA^{IA}

O modelo AURA^{IA} (*Avaliação com Uso de Recursos de Aprendizagem mediados por Inteligência Artificial*) propõe uma estrutura conceitual para a avaliação da aprendizagem em cursos técnicos de programação. Desenvolvido no contexto da EPT, o modelo aborda o desafio emergente de verificar a aprendizagem em cenários com uso crescente de LLMs pelos estudantes.

Partindo da concepção da avaliação como processo formativo, contínuo e orientado por dados, o modelo busca promover a autorregulação da aprendizagem e o desenvolvimento de competências técnicas com suporte pedagógico personalizado. Para isso, o modelo integra sistemas de gestão da aprendizagem (LMS), ferramentas de GenAI e mediação docente intencional, formando um fluxo que coleta, interpreta e retroalimenta evidências de aprendizagem com base nas interações prévias dos estudantes.

Seu principal diferencial está na geração de instrumentos avaliativos individualizados, derivados das próprias produções dos estudantes, como códigos-fonte submetidos em atividades práticas. A partir desses dados, a IA cria questões personalizadas que demandam interpretação, explicação ou modificação dos resultados anteriores. Dessa forma, pretende-se desestimular práticas de uso indiscriminado da GenAI e favorecer o desenvolvimento de uma atitude mais autônoma, ética e reflexiva no percurso formativo do estudante.

3.1. Estrutura funcional

Do ponto de vista funcional, o modelo é estruturado como um pipeline composto por etapas operacionais que se encadeiam em fluxo cíclico e adaptativo. Cada etapa recebe entradas da anterior e produz saídas para a seguinte, conforme descrito a seguir:

1. **Seleção de conteúdo:** o professor seleciona atividades previamente realizadas pelos estudantes em uma plataforma de gestão de aprendizagem, como o Google Classroom ou Moodle. Essas atividades podem incluir códigos-fonte, respostas discursivas, resolução de problemas, entre outros tipos de produção. A seleção define a base de referência sobre a qual serão geradas as avaliações personalizadas.
2. **Geração da avaliação:** a partir do conteúdo selecionado, o sistema utiliza LLMs para produzir instrumentos avaliativos individualizados. As questões geradas estão diretamente relacionadas às produções anteriores de cada estudante e podem incluir questões objetivas, discursivas ou verdadeiro/falso, com diferentes níveis de complexidade. Essa etapa garante que os itens de avaliação reflitam o histórico e o contexto específico de cada estudante.
3. **Aplicação da avaliação:** os instrumentos avaliativos podem ser disponibilizados em formato digital ou impressos, conforme as possibilidades técnicas e pedagógicas da instituição. A aplicação deve ocorrer preferencialmente de maneira síncrona (em sala de aula ou laboratório).
4. **Correção e análise:** uma vez submetidas as respostas, o sistema realiza a correção automatizada das questões com base em critérios definidos durante a geração do item. Além da atribuição de notas, o sistema produz relatórios individuais de desempenho, destacando acertos, erros e padrões de resposta relevantes para o diagnóstico pedagógico. O relatório pode ser revisado, editado e aprovado pelo professor.
5. **Interação com tutor virtual:** cada estudante acessa um ambiente de *feedback* para dialogar em chat com um tutor virtual baseado em GenAI. Com base no laudo pré-gerado pela IA e validado pelo professor, o tutor fornece explicações, esclarece dúvidas, solicita reinterpretações, estimulando a metacognição e a autorregulação da aprendizagem. Ao acompanhar o histórico do estudante, o sistema pode identificar lacunas e propor estratégias de estudo personalizadas. Por exemplo, ao detectar dificuldades recorrentes com estruturas condicionais, pode priorizar questões que exijam explicar ou corrigir trechos de código com esse tipo de estrutura. Futuramente, o modelo pode vir a permitir sessões multiusuário com participação simultânea de estudante, tutor e professor em um mesmo chat.

Esse fluxo de cinco etapas constitui a espinha dorsal operacional do modelo AURA^{IA}, promovendo uma avaliação contínua, personalizada e pedagógica orientada por dados. O ciclo pode ser iterado ao longo do processo formativo, permitindo que o sistema evolua em resposta ao progresso individual dos estudantes e às intervenções do professor. A Figura 2 ilustra a interação dos atores e componentes em torno do modelo.

3.2. Estrutura pedagógica

Do ponto de vista pedagógico, o modelo fundamenta-se nos princípios da avaliação formativa, da personalização da aprendizagem e da autorregulação. A personalização ocorre tanto na geração de instrumentos avaliativos quanto na estruturação do *feedback*: o tutor virtual interage com o estudante com base em seu histórico, oferecendo comentários explicativos e orientações individualizadas. O professor mantém papel central, validando sugestões da IA, acompanhando diagnósticos e intervindo com estratégias complementares. A arquitetura contempla dois modos de operação: supervisionado, com participação ativa do professor em todas as etapas; e semiautônomo, com maior automação e relatórios

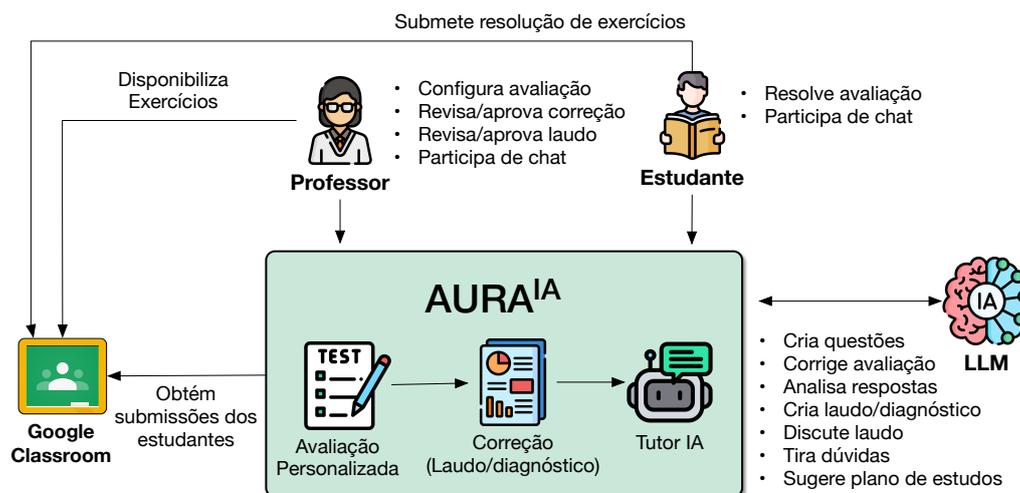


Figura 2. Arquitetura funcional do modelo AURA^{IA}

disponíveis para revisão docente. Essa flexibilidade favorece a adaptação a distintos contextos institucionais e estilos pedagógicos.

A articulação entre automação e mediação docente no modelo AURA^{IA} está alinhada com a abordagem Hybrid Human-AI Regulation (HHAIR) [Sardi et al. 2025], que defende a integração do potencial adaptativo da IA com a atuação ética e contextual do professor. Essa concepção reforça a importância da regulação humana para fomentar autorregulação e pensamento crítico [Wang 2024]. O modelo, portanto, não substitui a docência, mas a auxilia, oferecendo suporte diagnóstico e personalizado em ambientes educacionais frequentemente marcados por restrições de tempo e recursos.

4. Implementação Inicial

Até a submissão deste artigo, implementou-se um protótipo inicial do modelo AURA^{IA} para testes, validação e apresentação a estudantes, professores e pesquisadores interessados. O protótipo executa as etapas 1 e 2 da seção 3.1, sendo desenvolvido como aplicação web com *backend* em Python e *frontend* em Vue.js com TypeScript. O protótipo faz uso de APIs do Google Cloud para autenticação e acesso a dados do Google Classroom, além da API Python da OpenAI para interação com o GPT-4o.

Dadas as credenciais de login do professor, o sistema acessa suas turmas, exercícios previamente propostos e respectivas resoluções em forma de código-fonte submetidas pelos estudantes. Com isso, o professor pode configurar uma nova avaliação no AURA^{IA} selecionando a turma, o sub-conjunto de estudantes da turma, os assuntos e a quantidade de questões que devem compor a avaliação. Para cada estudante, o protótipo recupera aleatoriamente submissões de autoria do estudante e encaminha ao LLM requisições remotas contendo o código-fonte do estudante e *prompt* com contextualização sobre o conteúdo da disciplina (neste caso, disciplina de Lógica de Programação no primeiro ano de Curso Técnico em Informática Integrado ao Ensino Médio). O *prompt* inclui solicitação para criação de três questões baseadas no código fonte enviado: uma questão objetiva de múltipla escolha, uma questão do tipo verdadeiro/falso e uma questão dissertativa. As respostas do LLM são recebidas de forma estruturada em JSON, formatadas pela API OpenAI por meio de esquemas Pydantic, o que

garante uniformidade na estruturação das questões geradas. O professor realiza a curadoria das questões antes de liberá-las para aplicação, podendo excluir, ajustar ou reprovar automaticamente qualquer item gerado pela IA. O fragmento de código a seguir mostra um exemplo simples de requisição ao LLM. `Conjuntoquestoes` é uma classe Pydantic com as estruturas de formatação dos três tipos de questões e `source_code` é uma string contendo o código fonte de autoria do estudante, recuperado do Google Classroom.

```

1 completion = openai.beta.chat.completions.parse(
2     model="gpt-4o-2024-08-06",
3     messages=[
4         {"role": "system", "content": "Gere três questões com base no
5           código-fonte do estudante:\n"
6           "1) Uma questão objetiva (múltipla escolha).\n"
7           "2) Uma questão de verdadeiro ou falso.\n"
8           "3) Uma questão subjetiva (discursiva)."},
9         {"role": "user", "content": f"Aqui está o código-fonte do
10          estudante:\n{source_code}"},
11     ],
12     response_format=ConjuntoQuestoes,
13 )

```

É importante notar que as funcionalidades do protótipo ainda são bastante limitadas, e as próximas etapas da pesquisa preveem a implementação completa do modelo, bem como sua disponibilização para professores, visando testes e validação em contextos reais de ensino. Além disso, ainda há aspectos técnicos a explorar: testes com diferentes *prompts*; uso de LLMs locais para lidar com privacidade, custo e customização via *Retrieval-Augmented Generation* (RAG); melhorias na configuração das avaliações e adaptação a outras áreas do ensino técnico.

5. Trabalhos Relacionados

Diversas iniciativas recentes exploram o uso de GenAI para apoiar práticas educacionais em programação. Schacht, Barkur e Lanquillon [2023] apresentam um modelo que gera avaliações com base em materiais instrucionais, como livros e slides. Sarsa et al. [2022] mostram o potencial dos LLMs na criação de exercícios, soluções-modelo e explicações de código, contribuindo para material didático. Pesovski et al. [2024] incorporam GenAI a um LMS para personalizar conteúdos conforme objetivos dos docentes, observando maior engajamento dos estudantes. Logacheva et al. [2024] usam o GPT-4 para adaptar exercícios às preferências temáticas dos estudantes, promovendo maior motivação em contextos introdutórios. Por fim, Hu et al. [2025] desenvolvem o Socratic Playground, um tutor conversacional que promove o pensamento crítico com base em cinco modos interativos de aprendizagem. Apesar de fornecer feedback adaptativo, sua ênfase está no diálogo e na tutoria, e não em processos avaliativos estruturados.

O modelo AURA^{IA} busca preencher uma lacuna ao articular avaliação formativa, personalização epistemológica e mediação docente qualificada na EPT. Seu diferencial está na geração de instrumentos avaliativos a partir das produções reais dos estudantes, criando questões alinhadas à trajetória individual de aprendizagem. Ao combinar correção automatizada, tutoria virtual e estratégias de reforço personalizadas, o modelo integra avanços da GenAI sem perder a intencionalidade pedagógica nem a centralidade do professor.

6. Contribuições e Resultados Emergentes

O desenvolvimento do modelo AURA^{IA} baseia-se em hipóteses pedagógicas que orientam sua concepção e aplicação. Considera-se que a personalização epistemológica das avaliações, ancorada nas produções anteriores dos estudantes, promova maior engajamento e apropriação dos conteúdos, além de dificultar a delegação automatizada de tarefas à GenAI. Ao valorizar a interpretação, modificação e justificação de soluções autorais, busca-se mitigar o uso acrítico dessas ferramentas. Outro pressuposto é que o *feedback* individualizado, associado ao acompanhamento por tutor virtual, favoreça o desenvolvimento da autorregulação da aprendizagem, estimulando práticas de autoavaliação, planejamento e tomada de decisão. A mediação docente continua essencial para assegurar a intencionalidade pedagógica e a qualidade das devolutivas, ao passo que a presença da GenAI como mediadora da avaliação pode estimular a revisão de práticas tradicionais, fortalecendo uma cultura mais formativa e centrada na aprendizagem.

7. Considerações Finais e Trabalhos Futuros

Este artigo apresentou resultados parciais de uma pesquisa em andamento que busca enfrentar os desafios trazidos pela inteligência artificial generativa (GenAI) à avaliação em disciplinas introdutórias de programação, especialmente no contexto do uso crescente dessas ferramentas por estudantes da EPT. Até o momento, o trabalho envolveu revisão bibliográfica sobre o tema; formulação de um modelo conceitual de avaliação mediada por GenAI; aprendizado e aplicação de tecnologias de integração com LLMs por meio de APIs; implementação de um protótipo funcional; e apresentação preliminar da proposta a discentes e docentes de programas de pós-graduação *stricto sensu*.

A partir dos estudos realizados até aqui, percebe-se uma necessidade premente de revisar metodologias de avaliação, bem como de ensino e aprendizagem, diante da inserção massiva da GenAI na educação. Por isso, entendemos que não é mais adequado manter formas avaliativas dissociadas das transformações tecnológicas já incorporadas ao cotidiano. Ao propor o modelo AURA^{IA}, vislumbramos mudanças significativas no modelo tradicional de avaliação, geralmente centrado no produto final em vez do processo formativo. Nesta abordagem, busca-se avaliar o aprendizado ao longo do percurso, reconhecendo que não basta apresentar uma solução correta, como resoluções produzidas por GenAI, mas é fundamental compreender logicamente como se chega a ela. Espera-se que o modelo AURA^{IA} possa contribuir para essa transformação.

Consideramos consistente a abordagem proposta para qualificar os processos avaliativos no contexto da Educação Profissional e Tecnológica, especialmente em cursos técnicos e de graduação tecnológica voltados à área de computação. No entanto, sua consolidação requer ainda respostas a desafios técnicos, tais como a robustez do sistema, a integração com diferentes plataformas educacionais e o uso de LLMs de código aberto, bem como desafios pedagógicos relacionados à formação docente, à validação dos instrumentos gerados e à adesão às práticas de avaliação formativa mediada por IA. Os próximos passos da pesquisa incluem a ampliação do protótipo, a realização de testes piloto com turmas reais e a coleta e análise de dados empíricos sobre viabilidade técnica, usabilidade e impacto pedagógico, além da produção de publicações que aprofundem aspectos específicos do modelo AURA^{IA} e sua aplicação em diferentes contextos educacionais.

Disponibilidade de Artefatos

O *codebase* do protótipo está em <https://github.com/rodacki/auraia>.

Transparência no uso de IA

Ferramentas de IA (Writefull e GPT-4o) foram utilizadas para revisão textual e organização do código-fonte do protótipo (via Aider). As figuras são de autoria própria, sem uso de IA. As ideias, formulações teóricas e estrutura do artigo são integralmente de responsabilidade dos autores humanos.

Agradecimentos

O autor Paulo César Rodacki Gomes agradece ao Instituto Federal Catarinense (IFC) pelo apoio institucional concedido por meio de afastamento integral para realização de estágio pós-doutoral, conforme o Edital nº 121/2023 – ASSEG/GABI/IFC.

Referências

- Bido, Y., Wiese, I., and Nakamura, W. (2024). IAs generativas na educação: Usos, percepções, desafios e adaptações nas práticas pedagógicas do ponto de vista de professores do ensino fundamental, médio e superior. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação*, pages 1701–1714, Porto Alegre, RS, Brasil. SBC.
- 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(12):15403–15439.
- Currie, G. M. (2023). Academic integrity and artificial intelligence: is chatgpt hype, hero or heresy? *Seminars in Nuclear Medicine*, 53(5):719–730. Preclinical.
- Deus, W., Marcolino, A., Avellar, G., Oliveira, K., and Barbosa, E. (2024). Avaliando resoluções de exercícios introdutórios de programação na era das IAs generativas: Um estudo de caso com o ChatGPT. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação*, pages 1645–1657, Porto Alegre, RS, Brasil. SBC.
- Freire, P. (2014). *Pedagogia da autonomia: Saberes necessários à prática educativa*. Paz e Terra, São Paulo.
- Griffiths, D., Frías-Martínez, E., Tlili, A., and Burgos, D. (2024). A cybernetic perspective on generative AI in education: From transmission to coordination. *International Journal of Interactive Multimedia and Artificial Intelligence*, 8(5):15–24.
- Hermans, F. and Smit, M. (2018). Explicit direct instruction in programming education. In *Proceedings of the 29th Annual Workshop of the Psychology of Programming Interest Group, PPIG 2018, London, UK, September 5 - 7, 2018*, pages 32–41. Psychology of Programming Interest Group.
- Hu, X., Xu, S., Tong, R. J., and Graesser, A. (2025). Generative AI in education: from foundational insights to the Socratic Playground for learning. *ArXiv*, abs/2501.06682.
- Kiliç, S. (2024). Comprehensive AI assessment framework: Enhancing educational evaluation with ethical AI integration. *Journal of Educational Technology and Online Learning*, 7(4 - ICETOL 2024 Special Issue):521–540.

- Lishinski, A., Yadav, A., Enbody, R., and Good, J. (2016). The influence of problem solving abilities on students' performance on different assessment tasks in CS1. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education, SIGCSE '16*, page 329–334, New York, NY, USA. Association for Computing Machinery.
- Logacheva, E., Hellas, A., Prather, J., Sarsa, S., and Leinonen, J. (2024). Evaluating contextually personalized programming exercises created with generative ai. In *Proceedings of the 2024 ACM Conference on International Computing Education Research - Volume 1, ICER '24*, page 95–113, New York, NY, USA. Association for Computing Machinery.
- Maturana, H. R. and Varela, F. J. (2011). *A árvore do conhecimento: as bases biológicas da compreensão humana*. Palas Athena, São Paulo, 9 edition.
- Moorhouse, B. and Wong, K. (2023). Educators' assessment literacy in the age of AI: ethical considerations and pedagogical shifts. *AI and Education Journal*, 2(1):10–25.
- Pesovski, I., Santos, R., Henriques, R., and Trajkovik, V. (2024). Generative AI for customizable learning experiences. *Sustainability*, 16(7):1–23.
- Saini, A. K., Cope, B., Kalantzis, M., and Zapata, G. C. (2024). The future of feedback: Integrating peer and generative AI reviews to support student work.
- Sardi, J., Darmansyah, Candra, O., Yuliana, D. F., Habibullah, Yanto, D. T. P., and Eliza, F. (2025). How generative AI influences students' self-regulated learning and critical thinking skills? a systematic review. *International Journal of Engineering Pedagogy (iJEP)*, 15(1):pp. 94–108.
- Sarsa, S., Denny, P., Hellas, A., and Leinonen, J. (2022). Automatic generation of programming exercises and code explanations using large language models. In *Proceedings of the 2022 ACM Conference on International Computing Education Research - Volume 1, ICER '22*, page 27–43, New York, NY, USA. Association for Computing Machinery.
- Schacht, S., Kamath Barkur, S., and Lanquillon, C. (2023). Generative agents to support students' learning progress. In *Proceedings of the 5th International Conference Business Meets Technology*, pages 179–197, Valencia.
- Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, IL.
- Silvestre, A., Amaral, E., Holanda, M., and Canedo, E. (2023). Students' perception about ChatGPT's impact on their academic education. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 1260–1270, Porto Alegre, RS, Brasil. SBC.
- Sousa, H. and Cruz, D. (2024). Capacitando educadores com IA generativa: Implicações na educação. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação*, pages 1931–1941, Porto Alegre, RS, Brasil. SBC.
- Stockard, J., Wood, T. W., Coughlin, C., and Khoury, C. R. (2018). The effectiveness of direct instruction curricula: A meta-analysis of a half century of research. *Review of Educational Research*, 88(4):479–507.

Wang, Y. (2024). Cognitive and sociocultural dynamics of self-regulated use of machine translation and generative AI tools in academic EFL writing. *System*, 126:103505.

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., and Chiu, T. K. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21(1):40.