1️⃣What is RAG?

RAG stands for **Retrieval-Augmented Generation**.

- Imagine a **super smart assistant** (LLM like ChatGPT) that can answer questions.
- Normally, it only knows what it was trained on (its “memory”).
- But sometimes it **doesn’t know enough** or the info is too new.

RAG fixes this problem by letting the assistant **look up information from external sources** before answering.

Think of it like this:

> “I know a lot, but let me quickly check the internet/books/database to give you the most accurate answer.”

2️⃣How it works (in simple steps)

A RAG system has **two main parts**:

a) Retriever

- This part **searches through documents, databases, or knowledge sources**.
- For example: Wikipedia, PDFs, company manuals.
- It finds the most relevant pieces of information based on your question.

b) Generator (LLM)

- Once the retriever finds useful info, the **LLM uses it to generate the answer**.
- It combines what it already knows + the retrieved info.

So basically:

Question → Retriever finds info → LLM generates smart answer using that info.

3️⃣Why is RAG useful?

- Keeps answers **up-to-date**, even if the LLM wasn't trained on the newest info.
- Reduces mistakes because the model can **look at real sources**.
- Can handle **big knowledge bases** without the LLM needing to memorize everything.

4️⃣Example

Question: "What is the current population of India in 2025?"

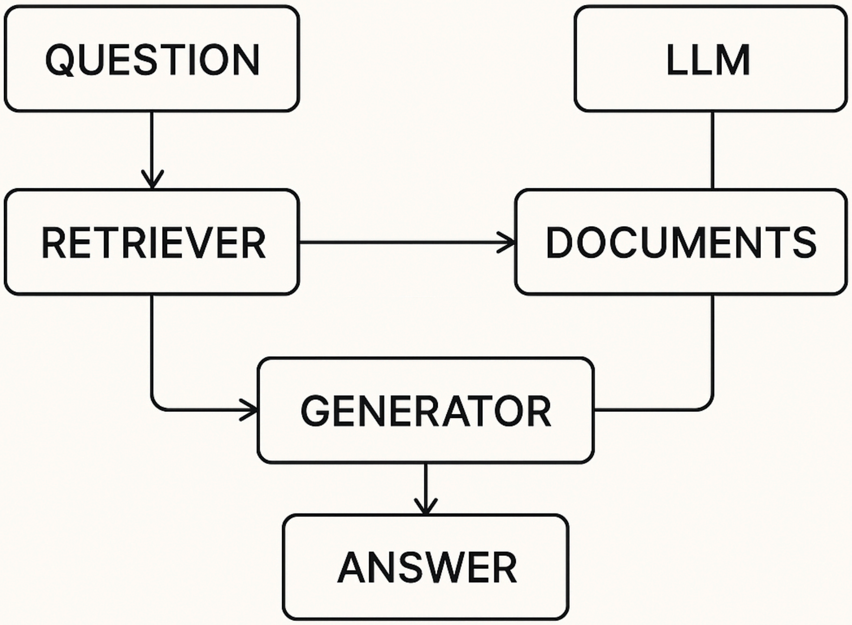
- LLM alone: Might only know data until 2021.
- RAG Agent: First searches a reliable source → finds 2025 estimate → LLM generates answer using that info.

5️⃣How developers build RAG Agents

1. **Collect your documents** (PDFs, articles, database entries).
2. **Use a retriever** to turn them into searchable vectors.
3. **Query the retriever** with the user's question.
4. **Feed retrieved info to LLM** to generate a precise answer.

✅ **In short:**
A **RAG Agent = Smart assistant + search engine**. It retrieves relevant info first and then gives you the best answer.

BUILDING RAG AGENTS WITH LLMs


QUESTION
LLM
RETRIEVER
DOCUMENTS
GENERATOR
ANSWER