
Retrieval-Augmented Generation (RAG)

Going Beyond the LLM

Thaís de Araújo de Medeiros
thais.araujo.707@ufrn.edu.br

Qual atriz brasileira ganhou o Globo de Ouro em 2025?



Não posso responder com certeza porque não tenho acesso a informações atualizadas de 2025 sem utilizar busca na web. Meu conhecimento é baseado em informações disponíveis até junho de 2024. Se quiser, posso falar sobre atrizes brasileiras que já foram indicadas ou que têm potencial para ganhar prêmios como o Globo de Ouro.



Qual atriz brasileira ganhou o Globo de Ouro em 2025?

Contexto:

O Brasil amanheceu com uma conquista inédita: Fernanda Torres tornou-se a primeira atriz brasileira a vencer o Globo de Ouro em 2025. Ela foi premiada por sua atuação em *Ainda Estou Aqui*, primeiro filme original do Globoplay, na categoria de melhor atriz em filme de drama.

Na cerimônia em Hollywood, Walter Salles celebrou o reconhecimento e destacou o impacto do filme no público brasileiro. Baseado no livro de Marcelo Rubens Paiva, o longa retrata a luta de Eunice Paiva para encontrar o marido, Rubens Paiva, desaparecido na ditadura militar.



Fernanda Torres ganhou o Globo de Ouro em 2025 na categoria de **melhor atriz em filme de drama** por sua atuação em *Ainda Estou Aqui*.



Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks

Patrick Lewis^{†‡}, Ethan Perez^{*},

Aleksandra Piktus[†], Fabio Petroni[†], Vladimir Karpukhin[†], Naman Goyal[†], Heinrich Küttler[†],

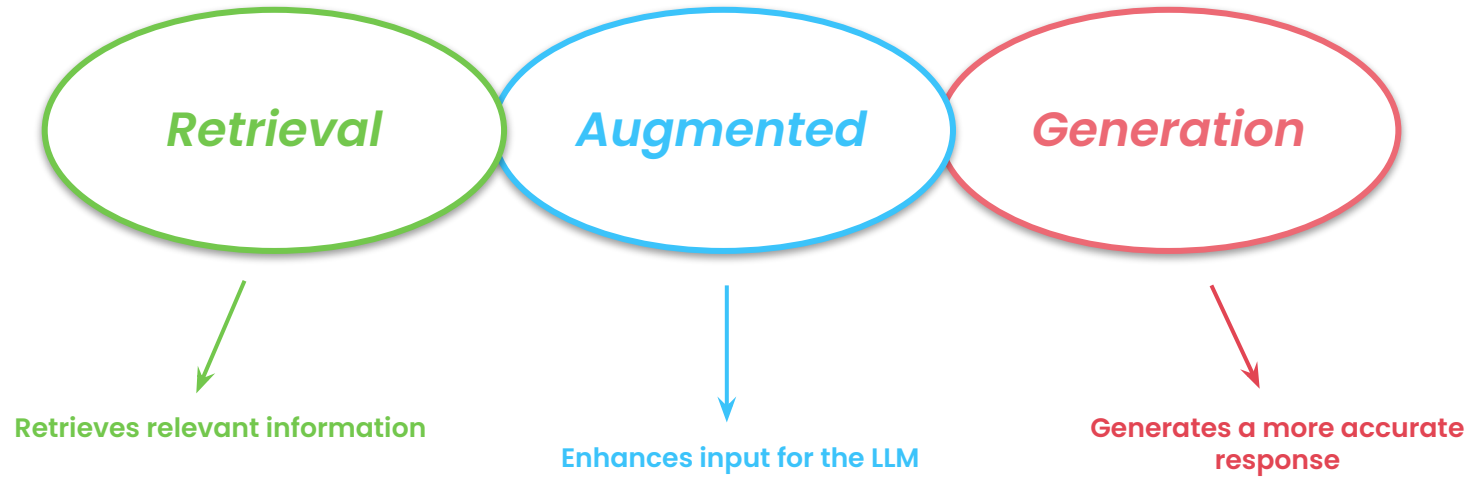
Mike Lewis[†], Wen-tau Yih[†], Tim Rocktäschel^{†‡}, Sebastian Riedel^{†‡}, Douwe Kiela[†]

[†]Facebook AI Research; [‡]University College London; ^{*}New York University;
plewis@fb.com

Abstract

Large pre-trained language models have been shown to store factual knowledge in their parameters, and achieve state-of-the-art results when fine-tuned on downstream NLP tasks. However, their ability to access and precisely manipulate knowledge is still limited, and hence on knowledge-intensive tasks, their performance lags behind task-specific architectures. Additionally, providing provenance for their decisions and updating their world knowledge remain open research problems. Pre-trained models with a differentiable access mechanism to explicit non-parametric memory have so far been only investigated for extractive downstream tasks. We explore a general-purpose fine-tuning recipe for retrieval-augmented generation (RAG) — models which combine pre-trained parametric and non-parametric memory for language generation. We introduce RAG models where the parametric memory is a pre-trained seq2seq model and the non-parametric memory is a dense vector index of Wikipedia, accessed with a pre-trained neural retriever. We compare two RAG formulations, one which conditions on the same retrieved passages across the whole generated sequence, and another which can use different passages per token. We fine-tune and evaluate our models on a wide range of knowledge-intensive NLP tasks and set the state of the art on three open domain QA tasks, outperforming parametric seq2seq models and task-specific retrieve-and-extract architectures. For language generation tasks, we find that RAG models generate more specific, diverse and factual language than a state-of-the-art parametric-only seq2seq baseline.

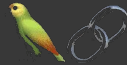
CONCEPT



HOW TO BUILD?



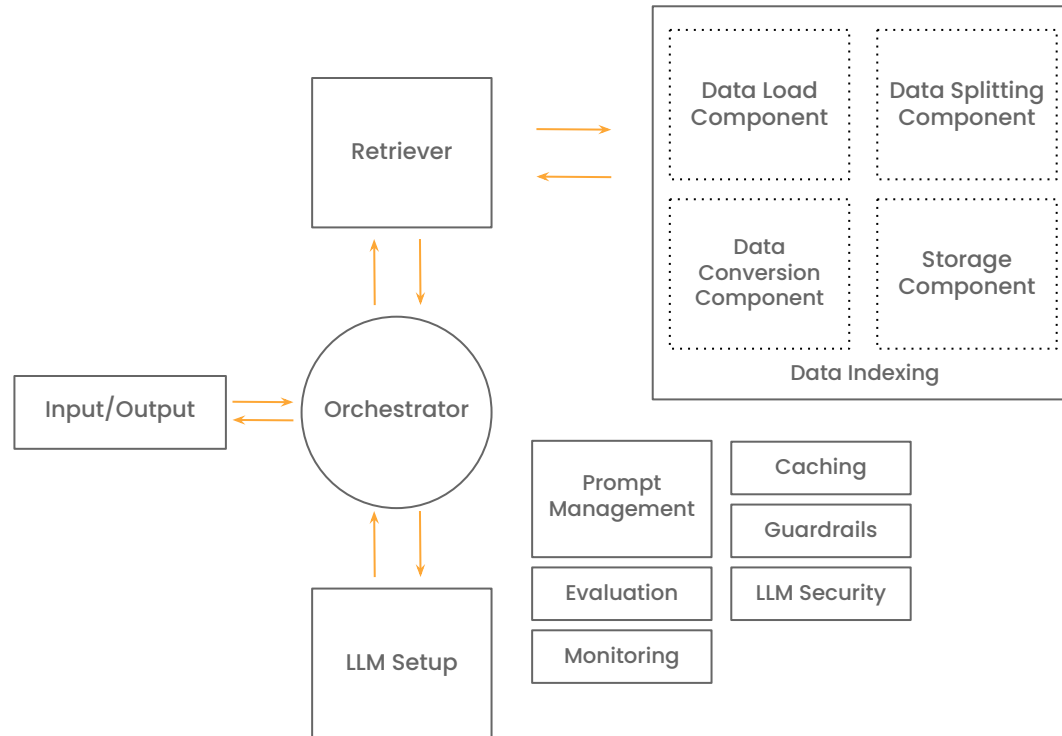
LlamaIndex



LangChain

OVERVIEW

RAG SYSTEM DESIGN

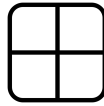


RETRIEVER

INDEXING PIPELINE



DATA LOAD
COMPONENT



DATA SPLITTING
COMPONENT



DATA CONVERSION
COMPONENT



STORAGE
COMPONENT

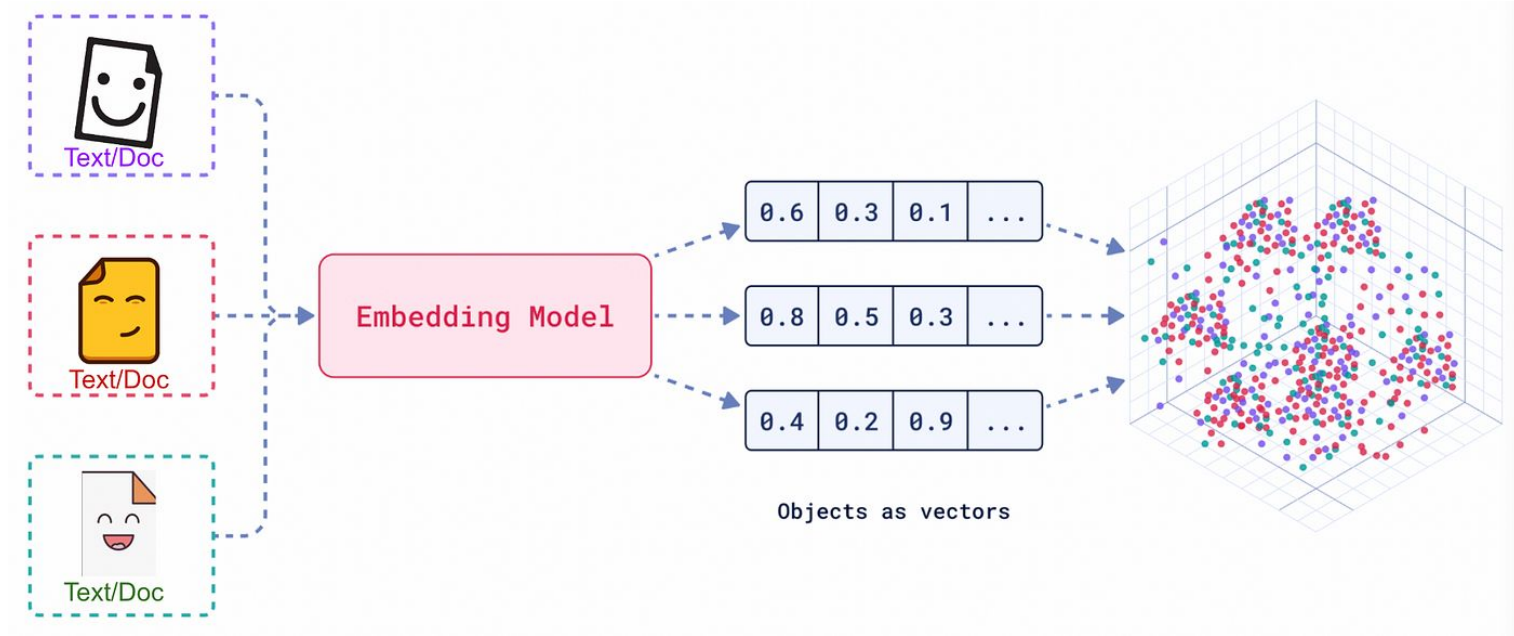
RETRIEVER: INDEXING PIPELINE

DATA LOAD AND DATA SPLITTING



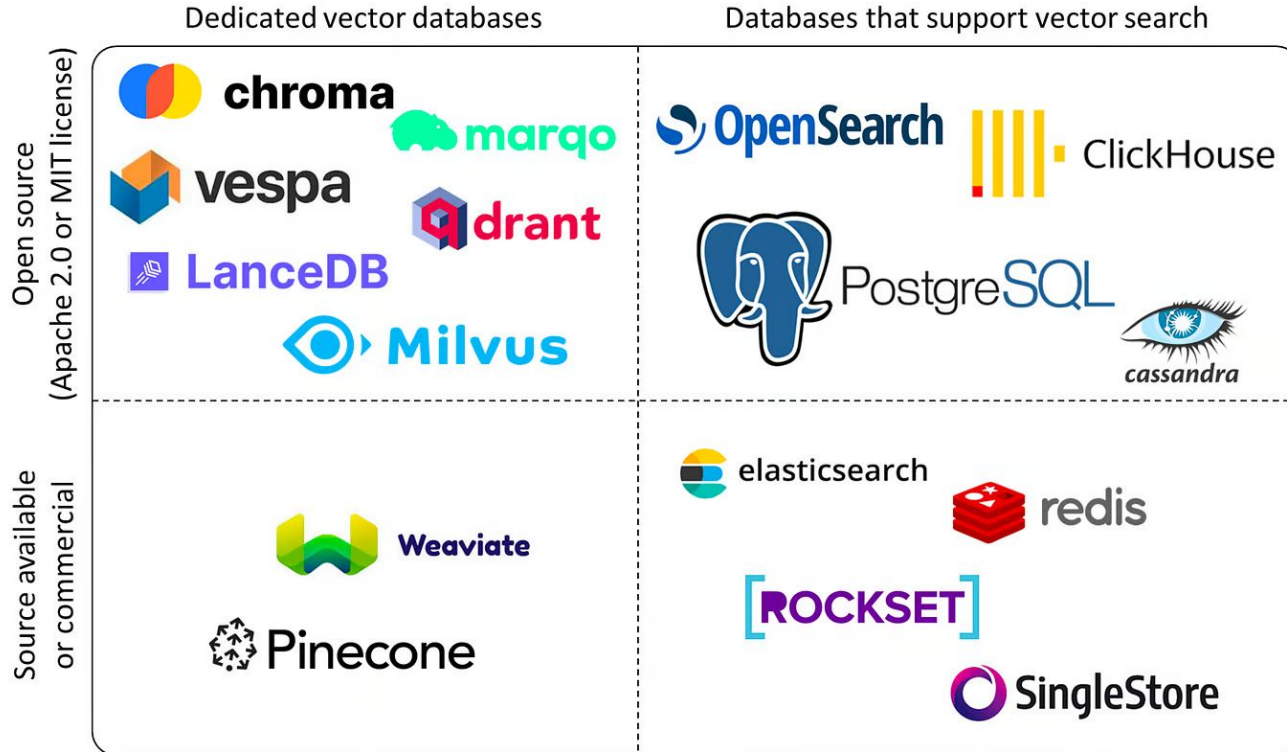
RETRIEVER: INDEXING PIPELINE

DATA CONVERSION

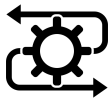


RETRIEVER: INDEXING PIPELINE

STORAGE



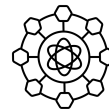
GENERATION **GENERATION PIPELINE**



RETRIEVER



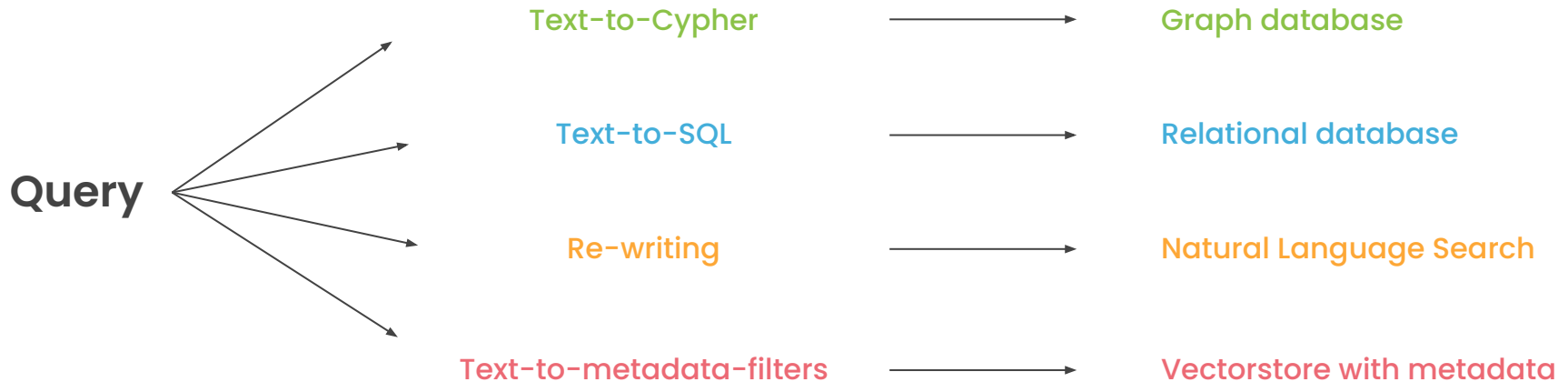
PROMPT MANAGEMENT



LLM COMPONENT

GENERATION RETRIEVER

1 Query analysis - use a model to transform or construct a search query



2

Information retrieval

GENERATION: RETRIEVER

QUERY ANALYSIS

Multi-query
Decomposition
Step-back
HyDE



QUERY RE-WRITING



QUERY CONSTRUCTION

Self Query
Text to SQL
Text-to-Cypher

GENERATION: RETRIEVER

INFORMATION RETRIEVER

COMMON RETRIEVER SYSTEMS



LEXICAL SEARCH INDEXES



VECTOR INDEXES



RELATIONAL DATABASES



GRAPH DATABASES

GENERATION **PROMPT MANAGEMENT**

1

**CONTEXTUAL
PROMPTING**

2

**FEW SHOT
PROMPTING**

3

**CONTROLLED
GENERATION
PROMPTING**

4

**CHAIN OF
THOUGHT
PROMPTING**

1

CONTEXTUAL PROMPTING

2

FEW SHOT PROMPTING

3

CONTROLLED GENERATION PROMPTING

4

CHAIN OF THOUGHT PROMPTING

Qual atriz brasileira ganhou o Globo de Ouro em 2025?

Contexto:

O Brasil amanheceu com uma conquista inédita: Fernanda Torres tornou-se a primeira atriz brasileira a vencer o Globo de Ouro em 2025. Ela foi premiada por sua atuação em *Ainda Estou Aqui*, primeiro filme original do Globoplay, na categoria de melhor atriz em filme de drama.

Na cerimônia em Hollywood, Walter Salles celebrou o reconhecimento e destacou o impacto do filme no público brasileiro. Baseado no livro de Marcelo Rubens Paiva, o longa retrata a luta de Eunice Paiva para encontrar o marido, Rubens Paiva, desaparecido na ditadura militar.

Fernanda Torres ganhou o Globo de Ouro em 2025 na categoria de **melhor atriz em filme de drama** por sua atuação em *Ainda Estou Aqui*.



1

CONTEXTUAL PROMPTING

2

FEW SHOT PROMPTING

3

CONTROLLED GENERATION PROMPTING

4

CHAIN OF THOUGHT PROMPTING

Responda à pergunta usando apenas o contexto fornecido. Siga o formato nestes exemplos:

Pergunta: Quem foi o primeiro brasileiro indicado ao Oscar?

Contexto:

Em 1945, o compositor brasileiro Heitor Villa-Lobos recebeu uma indicação honorária ao Oscar por sua contribuição para trilhas sonoras. Embora não tenha concorrido em uma categoria tradicional, seu trabalho influenciou a presença da música brasileira no cinema internacional.

Resposta: Heitor Villa-Lobos foi o primeiro brasileiro indicado ao Oscar, recebendo uma indicação honorária em 1945 por sua contribuição para trilhas sonoras.

Pergunta: Algum brasileiro já foi indicado ao Globo de Ouro?

Contexto:

Em 1999, a atriz brasileira Fernanda Montenegro foi indicada ao Globo de Ouro de Melhor Atriz em Filme de Drama por sua atuação no filme Central do Brasil. Apesar da indicação, o prêmio foi vencido por Cate Blanchett por seu papel em Elizabeth.

Resposta: Fernanda Montenegro foi indicada ao Globo de Ouro de Melhor Atriz em Filme de Drama em 1999 por sua atuação em Central do Brasil, mas não venceu.

Agora, responda à seguinte pergunta usando o contexto fornecido:

Pergunta: Qual atriz brasileira ganhou o Globo de Ouro em 2025?

Contexto:

O Brasil amanheceu com uma conquista inédita: Fernanda Torres tornou-se a primeira atriz brasileira a vencer o Globo de Ouro em 2025. Ela foi premiada por sua atuação em *Ainda Estou Aqui*, primeiro filme original do Globoplay, na categoria de melhor atriz em filme de drama.

Na cerimônia em Hollywood, Walter Salles celebrou o reconhecimento e destacou o impacto do filme no público brasileiro. Baseado no livro de Marcelo Rubens Paiva, o longa retrata a luta de Eunice Paiva para encontrar o marido, Rubens Paiva, desaparecido na ditadura militar.

Resposta:



ChatGPT

Fernanda Torres foi a primeira atriz brasileira a vencer o Globo de Ouro em 2025, recebendo o prêmio de Melhor Atriz em Filme de Drama por sua atuação em *Ainda Estou Aqui*.

1

**CONTEXTUAL
PROMPTING**

2

**FEW SHOT
PROMPTING**

3

**CONTROLLED
GENERATION
PROMPTING**

4

**CHAIN OF THOUGHT
PROMPTING**



1

CONTEXTUAL PROMPTING

2

FEW SHOT PROMPTING

3

CONTROLLED GENERATION PROMPTING

4

CHAIN OF THOUGHT PROMPTING

Use o contexto fornecido para responder à pergunta sobre a premiação do Globo de Ouro de 2025. Mostre seu raciocínio passo a passo, explicando como você utiliza as informações para chegar à conclusão.

Pergunta: Baseado nas informações fornecidas, qual atriz brasileira teve o melhor desempenho no Globo de Ouro de 2025 e por quê?

Contexto:

O Brasil amanheceu com uma conquista inédita: Fernanda Torres tornou-se a primeira atriz brasileira a vencer o Globo de Ouro em 2025.

Ela foi premiada por sua atuação em Ainda Estou Aqui, primeiro filme original do Globoplay, na categoria de Melhor Atriz em Filme de Drama.

Na cerimônia em Hollywood, Walter Salles celebrou o reconhecimento e destacou o impacto do filme no público brasileiro.

Baseado no livro de Marcelo Rubens Paiva, o longa retrata a luta de Eunice Paiva para encontrar o marido, Rubens Paiva, desaparecido na ditadura militar.

Responda à pergunta seguindo estas etapas:

1. Analise quem foi indicado ou vencedor.
2. Compare a relevância da atuação e o impacto do papel.
3. Considere a importância histórica do feito.
4. Avalie o contexto geral da premiação.
5. Chegue a uma conclusão clara.

1

CONTEXTUAL PROMPTING

2

FEW SHOT PROMPTING

3

CONTROLLED GENERATION PROMPTING

4

CHAIN OF THOUGHT PROMPTING

Resposta resumida:

A atriz brasileira que teve o melhor desempenho no Globo de Ouro de 2025, de acordo com o contexto fornecido, foi **Fernanda Torres**, pois ela se tornou a primeira atriz do Brasil a vencer um Globo de Ouro, conquistando o prêmio de Melhor Atriz em Filme de Drama pela atuação em *Ainda Estou Aqui*.

Explicação passo a passo (respeitando o contexto e evitando exposição de cadeias internas de pensamento):

1. Análise dos indicados ou vencedores

- No contexto, apenas Fernanda Torres é mencionada como vencedora do Globo de Ouro de 2025.

1

CONTEXTUAL PROMPTING

2

FEW SHOT PROMPTING

3

CONTROLLED GENERATION PROMPTING

4

CHAIN OF THOUGHT PROMPTING

2. Comparação da relevância da atuação e do impacto do papel

- O filme *Ainda Estou Aqui* retrata a história de Eunice Paiva em busca de seu marido desaparecido na ditadura militar, tema de grande peso emocional e relevância histórica. A performance de Fernanda Torres foi reconhecida internacionalmente a ponto de lhe garantir o prêmio de Melhor Atriz em Filme de Drama.

3. Consideração da importância histórica do feito

- Fernanda Torres tornou-se a primeira atriz brasileira a receber um Globo de Ouro, o que reforça a importância simbólica de sua conquista para o cinema brasileiro.

4. Avaliação do contexto geral da premiação


- O reconhecimento em uma premiação de grande prestígio como o Globo de Ouro destaca a qualidade artística do filme e a força da atuação de Fernanda Torres.

5. Conclusão

- Diante de todos os fatores – premiação, impacto do papel e reconhecimento inédito – a melhor performance brasileira no Globo de Ouro de 2025 é de Fernanda Torres.

GENERATION

LLM

 **deepseek-ai/DeepSeek-R1**

 Text Generation • Updated 13 days ago •  4.35M •  •  9.85k

 **meta-llama/Llama-3.2-3B-Instruct**

 Text Generation • Updated Oct 24, 2024 •  1.6M •  •  1.08k


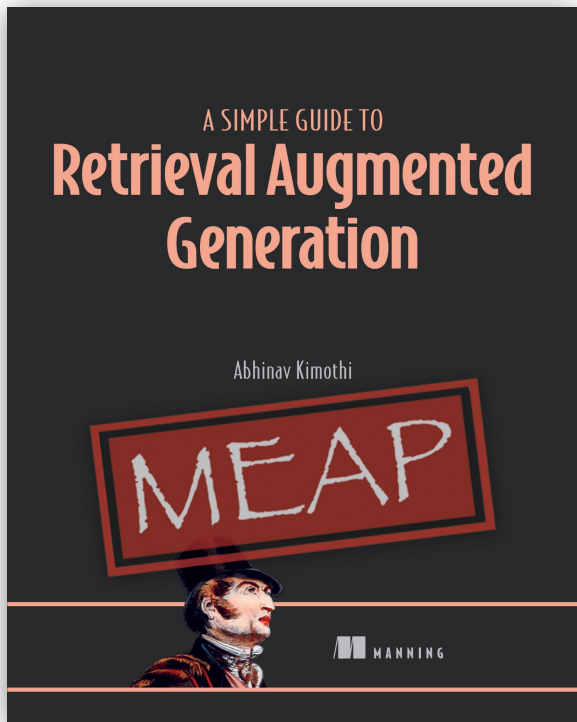
 **Qwen/Qwen2.5-VL-7B-Instruct**

 Image-Text-to-Text • Updated 6 days ago •  1.45M •  511

REFERENCES



LANGCHAIN. Retrieval. [n.d.]. Available at: <https://python.langchain.com/docs/concepts/retrieval/>. Accessed on: February 21, 2025.

OLLAMA. Library. [n.d.]. Available at: <https://ollama.com/library>. Accessed on: February 21, 2025.

WOLFE, Cameron R. The Basics of AI-Powered Vector Search. [n.d.]. Available at: <https://cameronrwolfe.substack.com/p/the-basics-of-ai-powered-vector-search>. Accessed on: February 21, 2025.

HANDS-ON