Data Collection and Preprocessing Phase

Date	28 November2024
Team ID	739684
Project Title	Silver screen Ai:Empowering movie conversations with RASA NLU
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	Silver Screen AI leverages the power of Rasa NLU to create an intelligent conversational platform that enhances the movie-watching experience. By facilitating natural language understanding, this AI assistant can answer queries, provide recommendations, and engage users in discussions about movies.Silver Screen AI: Focuses on building a conversational AI for movie enthusiasts using Rasa NLU to facilitate natural, engaging dialogues about movies and related information.
Data Collection Plan	Silver Screen AI: Search for movie-related datasets, user interaction logs, and metadata for training conversational agents.

Raw Data Sources Identified	Silver Screen AI: MovieLens, The Movie Database (TMDb) API, Open Movie Database (OMDb), user conversations (if available), and Kaggle datasets.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
MovieLens Dataset	A comprehensive dataset containing movie ratings, user preferences, and metadata, suitable for training recommendation systems.	`https://grouplens.org/datasets/movielens/`	CSV	25 mb	Public