In 2025, the AI landscape shifted from "chatting with documents" to specialized architectures designed to solve complex reasoning, memory, and high-precision data retrieval.Since you're a CS student in Chennai with a focus on DSA and DBMS, you'll appreciate that these trends aren't just about "better prompts"—they are structural innovations in how data is indexed and traversed.1. GraphRAG (Knowledge Graph RAG)While standard RAG treats documents as isolated text chunks, GraphRAG builds a structured knowledge graph ($G = (V, E)$) over the corpus.The Problem it Solves: Standard RAG fails at "global" questions like "What are the recurring themes across these 500 documents?" because it only retrieves specific local snippets.The Niche: It uses LLMs to extract entities (nodes) and relationships (edges) first. At query time, it traverses the graph to synthesize a high-level summary.DSA Connection: It's essentially performing a Community Detection algorithm on the graph to group related concepts before the LLM even sees the query.2. RAPTOR (Recursive Abstractive Processing for Tree-Organized Retrieval)RAPTOR addresses the "context window" limitation by organizing data into a hierarchy (a tree).The Niche: It recursively clusters and summarizes text chunks.Level 0: Raw text chunks.Level 1: Summaries of Level 0 clusters.Level 2: Summaries of Level 1 clusters.Why it's Niche: When you ask a broad question, RAPTOR retrieves high-level summaries from the top of the tree; when you ask a specific question, it retrieves detail from the leaves.3. Agentic RAG (Multi-Step Reasoning)Instead of a single "Retrieve -> Generate" flow, Agentic RAG gives the AI a "loop."The Niche: The model doesn't just fetch data once. It acts as an agent that can:Search for X.Realize X is missing detail on Y.Go back and search specifically for Y.Reflect on whether the answer is complete before showing it to you.Trend: Look into Self-RAG, where the model is trained to output special "reflection tokens" to critique its own retrieval quality.4. HyDE (Hypothetical Document Embeddings)HyDE flips the retrieval process on its head to solve the "vocabulary mismatch" problem.The Niche: When you ask a question, the AI first generates a fake (hypothetical) answer. It then uses that fake answer to search the database for real documents that look like it.Logic: Dense vectors for "questions" and "answers" often live in different parts of the vector space. By searching with a "fake answer," you're searching for documents in the same semantic neighborhood.5. Vibe Coding & Large Action Models (LAMs)Moving beyond text generation, 2025 saw the rise of models that do things.Vibe Coding: A niche developer trend where you describe the "vibe" or logic of an app, and the AI handles the entire file structure, environment variables, and deployment (e.g., platforms like Lovable or Replit Agent).LAMs: Models like Runner H or Anthropic Computer Use that don't just output code; they move the cursor, click buttons, and navigate UIs to complete tasks like "Book a flight from Chennai to Delhi using my saved card."