

Construye un RAG en VertexAI



¿Quién soy?



Mariia
Chizhikova
[@mariia.en.ia](https://mariia.en.ia)

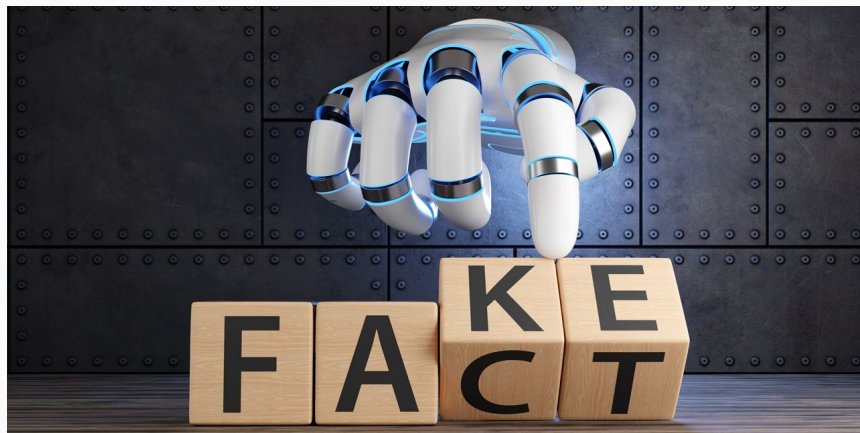
- Ingeniera de LLMs
- Divulgadora de IA

¿Qué es un RAG?



Retrieval Augmented
Generation

El problema



los cuñados estocásticos

01

Hace 2 años unos abogados de Estados Unidos fueron multados por 5000\$ por citar casos inexistentes

02

Un artículo de Microsoft recomendó el Banco de Alimentos de Ottawa como una atracción turística que visitar con estómago vacío

03

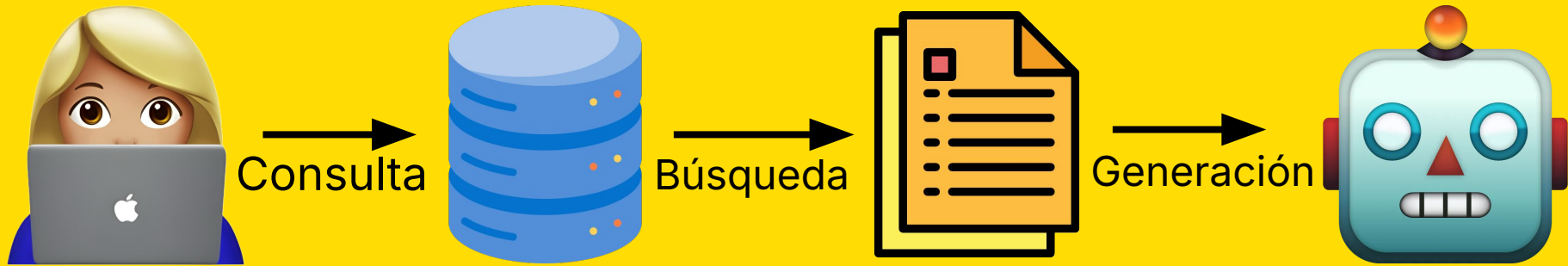
En Amazon se vendían libros que aconsejaban cocinar con setas venenosas

O simplemente...

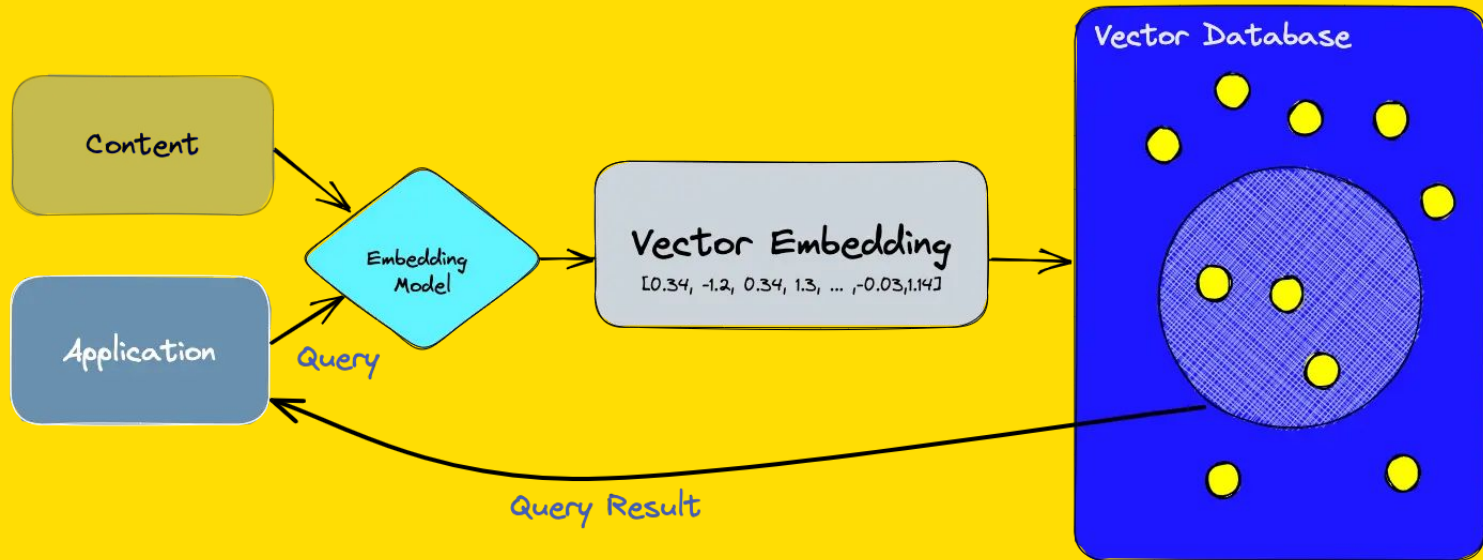


Yo es que tu
negocio no lo
conozco...

¿Y qué hacemos?



Base de datos vectorial



Bases de datos vectoriales



Pinecone



Pgvector



drant

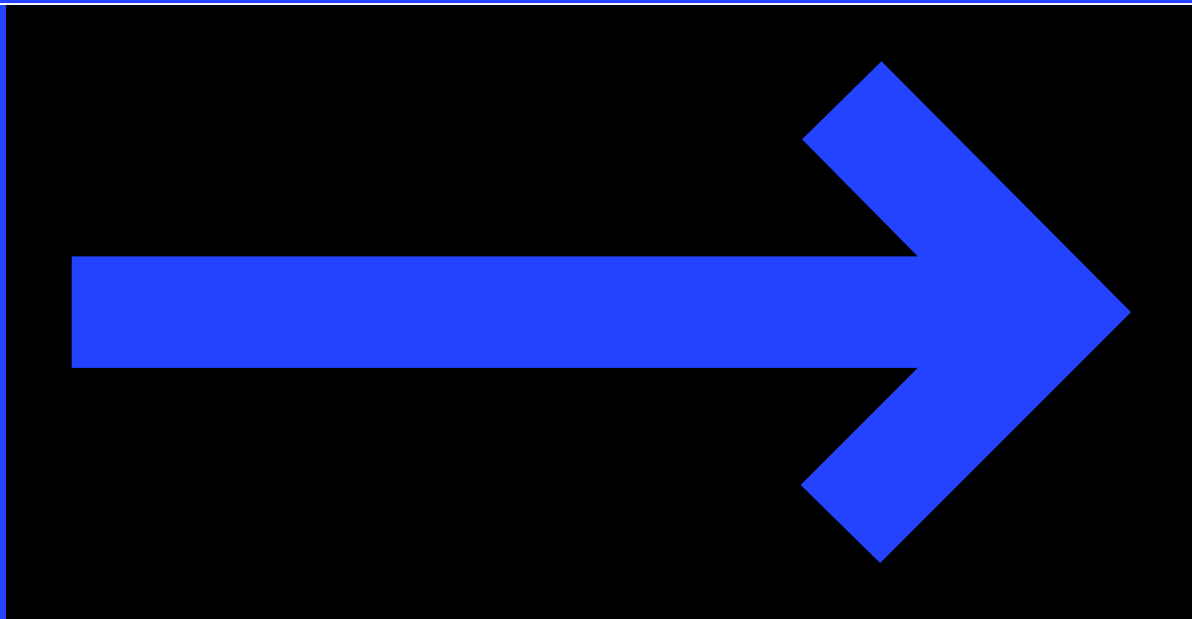


Chroma



Vertex AI

La plataforma unificada
de aprendizaje
automático de Google
Cloud



01

**Entrenamiento de
modelos de IA**

02

**MLOps y
escalabilidad**

03

IA generativa

El corpus

¿



?

¡Generemos uno!

01

Crear un json con prompts

02

Usarlo para generar los documentos

03

Hacer la RAG magia

Gracias!



Mariia
[@mariia.en.ia](https://mariia.en.ia)

