MCQ Generator using NLP & AI – Project Documentation

1. Introduction

The MCQ Generator using NLP & AI is a web-based application designed to automatically generate Multiple Choice Questions (MCQs) from textual content. The system supports both classic NLP-based question generation and AI-assisted MCQ generation with cost-aware and token-optimized design.

This project is intended for academic learning, NLP demonstrations, and educational use cases.

2. Problem Statement

Creating MCQs manually from large study materials is time-consuming and inefficient. Existing tools either rely entirely on AI (high cost, low explainability) or lack flexibility.

This project solves the problem by combining explainable NLP logic with optional AI support.

3. Objectives

- Automatically generate MCQs from documents and Google Drive links
- Provide both manual NLP-based and AI-based generation modes
- Optimize AI token usage and cost
- Keep the system beginner-friendly and deployable on free tiers

4. Features

- Upload PDF, TXT, and DOCX files
- Generate MCQs from Google Drive links
- Manual NLP-based MCQ generation using spaCy
- AI-based MCQ generation with summarization
- Answer reveal on button click
- PDF download of generated MCQs
- Modular and scalable backend architecture

5. System Architecture

Frontend: HTML, CSS, Jinja Templates
Backend: Flask (Python)
NLP Engine: spaCy
AI Engine: Google Gemini API
Deployment: Render with Gunicorn

6. Working Methodology

Step 1: User uploads files or provides a Google Drive link
Step 2: Text is extracted from the input source
Step 3: User selects Manual or AI mode
Step 4: Manual Mode → spaCy extracts nouns and sentences
Step 5: AI Mode → Text is summarized, then MCQs are generated
Step 6: Results are displayed and can be downloaded as PDF

7. AI Token Optimization Strategy

To reduce cost and improve performance, the application summarizes extracted text locally before sending it to the AI model. AI-generated MCQs are limited to a maximum of 15 questions.

8. Limitations

- Works best with text-based PDFs
- Google Drive files must be public
- English language only
- AI generation limited to 15 MCQs

9. Future Enhancements

- Support for more languages
- Difficulty-level based MCQs
- Export to CSV and DOCX
- Improved distractor generation

10. Conclusion

The MCQ Generator using NLP & AI demonstrates a balanced approach between classic NLP techniques and modern AI models. It is cost-aware, explainable, and suitable for academic and educational use.

11. Author

Ishu Agrawal
B.Tech Computer Science (AIML & IOT)