

POR MARCO ZECCHETTO, DESDE LAS VEGAS, EEUU.

Las aplicaciones en torno a la Inteligencia Artificial (IA) generativa, aquella que crea nuevo contenido a partir de otro existente, sigue sumando aplicaciones en empresas.

Este martes, Oracle, la multinacional de servicios tecnológicos, lanzó su primer agente basado en esta tecnología para Oracle Cloud Infrastructure (OCI), su plataforma de computación en la nube, el que permite a las empresas obtener información más rápida de sus bases de datos y optimizar procesos.

El anuncio lo realizó el vicepresidente senior de IA y gestión de datos en OCI de Oracle, Greg Pavlik, durante el Oracle CloudWorld 2024 que se está realizando Las Vegas, en Estados Unidos, quien dijo que el agente es parte de la estrategia de IA generativa de la compañía, la que está enfocada en desarrollar soluciones en todos los niveles -infraestructura, bases de datos, hasta aplicaciones SaaS (software como servicio)- que "repercutan en los resultados del negocio de las empresas".

A nivel de infraestructura, Pavlik se refirió a OCI Supercluster, una red de unidades de procesamiento de gráficos (GPU) de latencia "ultrabaja", que permite procesar grandes cantidades de datos a alta velocidad. "Esta capacidad es esencial para las empresas tecnológi-

Oracle CloudWorld 2024 se está realizando en Las Vegas, EEUU.



Oracle se suma a la ola de IA generativa y lanza su primer agente para agilizar búsquedas en sus bases de datos

- La multinacional de servicios tecnológicos presentó una herramienta que permite a las compañías obtener respuestas más rápidas desde sus bases de datos, con capacidad para evitar alucinaciones.*

cas que entrenan modelos de IA generativa a gran escala, optimizando el costo y el rendimiento", dijo.

Comentó que otro de foco de la estrategia es la colaboración con Meta y la startup Cohere, para mejorar sus modelos de lenguaje grandes (LLM, en inglés) para su funcionamiento en diferentes aplicaciones empresariales.

Agente de IA generativa

El OCI Generative AI Agents, como se denomina este primer agente, tiene capacidades de generación aumentada por recuperación (RAG), un método que combina la búsqueda de información en fuentes como bases de datos, con respuestas automáticas en lenguaje natural.

De esta forma, las empresas pueden obtener respuestas más rápidas y precisas desde sus propios documentos y bases de datos, evitar procesos manuales de programación y optimizar la búsqueda de información o la automatización de tareas. También tiene la capacidad de autoevaluarse para evitar alucinaciones.

Entre los usos del agente, Pavlik destacó la automatización de búsquedas en bases de datos judiciales para acelerar la investigación judicial, optimización de respuestas en centros de atención al cliente y facilitar la comprensión del historial de compras y tendencias de clientes.