

This is a sample document for testing the RAG system.

It contains some basic information about RAG and LLMs.

Retrieval Augmented Generation (RAG) is an AI framework that combines retrieval mechanisms with generative models to produce more accurate and

contextually relevant outputs.

Large Language Models (LLMs) like TinyLlama are powerful but can hallucinate.

RAG helps by providing external knowledge as context.