



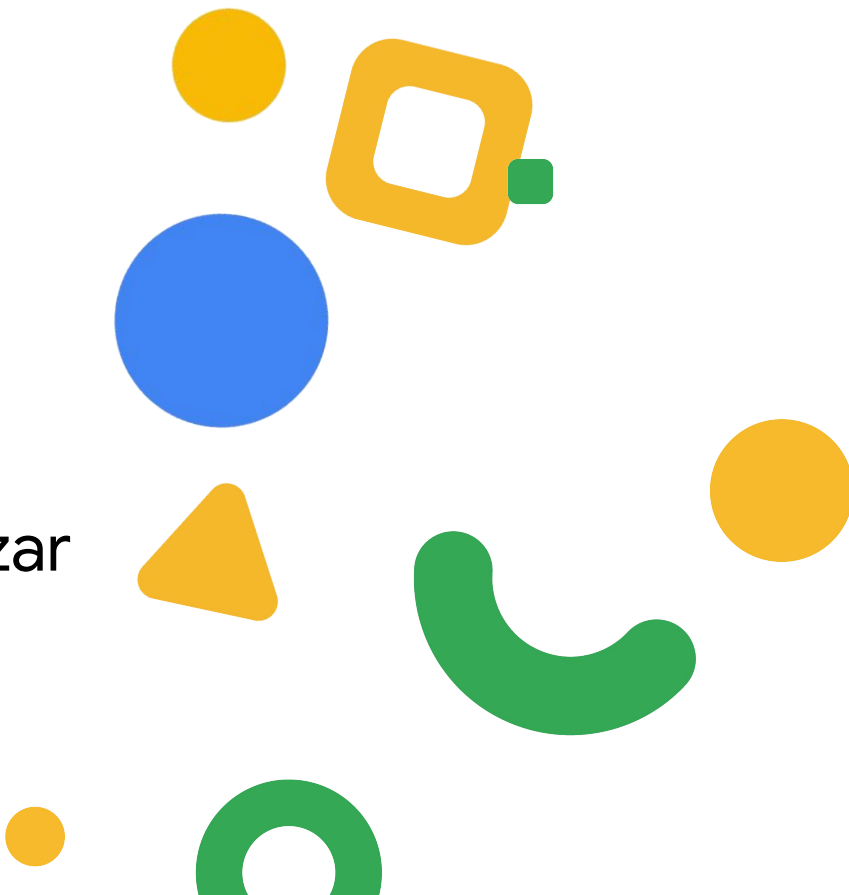
Marina de
Empresas



[Apuntate a resolver el reto](#)

Briefing Tech

Asistente multimodal para agilizar
la búsqueda de productos y
mejorar la atención al cliente.



Intro Challenge - MARINA DATAHUB

COFARES open competition for the best solution

[Apuntate a resolver el reto](#)

First bread crumb of the challenge:

Cofares dispone de una **página web B2C para las Farmacias denominada Área Privada**. En general es muy usada para llevar a cabo acciones que no están suficientemente bien desarrolladas en los ERPs de Farmacia, **destacando el buscador de artículos**.

Este buscador se usa mucho en las ocasiones en las que un paciente/consumidor intenta explicar al farmacéutico lo que necesita, situación habitual en los productos de venta libre y especialmente en los de parafarmacia. Por lo general, las indicaciones son vagas e imprecisas y el reto al que se enfrenta el farmacéutico es transformar eso en un artículo concreto que pueda ofrecerle al consumidor, que en muchas ocasiones acaba recurriendo a un buscador web para encontrar el producto.

Nuestro objetivo es que el farmacéutico entre siempre en el buscador del Área Privada y que este sea capaz de mostrar resultados precisos a partir de una descripción coloquial y en lenguaje natural, de forma que se evite el potencial riesgo que use un buscador y encuentre productos de competidores.

Para explorar diferentes soluciones a este reto, Cofares va a organizar un dataton en el DataHub de Marina de Empresas y Google en el que los participantes competirán por resolver este problema usando todo el catálogo de tecnologías y herramientas disponibles en Google Cloud Platform.

Solution steps and technologies - MARINA DATAHUB

COFARES open competition for the best solution

[Apuntate a resolver el reto](#)

Solution Steps and methodology:

Build high quality, scalable solution with **deep learning, generative ai, and statistical models** from Google.

- Business understanding
- Data understanding
- Data enrichment for both descriptions and images
- Data Transformation to embeddings space
- Create search engine or agent on top of the data
- Deploy search agent

Technologies:

Use Google technology and your AI/ML capabilities to design the best **future-proof solution.**

- Google Cloud Storage
- Bigquery
- Vector Search
- Vertex Agent Builder
- Embeddings APIs
- AlloyDB
- Gemini 1.5 Flash
- Colab & workbench
- Model Garden
- Cloud Run, GKE



Solution steps and technologies - MARINA DATAHUB

COFARES open competition for the best solution

Multimodal Data:

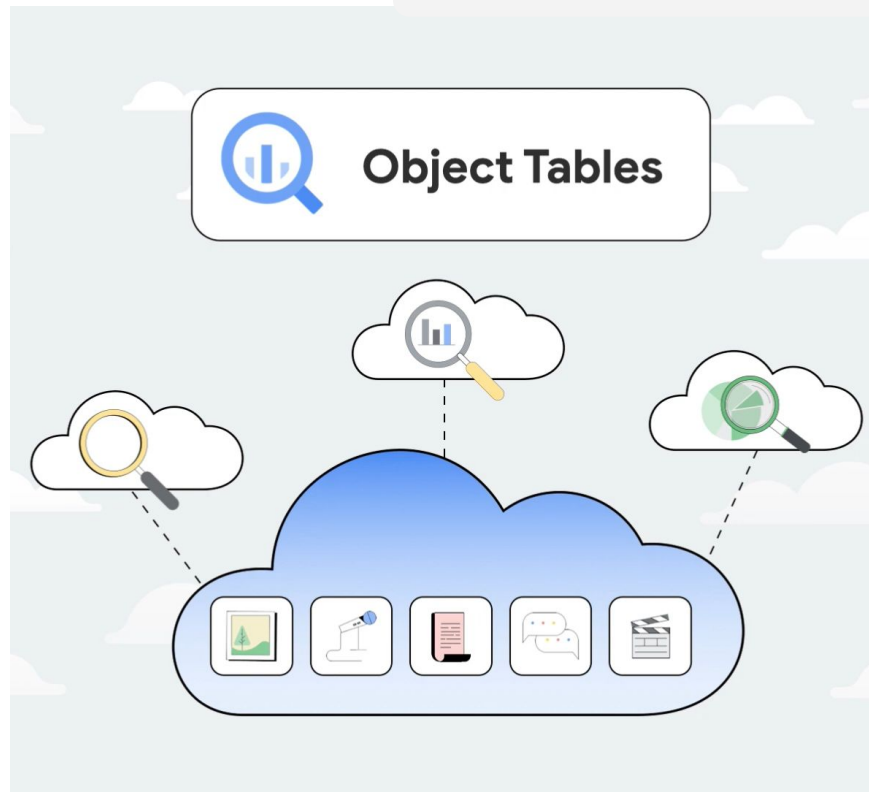
COFARES aportará un subconjunto de su catálogo de productos; información estructurada y desestructurada como imágenes o prospectos digitalizados.

El primer paso recomendable es crear una colección/estructura de datos multimodal (imágenes, texto...), esta puede ser generada como **Object/External Tables de BigQuery**, que son tablas que apuntan a objetos de datos no estructurados (imágenes, documentos...) almacenados en Cloud Storage.

A través de **Gemini 1.5** puedes extraer información de los datos cargados, información relevante, categorizada, clasificada, enriquecida de cada a que el buscador sea capaz de responder de forma precisa a estas descripciones vagas que el farmacéutico introduce en la interfaz.

Lo ideal es que estas respuestas estén ordenadas para **dar respuesta rápida y precisa a lo que el cliente necesita.**

[Apuntate a resolver el reto](#)



Solution steps and technologies - MARINA DATAHUB

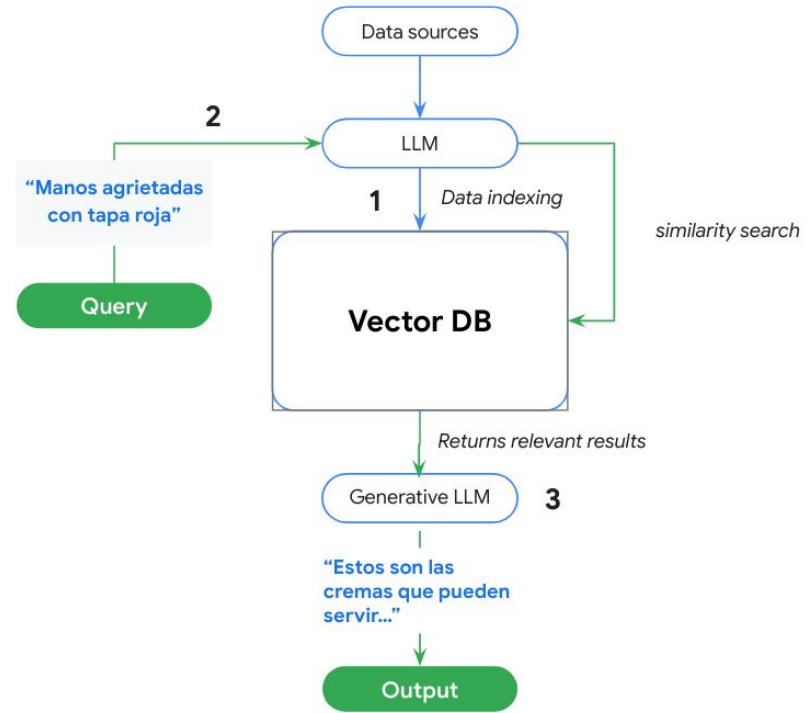
COFARES open competition for the best solution

[Apuntate a resolver el reto](#)

Vector Search is a key enabler of Gen AI applications; grounds LLM responses in relevant business data (aka RAG):

La tecnología de Búsqueda Vectorial (Vector Search) de Google permite usar la misma base fundacional sobre la que están contruidos productos como Google Search, YouTube o Play para implementar motores de recomendación que permiten explorar usando lenguaje natural ítems **semánticamente similares o semánticamente relacionados** dentro de catálogos con hasta billones de referencias.

A partir de la base de datos multimodal construida en la fase anterior, **se generarán los “embeddings”** para los diferentes ítems de la misma: una representación vectorial que incluye información visual y textual en un mismo espacio semántico para cada producto y permite la búsqueda mediante lenguaje natural con expresiones como: “Crema para manos agrietadas que tiene una tapa roja”, “productos similares al protector labial Acme”...



Solution steps and technologies - MARINA DATAHUB

COFARES open competition for the best solution

[Apuntate a resolver el reto](#)

User Interface:

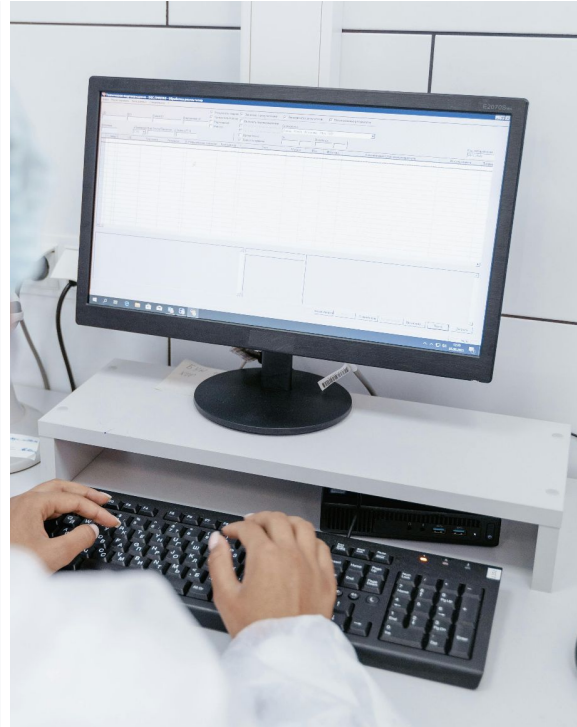
Integración del motor de recomendación basado en búsqueda vectorial dentro **dentro del Área Privada de Cofares**, incluyendo las siguientes capacidades:

Disponibilidad opcional para usuarios seleccionados, es decir, sólo para farmacias de confianza que actuarán como probadores de la tecnología.

Activación en función de la categoría de producto que se vaya a buscar, por ejemplo **sólo cuando se vaya a buscar parafarmacia**.

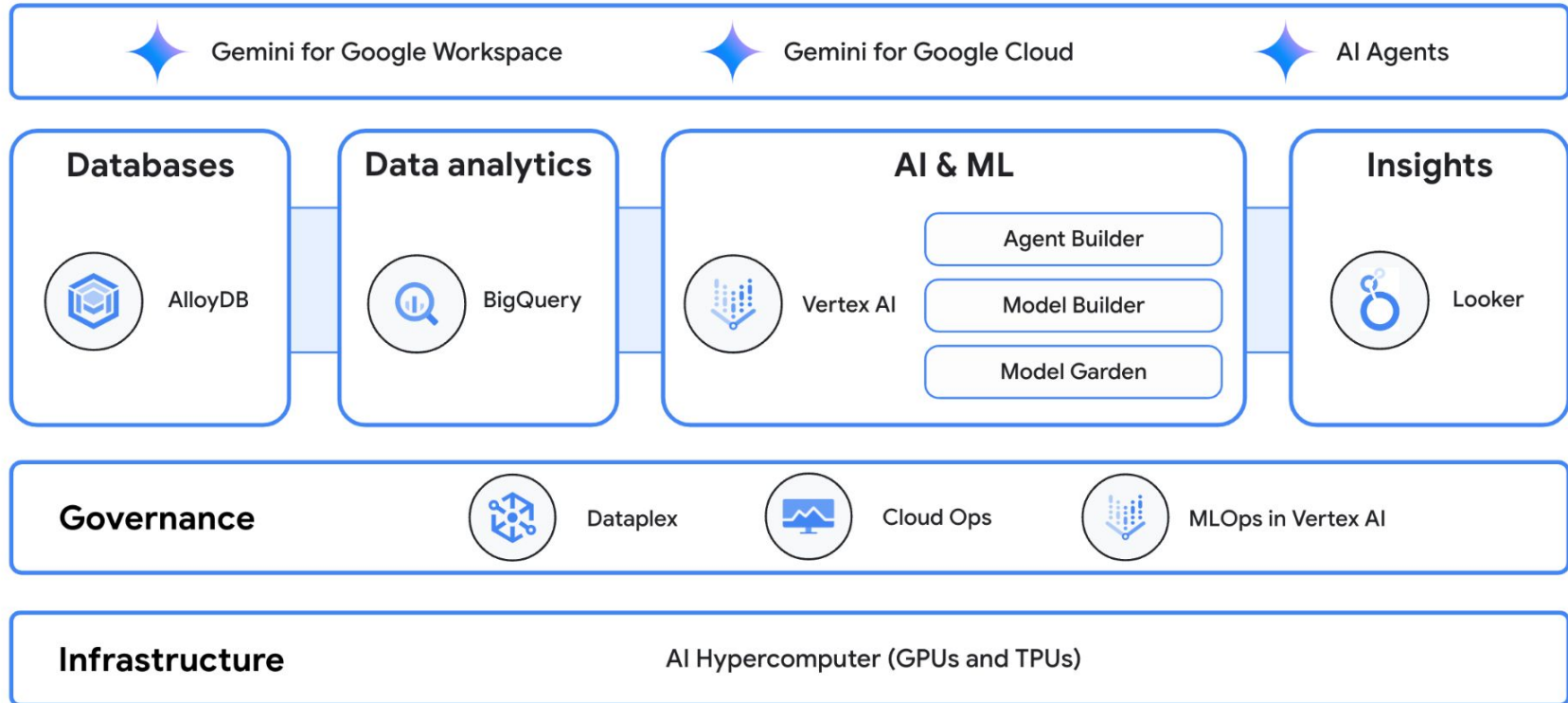
Disponibilización del buscador a través de un API que será usado en el mismo campo de búsqueda actual, de forma que no sea necesario un aprendizaje adicional por parte del usuario.

Interfaz gráfica sencilla que permita visualizar y probar el motor fuera del ERP de las farmacias



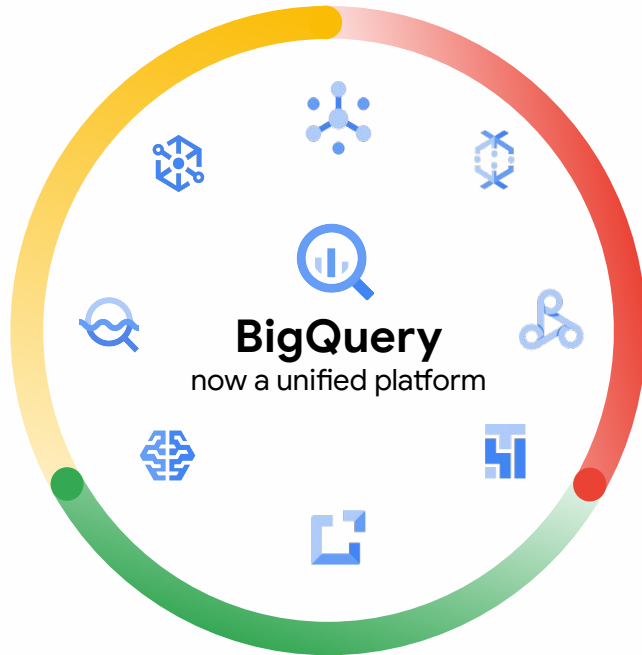
A unified platform from data to deployment and for all your predictive, generative, and agentic needs

[Apuntate a resolver el reto](#)



Simplifying and unifying data analytics

[Apuntate a resolver el reto](#)



- ➔ **Multi-engine**
- ➔ **Multi-storage**
- ➔ **Multi-format**
- ➔ **Multi-capability**
- ➔ **AI/ML/LLM integration**

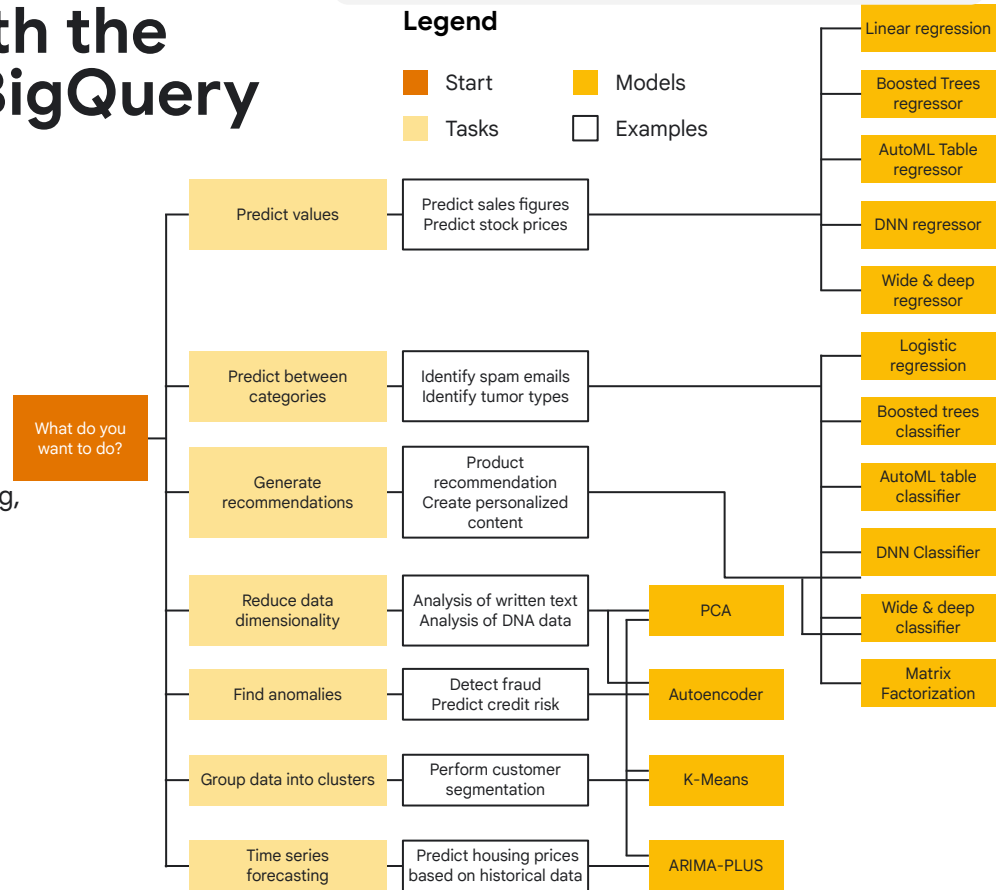
Bring AI to your data with the simplicity and scale of BigQuery

BigQuery ML: Democratize machine learning

→ **Data to AI**
Inference Engine for remotely hosted models and Google's pretrained models. Register your BigQuery ML models in the Vertex AI Model Registry

→ **Modeling capabilities**
ARIMA+, Explainable AI, Advanced Feature Engineering, Holiday modeling

→ **Model Ops**
Use Colab notebooks to perform ML workflows in BigQuery. Import tensorflow models for batch prediction or export models from BigQuery ML for online prediction



BigQuery ML for GenAI

Every data engineer can now be a ML engineer

[Apuntate a resolver el reto](#)

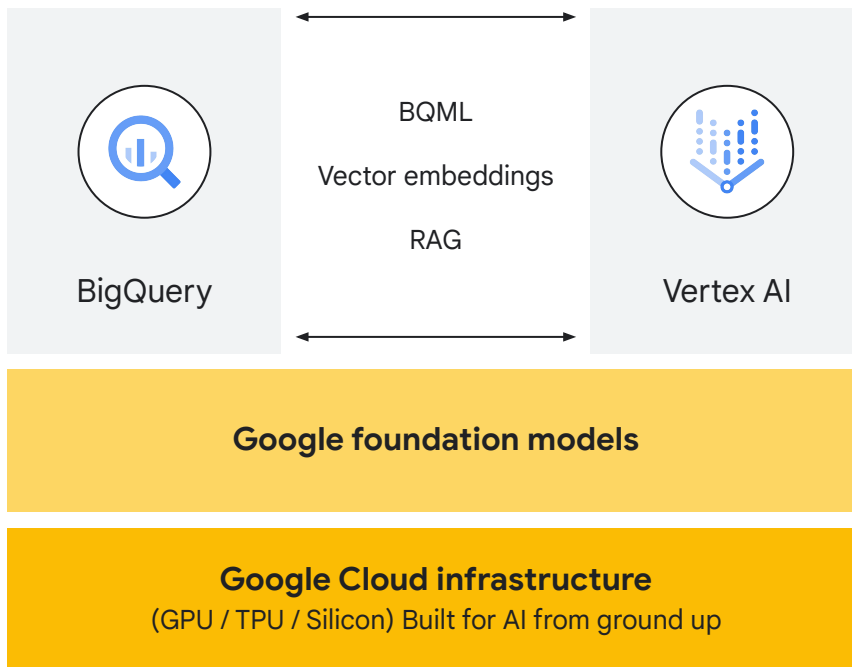
→ SQL LLM execution in BigQuery

- Summarization
- Sentiment Extraction
- Classification
- Data enrichment
- Entity Extraction
- Translation

→ Document and Speech AI in BigQuery

Secure and governed Doc & Speech AI over text and audio files directly in BigQuery

Unlock proprietary company data



Analyze unstructured data for answers at scale

BigQuery object tables offer SQL support for text, images, audio, video, documents, and conversations



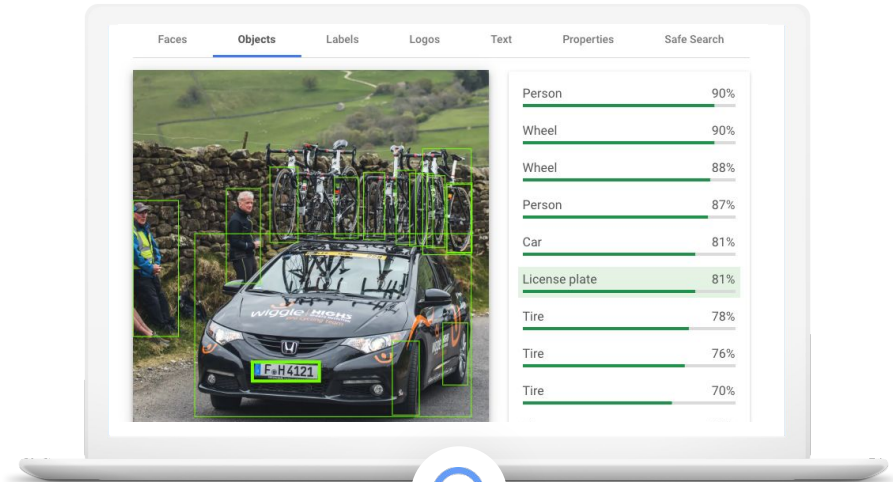
Easy

Object tables bring the simplicity and scale of SQL to unstructured data including images, audio, docs, and videos



Powerful

Google's leading AI technology in speech recognition, vision, translation, text processing accessible through BigQuery



Upload your images

Create BigQuery dataset and object table

Extract text from images

Run SQL queries

```
SELECT *, gcp_cloud_function(image_data) AS text  
FROM cars
```

BigQuery vector embeddings, search, and support for LLM frameworks

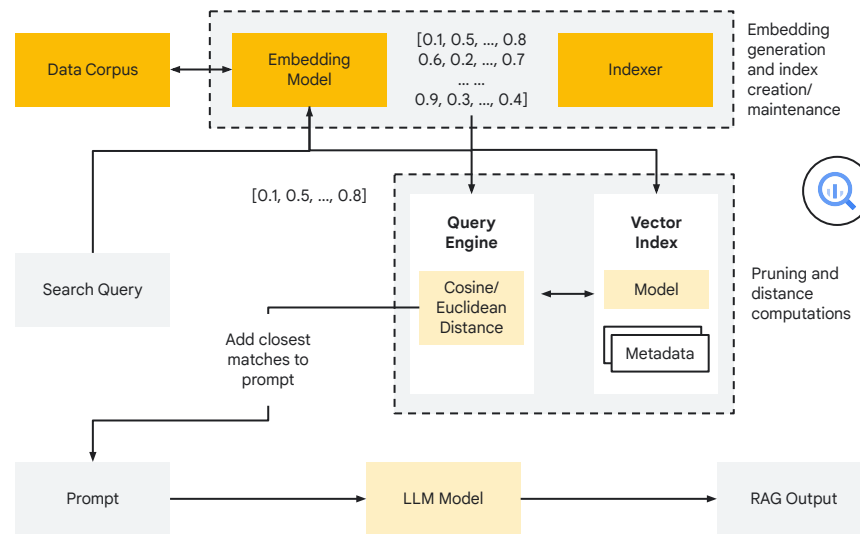
Supercharge LLMs with RAG and long term memory

- Use the scale of BigQuery to collect and standardize on vectors across multiple operational and vector databases and Cloud Storage
- Develop Retrieval Augmented Generation (RAG) design patterns directly in the data warehouse to combine the power of LLMs with your business data to improve accuracy
- Fully managed index keeps embeddings auto generated and in sync with Vertex for online prediction
- Support for popular LLM frameworks such as Langchain so your BigQuery data talks to your LLM

When do you need a vector embeddings and search?

- Retrieval augmented generation (RAG)
- Long-term memory for LLMs
- Semantic Search - search based on meaning vs keyword
- Similarity Search - Identify text similar to other text
- Recommendations

[Apuntate a resolver el reto](#)



RAG Workflow in BigQuery