



## Recepción de resúmenes CCG

### Título / Autores / Institución

#### TÍTULO DE LA PONENCIA

IA para la Vigilancia Volcánica: Sistema Inteligente de consulta y análisis de Información multifuente

#### AUTORES

Oscar Cadena, John Meneses, Juan Carlos Alvarado

#### INSTITUCIÓN

Servicio Geológico Colombiano

#### CORREO ELECTRÓNICO

ocadena@sgc.gov.co

### Estilo preferido

#### ESTILO DE PRESENTACIÓN

- Presentación Oral

### Categoría del resumen

#### ÁREA TEMÁTICA

Inteligencia Artificial

#### LINEAS TEMÁTICAS AI

Inteligencia Artificial

### Resumen

#### PALABRAS CLAVE

Vigilancia volcánica, sistema inteligente, consulta información, RAG, CAG, SQL

#### CONTENIDO DEL RESUMEN

La vigilancia de la actividad volcánica en Colombia requiere el análisis constante de múltiples fuentes de información, entre ellas datos sísmicos, geoquímicos, geodésicos y documentos técnicos generados periódicamente por los observatorios volcanológicos. En este contexto, presentamos un sistema innovador que aplica técnicas de Inteligencia Artificial (IA) para facilitar la consulta, análisis, interpretación y generación automática de respuestas redactadas en lenguaje natural a partir de datos estructurados (bases de datos SQL) y no estructurados (informes PDF, Word y texto plano).

El sistema combina una arquitectura híbrida basada en los enfoques RAG (Retrieval-Augmented



Generation) y CAG (Cache-Augmented Generation), integrando modelos de lenguaje natural con recuperación semántica de información desde una base vectorial. Una vez ingresada la pregunta a una interfase tipo chat, esta se clasifica automáticamente lo que para determinar si la consulta debe resolverse mediante generación de sentencias SQL (a una base de datos) o mediante análisis de documentos. En ambos casos, el sistema utiliza modelos LLM (Large Language Models) locales para generar respuestas precisas, contextualizadas y redactadas en lenguaje técnico apropiado para el monitoreo volcánico.

El sistema incorpora técnicas avanzadas como:

- Segmentación temática por secciones numeradas, para preservar la estructura lógica de los informes técnicos.
- Re-ranking que mejora la relevancia de la información recuperada al priorizar los más cercanos semánticamente a la consulta.
- Visualización en tiempo real con respuestas explicativas, tablas de datos y pasajes recuperados visibles para el analista.
- Extensibilidad hacia técnicas como Multi-Query Retrieval e HyDE (Hypothetical Document Embedding), que permiten enriquecer la recuperación de información en contextos complejos o con escasa información directa.

El sistema ha sido probado con informes oficiales del Servicio Geológico Colombiano y ha mostrado resultados prometedores en la generación automática de respuestas a preguntas técnicas del usuario. La propuesta representa un paso importante hacia la automatización de tareas analíticas en las geociencias, permitiendo que investigadores y profesionales del monitoreo volcánico consulten grandes volúmenes de datos de manera más eficiente y con respuestas generadas en lenguaje natural. A futuro, el sistema podría escalarse para integrar datos en tiempo real y herramientas de visualización geoespacial.