

Agentic RAG Pipeline – Test PDF

This document is designed to test an agentic Retrieval-Augmented Generation (RAG) pipeline. It contains text, a table, and a chart description. The goal is to verify whether the pipeline can parse, embed, retrieve, and reason over multimodal content.

Section 1: Concept Overview

An agentic RAG system combines document parsing, vector search, and large language models. The agent retrieves relevant chunks from a vector database and synthesizes an answer using both text and structured data.

Section 2: Sample Table

Component	Purpose	Technology
PDF Parser	Extract text/images/tables	PyMuPDF / Unstructured
Vector DB	Semantic search	ChromaDB
LLM Agent	Reasoning & summarization	Gemini

Section 3: Chart Description

Figure 1 (not shown) represents a bar chart comparing retrieval accuracy across different vector databases. ChromaDB shows higher accuracy compared to baseline keyword search.

Section 4: Example Queries

- What is an agentic RAG pipeline?
- Which components are used in the system?
- Summarize the role of the vector database.