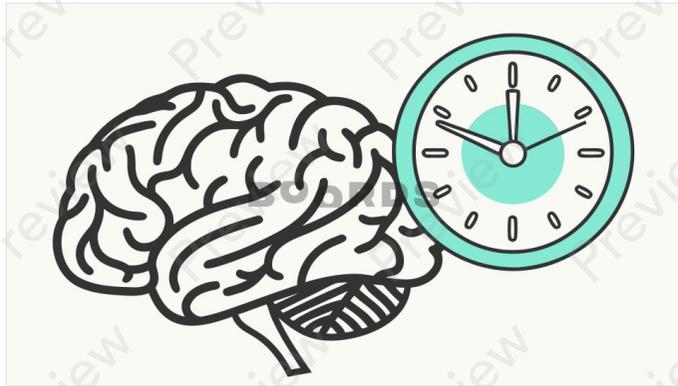


BOARDS



1 LLM Knowledge Limitation

- 📄 A stylized brain icon with a clock superimposed, indicating frozen knowledge. Text below reads 'A base LLM only knows what it learned during training — its knowledge is frozen at the cutoff.' This makes it unreliable for anything recent, private, or domain-specific.



2 Introducing RAG

- 📄 A diagram showing a question mark entering a box labeled 'RAG'. Text below reads 'Retrieval-Augmented Generation (RAG) solves this by retrieving relevant documents before generation.'



3 Vector Database Retrieval

- 📄 An animation of a query being embedded and then arrows pointing from a central database icon to several document icons. Text below reads 'The query is embedded and similar chunks are fetched from a vector database.'



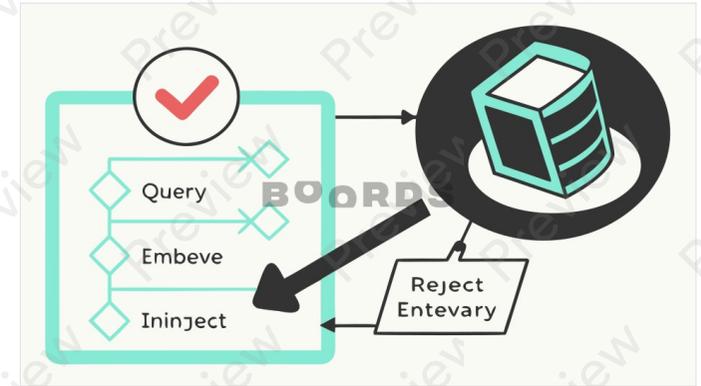
4 Prompt Injection

A visual representation of a prompt being constructed, with the original query and retrieved document chunks being combined. Text below reads 'Those chunks are injected into the prompt alongside the original query.'



5 LLM Generation with Context

A large language model icon receiving the combined prompt. Text below reads 'The LLM generates an answer using both its training knowledge and the retrieved context.'



6 RAG Solution

A checkmark icon next to a diagram summarizing the RAG process: Query -> Embed -> Retrieve -> Inject -> Generate. Text below reads 'RAG provides up-to-date, relevant, and accurate answers.'