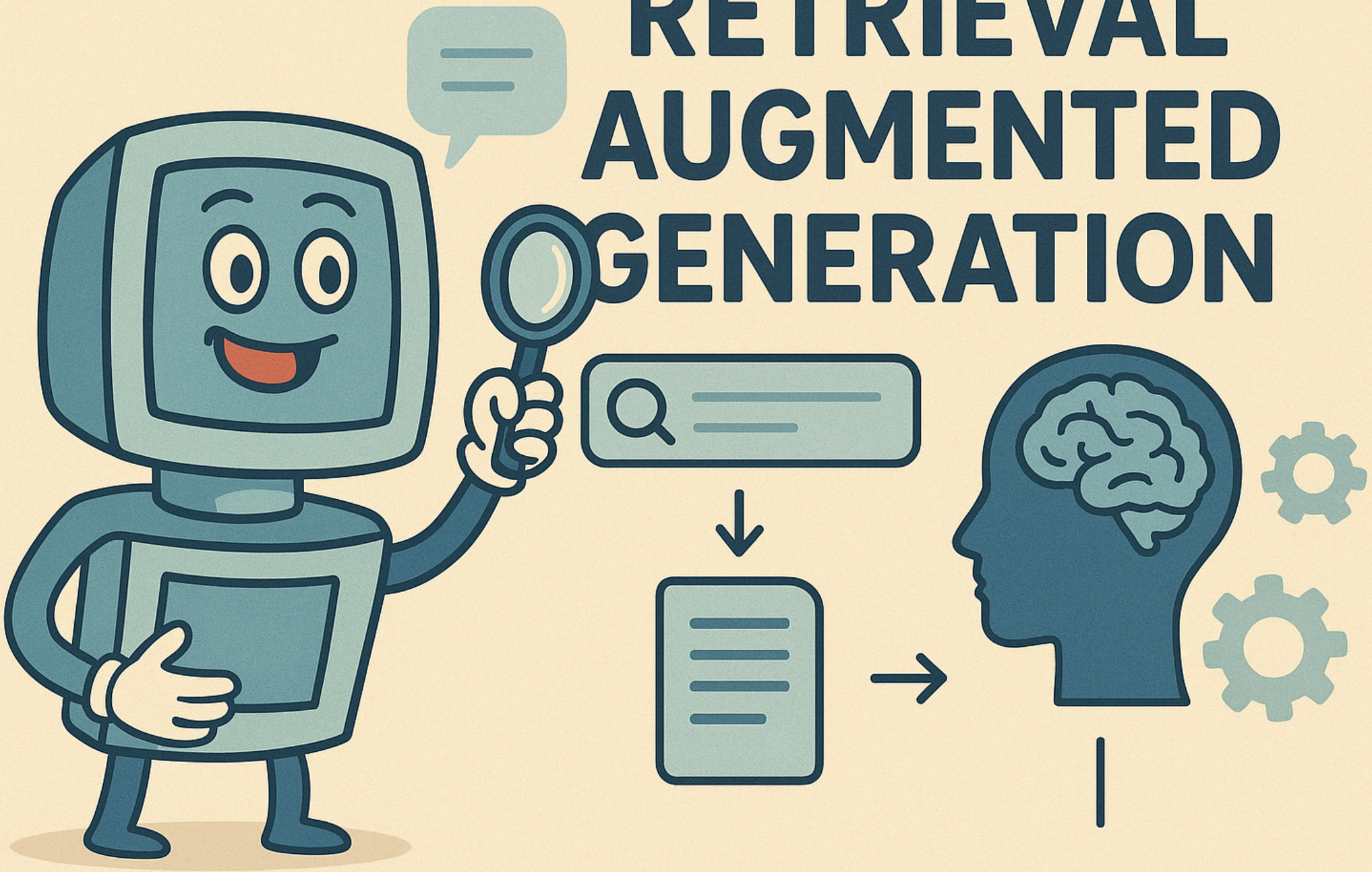
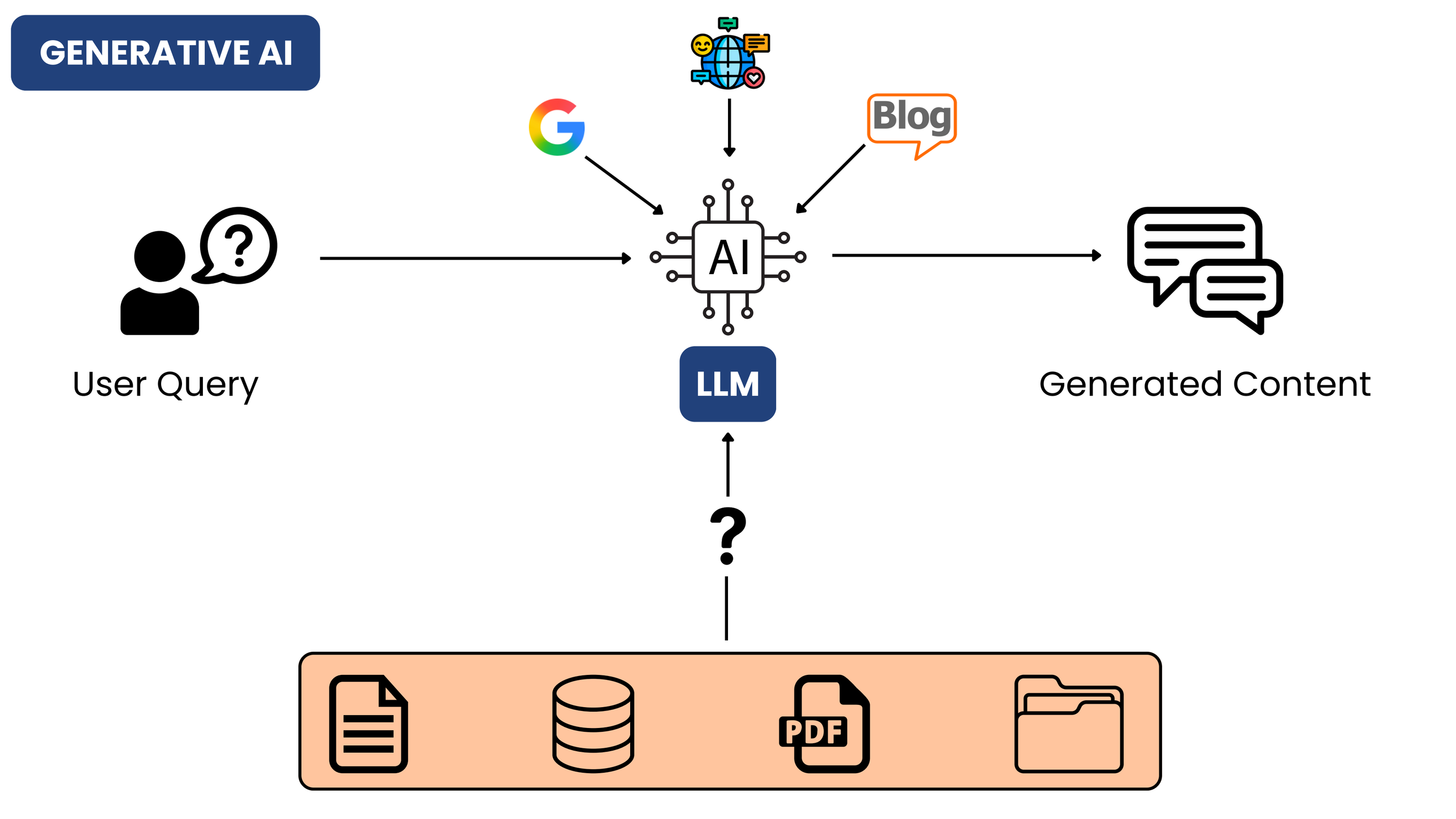
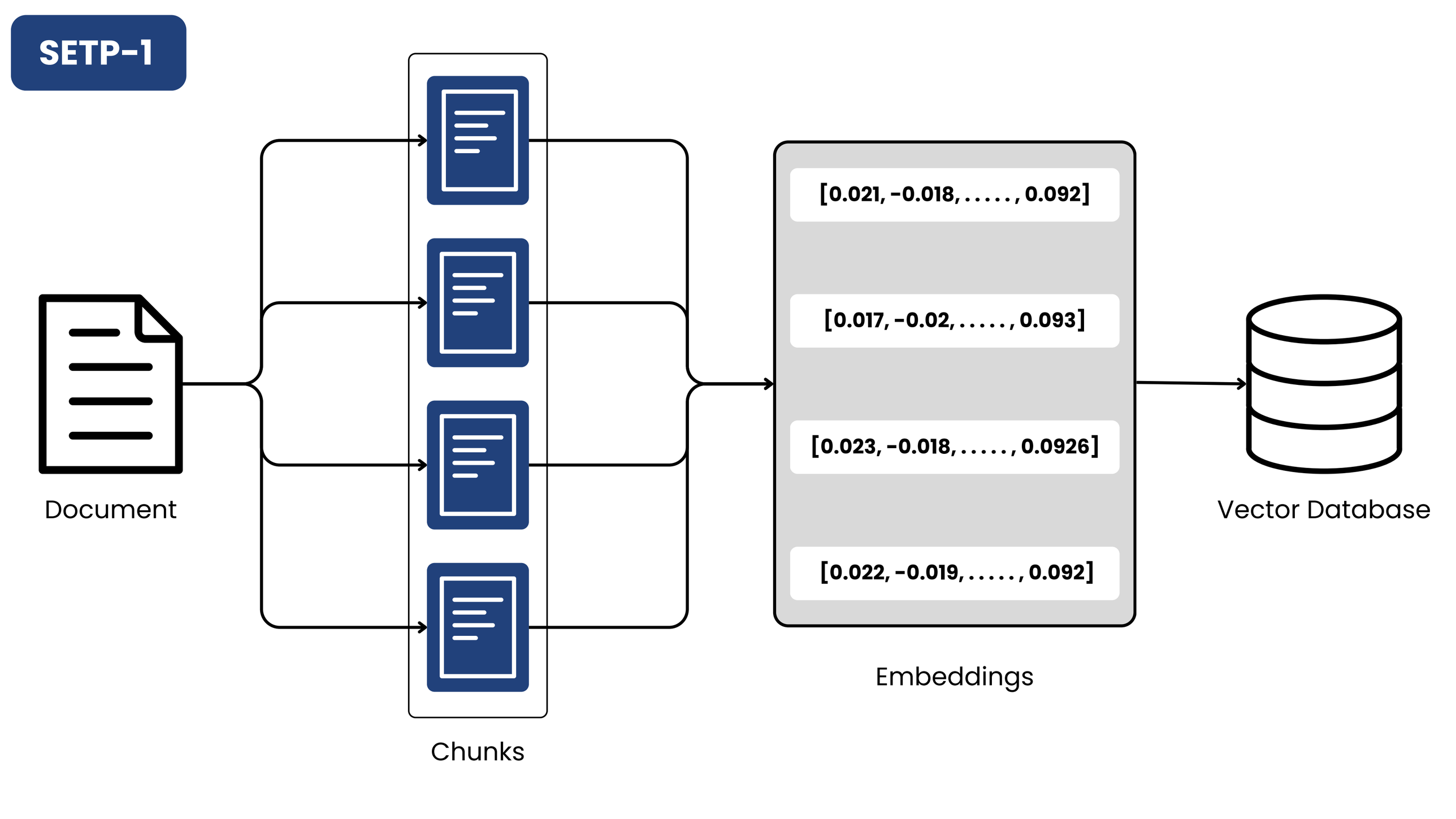
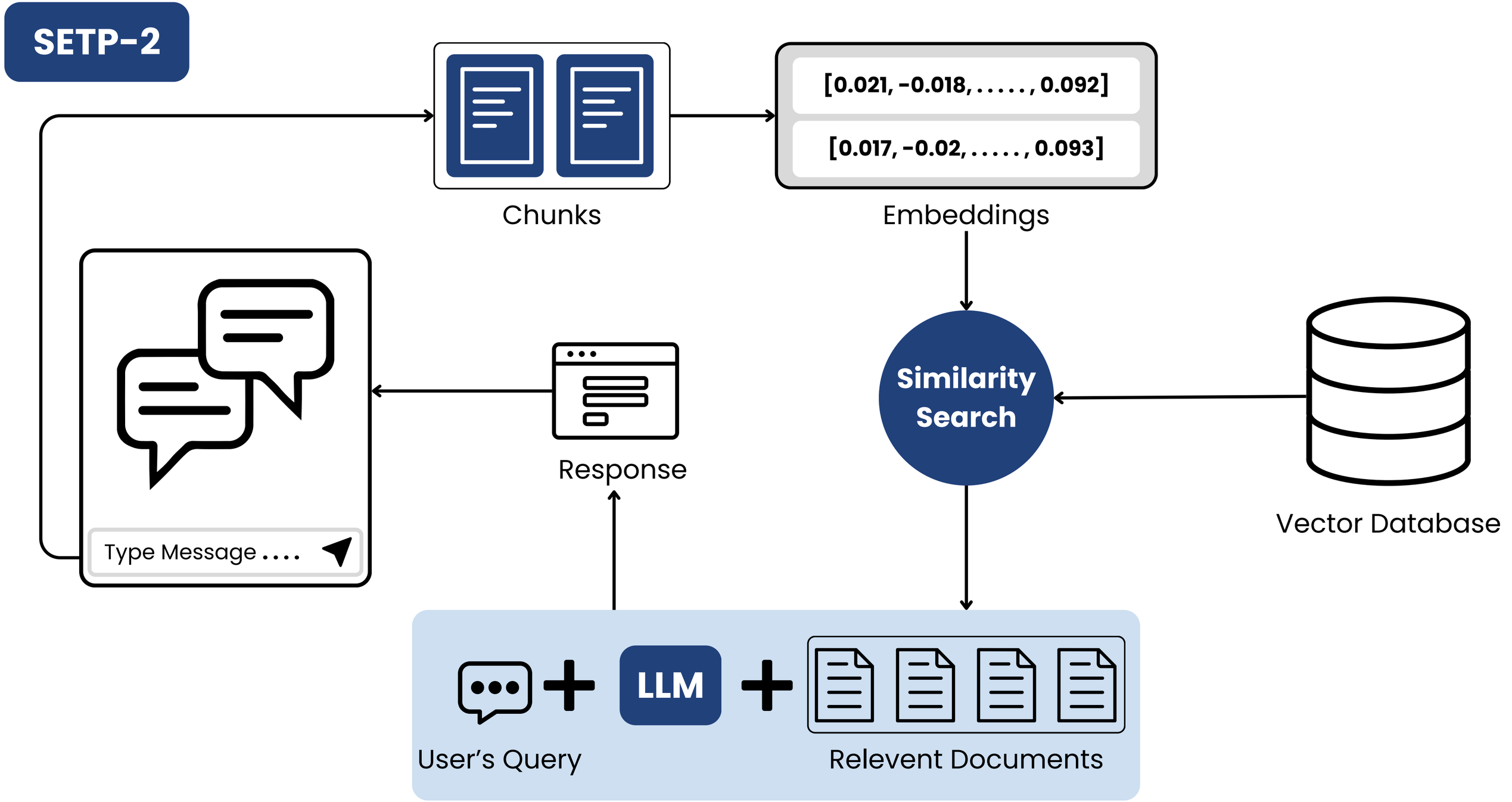
RETRIEVAL AUGMENTED GENERATION



GENERATIVE AI




SETP-1
Document
Chunks
[0.021, -0.018, ... , 0.092]
[0.017, -0.02, ... , 0.093]
[0.023, -0.018, ... , 0.0926]
[0.022, -0.019, ... , 0.092]
Embeddings
Vector Database


SETP-2
[0.021, -0.018, ... , 0.092]
[0.017, -0.02, ... , 0.093]
Chunks
Embeddings
Similarity Search
Vector Database
Response
Type Message . . . .
LLM
User's Query
Relevent Documents