

INTRO TO AI: PROJECT REPORT/PROPOSAL



Submitted By:

- Hassan Bin Saqib (231170)
- Awais Khan (231172)
- Kamran Ahmed (231232)
- Abdul Moez (231228)

BSAI-IV-F-23-A

Submitted To: Mam Umme Sadima

Submission Date: 23th April, 2025

Department of Creative Technologies (CT)
Air University Islamabad

Table of Contents

INFORMATION RETRIEVAL: PROJECT PROPOSAL	1
Submitted By:.....	1
Project Title:.....	3
“Deen-AI: An Islamic Knowledge Assistant Using Retrieval-Augmented Generation (Rag)”	3
Introduction	3
Objectives.....	3
Project Overview.....	4
Functionalities / Proposed Features	4
Reason for Selection	5
Scope	5
Tools & Technologies used	5
Project Timeline	6
Conclusion	6

Project Title:

“Deen-AI: An Islamic Knowledge Assistant/AI Agent using Retrieval-Augmented Generation (RAG)”

Introduction

In the modern digital age, accessing authentic Islamic knowledge can be difficult due to scattered sources and misinformation. With advancements in Artificial Intelligence and Natural Language Processing, we aim to build an Islamic assistant that perceives user Islam related queries, reason and generate accurate, reference-based responses as This assistant/agent will not only provide answers but also include proper references such as Quran Ayat or Hadith number to ensure authenticity. For this purpose, we'll train a RAG LLM Model on Dataset which will contain Quranic Verses and Hadith Books such as Sahih Bukhari & Sahih Muslim, and aim to respond according to the user's query.

Objectives

- We'll design and implement a retrieval-based question-answering system using RAG (Retrieval-Augmented Generation).
- We'll train the system on authenticated Islamic sources such as the Qur'an and Hadith books: Sahih Bukhari & Sahih Muslim.
- We'll generate responses that include direct references (Ayah or Hadith) along with proper citations like Surah, Ayah number, or Hadith number.
- We'll demonstrate how the agent's architecture: Combining perception (user query), reasoning (retrieval + generation), and action (response), can be used to solve real-world problems in AI.

- We'll keep the system scalable, allowing it to potentially expand into other religious or knowledge-based applications in the future.
-

Project Overview

This project will involve the development of a Quranic and Hadith-based assistant that takes a user's question in natural language and responds with a well-structured, AI-generated answer backed by retrieval of verified texts. The **core of the project** revolves around developing an **AI agent** that interacts with its environment by **perceiving** a user's question, **reasoning** over authenticated Islamic sources using the RAG (Retrieval Augmented Generation) architecture, and acting by **generating** a response. This mirrors the decision-making process of an intelligent agent. Initially, we'll focusing on giving response to user's queries in English, but there are possibilities to expand to Urdu or Arabic later.

The Islamic content will either be sourced via publicly available APIs or, if not feasible, scraped from reliable Islamic websites using scraping tools. Throughout the process, care will be taken to maintain the accuracy and integrity of the religious texts used.

Functionalities / Proposed Features

- Input: Natural language query from the user.
- Backend retrieval system to fetch relevant Qur'anic verses or Hadith.
- Generation of an answer using the retrieved content.
- References included in the final output (e.g., Surah Number, Ayat Number, Hadith Book, etc).
- Categorization of queries (e.g., prayer, fasting, behaviour, etc.) if time allows).
- If everything goes as planned, and we still have time then we'll make front-end for this and deploy on web for easier interaction.

Reason for Selection

We chose this idea because it allows us to explore the core functionalities of agents and retrieval systems in a meaningful or real-life context. There are only a few systems currently available like Islamic 360 that combine the power of modern AI with religious knowledge in a user-friendly and reference/evidence-based way. This project also gives us a chance to explore how **retrieval systems combined with Agents** (that we studied throughout this course as well), work in practice, alongside foundational NLP tools.

Scope

- The project will initially cover English-language queries related to general Islamic concepts.
 - The scope is limited to retrieving content from the Qur'an and authentic Hadith collections (Sahih Bukhari, Sahih Muslim).
 - If time allows, we may add additional books of Hadith or multilingual capabilities.
 - Deployment of the system on a web or application is optional but intended if everything runs smoothly and our time allows us.
-

Tools & Technologies used

We are in the learning phase and exploring tools side by side with implementation, so there might be minor changes depending on what fits best. However, the planned stack includes:

- **Programming Language:** Python
- **IDE:** Visual Studio Code or Google Colab
- **Retrieval:** FAISS / BM25 / Elasticsearch (TBD based on testing)
- **Text Embedding:** Sentence Transformers (all-MiniLM-L6-v2)
- **LLM Integration (if needed):** OpenAI, Gemini's free version or Local LLMs

- **Web Scraping (if needed):** Selenium + BeautifulSoup
- **APIs:** Quran.com API, HadithHouse API (they're functional for our needs)
- **Frontend (optional):** Deployed Web or Application (to be decided)

We'll prefer using APIs to fetch verified text but will switch to scraping methods if those APIs are unreliable or do not serve our use case.

Project Timeline (we're currently at Week 4)

Week	Milestone
Week 1	Finalize project scope, gather and clean initial datasets, explore APIs
Week 2	Start building basic retrieval system and structure data (Qur'an, Hadith)
Week 3	Implement embedding + retrieval (FAISS/BM25), test query results
Week 4	Integrate generation component, form RAG pipeline
Week 5	Polish outputs, include reference formatting, fix bugs
Week 6	Optional: Build frontend, deploy system, final testing

Conclusion

This project aims to practically apply our AI & Information Retrieval knowledge in a way that not only fulfills academic requirements but also serves a noble cause. It aligns closely with our AI curriculum while encouraging deeper exploration of tools like **embedding models, search indexes, and natural language generation**. If successfully implemented, this assistant can grow into a powerful platform that helps users access Islamic knowledge with clarity and credibility.