

Fallstudie



Airbnb

Stammdatenmanagement unterstützt das globale Mitarbeiter-Netzwerk von Airbnb

BRANCHE

Reise- und Hotelgewerbe

USE CASE

Stammdatenmanagement / Knowledge Graph

ZIEL

Einfach zu nutzende Suchmaschine der Datenquellen für zuverlässige Ergebnisse

HERAUSFORDERUNG

Eingeschränkte Produktivität und Mitarbeiterperformance durch Datensilos und intrinsische Wissensweitergabe

LÖSUNG

Auf Neo4j basierte Suchmaschine „Dataportal“ zur Verknüpfung verteilter Datenquellen

ERGEBNISSE

- Zentrale Anlaufstelle für Airbnb-Mitarbeiter für die Datenrecherche
- Einfache Integration von Neo4j in den Airbnb TechStack für einfache Programmierung

Angesichts steigender weltweiter Mitarbeiterzahlen fiel es Airbnb immer schwerer, die verteilten Daten effektiv zu managen, ohne Produktivität einzubüßen. Für das neue schnelle und benutzerfreundliche Stammdatenmanagementsystem „Dataportal“ kam daher Neo4j zum Einsatz.

Das Unternehmen

Airbnb verbindet Menschen mit einzigartigen Reiseerlebnissen und stellt auf seinem Online-Marktplatz Kurzzeitunterkünfte zum Mieten und Vermieten vor. Über Suchfilter können Nutzer in über 4 Millionen Angeboten in 65.000 Städten und 191 Ländern nach der passenden Unterkunft suchen. Für das dritte Quartal 2018 meldete Airbnb einen Umsatz von mehr als 1 Mrd. US-Dollar. Das Unternehmen wird auf 31 Mrd. US-Dollar geschätzt.

Die Herausforderung

Einst ein ehrgeiziges Startup, ist Airbnb heute der Inbegriff für die Online-Vermietung von Unterkünften. Mit dem Erfolg wuchs das Unternehmen rapide. Heute sind rund 3.500 Mitarbeitern an 20 Standorten weltweit bei Airbnb beschäftigt.

Eine stetig wachsende IT-Landschaft aus internen und externen Datenquellen – verteilt über verschiedene Plattformen – kann für jedes große und komplexe Unternehmen schnell unübersichtlich und restriktiv werden. Nach seinem ersten Jahr bei Airbnb war für Softwareingenieur John Bodley klar: Die Daten befanden sich ohne echten Kontext in unzugänglichen Silos.

Im Hive-Data-Warehouse waren mehr als 200.000 Tabellen über verschiedene Cluster, 10.000 Superset-Charts und Dashboards, 6.000 Experimente in Metrics, über 6.000 Tableau Workbooks und Charts und über 1.500 Knowledge Posts verteilt. Diese enorme Menge an unübersichtlichen Daten machte ein effizientes betriebliches Arbeiten unmöglich.

Bodley stellte fest, dass Mitarbeiter vor allem eigenes, intrinsisches Wissen nutzten, um Fragen zu beantworten. „Bei unseren regelmäßigen Mitarbeiterbefragungen, schnitten wir immer sehr schlecht ab, wenn es darum ging, wie leicht eine benötigte Information zu finden war“, erinnert sich Bodley. Er wusste, um jeden Mitarbeiter unabhängig von seiner Rolle und seiner Kenntnisse Zugang zu Informationen zu verschaffen, mussten die Daten demokratisiert werden. Zuverlässigkeit und Relevanz der Suchergebnisse spielten dabei eine zentrale Rolle.

Die Strategie

„Vereinfacht gesagt, ging es uns um die Suchfunktion“, so Bodley. „Wie können wir unsere Daten so modellieren, dass die Suche, das Ranking und die Relevanz besser werden?“

Das Team begann mit der Entwicklung von Dataportal, einem integrierten Self-Service-Datenraum, der eine ganzheitliche, kontextbezogene Sicht auf alle Airbnb-Daten ermöglicht.

Fallstudie



„Warum Neo4j als Graphdatenbank? Unsere Daten stellen einen Graphen dar, so dass es nur logisch ist, auch eine Graphdatenbank als Speicher zu verwenden. Wir wollten ein schnelles und performantes System und diese Lösung ist flexibel. Als weltweit führende Graphdatenbank ist Neo4j weitverbreitet. Und schließlich lässt sich die Datenbank sehr gut integrieren.“

– John Bodley
Software Engineer, Airbnb

Mitarbeiter sollten schnell und einfach durch die Datenlandschaft navigieren können – um jederzeit Antworten für ihre täglichen Aufgaben zu erhalten.

Gemeinsam machte das Team schnell vier wesentliche Grundanforderungen aus: Suche, Kontext & Metadaten, Mitarbeiterdaten und Teamdaten. Die Verknüpfung der Datenquellen einschließlich der jeweiligen Metadaten waren der Schlüssel, um Mitarbeitern ein voll funktionsfähiges Suchwerkzeug bereitzustellen.

Die Lösung

Für Bodley und sein Team war schnell klar: Die unterschiedlichen Daten (z. B. Tabellen, Dashboards, Berichte, Nutzer, Teams, Geschäftsergebnisse) sowie die dazugehörigen Kontext- und Verbindungsebenen lassen sich am Besten in einem Graphen abbilden. Das führte sie zur Neo4j Graphdatenbank.

„Es gab vier maßgebliche Gründe für Neo4j“, erzählt Bodley. „Erstens stellen unsere Daten einen Graphen dar, so dass es nur logisch ist, auch eine Graphdatenbank zu verwenden. Wir wollten ein schnelles und performantes System und diese Lösung ist flexibel. Als weltweit führende Graphdatenbank ist Neo4j weitverbreitet. Und schließlich lässt sich die Datenbank sehr gut integrieren.“

Dataportal ist als Suchmaschine der Airbnb Datenquellen konzipiert und soll durch schnelle, detaillierte und genaue Interaktionen letztendlich neue Erkenntnisse liefern. Neo4j ist hier der schnellste Weg, um Millionen von Datenverbindungen pro Sekunde zu durchsuchen.

Airbnb hatte seinen eigenen Tech-Stack, einschließlich Elasticsearch und Python. „Wir verwenden Flask als leichtes Python Web-Framework für die API. Diese ist kompatibel zu einer Reihe von Open-Source-Datentools bei Airbnb, darunter Airflow, The Knowledge Repo und Superset“, so Bodley. „Für die Web-App kommen React und Redux zum Einsatz.“

Neo4j ließ sich gut in die von Airbnb genutzten Programmiersprachen integrieren, und ermöglicht es gleichzeitig, die Suchergebnisse mittels Graph-Topologie anzureichern. Täglich werden Daten von Hive nach Neo4j verschoben, um Datensilos miteinander zu verknüpfen und Datenbeziehungen zu nutzen. So lassen sich schnelle, hochrelevante und kontextbezogene Suchergebnisse erzielen.

Die Ergebnisse

Mit Neo4j konnte Airbnb sein gesamtes Datenökosystem miteinander vernetzen. Selbst ganz neue Mitarbeiter können auf einfache Weise nach relevanten und zuverlässigen Informationen suchen.

Das Dataportal ist nun die zentrale Anlaufstelle von Airbnb für die Datenrecherche und wirkt sich insbesondere beim Auffinden von Mitarbeiter- und Teamdaten entscheidend auf die Performance aus. Mitarbeiter müssen nicht länger auf undokumentiertes Wissen zurückgreifen und können auf konsistent hohem Niveau arbeiten.

Dank der hohen Skalierbarkeit von Neo4j ist Dataportal in der Lage, das Unternehmenswachstum voranzutreiben, neue Mitarbeiter einzulernen und neue Projekte in Echtzeit zu integrieren.

Neo4j ist der führende Anbieter von Graphtechnologie. Die weltweit am häufigsten eingesetzte Graphdatenbank unterstützt Unternehmen wie [Deutsches Zentrum für Diabetesforschung e.V.](#), [NASA](#), [UBS](#) und [Daimler](#) darin, Zusammenhänge zwischen Menschen, Prozessen, Standorten und Systemen aufzudecken und datengestützte Vorhersagen zu treffen. Der Fokus auf Datenbeziehungen ermöglicht es, smarte Anwendungen zu entwickeln und die Herausforderungen vernetzter Daten zu meistern – von [Analytics und künstlicher Intelligenz](#) über [Betrugserkennung](#) und [Echtzeit-Empfehlungen](#) bis hin zu [Knowledge Graphen](#). Weitere Informationen unter [Neo4j.com](#).

Fragen zu Neo4j?

Kontakt:
info@neo4j.com
neo4j.com/contact-us