



Campus Recife

Departamento Acadêmico dos Cursos Superiores

Curso de Tecnologia em Análise e Desenvolvimento de Sistemas

RICHARDSON SILVA LIMA DOS SANTOS

**Sistema Automatizado de Criação de Avaliações Baseado nos
Descritores do SAEB utilizando LLM**

Recife

2026

RICHARDSON SILVA LIMA DOS SANTOS

**Sistema Automatizado de Criação de Avaliações Baseado nos Descritores do
SAEB utilizando LLM**

Trabalho de Conclusão de Curso apresentado
como requisito parcial para obtenção do título
de Tecnólogo em Análise e Desenvolvimento de
Sistemas, no Instituto Federal de Educação,
Ciência e Tecnologia de Pernambuco (IFPE),
Campus Recife.

Orientador: Prof. Dr. Henrique Correia Torres Santos

Recife

2026

Ficha catalográfica elaborada pela bibliotecária Danielle Castro da Silva CRB4/1457

S237s
2026

Santos, Richardson Silva Lima Dos

Sistema Automatizado de Criação de Avaliações Baseado nos Descritores do SAEB utilizando LLM / Richardson Silva Lima Dos Santos. --- Recife: O autor, 2026.

38f. il. Color.

Trabalho de Conclusão (Curso Superior Tecnológico em Análise e Desenvolvimento de Sistemas) –Instituto Federal de Pernambuco, Recife, 2026.

Inclui Referências.

Orientador: Prof. Dr. Henrique Correia Torres Santos.

1. Inteligência Artificial. 2. Large Language Models. 3. Modelo de Linguagem. 4. Geração de Questões. 5. Avaliação Educacional I. Título. II. Santos, Richardson Silva Lima Dos (orientador). III. Instituto Federal de Pernambuco.

CDD 607

RICHARDSON SILVA LIMA DOS SANTOS

**Sistema Automatizado de Criação de Avaliações Baseado nos Descritores do
SAEB utilizando LLM**

Trabalho de Conclusão de Curso
apresentado, como requisito
parcial para obtenção do título de Tec-
nólogo em Análise e Desenvolvimento de
Sistemas, do Instituto Federal de Educação,
Ciência e Tecnologia de Pernambuco –
Campus Recife.

Aprovado em: 09 de maio de 2026.

Banca Examinadora:

Prof. Dr. Henrique Correia Torres Santos
IFPE - Campus Recife

Prof. Msc. Paulo Abadie Guedes
IFPE - Campus Recife

Prof. Msc. Hilson Gomes Vilar de Andrade
Instituição

Recife

2026

AGRADECIMENTOS

Como o primeiro da minha família a concluir o ensino superior, expresso minha profunda gratidão à minha mãe e minha duas avós, que sempre esteve ao meu lado, me apoiando em tudo. Agradeço também a todos os amigos que fizeram parte dessa caminhada, ao meu orientador, cujo auxílio foi fundamental, e, acima de tudo, a mim mesmo, por ter enfrentado os desafios e perseverado nessa jornada.

RESUMO

A avaliação educacional desempenha papel fundamental na mensuração da aprendizagem e no direcionamento de políticas públicas. No contexto brasileiro, o Sistema de Avaliação da Educação Básica (SAEB) constitui o principal instrumento de avaliação em larga escala, exigindo que as questões aplicadas estejam alinhadas a descritores específicos de competências e habilidades. Entretanto, a elaboração manual dessas avaliações demanda tempo, conhecimento técnico e cuidado metodológico por parte dos docentes. Este trabalho apresenta o desenvolvimento de um sistema automatizado para criação de avaliações baseadas nos descritores do SAEB, utilizando Modelos de Linguagem de Grande Escala (LLMs). A solução proposta integra uma arquitetura em camadas composta por parametrização, geração, avaliação e saída, permitindo a geração de questões textuais personalizadas quanto ao nível de dificuldade e descritor selecionado. O sistema foi implementado com FastAPI no backend, Vue.js no frontend e integração com API de modelo de linguagem para geração automática de itens. A validação foi realizada por meio de testes piloto com quatro educadores do ensino médio, que avaliaram as questões geradas quanto à coerência, contextualização, clareza, nível de dificuldade e alinhamento ao descritor. Os resultados indicaram elevados índices de aceitação, com médias superiores a 4,0 em escala de 1 a 5, evidenciando a viabilidade da ferramenta como apoio ao planejamento pedagógico. Conclui-se que a aplicação de LLMs na geração automática de avaliações alinhadas ao SAEB apresenta potencial significativo para otimizar o trabalho docente, mantendo qualidade pedagógica e padronização avaliativa, desde que acompanhada de supervisão humana.

Palavras-chaves: Geração de questões; SAEB; avaliação educacional; Large Language Model; Modelo de Linguagem; IA.

ABSTRACT

Educational assessment plays a fundamental role in measuring student learning and guiding public policies. In Brazil, the Basic Education Assessment System (SAEB) is the main large-scale evaluation instrument, requiring test items to be aligned with specific skill descriptors. However, the manual development of such assessments demands time, technical knowledge, and methodological rigor from educators. This study presents the development of an automated system for generating assessments based on SAEB descriptors using Large Language Models (LLMs). The proposed solution adopts a layered architecture composed of parameterization, generation, evaluation, and output modules, enabling the creation of textual questions customized according to difficulty level and selected descriptor. The system was implemented using FastAPI for the backend, Vue.js for the frontend, and integration with a language model API for automatic item generation. Validation was conducted through pilot testing with four high school teachers, who evaluated the generated questions regarding coherence, contextualization, clarity, difficulty level, and alignment with SAEB descriptors. The results indicated high acceptance rates, with mean scores above 4.0 on a 1-5 scale, demonstrating the feasibility of the tool as support for pedagogical planning. It is concluded that the application of LLMs in the automatic generation of SAEB-aligned assessments presents significant potential to optimize teachers' workload while maintaining pedagogical quality and assessment standardization, provided that human supervision is maintained.

Keywords: Question Generation; SAEB; Educational Assessment; Large Language Model; Language Model; AI.

LISTA DE FIGURAS

Figura 1	Diagrama da arquitetura em quatro camadas do sistema	24
Figura 2	Imagem da tela de avaliação de perguntas	27
Figura 3	Imagem da tela de avaliação de perguntas	28
Figura 4	Imagem da tela de criação de perguntas	29
Figura 5	Imagem da tela de criação de avaliação	31
Figura 6	Avaliação gerada	32
Figura 7	Exemplo de questão de Língua Portuguesa gerada pelo sistema.....	33
Figura 8	Exemplo de questão de Matemática gerada pelo sistema	33

LISTA DE TABELAS

Tabela 1	Descritores do 3 ^o ano do ensino médio na disciplina de Matemática.....	21
Tabela 2	Requisitos funcionais e não funcionais do sistema	23
Tabela 3	Stack tecnológica utilizada no desenvolvimento.....	25
Tabela 4	Avaliação das questões pelos educadores (n=4)	34

LISTA DE ABREVIATURAS E SIGLAS

TCC	Trabalho de Conclusão de Curso
API	<i>Application Programming Interface</i>
SAEB	Sistema de Avaliação da Educação Básica
IA	Inteligência Artificial
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Problema.....	12
1.2	Importância da avaliação na educação	12
1.3	Desafios da elaboração manual das avaliações	12
1.4	Sistema de Avaliação da Educação Básica Sistema de Avaliação da Educação Básica (SAEB).....	13
1.4.1	Objetivos do SAEB	13
1.4.2	Público Avaliado	14
1.4.3	Componentes Avaliados.....	14
1.4.4	Metodologia	14
1.4.5	Importância e Impacto	14
1.5	Por que usaremos LLMs?	15
1.6	Objetivo	15
1.6.1	Objetivos Específicos	15
1.7	Organização do trabalho.....	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	Inteligência Artificial	17
2.2	Inteligência Artificial na Educação.....	17
2.3	Modelos de Linguagem de Grande Escala (LLMs)	18
2.4	Potenciais Benefícios dos LLMs na Educação	19
2.5	Trabalhos relacionados.....	20
3	METODOLOGIA	21
3.1	Levantamento e análise dos descritores SAEB	21
3.2	Definição dos requisitos do sistema	23
3.3	Arquitetura do sistema proposto	23
3.4	Implementação Técnica do Sistema.....	25
3.5	Procedimentos de validação das questões	27
3.6	Procedimentos de Criação de Perguntas.....	29

3.7	Procedimentos de Criação da Avaliação	30
4	RESULTADOS	33
4.1	Exemplos de questões geradas pelo sistema	33
4.2	Feedback dos educadores durante testes.....	33
4.2.1	Análise de coerência, contextualização e dificuldade	34
5	CONSIDERAÇÕES FINAIS	35
5.1	Limitações do trabalho	35
5.2	Contribuições	36
5.3	Trabalhos futuros	36
	REFERÊNCIAS	38

1 INTRODUÇÃO

1.1 Problema

A elaboração manual de avaliações alinhadas aos descritores do SAEB é um processo complexo, que demanda tempo, esforço e conhecimento especializado dos educadores. A construção de instrumentos avaliativos exige coerência com os objetivos de aprendizagem e critérios bem definidos, o que reforça a necessidade de apoio nesse processo (Luckesi 2011).

1.2 Importância da avaliação na educação

A avaliação do processo de ensino-aprendizagem constitui uma etapa fundamental na educação, permitindo identificar o nível de compreensão dos estudantes e orientar estratégias pedagógicas. Nessa perspectiva, a avaliação deve ser compreendida como um processo contínuo e formativo, voltado para a melhoria da aprendizagem e não apenas para a mensuração de resultados (Hoffmann 2014).

Além disso, a avaliação desempenha um papel essencial no processo educacional ao possibilitar a identificação das habilidades já consolidadas e das dificuldades apresentadas pelos estudantes. Esse processo permite que os educadores compreendam o nível real de aprendizagem da turma, favorecendo o planejamento de intervenções pedagógicas mais assertivas e personalizadas (Hoffmann 2014).

Ao orientar a tomada de decisões em sala de aula, a avaliação também contribui para a padronização e a objetividade na mensuração do desempenho escolar, configurando-se como um instrumento estratégico para a melhoria contínua da qualidade do ensino.

1.3 Desafios da elaboração manual das avaliações

A construção manual de avaliações alinhadas aos descritores do SAEB apresenta diversos desafios que impactam diretamente a prática docente. Um dos principais problemas é o tempo demandado para elaboração das questões, já que o professor precisa selecionar conteúdos, redigir enunciados coerentes e produzir alternativas. Além disso, é necessário garantir variação no nível de complexidade dos itens, de modo a contemplar diferentes graus de aprendizagem (Luckesi 2011).

Outro desafio relevante consiste em assegurar a fidelidade pedagógica das questões em relação às competências e habilidades avaliadas pelo SAEB. Esse processo exige não apenas domínio do conteúdo, mas também experiência metodológica e cuidado na formulação linguística, para evitar ambiguidades ou interpretações equivocadas (Luckesi 2011).

Essas limitações podem comprometer tanto a eficiência do processo avaliativo quanto a frequência de aplicação das avaliações, restringindo o potencial de monitoramento contínuo da aprendizagem dos estudantes (Hoffmann 2014).

1.4 Sistema de Avaliação da Educação Básica SAEB

O SAEB constitui a principal política pública de avaliação educacional em larga escala no Brasil, sendo conduzido pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (INEP). Trata-se de um conjunto de avaliações externas que tem como finalidade produzir um diagnóstico abrangente da educação básica brasileira, bem como identificar fatores associados ao desempenho dos estudantes (Brasil, INEP 2021).

Por meio da aplicação de testes cognitivos e questionários contextuais, realizados de forma periódica, o SAEB permite aferir os níveis de aprendizagem dos estudantes e compreender as condições em que o processo educativo ocorre. Esses instrumentos possibilitam não apenas mensurar o desempenho em áreas específicas do conhecimento, mas também analisar variáveis socioeconômicas, escolares e pedagógicas que influenciam os resultados educacionais (Brasil, INEP 2021).

Os dados produzidos pelo SAEB são fundamentais para que escolas, redes de ensino e gestores públicos avaliem a qualidade da educação ofertada. Além disso, seus resultados subsidiam a formulação, o monitoramento e o aprimoramento de políticas educacionais baseadas em evidências, contribuindo para a melhoria contínua do ensino no país (Brasil, INEP 2021).

1.4.1 Objetivos do SAEB

- ♦ Avaliar a qualidade, a equidade e a eficiência da educação básica brasileira
- ♦ Identificar desigualdades regionais e fatores associados ao desempenho dos estudan-

tes

- ♦ Produzir informações que auxiliem na tomada de decisões educacionais
- ♦ Subsidiar a formulação e o aprimoramento de políticas públicas educacionais

1.4.2 Público Avaliado

- ♦ Estudantes do 5º e 9º anos do Ensino Fundamental
- ♦ Estudantes da 3ª série do Ensino Médio
- ♦ Escolas públicas (avaliação censitária) e privadas (avaliação amostral)
- ♦ Professores e gestores escolares, por meio de questionários contextuais

1.4.3 Componentes Avaliados

- ♦ Língua Portuguesa, com ênfase em leitura e interpretação de textos
- ♦ Matemática, com foco na resolução de problemas
- ♦ Ciências da Natureza e Ciências Humanas (para etapas específicas, conforme edições mais recentes)

1.4.4 Metodologia

- ♦ Aplicação periódica, geralmente a cada dois anos
- ♦ Utilização da Teoria de Resposta ao Item (TRI) para mensuração do desempenho
- ♦ Construção de escalas de proficiência que permitem comparações ao longo do tempo
- ♦ Divulgação dos resultados em diferentes níveis: escola, município, estado e país

1.4.5 Importância e Impacto

- ♦ Integra a composição do Índice de Desenvolvimento da Educação Básica (IDEB)
- ♦ Permite a comparação do desempenho entre redes e regiões

- ♦ Auxilia na identificação de lacunas no processo de ensino-aprendizagem
- ♦ Orienta a definição de políticas públicas e estratégias de intervenção educacional

1.5 Por que usaremos LLMs?

Diante dos desafios inerentes à elaboração manual de avaliações, o uso de Modelos de Linguagem de Grande Escala (LLMs – Large Language Models) surge como uma alternativa inovadora e promissora. Esses modelos, treinados com vastos volumes de dados textuais, possuem a capacidade de compreender, interpretar e gerar linguagem natural de maneira contextualizada e coerente (Brown et al. 2020).

A aplicação de LLMs no contexto educacional possibilita a automação da geração de questões, reduzindo o tempo e o esforço exigidos dos educadores. Além disso, permite a personalização das avaliações, adequando o nível de complexidade, o formato das respostas e até as temáticas às necessidades de cada turma ou estudante (Kasneci et al. 2023).

1.6 Objetivo

Desenvolver um sistema automatizado capaz de gerar avaliações, compostas por questões textuais geradas por LLMs e fundamentadas nos descritores do SAEB, visando otimizar e padronizar a elaboração de instrumentos avaliativos que possibilitem uma análise eficaz do desempenho dos estudantes.

1.6.1 Objetivos Específicos

- ♦ Identificar e analisar os descritores e competências avaliados pelo SAEB nas disciplinas de Língua Portuguesa e Matemática.
- ♦ Aplicar técnicas de Processamento de Linguagem Natural (PLN), com o apoio de Modelos de Linguagem de Grande Escala (LLMs), para a geração automatizada de questões textuais alinhadas aos descritores do SAEB.
- ♦ Desenvolver avaliações estruturadas a partir de questões geradas com base nos descritores, visando a análise precisa das competências e habilidades dos estudantes.
- ♦ Oferecer uma interface intuitiva e acessível para que os educadores possam utilizar o sistema de forma prática e eficiente.

- ♦ Avaliar a eficácia e a relevância das avaliações geradas pelo sistema por meio da análise de educadores, identificando potenciais melhorias.
- ♦ Investigar a viabilidade, a qualidade e a aplicabilidade do sistema com educadores, buscando compreender suas contribuições para o aprimoramento do processo de avaliação educacional.

1.7 Organização do trabalho

- ♦ **Capítulo 1 – Introdução:** Apresenta a contextualização do problema, a relevância da avaliação educacional, os desafios da elaboração manual de avaliações, a justificativa para o uso de LLMs e os objetivos do trabalho.
- ♦ **Capítulo 2 – Fundamentação Teórica:** Discute os conceitos fundamentais relacionados ao SAEB, avaliação educacional, inteligência artificial aplicada à educação e técnicas de processamento de linguagem natural.
- ♦ **Capítulo 3 – Metodologia:** Descreve a abordagem adotada para o desenvolvimento do sistema, incluindo a arquitetura, a implementação das funcionalidades de geração de questões e os procedimentos para validação do sistema.
- ♦ **Capítulo 4 – Resultados e Discussão:** Apresenta os resultados obtidos, analisando a qualidade das questões geradas e a percepção de educadores sobre a ferramenta.
- ♦ **Capítulo 5 – Conclusão e Trabalhos Futuros:** Sintetiza as principais contribuições do trabalho, aponta limitações identificadas e sugere possibilidades de aprimoramento e pesquisas futuras.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Inteligência Artificial

A Inteligência Artificial (IA) refere-se a sistemas computacionais desenvolvidos para interagir com o ambiente por meio de capacidades que simulam habilidades humanas, como percepção visual, reconhecimento de fala e tomada de decisão. Esses sistemas são capazes de analisar informações disponíveis, aprender com dados e executar ações de forma autônoma, com o objetivo de alcançar determinados resultados de maneira eficiente. Nesse contexto, a IA pode ser compreendida como um conjunto de tecnologias que reproduzem comportamentos considerados inteligentes, permitindo que máquinas realizem tarefas que, tradicionalmente, exigiriam a intervenção humana. De acordo com (Luckin et al. 2016), tais sistemas apresentam características que se aproximam das capacidades cognitivas humanas, ao avaliarem informações e tomarem decisões com base em objetivos previamente definidos.

2.2 Inteligência Artificial na Educação

A Inteligência Artificial na Educação é definida como um campo interdisciplinar que aplica técnicas de IA para criar sistemas computacionais capazes de interagir com o mundo de maneira semelhante à inteligência humana, com o objetivo primordial de compreender e melhorar o processo de aprendizagem em qualquer contexto (Luckin et al. 2016). Seu papel não é substituir o professor, mas sim atuar como um poderoso catalisador para transformar e potencializar a educação, tornando-a mais personalizada, inclusiva, eficiente e alinhada com as demandas do século XXI.

Seu papel central reside na sua capacidade de personalizar a aprendizagem em escala. Através de três modelos computacionais centrais — o modelo pedagógico (conhecimento sobre ensino), o modelo de domínio (conhecimento sobre a matéria) e o modelo do aprendiz (conhecimento sobre o aluno) —, os sistemas de AIEd podem adaptar dinamicamente o conteúdo, o ritmo e o tipo de suporte oferecido a cada estudante. Isso se concretiza em Sistemas Tutores Inteligentes (ITS), que simulam o efeito comprovadamente eficaz do tutoring one-to-one, fornecendo feedback individualizado e oportuno sem a necessidade de um tutor humano para cada aluno (Luckin et al. 2016).

Além da personalização, a AIEd tem um papel crucial em facilitar a aprendizagem colaborativa. Através de técnicas como formação adaptativa de grupos, facilitação especializada por algoritmos, uso de agentes virtuais e moderação inteligente de discussões, a IA pode mediar e enriquecer as interações entre alunos, garantindo que a colaboração seja produtiva e gere melhores resultados de aprendizagem (Luckin et al. 2016).

O papel da AIEd se estende também à avaliação, promovendo uma mudança de paradigma. Em vez de avaliações pontuais ("stop-and-test"), a AIEd permite uma avaliação contínua e formativa integrada às atividades de aprendizagem. Ao analisar dados em tempo real, o sistema pode fornecer novos insights sobre como o aluno chegou a uma resposta, seu estado emocional (tédio, frustração, confiança) e seu processo cognitivo, oferecendo intervenções just-in-time (Luckin et al. 2016).

Para os professores, a AIEd atua como um assistente inteligente, assumindo tarefas rotineiras como correção e organização, o que libera tempo para que possam se concentrar nas interações mais complexas e humanas, como orientar, inspirar e mentorar. A IA empodera os educadores com dados ricos e acionáveis sobre seus alunos, por meio de Modelos Abertos do Aprendiz, permitindo-lhes tomar decisões pedagógicas mais informadas (Luckin et al. 2016).

2.3 Modelos de Linguagem de Grande Escala (LLMs)

De acordo com Brown et al. (2020) LLMs são modelos de inteligência artificial treinados em grandes volumes de texto para prever a próxima palavra em uma sequência. Eles são capazes de:

- ♦ Aprender de forma não supervisionada a partir de texto bruto.
- ♦ Realizar uma variedade de tarefas sem necessidade de ajuste fino (fine-tuning), apenas por meio de exemplos no contexto (few-shot ou one-shot).
- ♦ Generalizar para tarefas não vistas durante o treinamento.
- ♦ Gerar texto coerente, responder perguntas, traduzir, resumir, corrigir gramática, entre outras habilidades.

2.4 Potenciais Benefícios dos LLMs na Educação

1. Personalização do Aprendizado

- ♦ LLMs podem gerar explicações, exercícios e materiais adaptados ao nível e ritmo de cada aluno.
- ♦ Podem atuar como tutores virtuais 24/7, respondendo dúvidas em tempo real.

2. Apoio à Redação e Correção

- ♦ Corrigir gramática, estilo e coerência textual.
- ♦ Sugerir melhorias em redações, ensaios e trabalhos acadêmicos.
- ♦ Gerar exemplos de textos bem estruturados.

3. Tradução e Acessibilidade

- ♦ Traduzir materiais didáticos para diferentes idiomas.
- ♦ Auxiliar alunos com dificuldades de leitura ou com deficiência visual por meio de síntese de voz ou resumos.

4. Criação de Conteúdo Educativo

- ♦ Professores podem usar LLMs para criar questões, provas, planos de aula e explicações complementares.
- ♦ Geração de histórias, problemas de matemática, estudos de caso, etc.

5. Aprendizado de Idiomas

- ♦ Prática de conversação, correção de pronúncia (via integração com outras ferramentas) e exercícios contextualizados.

6. Estimulação do Raciocínio e Criatividade

- ♦ LLMs podem propor desafios, enigmas, analogias e problemas que exigem pensamento crítico.
- ♦ Auxiliar na geração de ideias para projetos e pesquisas.

7. Suporte a Alunos com Dificuldades

- ◆ Explicar conceitos complexos de maneira simplificada.
- ◆ Oferecer múltiplas formas de abordar um mesmo conteúdo.

2.5 Trabalhos relacionados

Diversos estudos recentes têm explorado a aplicação de técnicas de Inteligência Artificial, em especial modelos de linguagem, no contexto educacional. (Heilman e Smith 2010) apresentaram uma das primeiras abordagens sistemáticas para geração automática de questões, propondo algoritmos baseados em transformações linguísticas e ranqueamento. Embora relevante, tal abordagem era limitada pela dependência de regras linguísticas e pelo escopo restrito de aplicação.

Com o avanço dos Modelos de Linguagem de Grande Escala (LLMs), novas possibilidades emergiram. (Brown et al. 2020) demonstraram o potencial do GPT-3 na geração de textos coerentes, o que abriu espaço para pesquisas na criação automatizada de instrumentos avaliativos. (Mittal et al. 2024) e (Ogunleye et al. 2024) realizaram revisões sistemáticas destacando como a IA generativa pode apoiar práticas educacionais, desde a personalização de conteúdos até a elaboração de avaliações.

No contexto brasileiro, Santana e Magalhães (2024) apresentaram uma revisão da aplicação do Processamento de Linguagem Natural (PLN) em ambientes educacionais, ressaltando o potencial das ferramentas de IA para auxiliar docentes. Ainda que haja progressos, poucos trabalhos investigam diretamente a integração entre os descritores do SAEB e a geração automática de questões, o que evidencia a relevância e ineditismo da proposta deste estudo.

3 METODOLOGIA

3.1 Levantamento e análise dos descritores SAEB

As Matrizes de Referência do SAEB, publicadas pelo INEP em 2001, representam documentos fundamentais para compreender as expectativas de aprendizagem avaliadas nacionalmente. Este documento apresenta uma síntese dos descritores contidos nas matrizes de Língua Portuguesa e Matemática. Os descritores foram organizados por tópicos, mantendo a estrutura original das matrizes.

A análise foi realizada mediante consulta direta aos documentos oficiais disponibilizados pelo INEP:

1. Matriz de Referência de Língua Portuguesa - SAEB 2001
2. Matriz de Referência de Matemática - SAEB 2001

Exemplo referente aos descritores do 3º Ano do ensino médio da disciplina de matemática:

Tabela 1: Descritores do 3º ano do ensino médio na disciplina de Matemática

Código	Descrição
D1	Identificar figuras semelhantes mediante o reconhecimento de relações de proporcionalidade.
D2	Reconhecer aplicações das relações métricas do triângulo retângulo em problemas com figuras espaciais.
D3	Relacionar diferentes poliedros ou corpos redondos com suas planificações ou vistas.
D4	Identificar a relação entre as representações algébrica e geométrica de um sistema de equações.
D5	Resolver problema que envolva razões trigonométricas no triângulo retângulo.
D6	Identificar a localização de pontos no plano cartesiano.
D7	Interpretar geometricamente os coeficientes da equação de uma reta.

Continua na próxima página

Tabela 1 – Continuação

Código	Descrição
D8	Identificar a equação de uma reta a partir de dois pontos ou de um ponto e sua inclinação.
D9	Relacionar a interseção de retas com a resolução de sistemas de equações.
D10	Reconhecer equações do 2º grau com duas incógnitas que representam circunferências.
D11	Resolver problema envolvendo cálculo de perímetro de figuras planas.
D12	Resolver problema envolvendo cálculo de área de figuras planas.
D13	Resolver problema envolvendo área total e/ou volume de sólidos.
D14	Identificar a localização de números reais na reta numérica.
D15	Resolver problema com variação proporcional entre grandezas.
D16	Resolver problema envolvendo porcentagem.
D17	Resolver problema envolvendo equação do 2º grau.
D18	Reconhecer função a partir de tabela.
D19	Resolver problema envolvendo função do 1º grau.
D20	Analisar comportamento de funções em gráficos.
D21	Identificar gráfico correspondente a uma situação descrita.
D22	Resolver problema envolvendo P.A./P.G.
D23	Reconhecer gráfico de função do 1º grau pelos coeficientes.
D24	Identificar função a partir do gráfico.
D25	Resolver problemas de máximo e mínimo em funções quadráticas.
D26	Relacionar raízes de polinômio com fatoração.
D27	Identificar função exponencial a partir do gráfico.
D28	Identificar função logarítmica como inversa da exponencial.
D29	Resolver problema envolvendo função exponencial.
D30	Identificar funções trigonométricas e suas propriedades.
D31	Resolver sistema linear por representação algébrica ou gráfica.
D32	Resolver problemas de contagem (combinatória).
D33	Calcular probabilidade de eventos.
D34	Resolver problemas com dados em tabelas ou gráficos.

Continua na próxima página

Tabela 1 – Continuação

Código	Descrição
D35	Relacionar tabelas e gráficos.

Fonte: INEP (2021)

3.2 Definição dos requisitos do sistema

Para orientar o desenvolvimento da plataforma de forma estruturada e alinhada com as necessidades educacionais identificadas, foram definidos requisitos funcionais e não funcionais que guiarão a implementação do sistema. Os requisitos funcionais descrevem as funcionalidades específicas que o sistema deve oferecer, enquanto os requisitos não funcionais estabelecem critérios de qualidade e restrições técnicas que devem ser observadas.

A **Tabela 2** apresenta a categorização detalhada desses requisitos, organizados para garantir que a solução desenvolvida atenda adequadamente aos objetivos propostos de geração de questões alinhadas aos descritores SAEB e criação de avaliações personalizadas.

Tabela 2: Requisitos funcionais e não funcionais do sistema

Tipo de Requisito	Descrição
Requisitos Funcionais	Geração de questões alinhadas a descritores do SAEB Personalização por nível de complexidade e tema Criação de avaliações baseadas nas questões geradas Exportação de avaliações em formatos acessíveis (PDF)
Requisitos Não Funcionais	Usabilidade Escalabilidade Desempenho

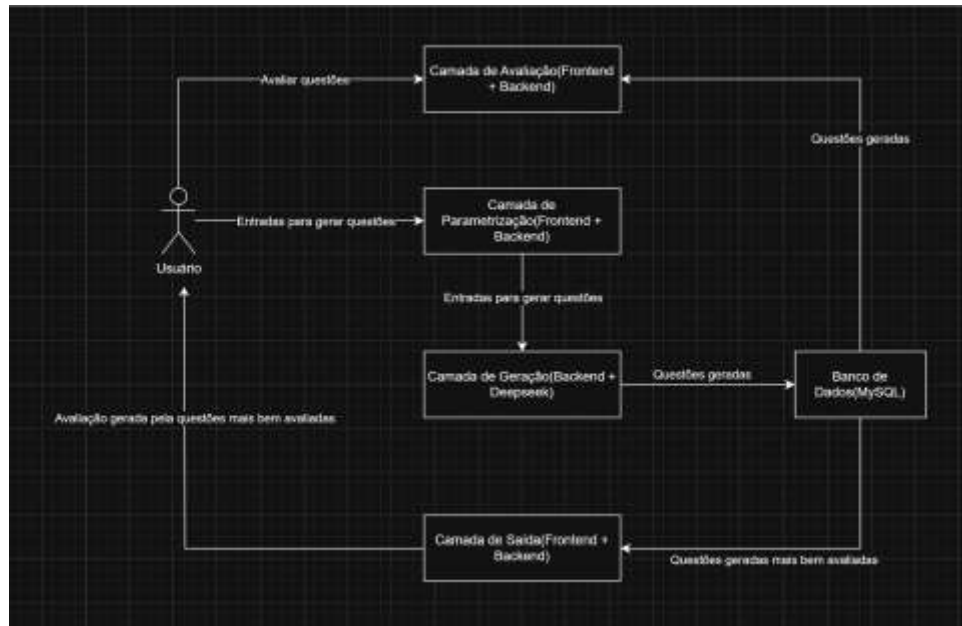
Fonte: Autor (2026)

3.3 Arquitetura do sistema proposto

A arquitetura do sistema foi estruturada em quatro camadas principais interconectadas, cada uma com responsabilidades específicas no fluxo de geração de avaliações.

Figura 1.

Figura 1: Diagrama da arquitetura em quatro camadas do sistema



Fonte: Autor (2026)

Camadas da Arquitetura

1. **Camada de Parametrização** - Esta camada é responsável pela coleta e estruturação das entradas do usuário para geração de questões, permite:
 - ◆ Seleção de descritores SAEB específicos.
 - ◆ Definição de parâmetros de geração (nível de dificuldade, tipo de questão).
 - ◆ Configuração de preferências e personalizações.
2. **Camada de Geração** - Esta camada é responsável pela transformação das entradas do usuário em questões educacionais, permite:
 - ◆ Processamento das entradas pelo modelo de linguagem (LLM).
 - ◆ Geração de questões alinhadas aos descritores SAEB selecionados.
 - ◆ Aplicação de algoritmos de garantia de qualidade e consistência.
3. **Camada de Avaliação** - Esta camada é responsável pela avaliação humana das questões geradas pela LLM, permite:
 - ◆ Avaliar quantitativamente com valor de 1 a 5.
 - ◆ Avaliar qualitativamente fazendo comentários nas questões.

4. **Camada de Saída** - Esta camada é responsável por gerar a avaliação e disponibilizar ao usuário, permite:

- ◆ Apresentação da avaliação em interface web.
- ◆ Opção de exportação em PDF.

3.4 Implementação Técnica do Sistema

O sistema foi arquitetado com uma stack tecnológica moderna e robusta, selecionada para atender aos requisitos de desempenho, escalabilidade e manutenibilidade. A **Tabela 3** sumariza as principais tecnologias empregadas no desenvolvimento.

Tabela 3: Stack tecnológica utilizada no desenvolvimento

Componente	Tecnologia e Finalidade
Linguagens Principais	Python e JavaScript
LLM	DeepSeek
Frontend	Vue.js com Composition API
Backend	Python com FastAPI
Banco de Dados	MySQL 8.0
Containerização	Docker e Docker Compose

Fonte: Autor (2026)

Geração Automática de Questões

Para a geração automática de questões, integrou-se a **DeepSeek API**, que fornece acesso a modelos de linguagem de grande escala (LLMs) especializados em tarefas de compreensão e produção textual em português. Esta integração permite:

- ◆ Geração contextualizada de questões baseadas nos descritores SAEB
- ◆ Adaptação do nível de complexidade conforme parâmetros educacionais
- ◆ Produção de enunciados claros e pedagogicamente adequados
- ◆ Geração de alternativas múltiplas com distratores plausíveis

Arquitetura de *Frontend*

A interface do sistema foi implementada em *Vue.js* com abordagem componentizada, garantindo:

- ♦ Experiência de usuário intuitiva e responsiva
- ♦ Componentes reutilizáveis e de fácil manutenção
- ♦ Interface de seleção de descritores com filtros avançados

Arquitetura de *Backend*

O backend foi construído em *FastAPI*, framework moderno que oferece:

- ♦ Alta performance no processamento de requisições
- ♦ Documentação automática da API via OpenAPI/Swagger
- ♦ Comunicação assíncrona entre os módulos do sistema

Armazenamento de Dados

Adotou-se o *MySQL* como sistema de gerenciamento de banco de dados relacional, proporcionando:

- ♦ Estruturação eficiente dos metadados das questões
- ♦ Gestão de usuários e perfis de acesso
- ♦ Armazenamento de histórico de gerações
- ♦ Consultas complexas com otimização de desempenho

Deploy e Containerização

O processo de implantação foi realizado em ambiente de containers *Docker*, assegurando:

- ♦ Portabilidade entre diferentes ambientes de execução

- ◆ Escalabilidade horizontal mediante aumento de demanda
- ◆ Padronização do ambiente de desenvolvimento e produção
- ◆ Isolamento de dependências e configurações

Esta stack tecnológica foi selecionada visando não apenas o atendimento aos requisitos atuais, mas também permitindo evolução futura do sistema mediante a incorporação de novas funcionalidades e integrações.

Todos os códigos estão disponíveis em: <https://github.com/MasterRick/TCC>

3.5 Procedimentos de validação das questões

O sistema incorpora uma abordagem multidimensional para validação da qualidade das questões geradas, combinando avaliação automática e humana para garantir adequação pedagógica, precisão conceitual e alinhamento com os descritores SAEB.

Figura 2: Imagem da tela de avaliação de perguntas

Avaliação de Questões
Use o formulário abaixo para avaliar as questões.

Descritor: D1
 Dificuldade: Fácil
 Disciplina: Matemática
 Turno: Ensino Fundamental
 Ano: 5º Ano

Identificador da Questão: 696
 Descritor: D1 - Identificar a localização/movimentação de objeto em mapas, croquis e outras representações gráficas.
 Dificuldade: Fácil

Enunciado:
 Maria desenhou um mapa simples do caminho de sua casa até o centro esportivo. No mapa, a seta representa a posição inicial dela em frente à sua casa. Para chegar ao centro esportivo, ela deve seguir estas instruções: Caminhe 3 quadradinhos para a direita, depois 2 quadradinhos para cima e, finalmente, 1 quadradinho para a esquerda. Observando o mapa, em qual quadradinho Maria terminará seu percurso? (Um mapa simples com uma grade de quadradinhos, onde a casa está no quadradinho inferior esquerdo e o centro esportivo em um quadradinho central. O aluno receberá uma imagem da grade com coordenadas ou letras, mas como é texto, considere-se uma grade com quadradinhos identificados por letras de A a E nas colunas e números de 1 a 5 nas linhas. A casa está em B1. O centro esportivo está em D3).

Alternativas:
 A. Quadradinho C4
 B. Quadradinho D3
 C. Quadradinho B3
 D. Quadradinho E2
 E. Quadradinho C2

Justificativa:
 A alternativa B é a correta porque, partindo de B1 (casa) e seguindo as instruções: 3 quadradinhos para a direita leva a E1, depois 2 quadradinhos para cima leva a E3, e finalmente 1 quadradinho para a esquerda leva exatamente ao quadradinho D3, que é a localização do centro esportivo no mapa. As outras alternativas resultam de erros na contagem dos quadradinhos ou na direção dos movimentos descritos.

AVALIAR QUESTÃO

Fonte: Autor (2026)

A validação ocorre em duas etapas sequenciais:

1. Validação Automática - Realizada imediatamente após a geração pela LLM
2. Validação Humana - Executada por especialistas em educação mediante interface específica

Validação Automática

Antes da disponibilização para avaliação humana, as questões passam por verificação automática:

- ♦ Verificação de Formatação: Validação da estrutura técnica da questão

Camada de Avaliação Humana

Esta camada proporciona a avaliação por especialistas em educação mediante interface ?? intuitiva que permite:

Figura 3: Imagem da tela de avaliação de perguntas

Avaliação de Questões

Use o formulário abaixo para avaliar as questões.

Identificador da Questão: 696
Descritor: D1 - Identificar a localização/
Dificuldade: Fácil

Enunciado:
Marta desenhou um mapa simples do centro da sua casa, quadrados para a direita, depois 2 quadrados para cima e, por fim, 1 quadrado para a esquerda. Ela desenhou o mapa da casa dela no quadrado inferior esquerdo e o centro do mapa para a direita do A e E nas colunas e número de 1 a 5 nas linhas.

Alternativas:
A. Quadrado D4
B. Quadrado D3
C. Quadrado D2
D. Quadrado E2
E. Quadrado C2

Justificativa:
A alternativa D é a correta porque, partindo de B3, vamos para a direita 2 quadrados e para cima 1 quadrado, então chegamos ao quadrado D3, que é a localização do centro do mapa.

Avaliação da Questão

Por favor, avalie cada aspecto da questão de 1 a 5 estrelas:
1 estrela = Ruim, 5 estrelas = Muito Bom

Coerência dos enunciados: ★★★★★

Contextualização adequada: ★★★★★

Nível de dificuldade apropriado: ★★★★★

Clareza das alternativas: ★★★★★

Alinhamento com descritor: ★★★★★

Comentário (opcional):

CANCELAR SALVAR AVALIAÇÃO

Fonte: Autor (2026)

- ♦ Avaliação Quantitativa: Atribuição de notas de 1 a 5 em critérios específicos:
 - Adequação ao descritor SAEB
 - Clareza do enunciado
 - Pertinência das alternativas
 - Nível de dificuldade apropriado
 - Originalidade e relevância contextual
- ♦ Avaliação Qualitativa: Comentários textuais detalhados sobre:

- Sugestões de melhoria no enunciado
- Análise de possíveis ambiguidades
- Recomendações de adaptação curricular
- Observações sobre adequação à série/ano

Processo de Consolidação

Após a avaliação humana, o sistema realiza:

- ♦ Cálculo da pontuação média baseada nas avaliações quantitativas
- ♦ Classificação das questões por índice de qualidade validada

3.6 Procedimentos de Criação de Perguntas

O processo de criação das perguntas, visível na **Figura 4**, foi estruturado de forma sistemática, garantindo alinhamento pedagógico aos descritores do SAEB e qualidade técnica na formulação dos itens.

Figura 4: Imagem da tela de criação de perguntas

Fonte: Autor (2026)

O processo ocorre conforme as seguintes etapas:

1. **Seleção do Descritor:** O usuário seleciona um ou mais descritores específicos do SAEB, que servirão como base conceitual para a geração da questão.

2. **Definição de Parâmetros:** São configurados parâmetros como nível de dificuldade (fácil, médio ou difícil), disciplina (Português ou Matemática), turma (Ensino médio ou Ensino fundamental), ano (3ANO, 5ANO ou 9ANO) e contexto temático.
3. **Construção do Prompt:** O sistema elabora automaticamente um prompt estruturado contendo:
 - ◆ Descrição do descritor selecionado;
 - ◆ Ano/série correspondente;
 - ◆ Nível de dificuldade desejado;
4. **Geração pela LLM:** O prompt é enviado ao modelo de linguagem, que retorna uma questão composta por:
 - ◆ Enunciado contextualizado;
 - ◆ cinco alternativas;
 - ◆ Indicação da alternativa correta;
 - ◆ Justificativa pedagógica.
5. **Validação Automática:** A questão gerada passa por verificação estrutural, garantindo:
 - ◆ Presença de apenas uma alternativa correta;
 - ◆ Formatação adequada ao padrão do sistema.
6. **Encaminhamento para Avaliação Humana:** Após a validação automática, a questão é disponibilizada para análise por especialistas, conforme descrito na seção anterior.

3.7 Procedimentos de Criação da Avaliação

A criação da avaliação, visível na **Figura 5**, consiste na organização estruturada das questões previamente geradas e validadas, compondo um instrumento avaliativo coerente e alinhado aos objetivos pedagógicos.

Figura 5: Imagem da tela de criação de avaliação

Fonte: Autor (2026)

O processo ocorre conforme as seguintes etapas:

1. Seleção de Questões: O usuário pode:

- ◆ Selecionar manualmente questões específicas;
- ◆ Filtrar por descritor, disciplina ou nível de dificuldade;
- ◆ Utilizar critérios de qualidade baseados na pontuação média atribuída na avaliação humana.

2. Definição da Estrutura da Avaliação: São configurados:

- ◆ Quantidade total de questões;
- ◆ Distribuição por descritor;
- ◆ Ordem de apresentação;
- ◆ Inclusão de capa.

3. Organização e Padronização: O sistema organiza automaticamente:

- ◆ Numeração das questões;
- ◆ Padronização tipográfica;
- ◆ Espaçamento e formatação adequada para impressão ou visualização digital.

4. Geração do Documento Final: A avaliação é disponibilizada em formato PDF para download e impressão. Exemplo na Figura 6.

Figura 6: Avaliação gerada

Avaliação

Aluno: -
Escola: -
Professor: -

Data: 13/04/2026

Quantidade de questões: 28

1. Peter Parker, o Homem-Aranha, está patrulhando a cidade. Ele parte do ponto P (Praça Central) e segue o trajeto desenhado no croqui: 3 quadradinhos para o leste, 2 quadradinhos para o norte, 1 quadradinho para o oeste e finalmente 1 quadradinho para o norte. Considerando que cada quadradinho do croqui representa uma quadra, em qual ponto, entre as opções, o Homem-Aranha termina seu trajeto?

A. 2 quadras a leste e 3 quadras ao norte da Praça Central.
B. 4 quadras a leste e 3 quadras ao norte da Praça Central.
C. 2 quadras a oeste e 3 quadras ao norte da Praça Central.
D. 2 quadras a leste e 3 quadras ao norte da Praça Central.
E. 2 quadras a leste e 1 quadra ao norte da Praça Central.

2. Observe as figuras geométricas tridimensionais abaixo e suas planificações. Marque a alternativa que apresenta apenas figuras que são poliedros e que, ao serem planificadas, resultam em figuras planas formadas somente por polígonos.

A. Cubo, pirâmide de base quadrada e cilindro.
B. Cone, esfera e cubo.
C. Cilindro, cone e pirâmide de base triangular.
D. Cubo, pirâmide de base triangular e prisma de base pentagonal.
E. Esfera, cone e cilindro.

3. Observe as figuras geométricas abaixo: um quadrado, um retângulo e um losango. Todas têm 4 lados. Escreva quais propriedades são comuns a essas três figuras e qual é a principal diferença entre elas, considerando o número de lados e os tipos de ângulos.

A. Todas têm 4 lados iguais e 4 ângulos retos. A diferença é que o losango tem lados de tamanhos diferentes.
B. Todas têm 4 lados e 2 pares de lados paralelos. A diferença é que apenas o quadrado e o retângulo têm os quatro ângulos retos, enquanto o losango não tem todos os ângulos retos.
C. Todas têm 4 lados e 4 ângulos. A diferença é que o quadrado tem lados iguais, o retângulo tem lados opostos iguais e o losango tem lados iguais, mas os ângulos podem não ser retos.
D. Todas têm 4 vértices e os lados são sempre retos. A diferença é que o quadrado tem diagonais desiguais, enquanto as outras têm diagonais iguais.
E. Todas são chamadas de triângulos por terem lados. A diferença é que o quadrado tem ângulos agudos e o retângulo tem ângulos obtusos.

Fonte: Autor (2026)

4 RESULTADOS

Os testes piloto foram conduzidos com 4 educadores do ensino médio, que utilizaram o sistema durante um período de quatro semanas.

O sistema demonstrou capacidade significativa na geração de questões textuais alinhadas aos descritores do SAEB, apresentando consistência estrutural e adequação pedagógica nos testes realizados.

4.1 Exemplos de questões geradas pelo sistema

Figura 7: Exemplo de questão de Língua Portuguesa gerada pelo sistema

Identificador da Questão: 465
Descritor: D1 - Localizar informações explícitas em um texto.
Dificuldade: Fácil

Enunciado:
Leia o texto abaixo e responda à pergunta. **Fonte:** A família Silva foi ao parque no domingo. Eles levaram uma cesta com sanduíches, frutas e suco. As crianças brincaram na grama e no balanço. O pai pagou toda a conta e ficou muito feliz. À tarde, todos comeram um sorvete antes de voltar para casa. **Pergunta:** O que a família levou para comer no parque?

Alternativas:
A. Eles levaram bolos e refrigerante.
B. Eles levaram pipoca e chocolate.
C. Eles levaram pizza e água.
D. Eles levaram apenas frutas.
E. Eles levaram sanduíches, frutas e suco.

Justificativa:
A alternativa E é a correta porque, no texto, está escrito explicitamente: "Eles levaram uma cesta com sanduíches, frutas e suco". As outras alternativas apresentam informações que não estão no texto.

Fonte: Autor (2026)

Figura 8: Exemplo de questão de Matemática gerada pelo sistema

Identificador da Questão: 553
Descritor: D8 - Estabelecer relações entre unidades de medida de tempo.
Dificuldade: Fácil

Enunciado:
O Homem-Aranha precisa calcular quanto tempo ele gastou em uma missão. Se sabe que patrulhou a cidade por 180 minutos. Quantas horas o Homem-Aranha ficou patrulhando? Explique como você chegou ao resultado.

Alternativas:
A. 1 hora e 30 minutos
B. 2 horas
C. 3 horas e 30 minutos
D. 3 horas
E. 4 horas

Justificativa:
A alternativa C é a correta porque 1 hora tem 60 minutos. Para saber quantas horas há em 180 minutos, divide-se 180 por 60, o que resulta em 3 horas.

Fonte: Autor (2026)

4.2 Feedback dos educadores durante testes

O feedback foi coletado por meio de questionários estruturados e perguntas abertas, com foco na usabilidade, relevância pedagógica, clareza textual e adequação das questões aos descritores do SAEB.

De modo geral, os participantes avaliaram positivamente a proposta da ferramenta, destacando:

- ♦ Facilidade de uso da interface;
- ♦ Coerência textual das questões geradas;
- ♦ Adequação das habilidades aos descritores selecionados;
- ♦ Potencial de economia de tempo no planejamento pedagógico.

Como pontos de melhoria, os educadores destacaram:

- ♦ Necessidade de ajustes no nível de dificuldade de algumas questões;
- ♦ Ausência de elementos visuais em determinados itens.

4.2.1 Análise de coerência, contextualização e dificuldade

A avaliação da qualidade das questões concentrou-se em cinco dimensões principais: coerência, contextualização, nível de dificuldade, clareza das alternativas e alinhamento com o descritor. A escala utilizada variou de 1 (muito insatisfatório) a 5 (muito satisfatório). Os resultados estão apresentados na **Tabela 4**.

Tabela 4: Avaliação das questões pelos educadores (n=4)

Dimensão de Análise	Média (1-5)
Coerência dos enunciados	3,8
Contextualização adequada	3,68
Dificuldade apropriada	3,41
Clareza das alternativas	3,61
Alinhamento com descritor	3,71

Fonte: Autor (2026)

Observou-se que 80,49% das questões obtiveram média de avaliação superior a 3, indicando desempenho satisfatório na maioria dos critérios analisados.

5 CONSIDERAÇÕES FINAIS

Este trabalho apresentou o desenvolvimento e a avaliação de um sistema baseado em Modelos de Linguagem de Grande Escala (LLMs) para a geração automática de questões alinhadas aos descritores do SAEB. Os resultados obtidos indicam que a ferramenta é capaz de produzir itens coerentes, contextualizados e pedagogicamente adequados, demonstrando potencial como instrumento de apoio ao planejamento docente.

A validação realizada com educadores evidenciou níveis satisfatórios de aceitação quanto à clareza textual, alinhamento aos descritores e adequação do nível de dificuldade, descritos na **Tabela 4**. Embora a supervisão humana permaneça indispensável, o sistema mostrou-se eficiente na redução do tempo dedicado à elaboração de avaliações, contribuindo para otimizar práticas pedagógicas.

Dessa forma, conclui-se que a aplicação de LLMs no contexto educacional representa uma alternativa promissora para apoiar a produção de instrumentos avaliativos, especialmente quando alinhada a diretrizes curriculares estruturadas, como as matrizes do SAEB.

5.1 Limitações do trabalho

- ♦ **Dependência da qualidade do modelo de linguagem:** o desempenho do sistema está diretamente vinculado à robustez e à acurácia do modelo utilizado, podendo ocorrer inconsistências ocasionais na formulação das questões.
- ♦ **Necessidade de supervisão docente:** embora o processo seja automatizado, a validação final por parte do professor é indispensável para assegurar correção conceitual, adequação pedagógica e alinhamento ao contexto específico da turma.
- ♦ **Escassez de estudos comparativos:** há carência de pesquisas sistemáticas que investiguem, de forma aprofundada, a eficácia da geração automática de questões no contexto específico do SAEB, o que limita análises comparativas mais amplas.

- ♦ **Ausência de suporte a elementos visuais:** a versão atual do sistema restringe-se à geração de questões exclusivamente textuais, não contemplando itens que envolvam gráficos, imagens ou representações visuais, o que reduz sua aplicabilidade em determinadas áreas do conhecimento.
- ♦ **Baixo número de educadores avaliadores:** a validação das questões geradas contou com a participação de 4 educadores, o que limita a diversidade de percepções pedagógicas e metodológicas sobre a qualidade do material produzido.

5.2 Contribuições

- ♦ Desenvolvimento de uma solução inovadora no contexto brasileiro para geração automática de questões alinhadas ao SAEB;
- ♦ Aplicação prática de Modelos de Linguagem de Grande Escala no campo da avaliação educacional;
- ♦ Apoio direto à prática docente, reduzindo o esforço e o tempo despendidos na elaboração de instrumentos avaliativos;
- ♦ Evidências empíricas iniciais sobre a viabilidade do uso de LLMs como ferramenta auxiliar no processo de ensino-aprendizagem.

5.3 Trabalhos futuros

- ♦ Expansão do sistema para contemplar outras matrizes de referência, incluindo avaliações estaduais e internacionais, ampliando sua aplicabilidade;
- ♦ Integração com plataformas de Ensino a Distância (EAD), permitindo a aplicação direta das avaliações em Ambientes Virtuais de Aprendizagem;
- ♦ Implementação de mecanismos de feedback automático para estudantes, possibilitando análises imediatas de desempenho e recomendações personalizadas de estudo;
- ♦ Realização de estudos em larga escala, envolvendo diferentes redes de ensino, a fim de avaliar o impacto pedagógico do sistema em contextos variados;

- ◆ Aprimoramento do controle de tokenização nos prompts, utilizando recursos da biblioteca Hugging Face Transformers para maior eficiência no consumo de APIs e melhor gestão de limites de contexto;
- ◆ Experimentação com modelos abertos disponíveis no Hugging Face Hub, visando comparar desempenhos, diversificar arquiteturas utilizadas e reduzir a dependência de serviços proprietários;
- ◆ Desenvolvimento de suporte à geração de elementos visuais, como gráficos, imagens e representações esquemáticas, possibilitando a criação de questões multimodais e ampliando a aplicabilidade do sistema em áreas que demandam interpretação visual.

Conclui-se que a integração entre Inteligência Artificial e avaliação educacional apresenta amplo potencial de transformação, desde que acompanhada de supervisão pedagógica criteriosa e de contínua validação empírica.

REFERÊNCIAS

- Brasil. INEP. *Sistema de Avaliação da Educação Básica (SAEB)*. Brasília, 2021. Acesso em: 19 mar. 2026. Disponível em: <<https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb>>.
- BROWN, T. B. et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 2020.
- CASELI, H. M.; NUNES, M. G. V. *Processamento de Linguagem Natural: Conceitos, Técnicas e Aplicações em Português*. 2. ed. BPLN, 2024. Disponível em: <<https://brasileiraspln.com/livro-pln/2a-edicao>>.
- CHOMSKY, N. Conditions on transformations. In: *A Festschrift for Morris Halle*. [S.l.]: Holt, Rinehart and Winston, 1973. p. 232–286.
- DeepSeek. *DeepSeek API Documentation*. [S.l.], 2026. Acesso em: 08 maio 2026. Disponível em: <<https://api-docs.deepseek.com>>.
- GRAESSER, A. C. et al. Autotutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, v. 48, n. 4, p. 612–618, 2005.
- HEILMAN, M.; SMITH, N. *Question Generation via Overgenerating Transformations and Ranking*. [S.l.], 2010. Disponível em: <<https://www.cs.cmu.edu/~ark/mheilman/questions/papers/heilman-smith-qg-tech-report.pdf>>.
- HOFFMANN, J. *Avaliar para promover: as setas do caminho*. 14. ed. Porto Alegre: Mediação, 2014.
- KASNECI, E. et al. Chatgpt for good? on opportunities and challenges of large language models for education. *Learning and Individual Differences*, 2023.
- LUCKESI, C. C. *Avaliação da aprendizagem escolar*. 22. ed. São Paulo: Cortez, 2011.
- LUCKIN, R. et al. *Intelligence Unleashed: An Argument for AI in Education*. Pearson, 2016. Acesso em: 19 mar. 2026. Disponível em: <<https://www.researchgate.net/publication/299561597>>.
- MITTAL, U. et al. A comprehensive review on generative ai for education. *IEEE Access*, v. 12, p. 142733–142759, 2024.
- OGUNLEYE, B. et al. A systematic review of generative ai for teaching and learning practice. *Education Sciences*, v. 14, p. 636, 2024.
- ROSS, J. R. *Constraints on Variables in Syntax*. Tese (Doutorado) — MIT, 1967.
- SANTANA, F. P.; MAGALHÃES, L. C. Aplicações do processamento de linguagem natural no ambiente educacional: Uma revisão sistemática da literatura. *Revista Foco*, v. 17, n. 1, p. e3921, 2024.
- WALKER, M. A.; RAMBOW, O.; ROGATI, M. Spot: A trainable sentence planner. In:

Proceedings of NAACL. [S.l.: s.n.], 2001.