iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XI **Month of publication:** November 2024

DOI: https://doi.org/10.22214/ijraset.2024.65247

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

Cognitive Query System Using Generative AI

Lakshay Wadhwani[1], Kushal Gupta[2], Akshat Kumar[3], Biraag S. Prabhakar[4], Dr. Archana Kumar[5]
[1, 2, 3, 4]*Student,* [5]*Professor, ADGIPS, Delhi India*

Abstract: ***The Cognitive Query System using Generative AI is a cutting-edge tool designed to transform the way we retrieve and analyse data from a wide range of sources, including text, images, and documents. Powered by the Google Gemini API and deployed through Streamlit, this system allows users to easily ask questions, process images, and explore document contents via an intuitive, interactive interface. The system comprises three core modules: the Question-Answering Module, which uses advanced natural language processing (NLP) techniques to provide contextually accurate, conversational responses to user queries, facilitating seamless, real-time interactions. The Image Processing and Querying Module enables users to upload images and ask questions related to their content, leveraging computer vision algorithms and AI to interpret and analyse visual data, making it easier to extract valuable insights. The Document Exploration Module supports various formats, including PDFs, Word documents, and spreadsheets, allowing users to quickly ask questions and navigate through sections without manual reading. Deployed online via Streamlit, the Cognitive Query System offers a smooth, real-time experience that does not require installation, making it ideal for remote and collaborative work environments. This innovative solution is designed to streamline workflows and effective outputs in the model.***

I. INTRODUCTION

Recent advancements in artificial intelligence, particularly in generative models, have revolutionized the way learners engage with educational content, offering new possibilities for interactive and personalized learning experiences. The Cognitive Query System using Generative AI, developed with the Google Gemini API, and deployed via Streamlit, is designed to enhance the learning process by enabling dynamic, interactive data retrieval across text, images, and documents. As the demand for reskilling and upskilling continues to rise, particularly among adult learners, traditional educational methods such as videos and quizzes, which often result in passive learning, are increasingly being supplemented with more engaging, active learning techniques. This system combines the capabilities of Cognitive AI and Generative AI to foster deeper understanding by providing contextually rich, reasoned responses to learners' queries, helping to bridge the gap between passive consumption and active knowledge creation.

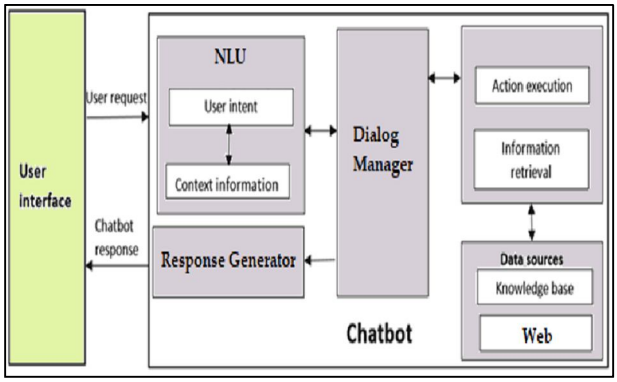
The integration of Generative AI, especially through LLMs, allows the system to provide dynamic explanations and respond to complex, context-specific queries, addressing the challenge of AI's previous inability to fully comprehend and explain the content it processes. This research explores the potential of combining Cognitive AI and Generative AI into a more efficient model for better learning and implementation of various approaches.

II. REVIEW

The integration of generative AI in chatbot systems has brought about a transformative shift in how users interact with educational content. By leveraging Cognitive AI and Generative AI, these chatbots provide dynamic, context-aware conversations that enhance user engagement and facilitate active learning. They go beyond simple question-answering by offering precise, insightful responses across a range of media formats, including text, images, and documents. Through a structured approach to task decomposition, these chatbots can break down complex skills into their components, providing learners with a deeper understanding of the material.

Powered by natural language processing (NLP) and computer vision, these AI-driven systems can address complex queries, guiding users through intricate concepts, and offering real-time, reasoned explanations. This ability not only fosters deeper cognitive engagement but also supports learners in navigating diverse content with ease. Whether for academic or professional development, such chatbot systems are scalable, accessible, and adaptable, making them powerful tools for personalized, interactive learning experiences. Their versatility positions them as valuable resources for a broad range of applications, from education to industry-specific training.

©IJRASET: All Rights are Reserved



(Fig. 1 General Chatbot Framework)

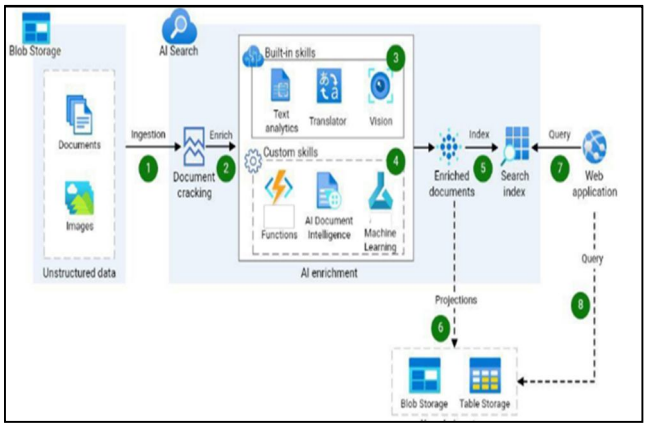
III. METHODOLOGY

Cognitive Search is a powerful cloud-based search service that enables you to search over vast amounts of structured and unstructured data. It involves several key steps:

1) *Data Ingestion and Enrichment:* Unstructured data such as documents and images are ingested from various sources and prepared for search. Document cracking breaks down complex documents into smaller, more manageable chunks. Custom skills and built-in skills like text analysis, translation, and vision further enrich the data by adding metadata and extracting key information. This enrichment process enhances the searchability and relevance of the data.
2) *Indexing:* The enriched documents are organized into an index, a searchable data structure. This index is optimized for efficient search and retrieval. It allows the search engine to quickly locate relevant information based on user queries.
3) *Query Processing:* Users submit queries, which are processed by the search engine. The search engine employs sophisticated algorithms to understand the intent of the query and retrieve relevant results from the index. It considers factors like keywords, semantic meaning, and context to provide accurate and relevant search results.
4) *Result Ranking and Presentation:* The search engine ranks the retrieved results based on relevance and presents them to the user. The results can be displayed in various formats, such as a list, a grid, or a knowledge graph. The presentation of results is tailored to the user's needs and preferences.

Cognitive Search offers several advanced features, including semantic search, faceting, geo-spatial search, and natural language processing. These features enable more sophisticated and accurate search experiences. Semantic search understands the meaning of queries and returns results based on context. Faceting allows users to filter search results based on specific criteria. Geo-spatial search enables searching for results based on geographic location. Natural language processing enables the search engine to understand natural language queries and return relevant results.

Cognitive Search is a versatile and powerful tool that can be used to build intelligent search applications. It offers a variety of features and benefits, including scalability, flexibility, and accuracy. By leveraging Cognitive Search, organizations can improve search efficiency, enhance user experience, and unlock valuable insights from their data.



(Fig.2 Cognitive Query System)

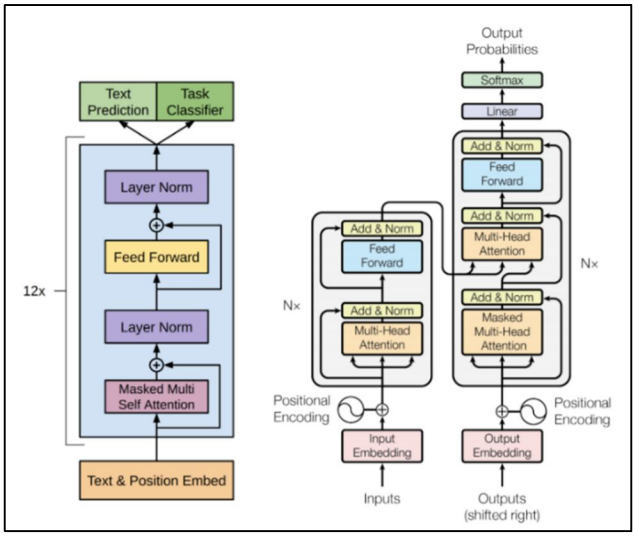
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

IV. SYSTEM ARCHITECTURE

The Cognitive Query System leverages advanced AI capabilities, specifically through the Google Gemini models, to handle diverse queries. Google Gemini is a suite of powerful generative AI models, designed for natural language processing tasks such as question-answering, image recognition, and document exploration. These models, powered by cutting-edge language understanding and generation techniques, enable the system to process and respond to user queries with high accuracy and context-awareness. The integration of Google's generative AI models allows the system to understand and interpret complex requests, providing meaningful, conversational answers across various domains, including text, images, and documents.

The system architecture is built on Streamlit, an interactive, open-source web framework, ensuring an intuitive user experience. To manage sensitive information like API keys securely, the system utilizes python-dotenv for environment variable management. The LangChain library acts as a bridge, integrating the Gemini models with various components for improved query handling, while chromadb and faiss-cpu are employed for efficient document and image retrieval through vector-based searches. LangChain-google-genai allows seamless integration with Google's Gemini models, enhancing the platform's ability to process complex queries related to text, images, and documents. Additionally, the LangChain-community module offers a range of tools for enhanced natural language processing and understanding.

For document processing, the system incorporates libraries like pdf2image, python-docx, and openpyxl to handle PDF, Word, and spreadsheet files, converting them into usable data formats. Pillow is utilized for image manipulation and analysis, while PyPDF2 supports PDF parsing and extraction of text from documents. Combined, these libraries empower the Cognitive Query System to process a variety of inputs—whether they are images, text documents, or structured files—through a cohesive and scalable architecture.



(Fig. 3 Google Gemini API Working)

V. PROBLEM FORMULATION

The task of efficiently querying and retrieving information from diverse data sources—such as images, documents, and text—presents a significant challenge. Traditional methods of information retrieval often require specific tools or expertise to handle different media types, resulting in a fragmented user experience. With the increasing demand for accessible, intelligent systems capable of processing complex, multi-format queries, there is a clear need for a unified approach that can seamlessly integrate multiple AI-driven functionalities into a single platform.

This problem becomes more complex when considering the limitations of current AI models, such as the inability to understand context fully or handle nuanced queries across various domains. Moreover, existing solutions often struggle with scalability, efficiency, and real-time data processing, especially when dealing with large volumes of data or high-resolution content. To address these issues, a system is needed that can process a wide array of inputs, from natural language queries to images and documents, while providing accurate, context-aware responses in an intuitive and accessible manner.

The goal of this project is to design and implement a generative AI-powered Cognitive Query System that leverages state-of-the-art AI models, such as Google Gemini, to enable users to interact with data in a more natural and efficient way.

©IJRASET: All Rights are Reserved

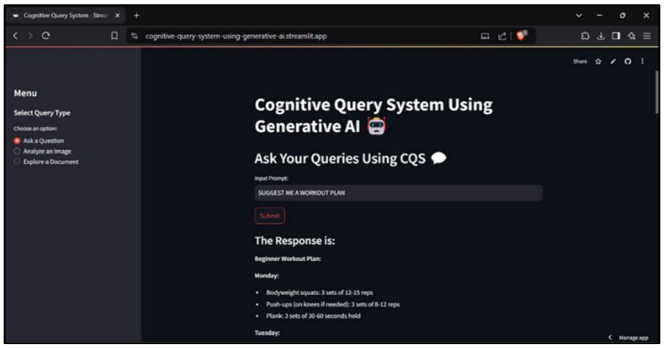
By integrating various AI tools into a cohesive platform, the system aims to provide a solution for information retrieval that is flexible, user-friendly, and capable of handling complex queries across multiple formats and domains.

VI. RESULT

The prototype of the Cognitive Query System, powered by Google Gemini models and built on Streamlit, has demonstrated promising results in terms of versatility, usability, and performance. The system successfully integrates multiple AI functionalities into a cohesive platform, allowing users to ask questions, analyse images, and explore documents, all within a single interface. By leveraging the Google Gemini generative AI models, the system provides accurate, contextually aware answers to a wide range of queries, showcasing the capability of large language models (LLMs) to interpret and respond to complex inputs across different media types.

The system's performance in handling image queries has been validated through real-time object recognition and analysis, which can detect and interpret text or objects within uploaded images. This functionality is particularly useful in scenarios where visual data needs to be processed quickly, such as in educational or business settings. In addition, the document exploration feature allows users to upload PDFs, Word documents, and spreadsheets, after which the system extracts relevant text and answers queries based on the content. This capability is especially valuable for professionals who need to retrieve insights from large volumes of text-based data.

The user interface, built with Streamlit, has been optimized for simplicity and interactivity, ensuring that users from various backgrounds—whether technical or non-technical—can navigate the platform with ease. The integration of various libraries, such as pdf2image, python-docx, openpyxl, and PyPDF2, has further enhanced the system's ability to handle diverse file types and formats, demonstrating the scalability and robustness of the architecture.



(Fig. 4 Working Prototype)

VII. CONCLUSION

The Cognitive Query System using Generative AI (Gemini) combines multiple AI-powered tools into a single platform, enabling users to ask questions, analyse images, and explore document content. Built on the Streamlit framework, it simplifies information retrieval, making it accessible and efficient for users across various sectors, such as education, business, and research. The system allows users to gain insights and answers without needing specialized knowledge, streamlining tasks like document exploration and object identification in images.

This integration of generative AI enhances productivity by providing an intuitive, user-friendly experience. Future updates could include real-time data processing, advanced image recognition, and support for audio/video analysis, further broadening its capabilities. Overall, the Cognitive Query System offers a flexible, AI-driven solution, setting the stage for future innovations in cognitive computing and transforming how users interact with information.

VIII. FUTURE SCOPE

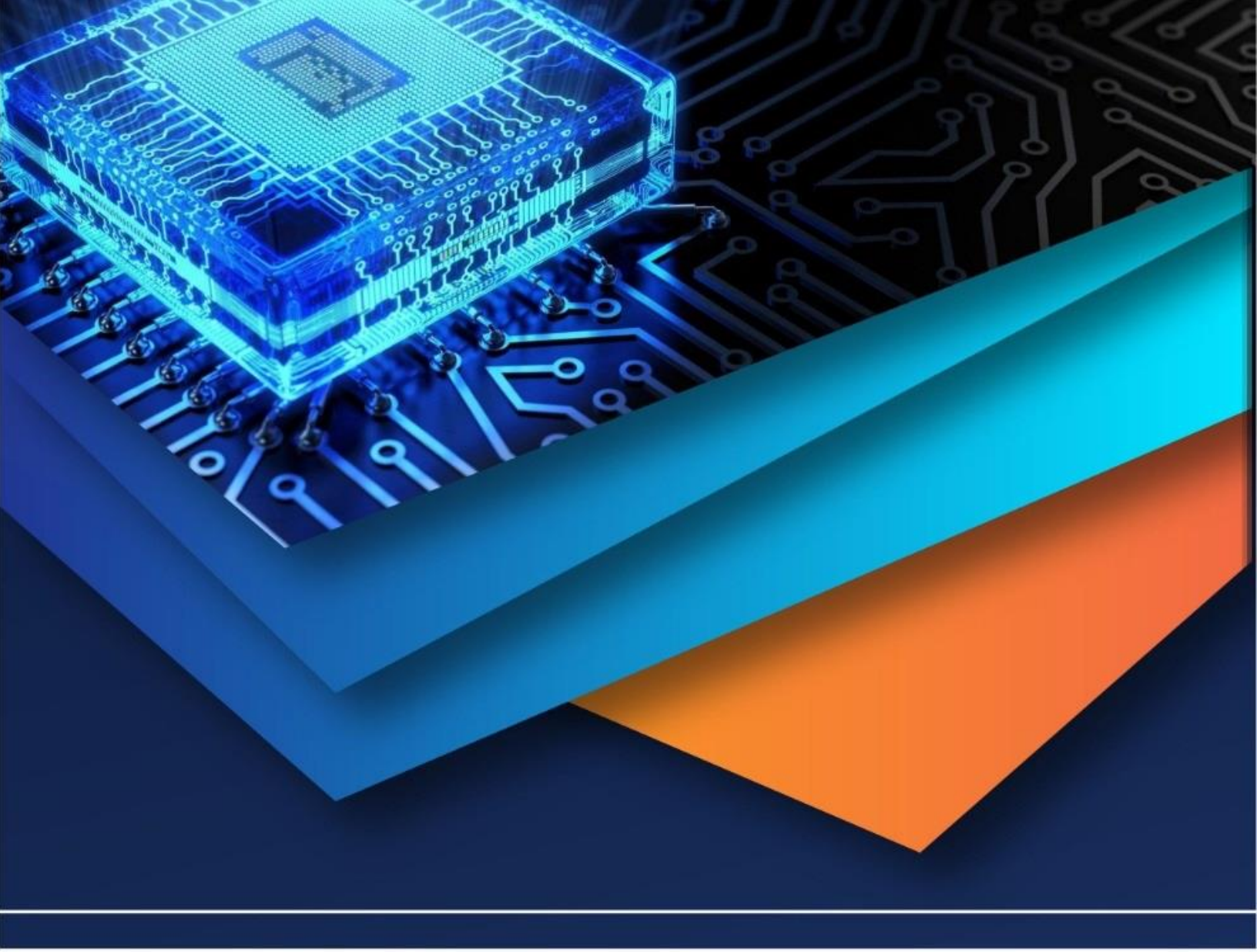
The Cognitive Query System has several future growth areas, including –

1) Advanced Model Integration, where future iterations could integrate newer generative AI models to improve accuracy and context-awareness, particularly for complex queries.
2) Cross-Platform Compatibility would allow the development of mobile and web-based versions, extending accessibility for users across multiple devices.

3) Enhanced Image Processing Capabilities could optimize image recognition models for high-resolution images, improving performance in fields like healthcare and inventory management.
4) Improved Document Parsing for Complex Structures would enable the system to handle complex layouts, such as tables and charts, benefiting sectors like legal and financial services.
5) Multilingual Query Support would make the system more globally applicable, particularly in education and industry.
6) Data Privacy and Security Enhancements would implement additional security protocols and privacy measures, such as encrypted storage and GDPR-compliant practices, ensuring sensitive information is protected.
7) Integration with External Databases and APIs could enhance the system's ability to provide more comprehensive, data-driven responses.
8) AI-Driven Analytics and Insights would incorporate AI-driven analytics for deeper insights, trend analysis, and predictive capabilities, aiding decision-making in business and research.

REFERENCES

[1] Vaid Devlin, J., et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of NAACL. Available at: https://arxiv.org/abs/1810.04805
[2] Vaswani, A., et al. (2017). Attention is All You Need. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS). Available at: https://arxiv.org/abs/1706.03762
[3] Zhong, V., et al. (2017). Seq2SQL: Generating Structured Queries from Natural Language using Reinforcement Learning. Available at: https://arxiv.org/abs/1709.00103
[4] Microsoft Project Scheduling and Gantt Chart Basics. Documentation and guides available at: https://support.microsoft.com/
[5] OpenAI - Introduction to Generative AI Models. Available at: https://openai.com
[6] Google Gemini API: Provides NLP, image recognition, and generative AI capabilities. Documentation available at Google's API documentation portal.
[7] Streamlit: The open-source framework for creating interactive web applications in Python. Available at: https://streamlit.io



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)