

Pakistani News RAG: Architecture and Logic

PAKISTANI NEWS RAG: ARCHITECTURE AND LOGIC

1. High-Level Flow

The system follows a standard RAG pipeline tailored for newspaper data. The flow is as follows:

1. **Ingestion**: CSV files from ``data/`` are loaded. Metadata (newspaper, date) is extracted from filenames.

2. **...**

2. Technical Logic

Chunking Strategy

In the
articl
embe
cont

Retrieval Mechanism

We u
sear
quer
prom

Pakistani News RAG: Architecture and Logic

3. The Math: Vector Embeddings

Embedding Model

We u

Vector Search (FAISS)

The r

4. Code Walkthrough

``data_loader.py``

* `loa
name

``vector_store.py``

* `bu

``rag_engine.py``

* `ge
that i

Pakistani News RAG: Architecture and Logic

5. Execution Steps

1. **Environment Setup**: Install dependencies via ``pip install -r requirements.txt``.

2. ****A**