Retrieval Augmented Generation

Retrieval quality determines whether RAG helps or hurts.

Failure Modes

RAG can introduce errors when the retrieved context is irrelevant.

Irrelevant passages distract the model and lower accuracy.