

# MedChat

Chatbot especializado en dominio médico

---

Josu Barrutia

10 Mayo, 2024

Facultad de Informática, UPV/EHU

## Descripción de la Tarea

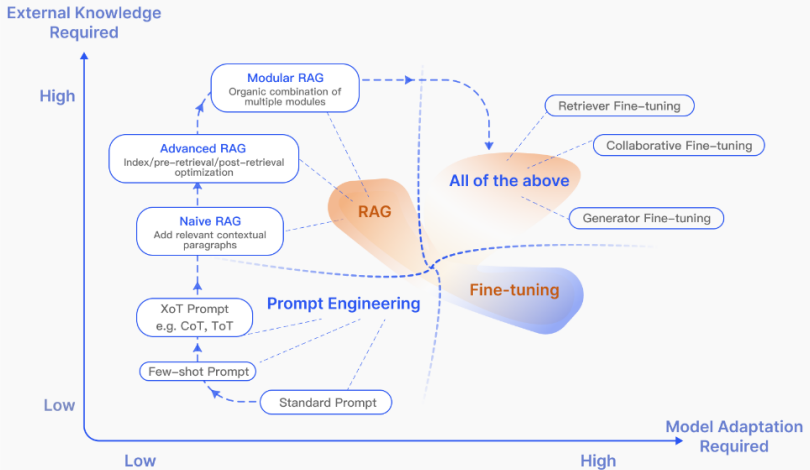
---

Implementar un chatbot especializado en el dominio médico

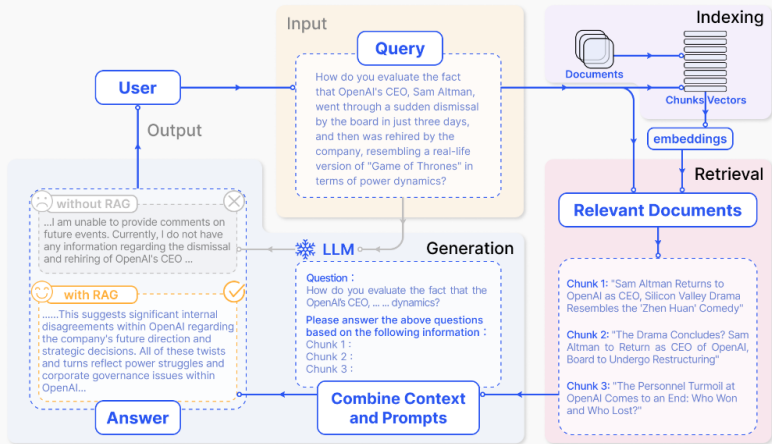
# Arquitectura

---

# Retrieval Augmented Generation



# Retrieval Augmented Generation

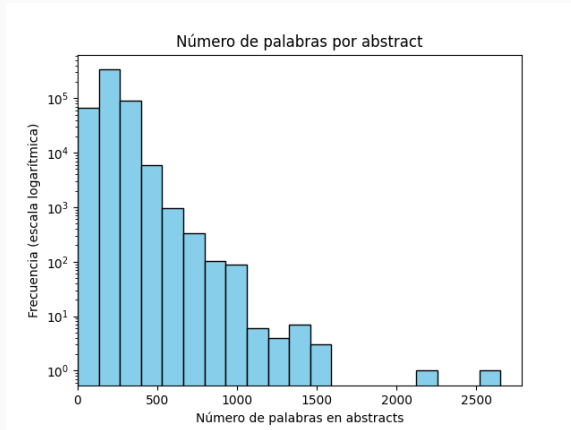


# Datos

---

# Conjunto de Datos

500.000 abstracts de artículos médicos de PubMed





# Creando el Chatbot

---

Dividimos por abstract.

Longitud máxima de 2500 palabras/tokens.

## jinaai/jina-embeddings-v2-base-es

- Spanish-English Bilingual embeddings
- Secuencia máxima de 8192 tokens
- Open-Source
- 137 millones de parámetros

## mistralai/Mistral-7B-Instruct-v0.2

- Instruct fine-tuned version de Mistral-7B-v0.2
- Secuencia máxima de contexto de 32k tokens
- 7 billones de parámetros
- Cuantizado

*`https://josubarru.github.io/`*

# Evaluación

---

MedMCQA : A Large-scale Multi-Subject Multi-Choice Dataset for Medical domain Question Answering

Preguntas?