Data Collection and Preprocessing Phase

Date	19 July 2024
Team ID	SWTID1720099578
Project Title	Inquisitive: A Multilingual AI Question Generator
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

Section	Description
Project Overview	Inquisitive is a multilingual AI question generator designed to create questions from user-provided text. The project aims to facilitate automated question generation in multiple languages, enhancing educational tools, content creation, and user engagement.
Data Collection Plan	Data will be collected from user inputs, which can be in various languages. These inputs will be processed in real-time to generate relevant questions.
Raw Data Sources Identified	**Source Name:** User-provided text input **Description:** Text input provided by users through the Streamlit interface, which can be in multiple languages. **Location/URL:** Streamlit application interface **Format:** Text **Size:** Variable (dependent on user input) **Access Permissions:** Public (any user can input text)

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
User-provided text input	his data source consists of text input provided by users in various languages through the Streamlit interface. The text can vary in length and complexity, and it forms the basis for generating questions using the generative AI model.	Streamlit application interface	Text	Variable (dependent on user input)	Public