

## DATA COLLECTION AND PREPROCESSING PHASE

<b>Date</b>	31 Jan 2025
<b>Team ID</b>	LTVIP2026TMIDS83843
<b>Project Name</b>	flavour fusion: ai-driven recipe blogging
<b>Maximum Marks</b>	2 Marks

### Data Collection Plan

Section	Description
<b>Project Overview</b>	<ol style="list-style-type: none"> <li>1. The project “flavour fusion: ai-driven recipe blogging” aims to generate structured recipe blogs using a pre-trained generative AI model.</li> <li>2. The system takes user input (recipe topic and word count) and produces structured output including Ingredients, Instructions, and Tips.</li> </ol>
<b>Data Collection Plan</b>	<ol style="list-style-type: none"> <li>1. This project does not use a traditional dataset.</li> <li>2. Data is collected dynamically through user input and processed using the pre-trained Gemini 2.5 Flash model via API integration.</li> </ol>
<b>Raw Data Sources Identified</b>	<ol style="list-style-type: none"> <li>1. Gemini 2.5 Flash API – Pre-trained generative model used for text generation.</li> <li>2. User Input – Recipe topic and word count entered via Streamlit interface.</li> </ol>

## Raw Data Sources

Source Name	Description	Location	Format	Access
Gemini 2.5 Flash API	Pre-trained generative large language model	Google AI Platform	Text Output	API Key Required
User Input	Topic and word count provided by user	Streamlit UI	Text	Public