



**FUNDAÇÃO UNIVERSIDADE FEDERAL DO VALE DO SÃO FRANCISCO
COLEGIADO DE ENGENHARIA DA COMPUTAÇÃO**

Av. Antônio Carlos Magalhães, 510, Country Club, Juazeiro-BA,
CEP: 48.902-300, Fone/Fax: +55 (74) 2102-7636, www.cecomp.univasf.edu.br

**FORMULÁRIO DE MATRÍCULA DE
TRABALHO DE CONCLUSÃO DE CURSO**

Título do projeto de pesquisa: Desenvolvimento de uma Solução de Geração Aumentada por Recuperação (RAG) para Automatizar a Consulta de Estatutos Universitários
Discente: Bruno Matheus Pereira Silva
Orientador(a): Prof.(a) Dr.(a) Jadsonlee da Silva
Coorientador(a) (se houver):
Resumo: Alunos e servidores enfrentam desafios significativos ao buscar informações em documentos institucionais extensos e complexos, como estatutos e regimentos universitários. A busca manual, além de consumir tempo, pode levar a interpretações incorretas. Este projeto propõe a criação de um sistema de inteligência artificial baseado em Modelos de Linguagem Grande (LLMs) para interagir com uma base de conhecimento privada composta pelos documentos da universidade. O objetivo é desenvolver um protótipo funcional fundamentado na arquitetura de Geração Aumentada por Recuperação (RAG), capaz de responder a perguntas de maneira precisa e confiável, citando as fontes originais dos documentos consultados.
Palavras-chave: Inteligência Artificial; Geração Aumentada por Recuperação (RAG); Modelos de Linguagem Grande (LLM); Sistemas de Resposta a Perguntas.
Qualificação do problema a ser abordado: O ambiente acadêmico é marcado pela produção contínua de regulamentos, estatutos e portarias que, embora essenciais, são extensos e de difícil consulta. Métodos tradicionais, como buscas por palavras-chave ou leitura manual, mostram-se insuficientes diante do volume e da complexidade dessas informações. O sistema proposto difere das ferramentas de busca convencionais por empregar processamento de linguagem natural para compreender o sentido das questões formuladas e localizar trechos relevantes dos documentos. Essa abordagem representa um avanço qualitativo no acesso à informação, tornando-o mais ágil, preciso e contextualizado.
Justificativa: Este projeto contribui para a aplicação prática e validação de técnicas de Processamento de Linguagem Natural (PLN), especialmente a arquitetura RAG, em um cenário real. Demonstra como Modelos de Linguagem Grande podem ser usados de maneira segura e controlada para resolver problemas específicos de acesso à informação. A implementação de um sistema de perguntas e respostas baseado nos estatutos e regulamentos da universidade democratiza o acesso à informação, reduz a carga de trabalho de setores administrativos e promove maior autonomia na resolução de dúvidas da comunidade acadêmica.

FUNDAÇÃO UNIVERSIDADE FEDERAL DO VALE DO SÃO FRANCISCO
COLEGIADO DE ENGENHARIA DA COMPUTAÇÃO
FORMULÁRIO DE MATRÍCULA DE TRABALHO DE CONCLUSÃO DE CURSO I

Objetivo a ser alcançado: Desenvolver e validar um protótipo de sistema de inteligência artificial que, utilizando a técnica de Geração Aumentada por Recuperação (RAG), seja capaz de interagir com os estatutos e regulamentos de uma universidade para responder a perguntas de forma precisa e contextualizada.

Objetivos Específicos:

1. Levantamento e pré-processamento de dados: coletar os documentos oficiais da universidade e prepará-los (dividindo-os em trechos, por exemplo) para uso na solução.
2. Implementação da arquitetura RAG: construir o pipeline técnico que inclua um modelo de embeddings, um banco de dados vetorial e a integração com um LLM (como a Gemini API ou GPT).
3. Desenvolvimento da interface: criar uma interface simples e intuitiva (utilizando Streamlit) para que o usuário possa interagir com o sistema, fazer perguntas e visualizar as respostas.
4. Avaliação e validação: realizar testes para avaliar a precisão e a eficácia do sistema, comparando as respostas geradas pela IA com as informações contidas nos documentos originais.

Juazeiro-BA, 15/09/2025.

Prof.(a) Dr.(a) Jadsonlee da Silva

Bruno Matheus Pereira Silva
Discente