# - [00:00 - 00:28]: Introduction to the continuation of the Lang Chain series focusing on the RAG pipeline (Retrieval Augmented Generation) and its importance in solving use cases using LLM models.

- \*\*[00:28 - 01:00]\*\*: Discussion on the components of a RAG pipeline, starting with loading data sources and the importance of being able to query different types of data sources.

- \*\*[01:00 - 01:15]\*\*: Mention of implementing the components from data ingestion to query vector store in the upcoming videos.

- \*\*[01:15 - 01:50]\*\*: Explanation of the process of loading, transforming, and embedding data, including the concepts of feature engineering, breaking data into smaller chunks, and converting chunks into vectors for storage in a vector database.

- \*\*[01:50 - 02:35]\*\*: Discussion on the importance of dividing data into smaller chunks based on context size for efficient querying in LLM models.

- \*\*[02:35 - 02:61]\*\*: Setting up the project in Versus code and creating an IPYNB file for the RAG implementation.

- \*\*[02:61 - 03:21]\*\*: Installation of IPy kernel for Jupyter notebook kernels and resolving any installation issues.

- \*\*[03:21 - 03:49]\*\*: Initiating the coding process by importing document loaders for reading from various sources like PDF, text files, and web pages.

- \*\*[03:49 - 04:53]\*\*: Demonstrating the process of loading a text file using a text loader and reading its content.

- \*\*[04:53 - 05:19]\*\*: Setting up environment variables and importing necessary libraries for further data loading processes.

- \*\*[05:19 - 06:26]\*\*: Loading data from a web page using a web-based loader and configuring parameters for parsing the content.

- \*\*[06:26 - 06:54]\*\*: Creating a loader for reading from a PDF file and loading the content into the system.

- \*\*[06:54 - 07:58]\*\*: Exploring different data ingestion techniques for various file types like Excel, README files, and directories.

- \*\*[07:58 - 08:20]\*\*: Transitioning from data loading to the transformation phase and the significance of transforming data for further processing.

- \*\*[08:20 - 09:30]\*\*: Introduction to text splitting techniques using Langchain's text splitter for dividing PDF documents into smaller chunks.

- \*\*[09:30 - 10:52]\*\*: Implementing text splitting to break down PDF documents into manageable chunks for processing.

- \*\*[10:52 - 11:25]\*\*: Discussion on the importance of converting text into vectors using embedding techniques like OpenAI embeddings.

- \*\*[11:25 - 12:35]\*\*: Introduction to vector stores and the process of storing text embeddings in databases like Chroma DB for efficient querying.

- \*\*[12:35 - 13:55]\*\*: Setting up the vector store using Chroma DB and storing document embeddings for retrieval.

- \*\*[13:55 - 15:19]\*\*: Querying the vector database for similar results based on a specific query and displaying the retrieved information.

- \*\*[15:19 - 16:38]\*\*: Demonstrating the retrieval of information from the vector database based on different queries and displaying the relevant results.

- \*\*[16:38 - 17:09]\*\*: Exploring the retrieval of specific information from research papers using the vector database.

- \*\*[17:09 - 17:35]\*\*: Implementing vector storage in Fyze database and comparing it with Chroma DB for storing and querying embeddings.

- \*\*[17:35 - 18:13]\*\*: Summarizing the process of loading, transforming, embedding, and querying data in the RAG pipeline and hinting at future discussions on retrievers and chains.

- \*\*[18:13 - 18:20]\*\*: Conclusion and closing remarks on the video content and upcoming topics in the series.