

Explorando el Potencial del Retrieval Augmented Generation (RAG)



En el mundo de la inteligencia artificial, el Retrieval Augmented Generation (RAG) se está destacando como una técnica revolucionaria que mejora la forma en que recuperamos y generamos información. Para comprender mejor esta tecnología, conversamos con [Javier Padrón](#), Head of Architecture de CleverIT.

En esta entrevista, Javier nos explica en detalle qué es el RAG, cómo funciona, sus beneficios, aplicaciones prácticas y la manera en que [CleverIT](#) está aprovechando esta tecnología para optimizar procesos y ofrecer soluciones innovadoras a sus clientes.

¿Qué es el RAG y cómo funciona?

J: El Retrieval Augmented Generation (RAG) es una técnica de inteligencia artificial que combina la recuperación de información con la generación de texto para responder preguntas o generar contenido. Utiliza datos basados en documentos para recuperar información relevante y luego utiliza un modelo de lenguaje para generar respuestas precisas y detalladas a partir de esa información.

¿En qué casos se puede aplicar?

J: RAG es especialmente útil en situaciones donde la precisión y la riqueza de información son cruciales, como en atención al cliente, asistencia médica, investigación legal y académica, y análisis financiero.

Dinos tres beneficios que todos deberían conocer sobre el RAG:

J: Tres beneficios principales que todos deberían conocer son:

- **Precisión mejorada:** Al combinar recuperación de datos y generación de texto, RAG proporciona respuestas más precisas basadas en información verificada.
- **Escalabilidad:** Puede manejar una gran variedad de consultas ajustando los recursos computacionales según la necesidad.
- **Actualización continua:** Al integrarse con archivos no estructurados, estos pueden ser actualizados en cualquier momento, por lo cual esta solución puede tener datos más recientes para generar respuestas.

Un dato curioso sobre el RAG:

J: Un dato interesante podría ser cuando en ChatGPT agregas archivos de texto y la solución es capaz de colocar en memoria los datos del mismo o responder preguntas específicas del mismo, estás usando la técnica de RAG.

¿Cómo puedo saber si el RAG sirve para mi empresa o no?

J: Si tu empresa desea automatizar tareas basados en documentos de texto donde se tienen grandes volúmenes de consultas o datos que requieren respuestas detalladas y precisas, es ideal para esta solución. RAG es particularmente útil donde la calidad de la información impacta directamente la toma de decisiones y la satisfacción del cliente.

¿De qué manera trabaja CleverIT con esta tecnología?

J: Dentro de proyectos internos como [Groowcity](#) hemos utilizado estas técnicas IA para la mejora y automatización de tareas. También tenemos experiencia en diversas nubes y especialmente en Azure para optimizar procesos de información y respuesta, integrando soluciones personalizadas con los sistemas existentes de los clientes para mejorar la eficiencia y precisión en servicios como soporte técnico, gestión de conocimiento y asistencia personalizada.

En conclusión, el RAG representa una poderosa herramienta para las empresas que buscan mejorar la precisión y eficiencia en la gestión de información y la generación de contenido. Con la experiencia de expertos como CleverIT, la implementación de RAG puede transformar la manera en que las organizaciones manejan sus datos y brindan servicios a sus clientes.