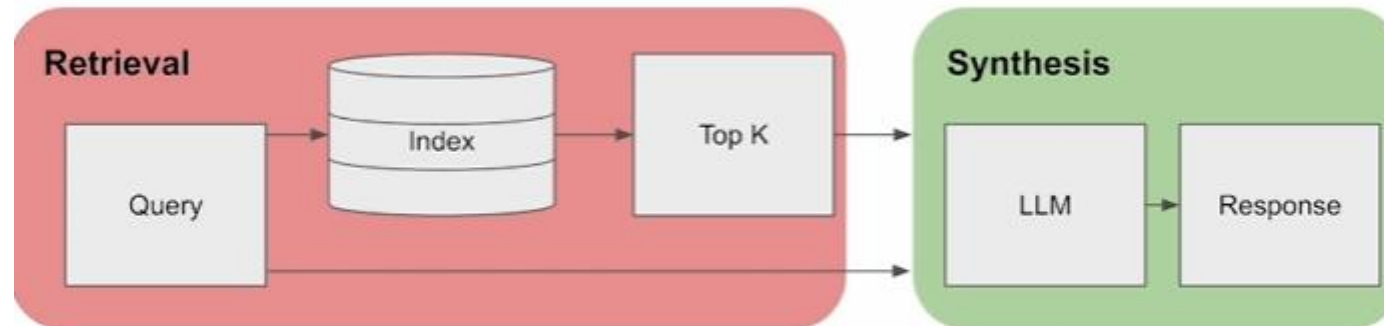
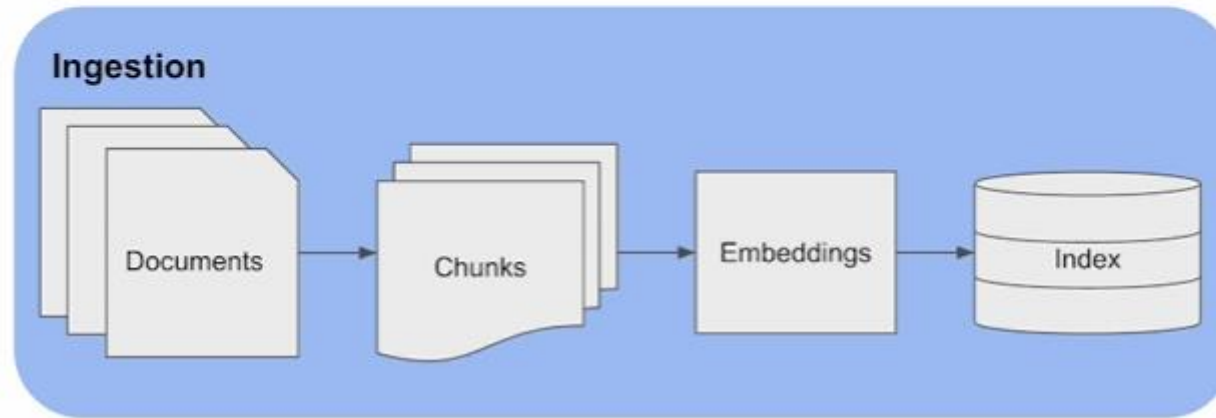
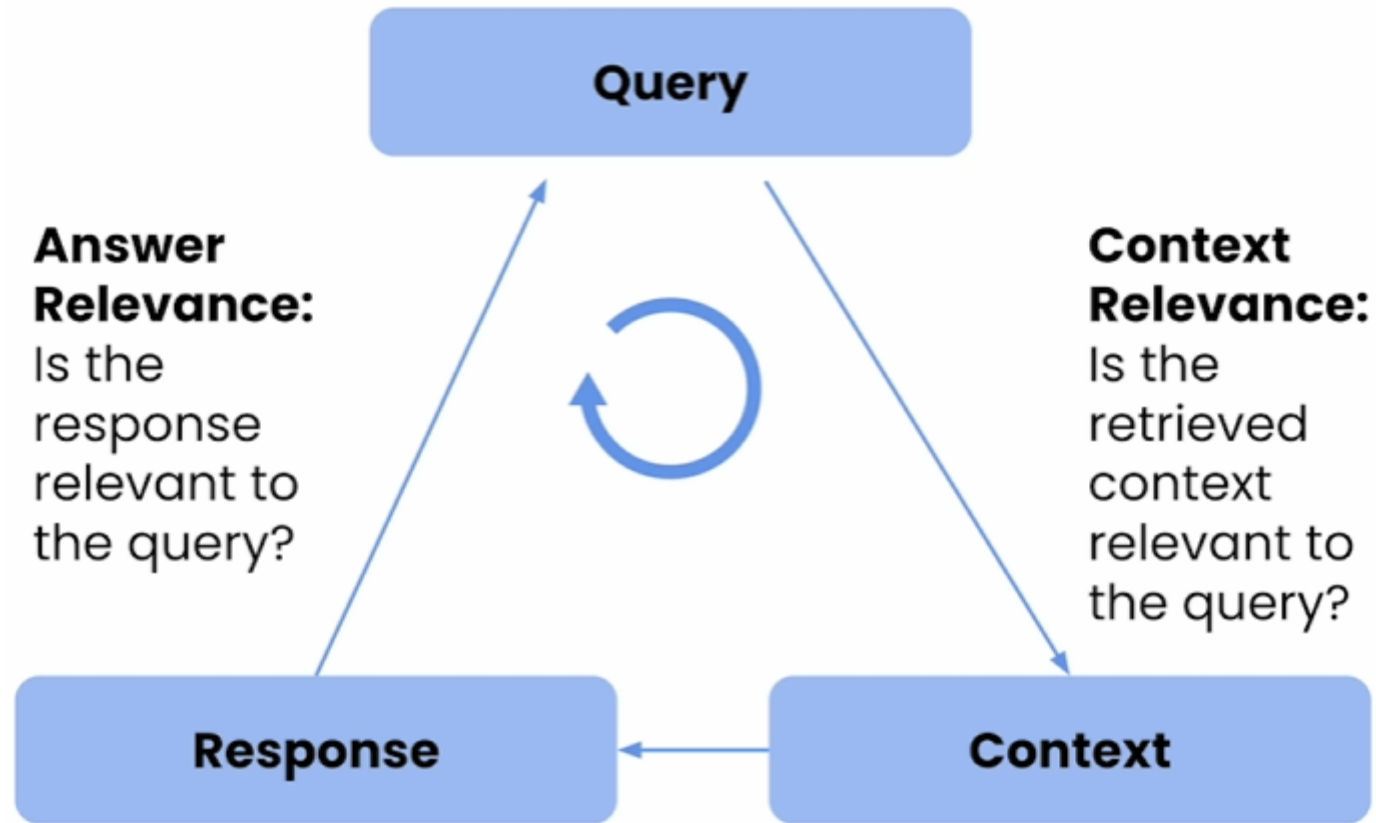


Basic RAG Pipeline

<https://www.deeplearning.ai/short-courses/building-evaluating-advanced-rag/>



The RAG Triad



Answer Relevance:
Is the response relevant to the query?

Context Relevance:
Is the retrieved context relevant to the query?

Groundedness: Is the response supported by the context?

Setup

Basic and advanced RAG Pipeline
with LlamaIndex

https://github.com/run-llama/llama_index

Evaluation Benchmark
with TruEra

<https://github.com/truera/trulens>

Sentence-window retrieval

Query: What are the concerns surrounding the AMOC?

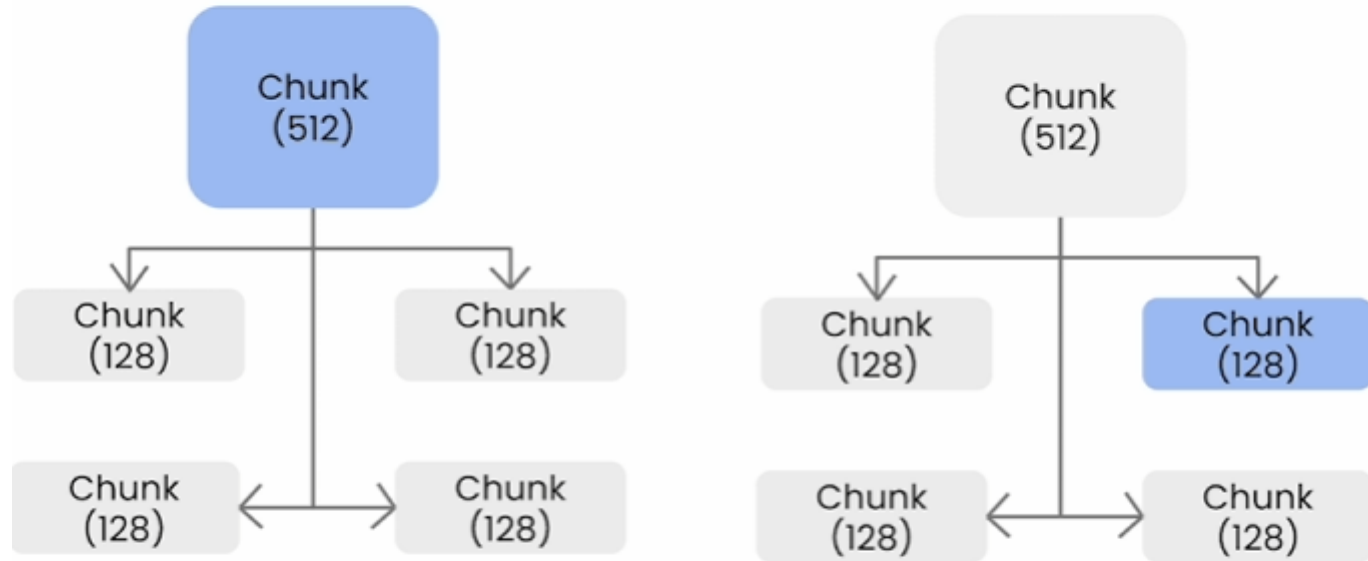
Continuous observation of the Atlantic meridional overturning circulation (AMOC) has improved the understanding of its variability (Frajka-Williams et al., 2019), but there is low confidence in the qualification of AMOC changes in the 20th century because of low agreement in quantitative reconstructed and simulated trends. Direct observational records since the mid-2000s remain too short to determine the relative contributions of internal variability, natural forcing and anthropogenic to AMOC change (high confidence). Over the 21st century, AMOC will very likely decline for all SSP scenarios but will not involve an abrupt collapse before 2100. 3.2.2.4 Sea Ice Changes Sea ice is a key driver of polar marine life, hosting unique ecosystems and affecting diverse marine organisms and food webs through its impact on light penetrations and supplies of nutrients and organic matter (Arrigo, 2014).

What the LLM sees

Embedding Lookup

What the LLM sees

Auto-merging retrieval



	Groundedness	Answer Relevance	Context Relevance	latency	total_cost
app_id					
Automerging Query Engine	1.00000	0.940	0.4350	2.25	0.000799
Sentence Window Query Engine	0.87800	0.925	0.3675	2.25	0.000868
Direct Query Engine	0.80125	0.930	0.2550	2.20	0.002911