TRANSLATION, RAG, AND VISCERAL REALISM

https://github.com/JessicaGarson/Translation-and-Vector-Search -with-LangChain-and-Elasticsearch





This talk inspired this meetup

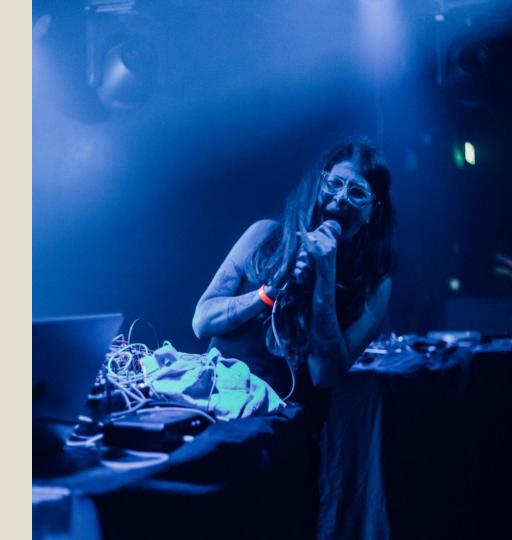
Hi, I'm Jessica Garson

I'm currently a senior developer advocate at Elastic.



In some circles I'm known as Messica Arson

I make loud aggressive music with code and make audio reactive visuals





What's your band?

It's an open source project that anyone including you can be a member of. VISCERAL REALISTS HAVE
PERFORMED IN NEW YORK CITY,
PHILADELPHIA, BALTIMORE,
SUZHOU, AND SHANGHAI.

BECOME VISCERALLY REAL.

MORE INFORMATION ABOUT
VISCERAL REALISM

- Visceral Realism is open source computer music.
- Visceral Realism cannot be performed by a solo performer.
- Visceral Realism cannot be boring.
- Visceral Realism cannot be merchandised.
 - 5. Practitioners of Visceral Realism perform under the name "Visceral Realists".
 - 6. Visceral Realists perform by executing computer code.
 - 7. Visceral Realists do not use Ableton Live or commercial DAWs.
 - 8. Visceral Realists show their work, either with a connected projector or via printed hand-outs.
 - Visceral Realists only use computers, microphones, and MIDI controllers. Audio hardware devices and effects are not permitted.

I had a problem

I saw some local press about the shows but couldn't understand much.

I saw some press that indicated some excitement around our shows

And I wondered if there was more?

trigger presents Visceral Realists/US/experimental sofolofo/KR AT/ambient Julia Dèng Hànzú/CN UK/electroacoustic Torturing Nurse/CN/harsh noise YAMAHAHA/CN/modular noise Shu Ride/CN/glitch 19.30 May 31 Fri 2024 #308 350 Wuning RD trigger

I heard that LangChain was good with Translation

I wrote some code which translates the text of these clippings from Chinese to English using LangChain and Elastic's vector database to learn more about the translated documents.



What version of Elasticsearch should I use?

This demo uses Elasticsearch version 8.16, but you can use any version of Elasticsearch that is higher than 8.0.



What version of Python should I use?

The version of Python that is used is Python 3.12.1 but you can use any version of Python higher than 3.9.





export OPENAI_API_KEY=" ... "





pip install jupyter pandas langchain nest_asyncio langchain-elasticsearch langchain-openai tiktoken elasticsearch datasets





jupyter notebook





A vector database allows you to find similar data quickly. It stores vector embeddings, a type of vector data representation that converts words, sentences, and other data into numbers that capture their meaning and relationships.



What are embeddings?

Embeddings leverage a machine learning model to translate text into numbers, allowing you to perform vector searches.

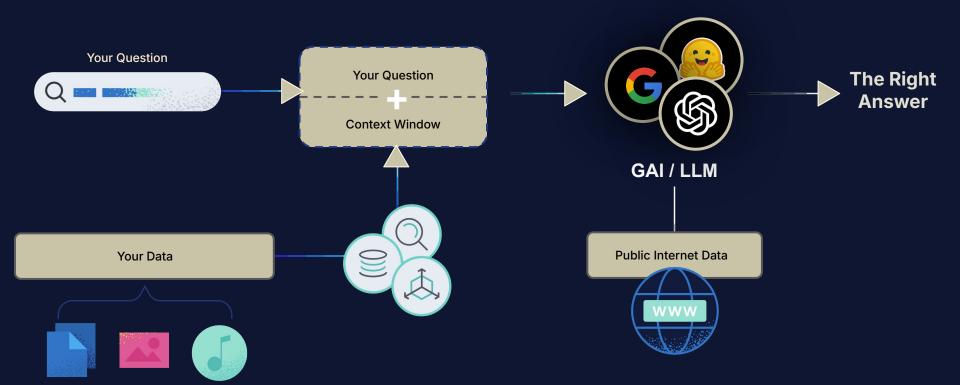


What is RAG?

Retrieval-augmented generation (RAG) integrates external information retrieval into generating responses by Large Language Models (LLMs).



Retrieval Augmented Generation



documents

images

audio



Vector Databases vs RAG



Let's take a look at the code



Please feel free to reach out!

I can be found at @JessicaGarson on most places on the internet.

- → Work email: jess.garson@elastic.co
- → Personal emai: jessica.garson@gmail.com
- → Bluesky: jessicagarson.bsky.social
- → Mastodon: jessicagarson@macaw.social

