

Associação Propagadora Esdeva
Centro Universitário Academia - UniAcademia
Curso de Bacharelado em Sistemas de Informação
Trabalho de Conclusão de Curso - Artigo

ENADEQUEST - UM SISTEMA DE GERAÇÃO DE QUESTÕES NO MODELO ENADE

Leandro Cunha Magalhães¹

Centro Universitário Academia - UniAcademia, Juiz de Fora, MG

Tassio Ferenzini Martins Sirqueira²

Centro Universitário Academia - UniAcademia, Juiz de Fora, MG

Linha de Pesquisa: Sistemas de Informação

RESUMO

ENADE(Exame Nacional de Desempenho Dos Estudantes) é uma avaliação aplicada pelo INEP com base na grade curricular nos conteúdos previstos nas diretrizes dos cursos ofertados. Com os resultados dos alunos é calculado os Indicadores de Qualidade de Educação Superior que mede a qualidade de ensino dos cursos. As provas do ENADE possuem um padrão de questões não costumeiro nas provas aplicadas até a conclusão do ensino básico, tornando uma das dificuldades dos alunos o entendimento para a interpretação cobrada nessas questões, exigindo que professores de ensino superior criem testes que auxiliem e preparem os alunos para a prova do ENADE. Visualizando esse cenário foi desenvolvido o software apresentado nesse trabalho, o ENADEQUEST é um software de geração de questões para os cursos de Sistemas de Informação e Engenharia de Software nos modelos cobrados no ENADE utilizando Inteligência Artificial Generativa com auxílio de um banco de dados com as provas aplicadas nos anos anteriores, visando auxiliar professores a criar questões que correspondem com os modelos cobrados nas provas do ENADE. O ENADEQUEST se mostra um software simples e intuitivo que de forma prática e com poucas ações gera questões que respeitem os padrões desejados e possibilita professores darem instruções que especifiquem disciplinas e temas que as questões devem abordar.

¹Discente do Curso de Sistemas de Informação do Centro Universitário Academia - UniAcademia. E-mail: leandro.2008_mg@hotmail.com

²Docente do Curso de Sistemas de Informação do Centro Universitário Academia - UniAcademia. Orientador

Palavras-chave: ENADE. Inteligência Artificial. SQL. Python. Educação. Prova.

ABSTRACT

ENADE (National Student Performance Exam) is an assessment conducted by INEP based on the curricular guidelines of courses offered. Using students' results, Higher Education Quality Indicators are calculated to measure the quality of teaching in the programs. The ENADE exams feature a style of questions that is uncommon compared to those typically encountered during basic education, making it challenging for students to understand and interpret the demands of these questions. This requires higher education professors to design tests that help prepare students for the ENADE exam. To address this issue, the software presented in this work, ENAQUEST, was developed. ENAQUEST is a question-generation tool tailored for Information Systems and Software Engineering courses, designed to follow the formats required by ENADE. It uses Generative Artificial Intelligence and a database of previous exam questions to assist professors in creating questions that align with ENADE's standards. ENAQUEST is a simple and intuitive software that generates questions efficiently and with minimal effort, ensuring compliance with desired patterns. It also allows professors to specify subjects and themes to be addressed in the generated questions.

Keywords: ENADE. Artificial Intelligence. SQL. Education. Exam.

1 INTRODUÇÃO

Durante o período acadêmico em uma instituição de ensino superior é realizado o ENADE (Exame Nacional de Desempenho do Estudante), aplicado pelo INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira) desde 2004. O exame tem como objetivo avaliar os concluintes dos cursos de graduação com base nas diretrizes previstas para cada área de formação. A partir dos resultados obtidos pelos alunos, uma nota é atribuída à instituição, com o propósito de classificar a qualidade de ensino do curso avaliado. As notas variam de 1 (pior desempenho) a 5 (melhor desempenho) (Schlickmann, Roczanski e Azevedo 2008).

O desempenho dos estudante pode ser analisado a partir de duas perspectivas principais: a primeira refere-se ao alinhamento entre os conteúdos programáticos das disciplinas ofertadas pela instituição e os conhecimentos exigidos no exame; a segunda trata da preparação dos alunos para a interpretação e resolução das questões que compõem a avaliação.

O ENADEQUEST foi desenvolvido para atuar em ambas as suposições, utilizando inteligência artificial generativa para criar novas questões com base nas provas dos cursos de Sistemas de Informação e Engenharia de Software aplicadas em anos anteriores. O sistema segue os modelos de questões do ENADE, auxiliando os alunos a se prepararem para a prova.

A prova é elaborada seguindo o "Guia de Elaboração e Revisão de Itens - Banco Nacional de Itens"(Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) 2024) que descreve todos os critérios que o ENADE deve cumprir, também é possível identificar um paralelo dos tipos de questões cobradas e a Taxonomia de Bloom (Galhardi e Macorin 2013), que auxilia na avaliação das práticas educacionais do ensino superior. Aplicada às questões do ENADE, a taxonomia pode ser associada à categorização proposta na revisão de 2002 por Krathwohl, com as seguintes categorias (Bogéa 2024):

O Guia de Elaboração e Revisão de Itens descreve uma série de critérios que uma questão do ENADE deve seguir, como estrutura, diretrizes e matriz da prova, estrutura dos itens(questões), sua elaboração dividida em texto-ase, enunciado, opções de resposta: gabarito e distratores (para questões de múltipla-escolha), formulação de tipo de questões específicas, padrão de resposta esperado (para questões discursivas), além dos objetivos formatação das questões e cuidados na hora de criar uma questão que valorize e recompense os alunos que adquiriram o conhecimento esperado.

A Taxonomia de Bloom propõe diferentes habilidades cognitivas que consolidam os estudos de um aluno. A primeira delas é "lembrar", que envolve reconhecer, compreender, relembrar e expressar informações. Já "aplicar"consiste em utilizar informações factuais para apresentar teorias, modelos e estruturas em contextos reais. A habilidade de "analisar"refere-se à construção de relações complexas a partir de elementos factuais únicos, além de reconstruir essas relações e identificar necessidades. "Avaliar"descreve a capacidade de realizar julgamentos complexos sobre a natureza do contexto, das informações e dos processos, estabelecendo novas conclusões não representadas nas informações originais. Por fim, "criar"significa gerar novas representações de estruturas de conhecimento, combinando conjuntos complexos de informações em contextos originais.

Com essas habilidades, é possível estabelecer um paralelo com as questões da prova, que são elaboradas a partir de diferentes modelos apresentados. Todas as questões possuem elementos como texto complementar, imagens, gráficos ou tabelas para contextualizar o assunto abordado, sendo que a aplicação da taxonomia tem maior impacto na formulação da pergunta-

chave e nas opções de resposta.

As questões do ENADE dividem-se entre discursivas e de múltipla escolha. As questões discursivas, contêm um text-base, e desafiam o aluno a elaborar um texto explicativo que demonstre conhecimento técnico sobre determinado tema ou apresente a solução para um problema proposto. Já as questões de múltipla escolha seguem formatos específicos que exigem habilidades distintas. Entre elas, estão aquelas que avaliam o domínio teórico do estudante, por meio de frases afirmativas com lacunas que devem ser preenchidas pelas alternativas corretas. Há também questões que exigem interpretação, nas quais o aluno deve relacionar as alternativas às ideias ou contextos apresentados no enunciado. Outro formato, conhecido como "Asserção-Razão", apresenta duas afirmativas no final do enunciado, que podem ou não estar relacionadas, cabendo ao estudante identificar a relação correta entre elas. Por fim, há questões que envolvem a classificação de afirmativas como verdadeiras ou falsas, com as alternativas representando combinações possíveis dessas classificações.

Esses formatos de questões cobrados no ENADE tornam-se um elemento fundamental na preparação dos concluintes de cursos superiores, visando seu bom desempenho na prova. Por meio de simulados e exercícios, os estudantes podem ser capacitados a identificar corretamente as diferenças entre os tipos de questões, bem como os objetivos específicos de cada uma, aprimorando assim sua interpretação e raciocínio lógico.

A inteligência artificial está cada vez mais presente no cotidiano, sendo utilizada tanto em atividades pessoais, como pesquisas, correções textuais e geração de imagens, quanto em contextos profissionais, como criação de conteúdos, aceleração de códigos em TI e previsão de padrões complexos. Combinada ao *Machine Learning*, a IA facilita previsões climáticas, financeiras, diagnósticos médicos, estudos científicos e automação industrial.

Um cenário em que a IA está particularmente presente é o da educação. Exemplos de sistemas que utilizam IA para esse fim são o site Prova Fácil(Prova Fácil 2024), que busca facilitar o trabalho dos professores não apenas na geração de provas e questões, mas também no planejamento e na análise dessas avaliações, associando-se mais a um sistema de gestão do que à geração. Outro sistema com o mesmo propósito é o EducaBras(Educabras 2024), que possui um gerador de provas com base em um banco de questões organizadas em categorias, como aulas e vestibulares.

Diferente dos trabalhos semelhantes mencionados (Prova Fácil e EducaBras), o ENA-QUEST tem como base a criação de questões utilizando IA, buscando inovação em questões

variadas que atendem às demandas solicitadas. O ENAQUEST foca no auxílio ao professor na geração de questões no modelo ENADE, voltadas para cursos superiores, enquanto o EducaBras disponibiliza questões pré-definidas para a geração de novas questões, e o Prova Fácil busca auxiliar na gestão de provas, não se limitando apenas à criação de questões.

A utilização de bancos de questões e ferramentas geradoras de questões e provas não é algo novo; no entanto, o ENAQUEST busca preencher lacunas identificadas em sistemas semelhantes, como o auxílio na preparação de provas pelo professor e a utilização de IA generativa para criar questões inovadoras. O ENADEQUEST foi criado para simplificar a elaboração de questões de ensino superior em cursos de Sistemas de Informação e Engenharia de Software. Seu objetivo é elaborar questões no formato ENADE, usando Inteligência Artificial. Para atingir esses objetivos, foram adotadas as seguintes tecnologias no desenvolvimento da aplicação:

O React, uma biblioteca de JavaScript criada pelo Facebook em 2011, para facilitar a atualização constante de interfaces complexas. Ele introduz o conceito de componentização, que permite configurar e reutilizar componentes em diferentes partes da aplicação. No ENADEQUEST, o React foi utilizado para a criação visual da interface e para gerenciar a comunicação entre o *back-end* e as APIs da aplicação.

O Python é bastante conhecido por sua simplicidade de escrita e é uma das linguagens mais usadas na engenharia de dados. Ele oferece diversas bibliotecas que auxiliam na manipulação e tratamento de dados, no processamento de arquivos em múltiplos formatos (como .json, .csv e .pdf) e na integração com bancos de dados. No ENADEQUEST, o Python foi utilizado para adequar os dados das provas anteriores do ENADE, preparando-os para serem utilizados no sistema.

ETL (*Extract, Transform, Load*) é um processo amplamente utilizado na engenharia de dados. O termo refere-se às três etapas principais desse processo: *extract* (extração), que consiste em coletar dados de diferentes fontes, como arquivos JSON e CSV ou APIs; *transform* (transformação), que envolve a limpeza e padronização dos dados para organizá-los e prepará-los para análise; e *load* (carga), que corresponde ao armazenamento desses dados em bancos de dados ou *data warehouses*. Esse armazenamento pode ser realizado em lotes (*batch*) ou de forma contínua (*streaming*), dependendo da necessidade. No ENADEQUEST, o procedimento ETL foi fundamental para tratar os dados das avaliações anteriores.

O PostgreSQL, criado em 1986, é um sistema gerenciador de banco de dados relacional de código aberto, conhecido por sua capacidade de lidar com bancos de dados complexos e de

alto desempenho. Com uma comunidade ativa e amplo suporte, foi a solução escolhida para armazenar os dados utilizados na aplicação.

O Google anunciou, em dezembro de 2023, o Google Gemini, uma suíte de ferramentas focada em *deep learning* e modelos generativos. Essa tecnologia pode ser usada em navegadores, aplicativos móveis ou pela API, que integra diferentes aplicações. No ENADEQUEST, o Google Gemini foi escolhido como a ferramenta de Inteligência Artificial que gera as questões.

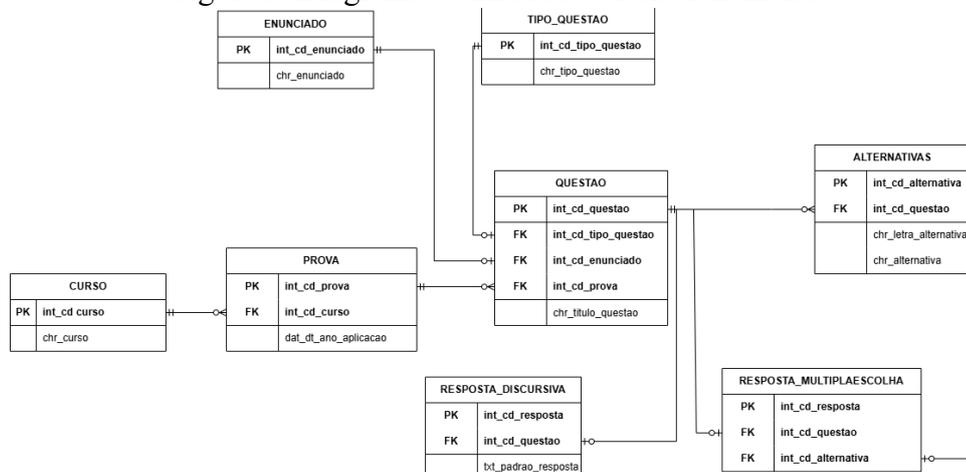
2 DESENVOLVIMENTO

O ENADEQUEST emprega duas etapas para criar questões nos modelos apresentados no ENADE. A primeira etapa consiste em organizar as questões de provas já aplicadas nas instituições. A segunda etapa consiste em elaborar um roteiro (também conhecido como prompt) que contém as instruções adequadas para que a inteligência artificial (no caso o Google Gemini) possa gerar questões adequadamente. A utilização de questões de provas anteriores é adicionada ao prompt para contextualizar a IA com exemplos práticos de algumas questões aplicadas em provas anteriores. Sendo assim, o desenvolvimento da aplicação foi dividido em duas etapas de desenvolvimento macro: o fluxo de dados e o fluxo de execução da aplicação.

2.1 BANCO DE DADOS

Analisando a prova do ENADE foi desenvolvido o diagrama relacional representado na Figura 1 onde cada questão é armazenada em um banco de dados PostgreSQL.

Figura 1: Diagrama de Entidade e Relacionamento



Fonte: elaborado pelo autor

Cada tabela do diagrama possui uma função específica para armazenar as informações

da prova, conforme descrito na Quadro 1:

Quadro 1: Descrição das tabelas do banco de dados

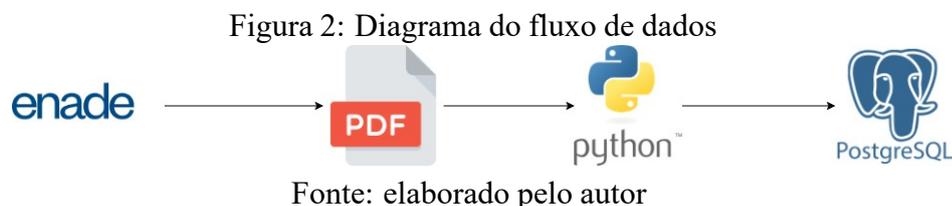
Nome da tabela	Descrição do dado armazenado
CURSO	Nome do curso específico da prova aplicada.
PROVA	Ano de aplicação da prova e responsável por relacionar o curso com a questões aplicadas.
TIPO_QUESTAO	Tipo de questão apresentada (discursiva ou múltipla escolha).
ENUNCIADO	Enunciado da questão.
ALTERNATIVAS	Alternativas de respostas para questões do tipo múltipla escolha.
QUESTAO	Relaciona todos os elementos de uma questão.
RESPOSTA_MULTIPLAESCOLHA	Alternativa correta para as questões do tipo múltipla escolha.
RESPOSTA_DISCURSIVA	Padrão de resposta esperado para questões do tipo discursiva.

Fonte: elaborado pelo autor

2.2 FLUXO DE DADOS

O fluxo de dados é o processo lógico que segue desde a aquisição dos dados, passando pelo tratamento e, finalmente, armazenando-os em um banco de dados. A Figura 2 apresenta o fluxo de dados para armazenar as provas aplicadas nos anos anteriores do ENADE que servirá como uma base de exemplo para treinar a IA durante a geração de novas questões. Primeiro é feito o download das provas já aplicadas que estão disponíveis no site do gov.br (Brasil. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) s.d.). Cada prova é dividida em três arquivos. O primeiro é a prova aplicada, o segundo é o gabarito para questões de múltipla escolha e o terceiro é o padrão de resposta esperado para questões discursivas. Ao baixar os arquivos em PDF, eles passam por um processo ETL utilizando Python, no qual cada questão da avaliação é estruturada de acordo com um padrão de banco de dados relacional, conforme ilustrado na Figura 1.

Ao ler o arquivo da prova no formato PDF utilizando Python, o arquivo é convertido para um texto contendo apenas caracteres. Contudo, a prova não é composta apenas por textos, mas



também por elementos visuais, como imagens, tabelas, linhas e gráficos, que, ao serem convertidos, resultam em sequências de caracteres. Isso impossibilita a preservação da visualização original dessas informações. Dessa forma, o processo de separação das questões do arquivo é realizado em algumas etapas. Inicialmente, é criada uma variável do tipo lista em Python, onde cada linha do arquivo é armazenada em uma posição da lista. Durante esse armazenamento, verificam-se os caracteres iniciais de cada linha, uma vez que o início de uma questão segue um padrão específico: o texto "questão" para questões de múltipla escolha e "questão discursiva" para questões discursivas. Dessa forma, o conteúdo de uma questão é separado a partir da primeira ocorrência do padrão identificado até a última linha antes da identificação do próximo padrão.

Após identificar uma questão, dois processos são executados. Em questões discursivas, a primeira linha contém apenas o título da questão (formado pelo texto "questão discursiva" seguido do número da questão) e todas as linhas subsequentes, até a próxima questão, são consideradas como o enunciado. Em questões de múltipla escolha, há uma terceira etapa, que consiste na separação de cada alternativa de resposta. As alternativas são identificadas pelo padrão de iniciação da linha: uma letra maiúscula (A, B, C, etc.), seguida de dois espaços e, finalmente, do texto da alternativa.

Com base nesses padrões textuais identificados, é possível utilizar expressões regulares que verificam, de forma precisa, os padrões de caracteres descritos. Além disso, estruturas de repetição percorrem o arquivo de linha a linha, permitindo uma separação adequada dos dados. Após identificar cada elemento de uma questão, ela é armazenada em uma variável do tipo lista, utilizando um padrão chave-valor. Dessa forma, é possível, com facilidade, identificar as partes específicas da questão, divididas em número da questão, enunciado e alternativas (quando aplicável)

Após separar corretamente as questões, é realizado um segundo procedimento em Python que percorre a variável lista para determinar o tipo de pergunta (seja múltipla escolha ou discursiva), seu número, enunciado e alternativas. Os dados são, então, armazenados adequadamente no banco de dados, conforme o diagrama apresentado na Figura 1.

O processo ETL compreende o objetivo de armazenar questões que servirão de exemplo para a criação de uma questão pela IA, porém não é capaz de extrair das provas anteriores informações como a disciplina e sub-disciplina abordada na questão e nível de dificuldade. É possível visualizar como é realizada a divisão das questões para armazenamento no banco de dados na Figura 3 para questões discursivas e Figura 4 para questões de múltipla escolha.

Figura 3: Divisão das questões de Multipla Escolha no processo ETL

Número da questão

QUESTÃO 15

Os projetos são um esforço temporário empreendido para criar um produto, serviço ou resultado único.
PMI. Guia PMBOK: Um Guia para o Conjunto de Conhecimentos em Gerenciamento de Projetos. 6. ed. Pennsylvania: PMI, 2017 (adaptado).

Com base na definição de projetos apresentada, avalie as afirmações a seguir.

- I. A criação de um produto único, ou um item do produto, pode representar um possível projeto.
- II. Um serviço único, ou uma capacidade de realizar um serviço, pode representar um possível projeto.
- III. A operação de um software, mesmo que repetidas vezes, pode representar um possível projeto.
- IV. Uma combinação única de um ou mais produtos, serviços ou resultados, pode representar um possível projeto.

É correto apenas o que se afirma em

A I e III.

B II e III.

C II e IV.

D I, II e IV.

E I, III e IV.

Alternativas da questão

Enunciado da questão

Fonte: elaborado pelo autor

Figura 4: Divisão das questões Discursivas no processo ETL

Número da questão

QUESTÃO DISCURSIVA 03

Dados são uma descrição elementar de coisas, eventos, atividades ou transações que são registradas, classificadas e armazenadas, mas não são organizados para carregar qualquer significado específico. Quando os dados são organizados e processados de tal forma que apresentam uma mudança (qualitativa ou quantitativa) tornando-se úteis, temos uma informação. Informações podem ser analisadas para gerar conhecimento tácito ou explícito, experiência, aprendizado acumulado, conforme se aplicam a um problema ou atividade atual.
TURBAN, E. et al. Administração da Tecnologia da Informação: Teoria e Prática. Rio de Janeiro: Elsevier, 2005 (adaptado).

A partir das diferenças explicitadas no texto sobre dados, informação e conhecimento, faça o que se pede nos itens a seguir.

- a) Cite cinco exemplos de dados. (valor: 2,5 pontos)
- b) Cite três exemplos de informação. (valor: 3,0 pontos)
- c) Cite três exemplos de conhecimento. (valor: 4,5 pontos)

Enunciado da questão

Fonte: elaborado pelo autor

Para organização do banco de dados foi adotado o um modelo de nomenclatura informando o tipo do dado armazenado, em sequência se houver uma função específica do campo caso ele seja um código de identificação, por exemplo, e por último o nome que descreve qual o contexto do dado armazenado e seguindo um padrão *snake case* mantendo o nome do campo em letras minúsculas e dividindo cada palavra por um *underline*, como, por exemplo, o campo "int_cd_questao" da tabela "QUESTOES" onde "int" é uma abreviação do tipo de dado inteiro, "cd" é a abreviação informando que o campo armazena um código, e "questao" é o nome do

campo, por tanto é possível associar que aquele campo armazena um código da questão e é do tipo inteiro.

2.3 FLUXO DE EXECUÇÃO

A Figura 5 apresenta a ordem de execução da aplicação. O código react executa scripts SQL para apresentar exemplos de questões já aplicadas no ENADE. Em seguida, o resultado da consulta é adicionado e um prompt é enviado para o Google Gemini através da API. O retorno da IA é tratado no código JavaScript e exibido para o usuário no *front-end* da aplicação.



Fonte: elaborado pelo autor

Para retornar as questões já armazenadas no banco de dados, são utilizados dois modelos de código SQL: um para questões de múltipla escolha e outro para questões discursivas. Na chamada da API do Google Gemini, o prompt utilizado foi construído em quatro etapas. A primeira parte apresenta um breve resumo da função da IA, bem como as regras para a geração de questões, como o uso do modelo Trapezoidal no formato das questões. Onde as alternativas de resposta devem ser organizadas da de menor tamanho para a de maior tamanho para evitar a indução do aluno a selecionar uma alternativa específica. Além disso, deve usar distratores plausíveis (Romão e Sá 2019) para as questões de múltipla escolha, onde as alternativas erradas sejam coerentes com erros comuns dentro do contexto apresentado. A segunda etapa é opcional, sendo necessária somente quando o usuário fornece um texto base para a questão que deve ser gerada. Na terceira etapa, é possível identificar qual a disciplina tema da questão e seu nível de dificuldade, onde foi adotada as seguintes explicações de como a questão deve ser criada com base na dificuldade escolhida, já que na extração das questões de provas anteriores não existe a informação do nível de dificuldade. "Fácil" abrange contextos introdutórios e compreensão básica de conceitos. "Média" é a dificuldade usada para cobrar aplicação prática de conceitos em cenários relativamente simples, buscando resolver problemas reais. Já a dificuldade "difícil" requer interpretação de contextos complexos e o conhecimento profundo dos conceitos das

disciplinas e a capacidade de integrar vários conceitos para resolver problemas. Além disso, há um texto explicativo sobre o tipo de questão a ser criada, dentre as opções discursivas ou múltipla escolha (como interpretação, asserção-razão, resposta múltipla e complementação simples). Por último, a fase final apresenta exemplos de questões que foram aplicadas em edições anteriores do ENADE, resultantes da análise SQL realizada.

Para a elaboração de questões que atendam às necessidades de avaliação do aluno de forma adequada, são utilizadas três habilidades que definem e descrevem a linha de raciocínio e o nível de cobrança requerido. "Interpretar" exige a capacidade de compreender e extrair informações de textos, gráficos, tabelas, diagramas ou enunciados complexos, bem como de coletar informações explícitas ou implícitas para que o estudante possa resolver os problemas apresentados em uma questão. "Raciocinar logicamente" implica estabelecer conexões entre ideias e identificar relações de causa e efeito para propor soluções coerentes, utilizando princípios lógicos e matemáticos, com aplicação do pensamento crítico e dedutivo. "Recordar" demanda a aplicação de conhecimentos prévios, como conceitos teóricos, fórmulas, leis ou definições adquiridos ao longo da formação acadêmica do aluno.

Complementando as habilidades, também são consideradas competências que auxiliam na avaliação das capacidades de interação e resolução de problemas em cenários reais, na tomada de decisões, no trabalho em equipe, entre outros aspectos. Essas competências envolvem elementos cognitivos, sociais, éticos e práticos que orientam a elaboração das questões.

As competências incluem a adaptação, que descreve a capacidade de flexibilizar o raciocínio em diferentes contextos, enfrentando mudanças constantes e desafios no ambiente acadêmico ou profissional. A autorregulação destaca-se como a habilidade de gerenciar o próprio aprendizado, emoções e comportamentos, promovendo autonomia na resolução de problemas e eficiência no desenvolvimento de atividades. A comunicação é essencial para transmitir informações de forma clara e precisa, seja oralmente ou por escrito, permitindo a apresentação eficaz de ideias e o entendimento de diferentes perspectivas.

A liderança é uma competência fundamental para fortalecer o trabalho em equipe, inspirando, motivando e coordenando pessoas. O pensamento analítico permite dividir problemas complexos em partes menores, promovendo a atenção aos detalhes e soluções práticas. Já o pensamento científico baseia-se no método científico para formular hipóteses, interpretar resultados e aplicar ideias inovadoras em contextos de pesquisa.

O pensamento criativo é associado à geração de ideias inovadoras e à proposição de

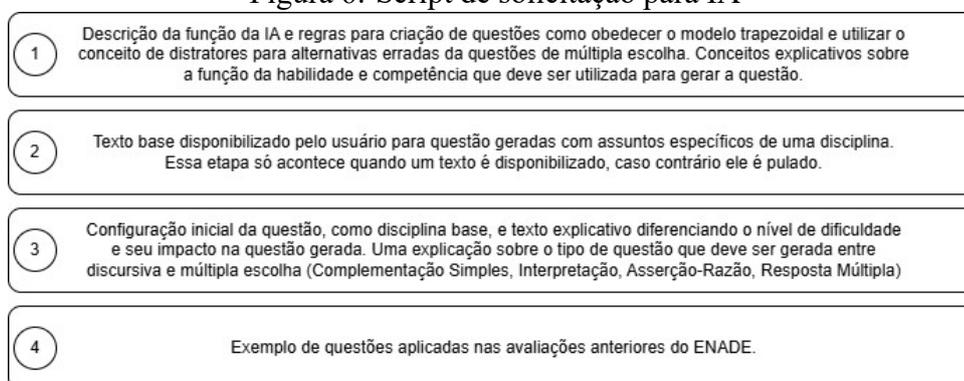
melhorias, enquanto o pensamento crítico foca na capacidade de tomar decisões fundamentadas em evidências consistentes, analisando argumentos e identificando falhas. O pensamento digital refere-se ao domínio de ferramentas e tecnologias digitais para resolver problemas e desenvolver soluções eficientes.

A proatividade reflete a iniciativa para resolver problemas e aproveitar oportunidades de forma eficaz. A competência de resolução de problemas destaca a habilidade de analisar, identificar e solucionar questões de maneira lógica e criativa. A dimensão sociocultural envolve a capacidade de interagir com pessoas de diferentes origens culturais e sociais, promovendo respeito, inclusão e colaboração em contextos globais.

A tomada de decisão é definida pela escolha da melhor solução, com base em evidências e valores, enquanto a ética reflete a importância de agir com responsabilidade e integridade, seguindo princípios éticos e morais.

A Figura 6 ilustra as etapas de construção do prompt aplicando todas as regras que instruem corretamente a criação de uma questão pela inteligência artificial.

Figura 6: Script de solicitação para IA



Fonte: elaborado pelo autor

3 UTILIZAÇÃO

A utilização do ENAQUEST pode ser realizada de forma intuitiva seguindo os seguintes passos:

3.1 Inserção de um texto base

No primeiro momento o usuário tem a possibilidade de digitar ou inserir uma cópia de um texto para a questão de deverá ser gerada, ou pode optar em realizar o upload de um arquivo

em formato .pdf ou .txt já armazenado no dispositivo, essa etapa não é obrigatória para a geração da questão, mas auxilia nas especificações do assunto principal da questão.

A Figura 9 mostra um exemplo de texto inserido manualmente para a geração da questão, onde é possível digitar um texto manualmente ou utilizar teclas de atalho para colar um texto retirado de outra fonte, já a Figura 10 demonstra a função de adicionar um arquivo de texto externo.

Figura 7: Exemplo de texto base digitado

Selecione a configuração da questão a ser gerada:

Os bancos de dados são sistemas organizados que permitem o armazenamento, gerenciamento e recuperação eficiente de informações. Eles podem ser classificados em diferentes modelos, como relacional, **NoSQL**, e orientado a objetos.

No modelo relacional, os dados são estruturados em tabelas que possuem colunas (atributos) e linhas (**tuplas**), conectadas por meio de relacionamentos definidos por chaves primárias e estrangeiras. Esse modelo é amplamente utilizado por sua simplicidade e capacidade de realizar consultas complexas por meio da linguagem SQL (**Structured Query Language**).

Por outro lado, bancos de dados **NoSQL** oferecem maior flexibilidade para lidar com dados sem estrutura rígida, como documentos JSON, valores chave ou grafos. Eles são especialmente úteis para aplicações que exigem escalabilidade horizontal, como redes

Subir arquivo

Disciplina: Cadastrar Disciplina Tipo:

Dificuldade: Habilidade: Competência:

Gerar Questão

Fonte: elaborado pelo autor

Figura 8: Exemplo de texto base como arquivo

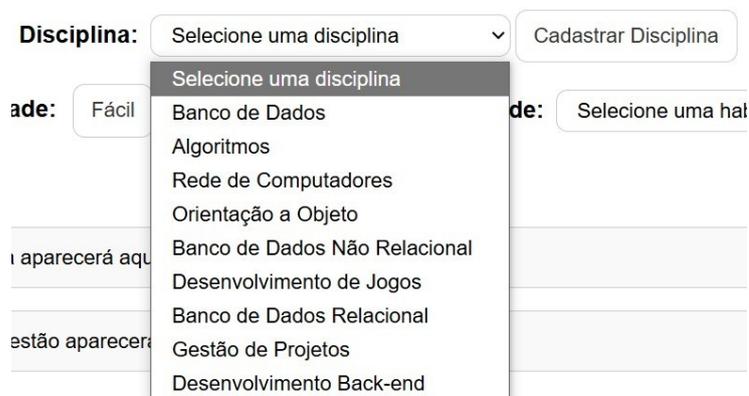
Se desejar cole um texto base aqui

Subir arquivo Texto Base Sobre Banco de Dados.txt Cancelar arquivo

Fonte: elaborado pelo autor

A Figura 11 exibe a lista de disciplinas cadastradas no momento e disponíveis para seleção como tema da questão, e ao lado é possível identificar um botão que ao ser selecionado abre um modal que permite o usuário cadastrar novas disciplinas conforme mostra a Figura 12.

Figura 9: Lista de seleção da disciplina tema da questão



The image shows a web form with a dropdown menu for selecting a discipline. The dropdown is open, displaying a list of options. The options are: Seleção de uma disciplina, Banco de Dados, Algoritmos, Rede de Computadores, Orientação a Objeto, Banco de Dados Não Relacional, Desenvolvimento de Jogos, Banco de Dados Relacional, Gestão de Projetos, and Desenvolvimento Back-end. The form also includes a 'Cadastrar Disciplina' button and a 'Fácil' difficulty level selector.

Fonte: elaborado pelo autor

Figura 10: Modal de cadastro de uma nova disciplina

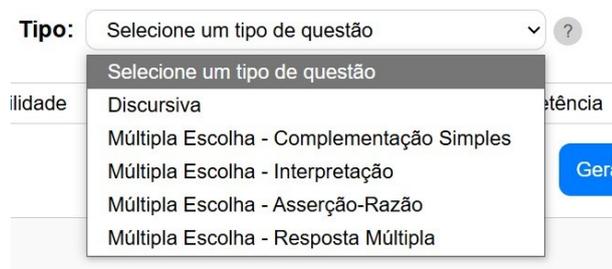


The image shows a modal window titled 'Cadastrar Nova Disciplina'. It contains a text input field labeled 'Nome da nova disciplina' and two buttons: 'Salvar' and 'Cancelar'.

Fonte: elaborado pelo autor

A Figura 13 mostra os tipos de questões que podem ser selecionada, e ao lado da caixa de seleção a um botão com o símbolo de interrogação que ao ser clicado é possível visualizar um exemplo de cada tipo de questão listado conforme a Figura 14.

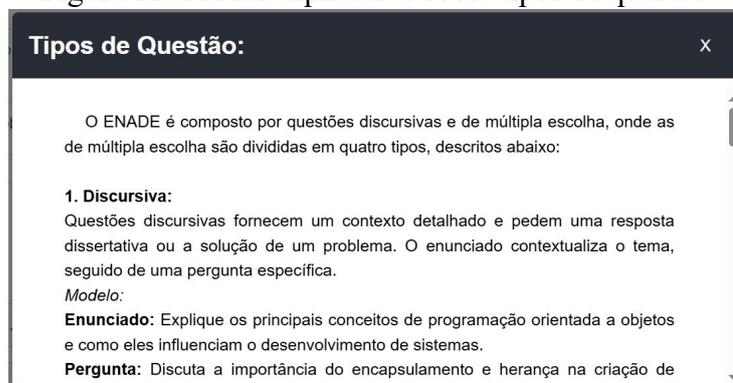
Figura 11: Lista da seleção do tipo de questão



The image shows a web form with a dropdown menu for selecting a question type. The dropdown is open, displaying a list of options: Seleção de um tipo de questão, Discursiva, Múltipla Escolha - Complementação Simples, Múltipla Escolha - Interpretação, Múltipla Escolha - Aserção-Razão, and Múltipla Escolha - Resposta Múltipla. The form also includes a question mark icon and a 'Gerar' button.

Fonte: elaborado pelo autor

Figura 12: Modal explicativo sobre tipos de questão



Fonte: elaborado pelo autor

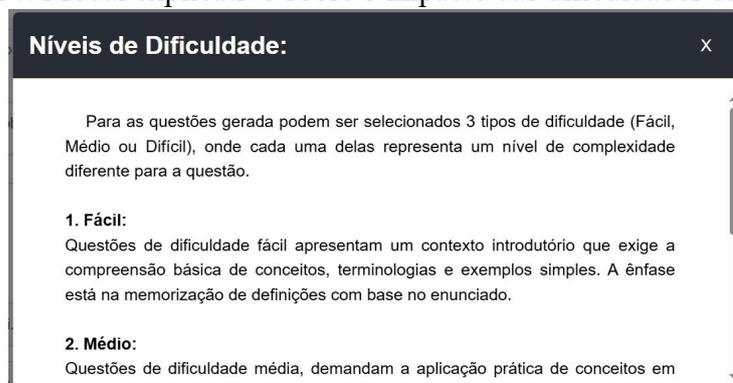
A Figura 15 traz um exemplo da seleção do nível de dificuldade da questão que deve ser gerada, ao lado um botão com o símbolo de interrogação pode ser clicado para exibir um modal que explica como cada nível de dificuldade influencia na criação de uma nova questão conforme a Figura 16.

Figura 13: Seleção do nível de dificuldade da questão



Fonte: elaborado pelo autor

Figura 14: Modal explicativo sobre o impacto das dificuldades da questão



Fonte: elaborado pelo autor

Na Figura 17 é demonstrado uma lista de habilidades que devem ser selecionadas para moldar o objetivo de cobrança da questão, e ao lado um botão de interrogação que exibe um modal explicando o objetivo de cada habilidade conforme a Figura 18.

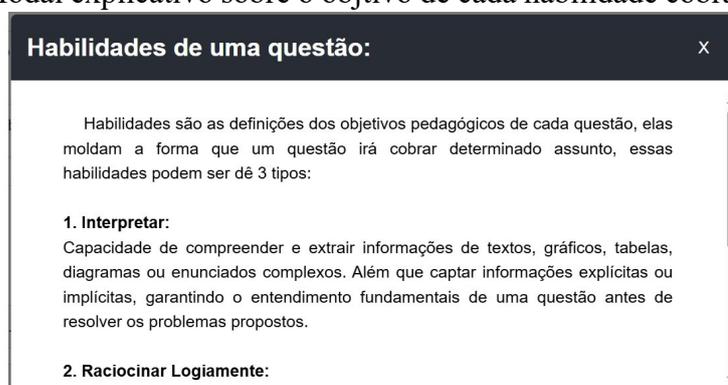
Figura 15: Lista da seleção das habilidades cobras na questão

Habilidade: ?

- Selecione uma habilidade
- Interpretar
- Raciocinar Logicamente
- Recordar

Fonte: elaborado pelo autor

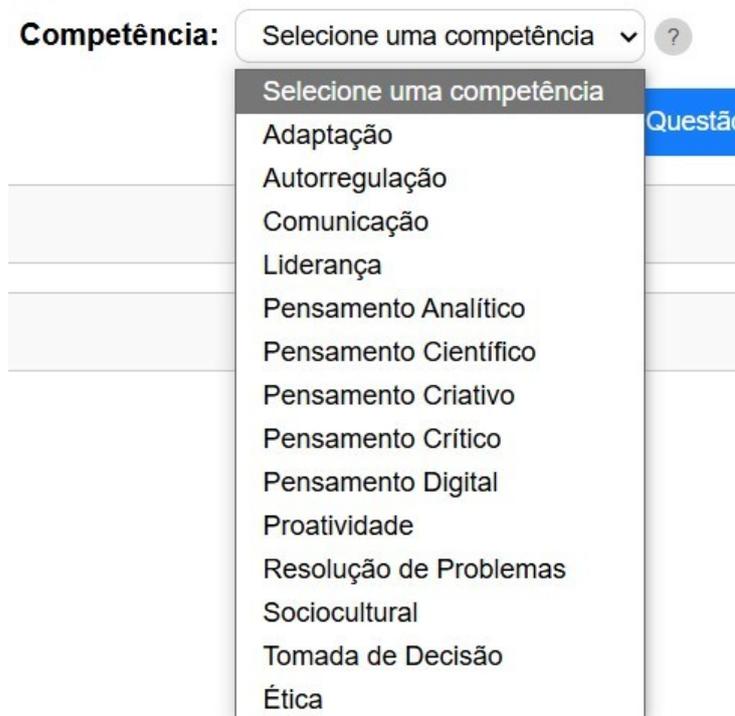
Figura 16: Modal explicativo sobre o objetivo de cada habilidade cobrada na questão



Fonte: elaborado pelo autor

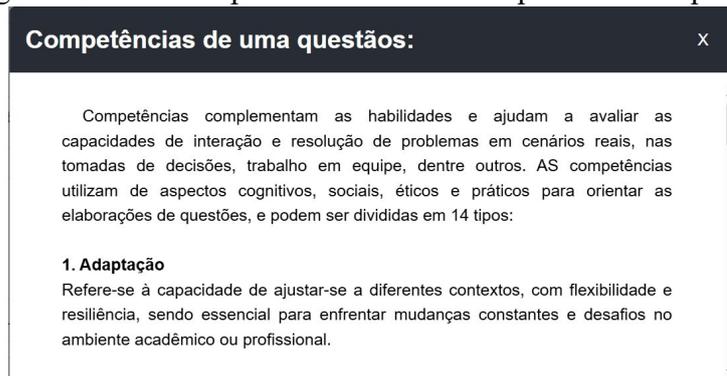
A Figura 19 exibe uma lista das possíveis competências de seleção que impactam no cenário de cobrança da questão que vai ser gerada, e ao lado um botão de interrogação que ao ser clicado exibe um modal explicando o objetivo e impacto de cada competência conforme a Figura 20.

Figura 17: Lista da seleção de competências da questão



Fonte: elaborado pelo autor

Figura 18: Modal explicativo sobre as competências da questão



Fonte: elaborado pelo autor

A Figura 21 representa o resultado de uma questão gerada conforme as configurações realizadas, é possível visualizar a divisão da questão em duas etapas, a primeira sendo a questão completa com enunciado e alternativas, e uma segunda etapa exibindo a alternativa correta e sua justificativa no caso de questões de múltipla escolha, e para questões discursivas o padrão de resposta esperado é exibido, conforme a Figura 22.

Figura 19: Exemplo de uma questão de múltipla escolha gerada

Enunciado:

Uma empresa de e-commerce está construindo um novo sistema de pedidos que precisa lidar com um grande volume de dados, principalmente relacionados a produtos, clientes e histórico de compras. O sistema deve ser escalável para atender ao crescimento do negócio e permitir consultas complexas sobre os dados armazenados.

Pergunta:

****Asserção 1:**** Bancos de dados NoSQL podem ser uma escolha adequada para o sistema de pedidos da empresa de e-commerce, pois oferecem escalabilidade horizontal e flexibilidade para armazenar diferentes tipos de dados.

PORQUE

A) As asserções I e II são proposições verdadeiras, e a II é uma justificativa correta da I.

****Justificativa:****

A asserção I é verdadeira porque bancos de dados NoSQL são projetados para lidar com grandes volumes de dados e oferecem escalabilidade horizontal, característica importante para sistemas que precisam lidar com o crescimento do volume de dados. A asserção II também é verdadeira, pois os bancos de dados NoSQL são frequentemente utilizados para armazenar dados não estruturados, como dados JSON, que são comuns em sistemas de e-commerce. A asserção II justifica a asserção I, pois a flexibilidade para lidar com dados não estruturados é um dos fatores que torna os bancos de dados NoSQL adequados para sistemas de e-commerce.

Fonte: elaborado pelo autor

Figura 20: Exemplo de questão discursiva gerada

Enunciado:

Uma empresa de e-commerce está desenvolvendo um sistema para gerenciar seus produtos e pedidos. A equipe de desenvolvimento precisa escolher o modelo de banco de dados mais adequado para a aplicação. As informações sobre os produtos incluem nome, descrição, preço, categoria, estoque e imagens. Os pedidos, por sua vez, contêm detalhes como data, cliente, itens comprados e endereço de entrega.

A equipe deve considerar que o sistema precisa ser escalável para lidar com um grande volume de dados e transações, além de permitir consultas complexas para análise de vendas e estoque.

Pergunta:

Com base nas características da aplicação e considerando os modelos de bancos de dados relacionais e NoSQL, justifique a escolha do modelo mais apropriado para o sistema.

O modelo de banco de dados relacional seria mais apropriado para o sistema de e-commerce. Apesar de oferecer flexibilidade e escalabilidade horizontal, o NoSQL pode apresentar dificuldades para realizar consultas complexas que envolvam múltiplas tabelas e relacionamentos entre os dados, como a análise de vendas e estoque. O modelo relacional, com sua estrutura rígida e linguagem SQL, permite a definição de relações claras entre os dados, facilitando a realização de consultas complexas, o que é crucial para a análise de vendas e estoque. Além disso, a linguagem SQL é amplamente utilizada e compreendida, o que facilita a manutenção e o desenvolvimento do sistema.

Por outro lado, o modelo NoSQL pode apresentar vantagens em relação à escalabilidade horizontal, o que pode ser importante para lidar com um grande volume de dados e transações. No entanto, a capacidade de resposta para consultas complexas e a complexidade de implementação de transações atômicas em ambientes distribuídos podem ser desafios em relação ao modelo relacional.

Fonte: elaborado pelo autor

4 DESAFIOS ENFRENTADOS

O desenvolvimento do software deste trabalho enfrentou diversos desafios técnicos e conceituais, desde a ideia inicial até a implementação do produto mínimo viável. Etapas como esboços, testes e provas de conceito evidenciaram barreiras cruciais que exigiram atenção e esforço ao longo do processo.

Um dos principais desafios enfrentados foi o tratamento dos dados das provas, que estão disponíveis em formato PDF. Ao utilizar bibliotecas em Python para leitura desses arquivos, os dados não são traduzidos para texto de forma estruturada. Isso ocorre porque as provas apresentam diferentes padrões de organização ao longo do documento. Por exemplo, algumas páginas possuem questões dispostas uma abaixo da outra, enquanto outras são divididas verticalmente, com as questões distribuídas entre o lado esquerdo e o direito dessa divisão. Essas diferenças de estrutura geraram dificuldades na organização correta das questões durante o processamento via código. No entanto, é importante destacar que essas inconsistências representam uma mi-

norria das questões, o que possibilitou a desconsideração desses casos em uma primeira versão funcional do software.

Outro grande desafio foi a interação com modelos de linguagem, como a IA utilizada no desenvolvimento do ENADEQUEST. Essa interação depende de prompts cuidadosamente elaborados, que precisam ser descritivos e objetivos, evitando ambiguidades que possam comprometer a geração de respostas adequadas. Construir prompts que resultem em respostas satisfatórias e alinhadas ao objetivo do software se mostrou um ponto de grande atenção ao longo do desenvolvimento.

A IA utilizada, acessada pela API gratuita do Google Gemini(Google AI 2024), apresentou inconsistências na geração de respostas, exigindo ajustes e novas tentativas. Uma alternativa seria adotar modelos mais robustos, como versões avançadas do ChatGPT(OpenAI 2024), que, apesar de serem pagos, oferecem maior capacidade e consistência.

5 CONSIDERAÇÕES FINAIS

A evolução tecnológica não apenas facilita a solução de problemas, como também requer que se aprenda a lidar com essas ferramentas e utilizá-las de forma eficiente. A inteligência artificial, por exemplo, pode ser aplicada para otimizar tarefas, o que resultará em ganhos significativos de produtividade. O ENADEQUEST teve como objetivo simplificar o trabalho dos professores, acelerando o processo de elaboração de avaliações que avaliem os alunos em temas relevantes das áreas de seus cursos.

Durante o desenvolvimento e os testes da aplicação, foi possível identificar pontos importantes para melhorias futuras. Dentre eles, destacam-se: a utilização de outras ferramentas de inteligência artificial com modelos de processamento diferentes, que possam oferecer resultados variados; a otimização dos prompts de solicitação enviados à IA, para melhorar os resultados obtidos; a classificação das questões aplicadas em edições anteriores do ENADE, para fornecer exemplos específicos e relevantes para os temas abordados nas questões criadas; a funcionalidade de salvar questões e criar um sistema de pesquisas para que questões possam ser reutilizadas e tenham um local único de armazenamento; capacidade de gerar mais de uma questão por vez, ampliando a capacidade de geração da ferramenta; e, por fim, a possibilidade de incorporar imagens às questões, enriquecendo o conteúdo produzido pela aplicação.

O ENADEQUEST (disponível em <https://github.com/leandromaga/enadequest>) demonstra de forma prática o uso da Inteligência Artificial como apoio às atividades educacionais, agi-

lizando a geração de questões sem a necessidade de redigir prompts diretamente. A aplicação oferece configurações simplificadas e acessíveis, tornando o processo mais intuitivo.

Em constante evolução, o ENADEQUEST incorpora melhorias a partir de sua utilização, buscando resolver problemas de forma eficiente e proporcionar uma experiência produtiva ao usuário. Assim, afirma-se como uma ferramenta relevante para integrar tecnologias inovadoras na educação.

Referências

- Bogéa, Tami (2024). “REVISITANDO A TAXONOMIA DE BLOOM: OS CÍRCULOS TAXONÔMICOS DE SIMON PAUL ATKINSON”. Em: *Revista Presença* 10.24, pp. 81–94.
- Brasil. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) (s.d.). *Provas e gabaritos do Enade*. Acesso em: 14 nov. 2024. URL: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enade/provas-e-gabaritos>.
- Educabras (2024). *Educabras - Plataforma Educacional Online*. [Acesso em: 18-Nov-2024]. URL: <https://www.educabras.com/>.
- Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), Instituto Nacional de (2024). *Guia de elaboração e revisão de itens - ENADE*. Acesso em: 5 dez. 2024. URL: https://download.inep.gov.br/bni/enade/guia_de_elaboracao_e_revisao_de_itens.pdf.
- Galhardi, Antonio César Azevedo e Marília Macorin (2013). “Avaliações de aprendizagem: o uso da taxonomia de Bloom”. Em.
- Google AI (2024). *Gemini API Documentation*. [Acesso em: 18-Nov-2024]. URL: <https://developers.google.com/ai/gemini>.
- OpenAI (2024). *ChatGPT API Documentation*. [Acesso em: 18-Nov-2024]. URL: <https://platform.openai.com/docs>.
- Prova Fácil (2024). *Sistema para Universidades | Prova Fácil*. [Acesso em: 18-Nov-2024]. URL: <https://provafacilnaweb.com.br/>.
- Romão, Gustavo Salata e Marcos Felipe Silva de Sá (2019). “Como elaborar questões de múltipla escolha de boa qualidade”. Em: *Femina* 47.9, pp. 561–4.
- Schlickmann, Raphael, Carla Regina Magagnin Roczanski e Paola Azevedo (2008). “Provão X ENADE: uma análise comparativa”. Em.