### **ANÁLISIS Y RESUMEN DE TECNOLOGÍAS CLAVE**

#### **1. INTRODUCCIÓN**

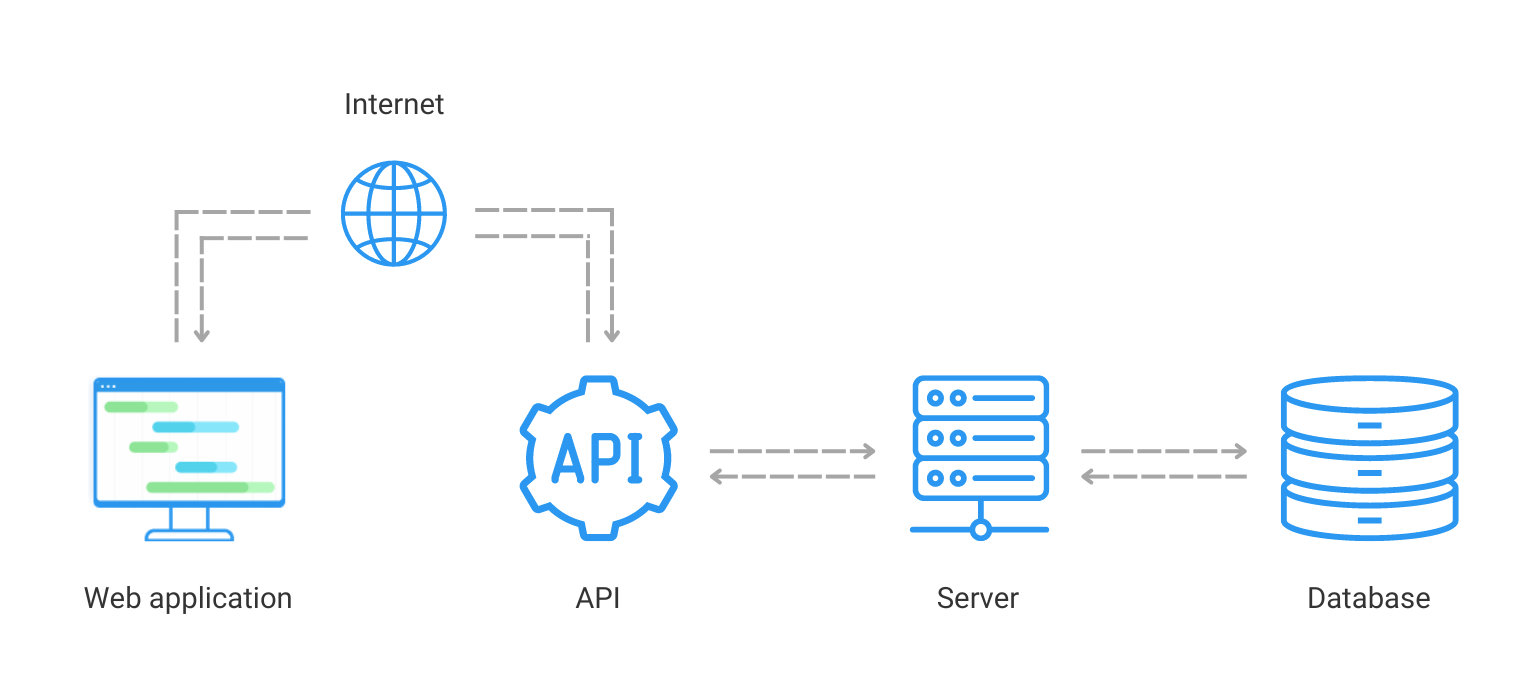
El presente informe tiene como objetivo principal exponer y resumir de manera clara y concisa una serie de tecnologías fundamentales que se utilizaran en el proyecto **DATACENSUS**, esto con el fin de proporcionar un marco de comprensión técnica sobre sus funcionalidades, interconexiones y aplicaciones.

#### **2. ANÁLISIS DE TECNOLOGÍAS**

**2.1. API (Interfaz de Programación de Aplicaciones)**

Una **API** es un conjunto de protocolos y herramientas que actúan como intermediarios entre dos sistemas de software, permitiendo la comunicación e intercambio de datos de forma estandarizada. Su función es abstraer la complejidad interna de un servicio, exponiendo únicamente las funcionalidades necesarias para la interacción. Esto facilita la integración de sistemas heterogéneos y la creación de arquitecturas modulares.

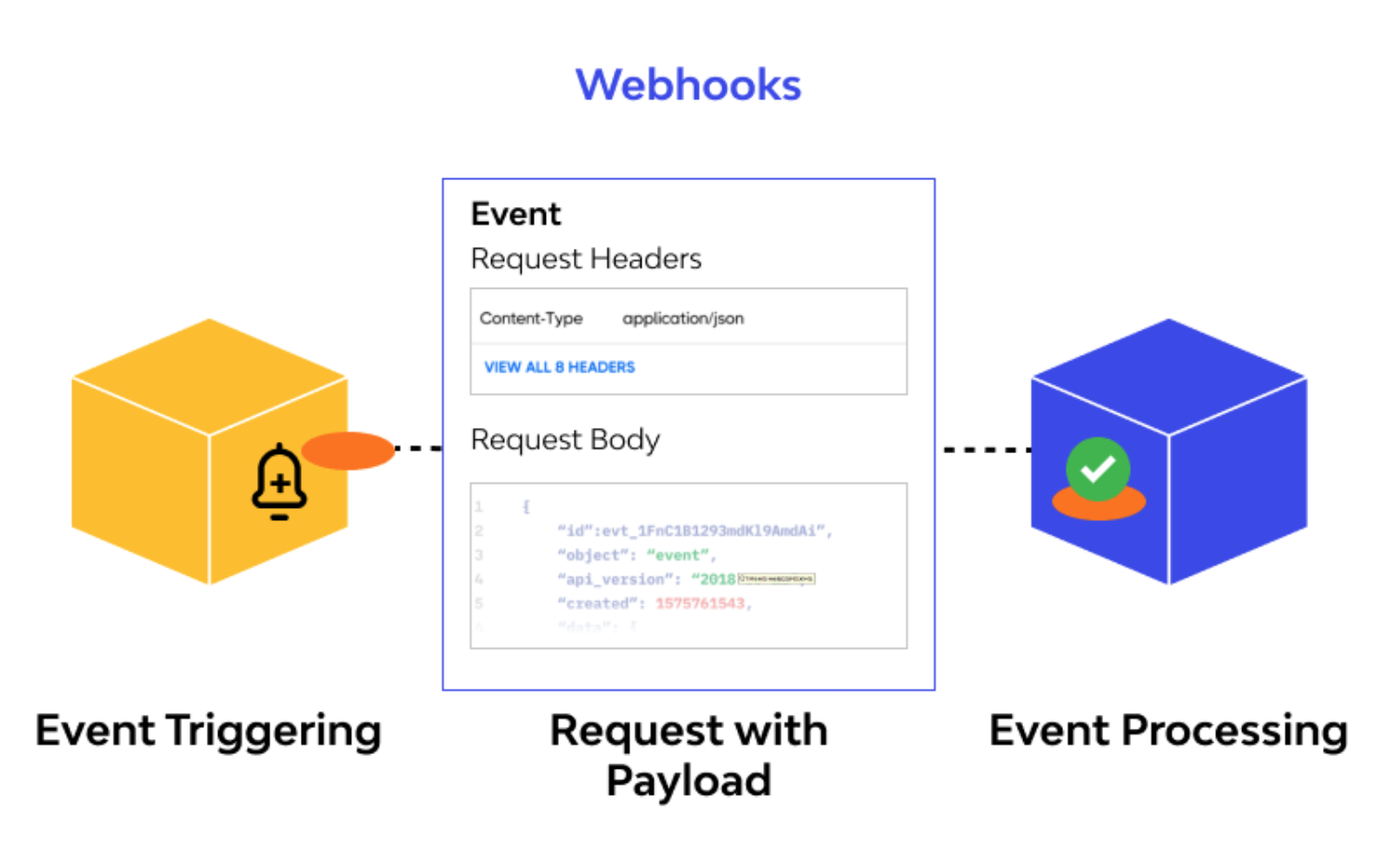
Caso de uso: Un intermediario que permite a las aplicaciones comunicarse, como cuando una aplicación de viajes usa la API de Google Maps para mostrar un mapa.

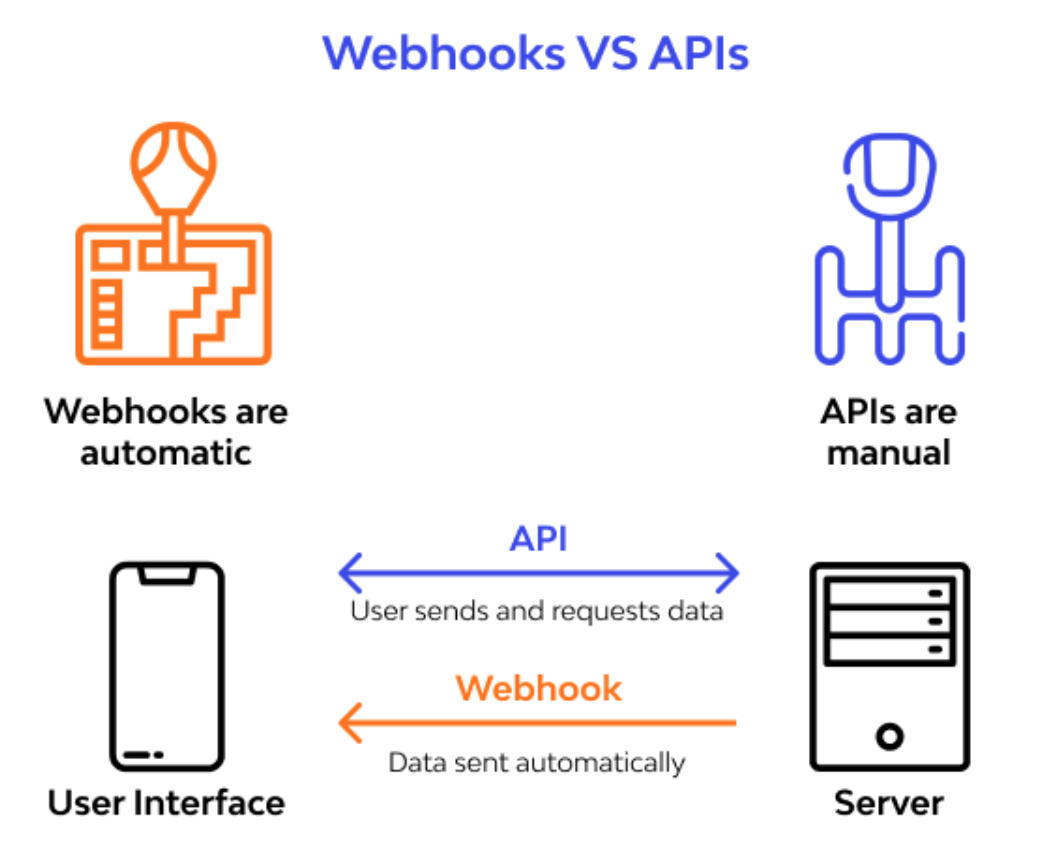


**2.2. Webhook**

Un **Webhook** es un método de comunicación asíncrono que permite a una aplicación notificar a otra sobre la ocurrencia de un evento en tiempo real. A diferencia de la consulta periódica (polling), el webhook opera mediante una llamada HTTP inversa, donde la aplicación fuente envía un "aviso" a una URL predefinida en la aplicación destino. Esto optimiza el uso de recursos al eliminar la necesidad de peticiones continuas.

Ej: Un sistema que notifica a una aplicación sobre un evento en tiempo real, como cuando Transbank le avisa a tu tienda online que un pago fue completado.

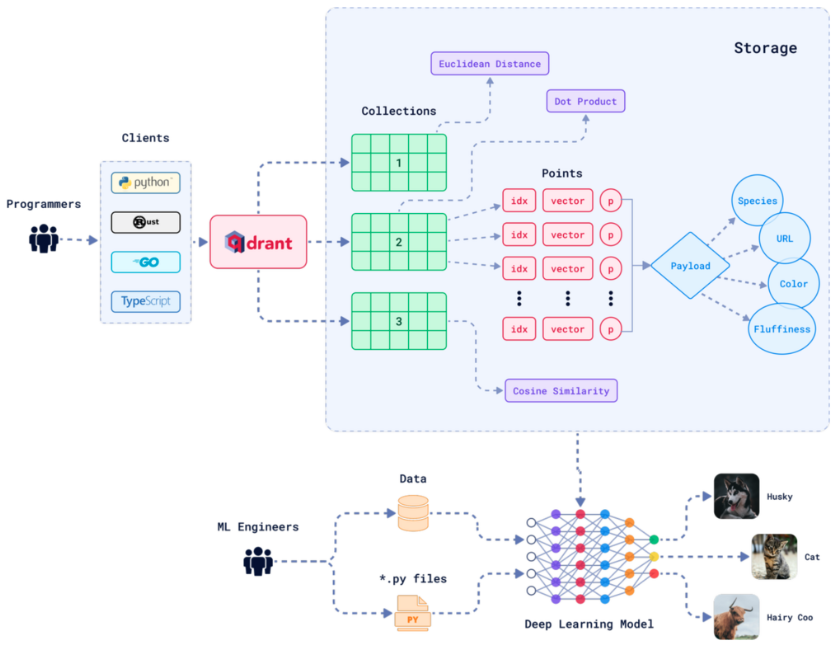


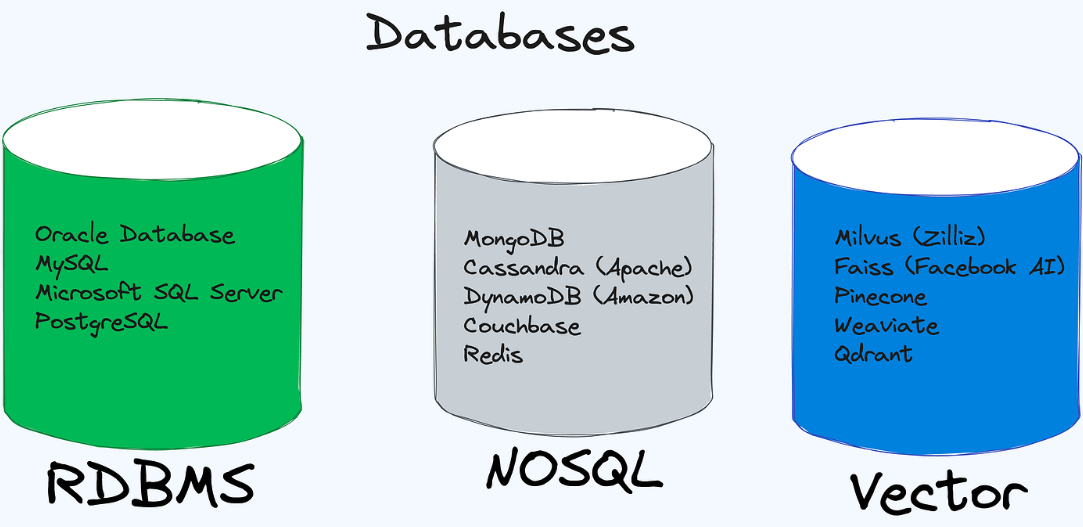


**2.3. Bases de Datos Vectoriales**

Las **bases de datos vectoriales** están especializadas en el almacenamiento y la búsqueda eficiente de datos representados como **vectores** de alta dimensión. A diferencia de las bases de datos relacionales o NoSQL, su principal fortaleza radica en la capacidad de realizar búsquedas de similitud (o "vecino más cercano"), lo que las hace indispensables para aplicaciones de búsqueda semántica, reconocimiento de patrones y sistemas de recomendación.

Caso de uso: Almacenan y buscan datos por su similitud, útil para encontrar canciones que suenen parecidas o cuando interactúas con un modelo de lenguaje grande (ChatGPT).





**2.4. LLM (Grandes Modelos de Lenguaje)**

Los **LLMs** (Large Language Models) son modelos de inteligencia artificial basados en arquitecturas de *transformer* que han sido entrenados con enormes volúmenes de datos textuales. Su capacidad inherente para comprender, generar y procesar lenguaje natural les permite realizar una amplia gama de tareas, desde la redacción de texto hasta la traducción y el resumen de información. Representan el núcleo de la IA generativa actual.

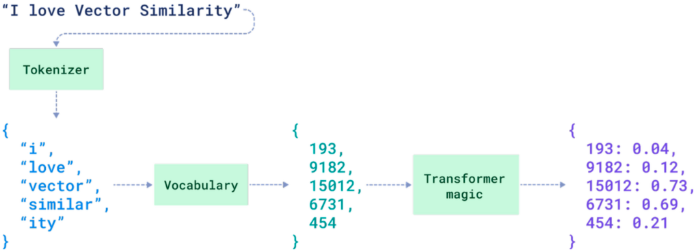
Caso de uso: Un modelo de IA que entiende y genera lenguaje, como el cerebro detrás de ChatGPT que escribe textos coherentes.

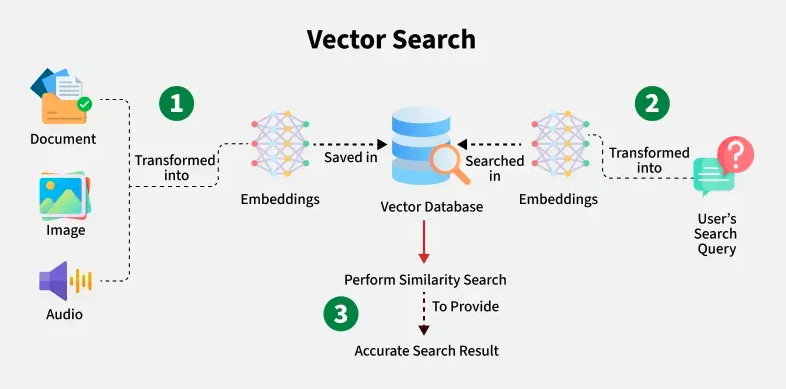
****

**2.5. Embeddings**

Un **embedding,** (incrustación) es la representación numérica de objetos de datos (como palabras, frases o imágenes) en un espacio vectorial. Este proceso de codificación captura la semántica y el contexto de los datos, de modo que elementos con significados similares se ubican cerca uno del otro en el espacio vectorial. Los embeddings son la base que permite a los algoritmos de IA procesar y comparar datos no estructurados.

Caso de uso: Transforma texto en números para que las computadoras entiendan su significado, permitiendo que un algoritmo sepa que "rey" y "príncipe" son palabras similares.

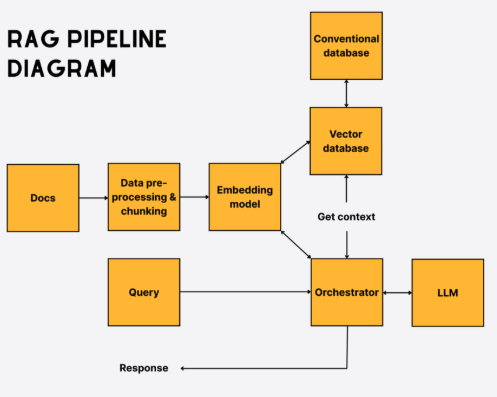
****

****

**2.6. RAG (Recuperación Aumentada por Generación)**

**RAG** (Retrieval-Augmented Generation) es una arquitectura híbrida que potencia un **LLM** al integrar un sistema de recuperación de información. El proceso implica dos fases: primero, se recupera información relevante de una fuente externa (ej. una base de datos vectorial) basada en la consulta del usuario; segundo, esta información recuperada se utiliza para contextualizar y mejorar la respuesta generada por el LLM. Esta técnica reduce las "alucinaciones" del modelo y asegura la precisión y actualidad de la información.

Caso de uso: Mejora un modelo de IA usando datos externos, como un chatbot que busca en documentos de tu empresa para responder preguntas específicas sobre políticas internas.

****

**2.7. n8n**

**n8n** (Node-based Workflow Automation) es una plataforma de automatización de flujos de trabajo de código abierto. Permite a los usuarios crear, conectar y automatizar tareas entre múltiples aplicaciones y servicios mediante una interfaz visual y basada en nodos. Su flexibilidad y naturaleza *low-code* lo convierten en una herramienta eficaz para la integración de sistemas y la optimización de procesos de negocio sin la necesidad de programación avanzada.

Caso de uso: Una herramienta para automatizar flujos de trabajo sin código, como un robot que automáticamente envía un email de bienvenida a cada nuevo suscriptor de tu blog.

#### 

#### 

#### 

#### 

#### 

#### 

#### 

#### 

#### 

#### **3. CONCLUSIÓN**

Las tecnologías analizadas en este informe demuestran un ecosistema interconectado en el que cada componente desempeña un rol crítico. Desde la comunicación básica entre sistemas (**APIs** y **Webhooks**), pasando por la gestión y comprensión de datos no estructurados (**Bases de Datos Vectoriales** y **Embeddings**), hasta la creación de sistemas inteligentes y contextualizados (**LLMs** y **RAG**), este conjunto de herramientas y metodologías impulsa la innovación y la eficiencia en el desarrollo de software moderno, el cual es clave en el funcionamiento correcto de **DATACENSUS**..