

Estudio de caso



Airbnb

La gestión de datos maestros empodera una red global de empleados de Airbnb

INDUSTRIA

Turismo y hostelería

CASO DE USO

Gestión de datos maestros / Grafo de conocimientos

OBJETIVO

Crear un motor de búsqueda de recursos de datos que sea fácil de navegar con resultados fiables

RETO

Los empleados estaban supeditados a datos almacenados de forma aislada y a la adquisición de conocimientos a través de otros empleados, lo que reducía la productividad.

SOLUCIÓN

Creación del Dataportal usando Neo4j para conectar recursos de datos diversos.

RESULTADOS

- Los empleados de Airbnb disponen de un punto único de acceso a todos los recursos de datos disponibles.
- Neo4j se integra fácilmente con la pila de tecnología existente de Airbnb para facilitar la programación.

A medida que la plantilla de Airbnb se expandía rápidamente por todo el mundo, se dificultaba su productividad en un panorama de datos de complejidad creciente. Acudieron a Neo4j para que les ayudara a crear un sistema de gestión de datos maestros rápido y fácil de usar, conocido como Dataportal.

La empresa

Airbnb posibilita experiencias de viaje únicas, con un mercado de intercambio online para alquiler de alojamiento a corto plazo. Utilizando diversos filtros de búsqueda, un usuario puede explorar fácilmente los más de 4 millones de alojamientos inscritos en el sitio, disponibles en 65.000 ciudades de 191 países. Airbnb anunció unos ingresos de más de 1.000 millones de dólares en el tercer trimestre de 2018, y el valor de la empresa asciende a 31.000 millones de dólares.

El reto

Desde sus comienzos como frágil startup, Airbnb está hoy en día entre los nombres más conocidos del mercado online de alojamiento. Con el éxito vino la rápida expansión de su plantilla, actualmente compuesta de 3.500 empleados ubicados en 20 oficinas internacionales.

En cualquier organización grande y compleja, los recursos de datos internos y externos en aumento constante se vuelven restrictivos e inmanejables, especialmente si están diseminados en varias plataformas. Tras un año en la empresa, el ingeniero de software John Bodley llegó a la conclusión de que los datos de Airbnb estaban almacenados en ubicaciones restringidas, eran inaccesibles o carecían de un contexto adecuado.

La enorme cantidad de datos dispersos menoscaba su ventaja operativa: tenían más de 200.000 tablas en múltiples agrupaciones en su almacén de datos Hive principal, 10.000 gráficos y dashboards Superset, 6.000 experimentos en métricas, más de 6.000 libros de trabajo y gráficos Tableau y más de 1.500 puntos de conocimientos.

Bodley también se percató de que los empleados dependían unos de otros para obtener respuestas a sus preguntas, algo que en última instancia reducía la productividad. “En una encuesta periódica entre los empleados”, explica, “la afirmación ‘La información que necesito para hacer mi trabajo es fácil de encontrar’ siempre recibía una calificación muy baja”.

Sabía que necesitaban democratizar los datos para que cualquier empleado, independientemente de su nivel de conocimiento al respecto, estuviera capacitado para encontrar recursos, con plena confianza en la relevancia y la fiabilidad de los resultados.

La estrategia

“A un nivel muy alto, si simplemente queremos buscar algo”, cuestionó Bodley, “¿cómo podemos enmarcar los datos de manera significativa para la búsqueda, la clasificación y la relevancia?”

Estudio de caso



“¿Por qué elegimos Neo4j como nuestra base de datos de grafos? Nuestros datos representan un grafo, o sea que para almacenarlos resultaba lógico utilizar una base de datos de grafos. Porque es ágil. Queríamos un sistema realmente rápido y eficaz. Porque es popular. Es la base de datos de grafos número uno del mundo. Y porque se integra realmente bien.”

– John Bodley

Ingeniero de Software, Airbnb

Su equipo inició el desarrollo de Dataportal, un espacio de datos de autoservicio integrado que presenta una visión contextual y global de los datos de Airbnb, para que los empleados naveguen de forma fácil y rápida siempre que necesiten acceso o respuestas para sus necesidades de trabajo cotidianas.

Bodley y su equipo determinaron que la herramienta requería cuatro características principales: búsqueda, contexto y metadatos, datos centrados en la plantilla y datos centrados en los equipos. La conexión de las relaciones entre cada recurso de datos, y sus metatipos correspondientes, sería la clave para proporcionar los enlaces de datos necesarios para una herramienta plenamente funcional y disponible para los empleados.

La solución

Con diversidad de recursos (tablas de datos, dashboards, informes, usuarios, equipos, resultados empresariales, etc.), cada uno con sus propios niveles de contexto y conexión, Bodley y su equipo no tardaron en darse cuenta de que su ecosistema de datos se representaba mejor en forma gráfica. Eso desembocó en la base de datos de grafos de Neo4j.

“Por cuatro razones principales”, dice Bodley. “Una, por lógica. Nuestros datos representan un grafo, o sea que para almacenarlos resultaba lógico utilizar una base de datos de grafos. Porque es ágil. Queríamos un sistema realmente rápido y eficaz. Porque es popular. Es la base de datos de grafos número uno del mundo.... Y porque se integra realmente bien.”

En lo relativo a la velocidad, el Dataportal está pensado para ser un motor de búsqueda de recursos de datos, donde las interacciones rápidas, detalladas y precisas contribuyen a fomentar la exploración. Neo4j ofrece la manera más rápida de buscar en millones de conexiones de datos por segundo.

En lo relativo a la integración, Airbnb tenía su propia pila de tecnología, que incluía Elasticsearch y Python. “Utilizamos Flask como un marco web Python ligero para la API, que es coherente con varias herramientas de datos de código abierto de Airbnb como Airflow, The Knowledge Repository y Superset”, detalla Bodley. “La aplicación web de una sola página hace uso de React y Redux.”

Neo4j se integró bien con todos los lenguajes de programación preferidos de Airbnb, además de permitir enriquecer las clasificaciones de búsqueda aprovechando la topología de grafos. Introducen a diario datos de Hive en la base de datos de grafos de Neo4j; así conectan sus datos aislados desde una perspectiva de relaciones, para facilitar resultados de búsqueda contextuales rápidos y muy relevantes.

Los resultados

Con Neo4j, Airbnb consiguió aunar todo su ecosistema de datos y hacerlo apto para búsquedas, relevante y fiable, incluso para empleados recién llegados e inexpertos.

En lugar de depender de los conocimientos de otros empleados dispersos por el mundo, dependencia que merma la producción de trabajo de calidad, el Dataportal es el único recurso de Airbnb para encontrar todos los datos relevantes, especialmente información de empleados y equipos, de importancia crucial para el rendimiento diario.

Y dada la alta escalabilidad de Neo4j, el Dataportal también está preparado para facilitar el crecimiento futuro de la compañía, conectando instantáneamente nuevas contrataciones con nuevos proyectos en tiempo real.

Neo4j es el líder en tecnología de base de datos de grafos. Siendo la base de datos de grafos más utilizada en el mundo, ayudamos a marcas globales - que incluyen [Comcast](#), [NASA](#), [JBS](#) y [Volvo Cars](#) - a revelar y predecir como las personas, procesos y sistemas se interrelacionan. Usando el enfoque en las relaciones, aplicaciones desarrolladas usando Neo4j afrontan problemas de conexión de datos tales como [análisis e inteligencia de datos](#), [detección de fraude](#), [recomendaciones en tiempo real](#) y [grafos de conocimiento](#). Más información en [neo4j.com](#).

¿Tiene preguntas sobre Neo4j?

Contacte con nosotros en todo el mundo:

info@neo4j.com

neo4j.com/contact-us