

AI-Enhanced Document and Web Content Query System Integrated with Telegram Bot

Deepthi M, Ashwini, Gayathri BS, Jeevitha S, Harsha B R

Global Academy of Technology, Bengaluru, Karnataka, India

Abstract: *This paper presents the development of an AI-enhanced document and web content query system integrated with a Telegram bot interface. The solution aims to simplify information retrieval from large-scale unstructured data sources, such as PDF documents and websites, by leveraging advanced Natural Language Processing (NLP) and vector similarity search techniques.*

Traditional search systems often lack contextual understanding and user-friendliness, leading to suboptimal query responses. This system incorporates Google's Generative AI for semantic embedding, FAISS for efficient similarity indexing, and Streamlit for an intuitive frontend. It enables users to interact with the system through a Telegram chatbot, making the solution mobile-friendly and accessible.

By integrating conversational AI and embedding-based search, this system delivers context-aware responses, addressing the limitations of keyword-based queries and offering an efficient method for intelligent content retrieval. The project demonstrates a novel approach to scalable, user-centric, and AI-powered query systems applicable to education, research, and enterprise documentation management.

This paper details the complete system pipeline from data acquisition and preprocessing to embedding generation and response delivery and evaluates its effectiveness across various user queries. The integration of AI-powered conversation with real-time document querying marks a significant leap in building intelligent, responsive, and mobile-accessible content discovery systems..

Keywords: AI-enhanced document

