

## 1. Fundamentals of RAG (Retrieval-Augmented Generation)

### What you must know

- What RAG is and **why it is used**
- Difference between
  - Normal LLM chatbot (just model knowledge)
  - RAG-based chatbot (model + external knowledge)
- Real-world examples of RAG use cases

### Key points to explain

- LLMs are **static** (trained once)
- RAG allows **dynamic, updated, domain-specific knowledge**
- Why RAG is better than fine-tuning for medical/psychology data

## 2. How RAG Works – End-to-End Architecture

### You should clearly explain this flow:

1. User asks a question
2. Query is converted to **embeddings**
3. Similar content is retrieved from **vector database**
4. Retrieved context is added to the prompt
5. LLM generates a grounded response

### Cover components

- Embedding models
- Vector databases (Pinecone, FAISS, Supabase Vector, Weaviate)
- Retriever
- Generator (LLM)

### 3. Building a RAG-Based Chat System (Technical Stack)

#### Topics to cover

- Data sources:
  - Psychology guidelines
  - Therapist notes (anonymized)
  - DSM-5 summaries
  - FAQs / treatment protocols
- Chunking strategies (why chunk size matters)
- Metadata usage (age group, disorder type, severity, language)

#### Tech stack example

- Frontend: React / Next.js
- Backend: Node.js / Python (FastAPI)
- LLM: OpenAI / Claude / Gemini
- Vector DB: Supabase Vector / Pinecone
- Orchestration: LangChain / LlamaIndex

### 4. Psychology-Focused Chatbot Design (Domain-Specific Logic)

#### Very important topic for your use case

Cover:

- Difference between:
  - Mental health support chatbot ❌
  - Clinical decision-support assistant ✅
- Role of the chatbot:
  - Assist psychologists, **not replace them**
- Use cases:
  - Session summarization

- Suggesting therapy techniques
- Risk flagging (depression, self-harm indicators)
- Treatment history retrieval

### **Prompt design**

- Safe system prompts
- Non-diagnostic responses
- Empathetic but professional tone

## **5. Safety, Ethics, Privacy & Compliance (Critical for Psychology)**

### **You MUST cover this**

- Data privacy (HIPAA-like thinking even if not legally required)
- Anonymization of patient data
- Access control (psychologist vs admin)
- Hallucination prevention using RAG
- Crisis handling:
  - Self-harm detection
  - Emergency escalation logic
- Logging & audit trails