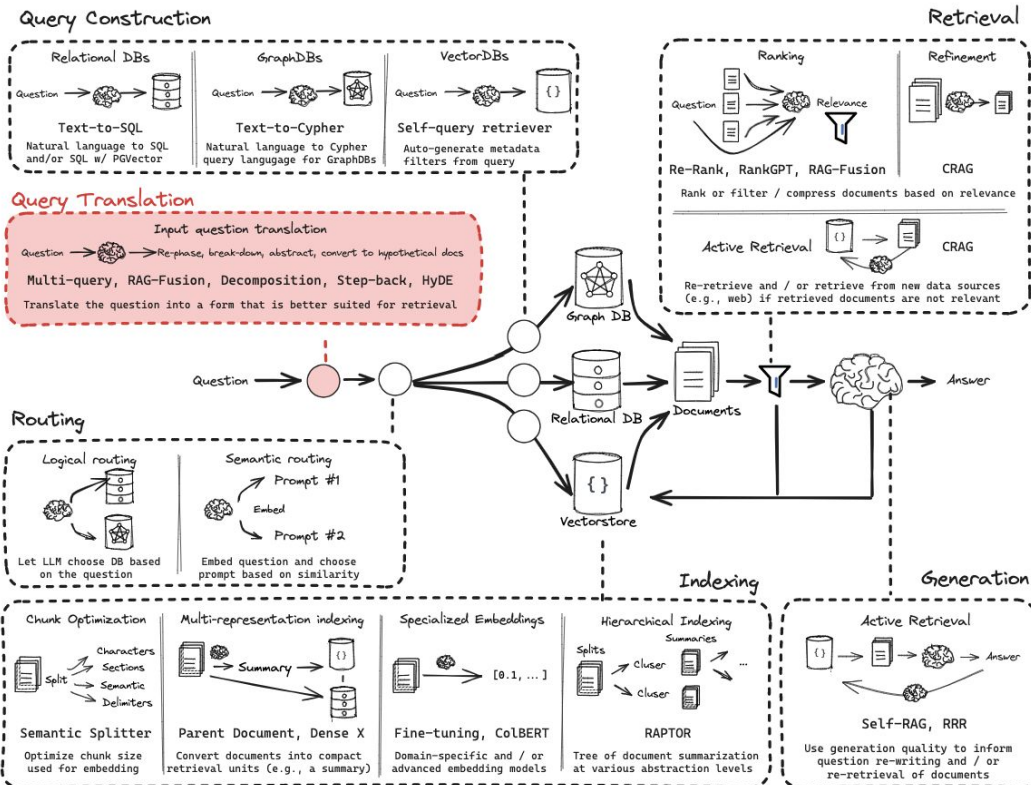


RAG from scratch: Query Translation (Multi-Query)

Lance Martin
Software Engineer, LangChain
[@RLanceMartin](https://twitter.com/RLanceMartin)

Query Translation



Problem statement



anton ✓
@abacaj



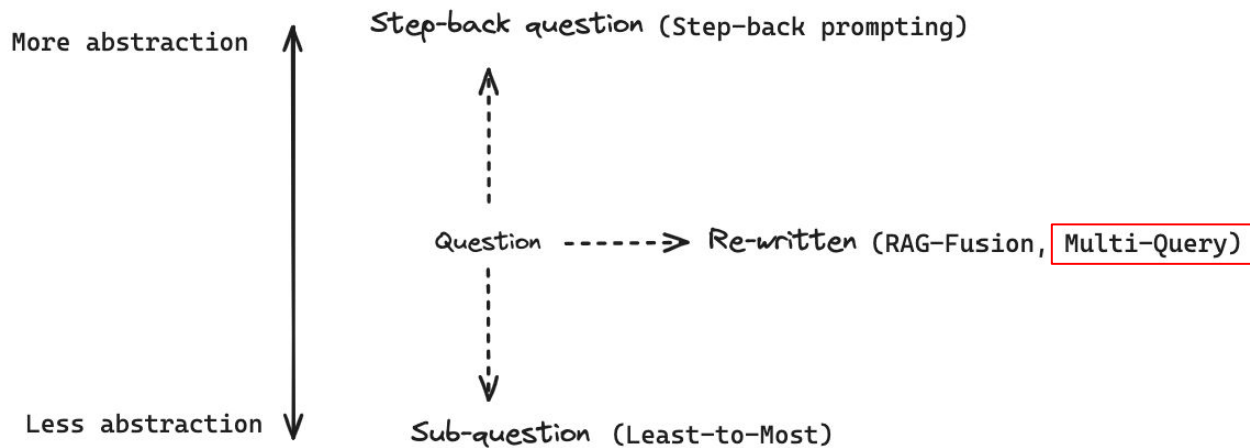
Semantic search on embeddings is hard to get right. Embedding long documents is a challenge. User queries are a challenge, if a user provides an ambiguous query they'll get ambiguous matches

LLMs just follow what was in the context and hallucinate answers as a result

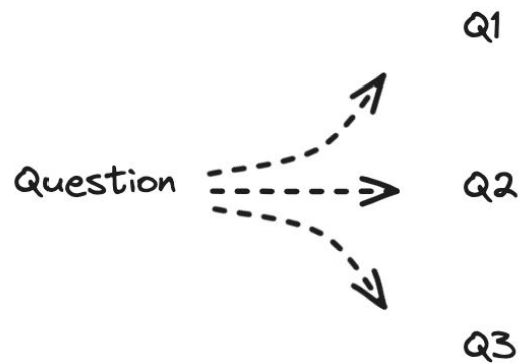
11:06 AM · Apr 28, 2023 · **113.7K** Views



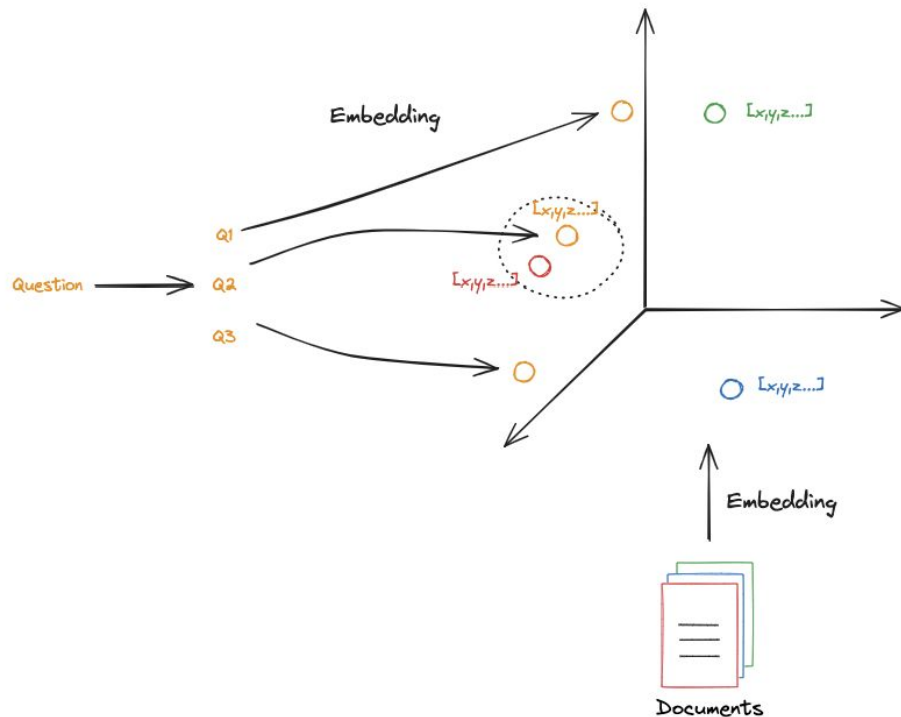
General approaches to transform questions



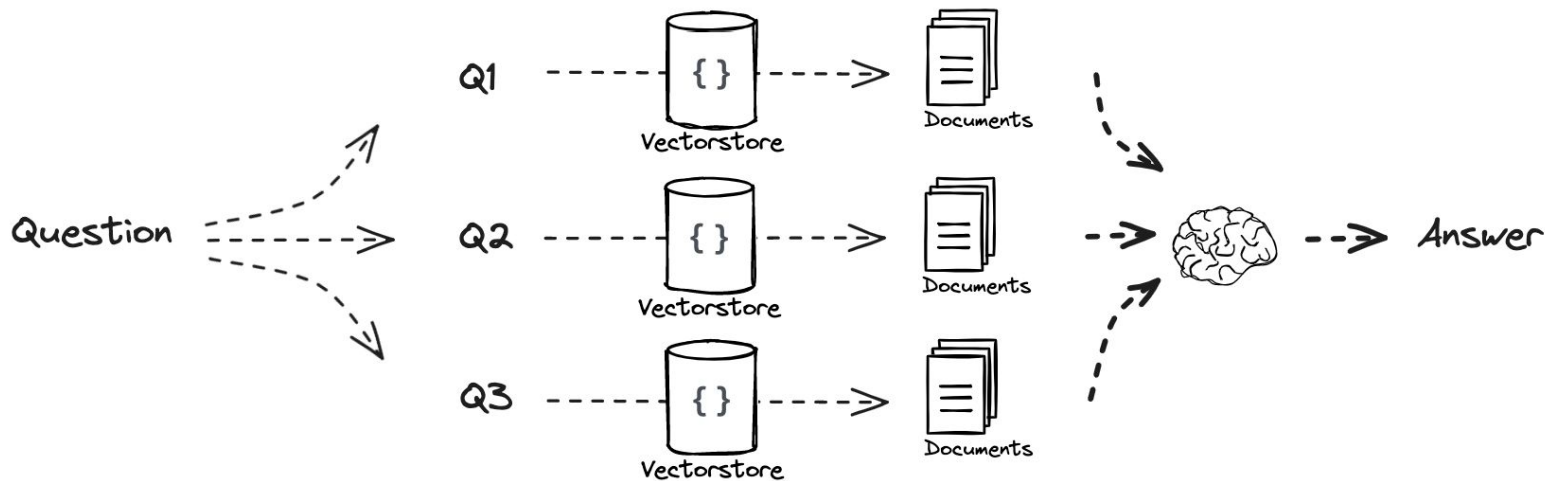
Transform a question into multiple perspectives



Intuition: Improve search



Use this with parallelized retrieval



Code walk-through