

VITRALIS: Uma Arquitetura LLM Baseada em RAG para Geração de Planos Educacionais Individualizados (PEIs)

Mikeyas Brito¹, Francisco Emerson Sousa¹,
Gabriel R. S. Matos¹, Stephany R. Gonçalves¹, Maria Clara P. de Sousa¹,
Marciel B. Pereira^{1,2}, Amanda Venceslau^{1,3}, Sttiwe Washington

¹Laboratório de Engenharia de Software e
Sistemas (EngineLab/UFC) – Crateús-CE

²Universidade Federal do Ceará (UFC) - Campus de Crateús

³Instituto UFC Virtual – Universidade Federal do Ceará (UFC)

Abstract. *This article introduces the VITRALIS assistant tool, based on Retrieval-Augmented Generation (RAG), to assist educators in the automated construction of Individualized Education Plans (IEPs). The proposed solution integrates Large Language Models (LLMs) with a technical-pedagogical knowledge base, allowing the retrieval of educational strategies contextualized to the specific needs of each student. The methodology describes the data ingestion flow, the semantic search process in vector databases, and the generation of educational guidelines aligned with the Brazilian National Curriculum Base (BNCC). In this way, the tool contributes to reducing development time, increasing pedagogical consistency, and promoting the personalization and monitoring of the educational process.*

Resumo. *Este artigo apresenta a ferramenta assistente VITRALIS^{1 2}, baseada em RAG para auxiliar educadores na construção automatizada de Planos Educacionais Individualizados (PEIs). A solução proposta integra LLMs a uma base de conhecimento técnico-pedagógica, permitindo a recuperação de estratégias educacionais contextualizadas às necessidades específicas de cada aluno. A metodologia descreve o fluxo de ingestão de dados, o processo de busca semântica em bancos de vetores e a geração de diretrizes educacionais alinhadas à BNCC. Dessa forma, a ferramenta contribui para a redução do tempo de elaboração, aumento da consistência pedagógica, além de favorecer a personalização e o acompanhamento do processo educacional.*

1. Introdução

O Plano de Ensino Individualizado (PEI) é um documento pedagógico que define estratégias, objetivos e adaptações específicas para atender às necessidades educacionais de um aluno [Mascaro 2018]. A educação inclusiva no ensino básico está fortemente associada à elaboração de PEIs consistentes e pedagogicamente bem fundamentados. Embora concebidos para orientar práticas pedagógicas personalizadas, esses programas frequentemente assumem, na prática, um caráter burocrático, comprometendo sua função

¹Repositório: <https://github.com/mikeyasbrito/RAG-Article>

²Vídeo de demonstração: <https://youtu.be/1I3In1PZTTQ>

pedagógica e onerando o trabalho docente. A necessidade de obter e integrar laudos médicos, históricos escolares e diretrizes curriculares impõe uma carga cognitiva exaustiva sobre os responsáveis pedagógicos, o que pode comprometer a precisão e a personalização desses planos [de Albuquerque et al. 2021].

Com o objetivo de automatizar o processo de construção de PEIs, apresenta-se o VITRALIS, uma ferramenta assistente que utiliza Modelos de Linguagem de Grande Escala (LLMs) como uma alternativa para mitigar as limitações envolvidas nesse processo, particularmente associada à técnica de Geração Aumentada por Recuperação (*Retrieval-Augmented Generation* – RAG). Diferente de modelos puramente generativos, o RAG ancora as respostas em bases de conhecimento externas, oferecendo fundamentação técnica [Lewis et al. 2020]. Ao integrar dados específicos do aluno à BNCC, o sistema não substitui o olhar docente, mas atua como um suporte à decisão, reduzindo o esforço operacional e devolvendo ao professor o tempo necessário para o que realmente importa: o cuidado pedagógico.

Este artigo possui a seguinte estrutura. Na Seção 2, abordam-se os trabalhos relacionados. A Seção 3 descreve a ferramenta proposta, sua arquitetura, módulos e funcionamento. E por fim, apresentam-se as considerações finais, limitações, e melhorias futuras da ferramenta na Seção 4.

2. Trabalhos Relacionados

O uso de LLMs na educação tem avançado rapidamente, deixando de ser apenas uma promessa para se tornar uma ferramenta prática de tutoria e gestão. No campo da aprendizagem ativa, projetos como o OBI-UAN [Bastos et al. 2025] demonstram como LLMs podem oferecer feedbacks personalizados em tempo real, guiando alunos na resolução de problemas complexos de programação através de comparação semântica. Essa ferramenta destaca-se como um ponto de partida referencial para o entendimento de como modelos de linguagem podem ser aplicados na resolução de problemas educacionais complexos.

Para garantir que essas interações entre docentes, alunos e modelos sejam fundamentadas em dados reais e não apenas em conhecimento genérico do modelo, a técnica de RAG tem se apresentado como alternativa. O sistema UniChat [Bordin et al. 2025], por exemplo, emprega RAG para consultar documentos institucionais, apresentando uma taxa de sucesso de 76,7% na recuperação de informações durante testes práticos. A avaliação demonstrou que a técnica é altamente assertiva para consultas literais, embora ainda enfrente desafios semânticos com siglas e sinônimos contextuais. Apesar dessas limitações, essa ancoragem em bases de dados vetoriais é o que permite mitigar alucinações dos LLMs, demonstrando seu potencial para reduzir a sobrecarga administrativa e atuar de forma fidedigna no contexto educacional.

Além da tutoria e do suporte informativo, os LLMs têm sido explorados na automação de planejamentos estruturados. O trabalho de [Laranjeira and Bezerra 2025] propõe um sistema para a criação de planos de aula focados em pensamento computacional, alinhados à BNCC e à realidade das escolas públicas. De forma semelhante, o sistema LearnMate [Wang et al. 2025] utiliza a IA generativa para construir planos de aprendizagem personalizados com base nas metas e no ritmo de cada usuário. Ambos os projetos demonstram que a tecnologia pode reduzir significativamente a carga de trabalho docente, embora foquem predominantemente no ensino regular e em cronogramas de

estudo padronizados.

Contudo, apesar desses avanços, o enfoque na educação inclusiva e especial permanece como uma lacuna aberta na literatura atual. Não foram identificados trabalhos que tratem da geração automatizada de PEIs utilizando dados clínicos e pedagógicos reais para ancorar a geração via RAG. Este é o diferencial do VITRALIS: ao aliar LLMs e RAG para processar o contexto de laudos e relatórios específicos, essa solução introduz uma abordagem cooperativa que facilita o desenvolvimento de PEIs assertivos. O objetivo é assegurar que o planejamento não seja apenas uma tarefa burocrática, mas uma estratégia fundamentada que ajude efetivamente no desenvolvimento de cada estudante.

3. VITRALIS

A seguir, será apresentada a arquitetura da ferramenta, detalhando sua concepção e funcionamento.

3.1. Arquitetura da VITRALIS

A arquitetura da ferramenta VITRALIS foi concebida de forma modular, estruturada em camadas interconectadas para construir um pipeline de RAG alinhado às práticas modernas. Toda a lógica de negócio, processamento de dados e orquestração do sistema foi desenvolvida em Python 3.10+, utilizando o framework LangChain³ como o maestro da aplicação. Como ilustrado na Figura 1, o fluxo de dados e as tecnologias empregadas dividem-se em três módulos principais:

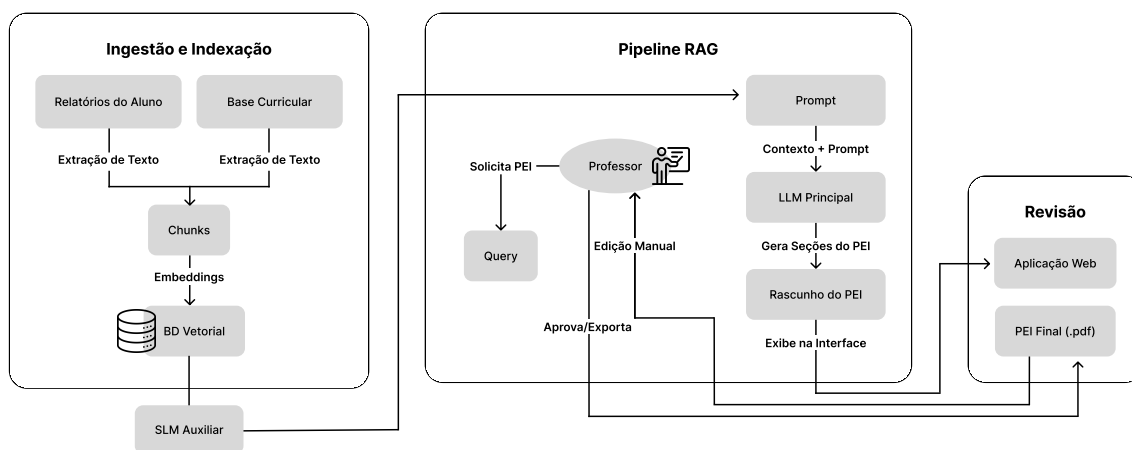


Figura 1. Arquitetura da ferramenta dividida por módulos.

3.1.1. Ingestão e Indexação

Este módulo é responsável pelo processamento e armazenamento dos documentos que fundamentam o sistema, como os relatórios do aluno e as competências da BNCC. O fluxo inicia-se com a extração de texto e sua respectiva fragmentação (*chunking*), processo padronizado pelo LangChain. Em seguida, a API do *Google Embeddings*⁴ atua

³<https://www.langchain.com/>

⁴<https://ai.google.dev/gemini-api/docs/embeddings>

mapeando esses fragmentos para criar representações vetoriais. O armazenamento permanente desses vetores e dos índices dinâmicos de sessão ocorre no *Weaviate*⁵, que atua como o banco de dados vetorial (*BD Vetorial*).

Para lidar com os dados sensíveis dos relatórios, o sistema utiliza uma inteligência distribuída de forma híbrida: um Modelo de Linguagem de Pequeno Porte (SLM), o Phi-3 (3.8b), é executado localmente via Ollama⁶, agindo como um filtro clínico e pedagógico para realizar a triagem e anonimização prévia dos perfis.

3.1.2. Pipeline RAG

A operação contínua do sistema tem início quando o professor solicita a criação de um PEI através da interface web. A requisição consulta o BD Vetorial por similaridade semântica para buscar as informações mais relevantes. Depois, o LangChain, atuando como orquestrador, encapsula essas informações, injetando as variáveis recuperadas em tempo real para formar o conjunto de “Contexto + Prompt”. Uma vez que os perfis já estão seguros e anonimizados pelo Ollama, o LangChain roteia a instrução para o LLM Principal, o modelo Gemini 2.5 Pro⁷ da Google, que assume o controle da síntese pedagógica e gera as seções primárias do PEI.

3.1.3. Revisão

A última etapa consolida a interação do professor com o conteúdo gerado pela IA. Todo o processo é operado na Aplicação Web (Frontend) construída com Streamlit⁸, que exibe o resultado do LLM em um *dashboard* único como um rascunho do PEI. Através dessa interface, o professor tem total liberdade para realizar a edição manual e personalizar o conteúdo gerado. Após a revisão, o usuário aprova o relatório e o sistema exporta o PEI finalizado e revisado em formato .pdf ou .json.

3.2. Interação com o VITRALIS

A Figura 2 apresenta o fluxo inicial da aplicação, dividido em dois passos principais. No **Passo 1**, o usuário interage com a **Tela Inicial**, onde o principal elemento é o menu lateral. Nesse painel, o professor pode selecionar o aluno de interesse e realizar o *upload* de documentos relevantes, como relatórios médicos e pedagógicos. Após o envio dos documentos, o sistema direciona o usuário para o **Passo 2**, correspondente à **Tela de Geração do PEI**. Nessa etapa, o professor pode iniciar o processo de geração por meio de um botão específico. Uma vez acionado, o sistema executa o *pipeline* de processamento, exibindo ao usuário o progresso das etapas, incluindo ingestão, análise e geração do conteúdo.

Após a conclusão do processamento, o sistema apresenta o rascunho do PEI ao usuário, conforme ilustrado na Figura 3. O conteúdo gerado é organizado em seções estruturadas, incluindo Perfil do Aluno, Objetivos Pedagógicos e Estratégias e Recursos, refletindo o padrão adotado em documentos de PEI. Nessa etapa, o professor pode revisar

⁵<https://weaviate.io/>

⁶<https://ollama.com/library/phi3:3.8b>

⁷<https://ai.google.dev/gemini-api/docs/models/gemini-2.5-pro?hl=pt-br>

⁸<https://streamlit.io/>

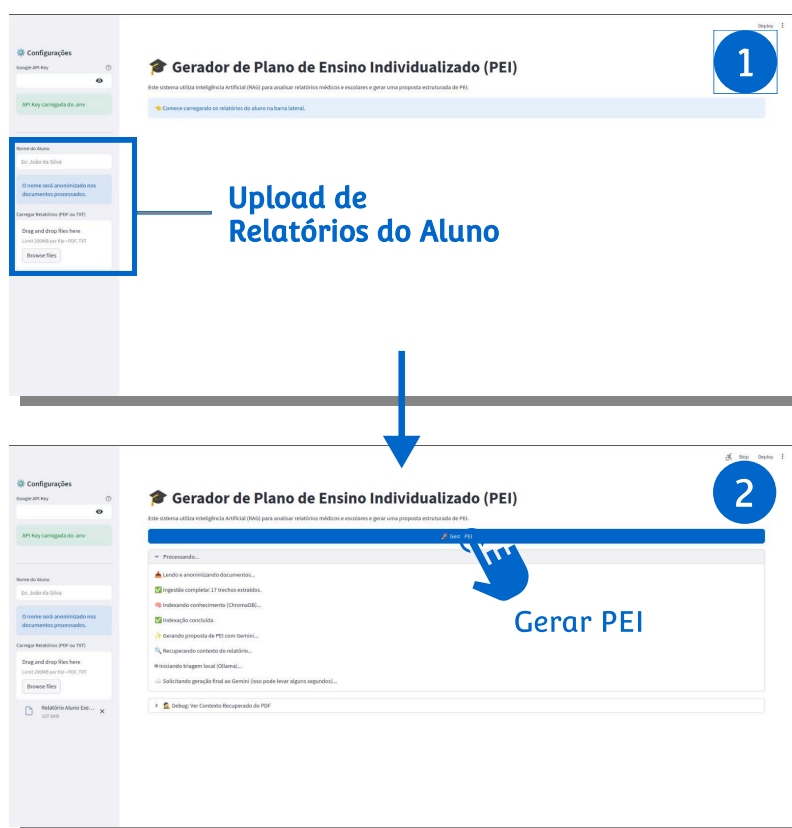


Figura 2. Fluxo de uso do VITRALIS: Tela Inicial e Tela da Geração

o conteúdo gerado, verificando a adequação das informações e realizando ajustes conforme necessário. Além disso, o sistema permite a exportação do rascunho nos diferentes formatos, possibilitando sua posterior edição e integração com outros sistemas.

4. Considerações Finais

Este artigo apresentou o VITRALIS, uma ferramenta assistente baseada em RAG que integra LLMs e uma interface web interativa para apoiar educadores na elaboração de PEIs. A proposta combina informações específicas dos estudantes, como relatórios pedagógicos e clínicos, com diretrizes educacionais, permitindo a geração automatizada de documentos estruturados e personalizados para auxiliar o processo de planejamento inclusivo.

Embora a ferramenta apresente uma implementação funcional e demonstre potencial para apoiar a elaboração de PEIs, este trabalho possui como limitação a ausência de uma avaliação experimental aprofundada no ambiente educacional. Como trabalhos futuros, pretende-se realizar estudos com profissionais da área para analisar aspectos como usabilidade, utilidade percebida e qualidade pedagógica dos documentos gerados. Além disso, busca-se investigar diferentes estratégias de validação das saídas dos modelos de linguagem, visando aumentar a confiabilidade, consistência e aplicabilidade da solução em contextos reais de ensino inclusivo.

Referências

Bastos, Z., Marques, R., Rudan, G., Duarte, M., and Franco, W. (2025). Obi-uan: Um agente para auxiliar nos estudos da olimpíada brasileira de informática. In *Simpósio*

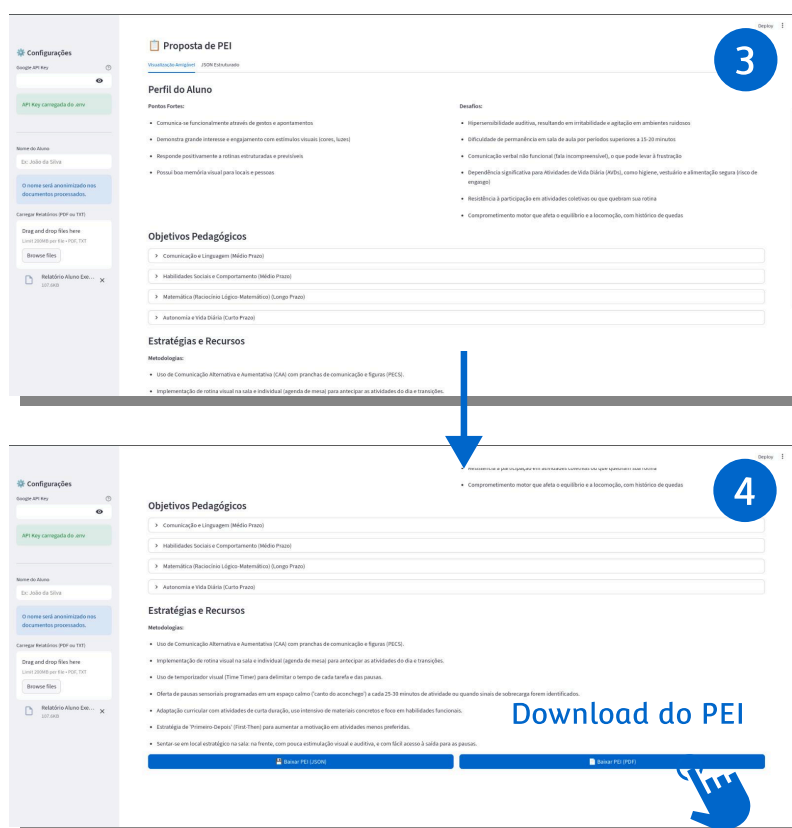


Figura 3. Fluxo de uso do VITRALIS: Rascunho do PEI

Brasileiro de Banco de Dados (SBBD), pages 94–99. SBC.

Bordin, A. C. B., Lunardi, G. M., Piveta, E. K., and Emmendorfer, L. (2025). Unichat: Arquitetura e avaliação de um agente rag de baixo custo para o contexto universitário. In *Simpósio Brasileiro de Banco de Dados (SBBD)*, pages 886–892. SBC.

de Albuquerque, C. C. A. C., da Silva Caldeira, M. C., Alves, A. L., et al. (2021). A construção de planos de ensino individualizados (peis) em um colégio de aplicação. *Cadernos do Aplicação*, 34(2):295–312.

Laranjeira, M. L. and Bezerra, P. T. (2025). Gerador automático de planos de aula sobre tecnologia baseado em llms. In *Workshop sobre Educação em Computação (WEI)*, pages 515–526. SBC.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., et al. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, 33:9459–9474.

Mascaro, C. A. (2018). O plano educacional individualizado e o estudante com deficiência intelectual: estratégia para inclusão. *Revista Espaço Acadêmico*, pages 12–22.

Wang, X. J., Lee, C. P., and Mutlu, B. (2025). Learnmate: Enhancing online education with llm-powered personalized learning plans and support. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pages 1–10.