

Étude de cas



Airbnb

La gestion des données de référence donne davantage de possibilités à un réseau mondial d'employés d'Airbnb

SECTEUR

Voyage & Hébergement

CAS D'USAGE

Gestion des données de référence / graphe de connaissances

OBJECTIF

Créer un moteur de recherche de ressources de données dans lequel il est simple de naviguer, avec des résultats fiables.

DÉFI

Les employés étaient freinés par des données cloisonnées et des connaissances parcellaires, ce qui réduisait la productivité

SOLUTION

Création de Dataportal en utilisant Neo4j pour relier des données disparates

RÉSULTATS

- Les employés d'Airbnb disposent d'un portail unique pour chaque source de données disponible
- Neo4j s'est intégré aisément à la pile technologique existante d'Airbnb, ce qui facilite la programmation

Face au développement rapide de ses effectifs dans le monde et un panorama de données toujours plus difficile à comprendre, la productivité d'Airbnb était freinée. L'entreprise a demandé à Neo4j de l'aider à créer un système de gestion des données de référence rapide et convivial, baptisé Dataportal.

L'entreprise

Airbnb connecte les gens à des expériences de voyage uniques grâce à une place de marché en ligne pour la location d'un logement de courte durée. En utilisant une variété de filtres de recherche, l'utilisateur peut facilement naviguer sur le site parmi plus de 4 millions d'annonces d'hébergement en location qui couvrent 65 000 villes et 191 pays. Au troisième trimestre 2018, Airbnb a annoncé un chiffre d'affaires de plus de 1 milliard de dollars et une valeur de 31 milliards de dollars.

Le défi

Airbnb, ancienne startup en difficulté, a désormais acquis son renom sur le marché de l'hébergement en ligne. Le succès de l'entreprise s'est accompagné d'une expansion rapide de ses activités et de ses effectifs qui comptent actuellement 3 500 employés répartis dans 20 bureaux dans le monde.

Dans chaque grande organisation complexe, le panorama des ressources de données internes et externes ne cesse de croître, en particulier les données disséminées sur diverses plateformes, et le tout finit par devenir ingérable et restrictif. Après un an chez Airbnb, John Bodley, ingénieur logiciel, a compris que les données d'Airbnb étaient complètement cloisonnées, inaccessibles ou sans contexte approprié.

Avec plus de 200 000 tableaux dans leur principal entrepôt de données Hive répartis sur plusieurs clusters, 10 000 graphiques et tableaux de bord Superset, 6 000 expériences sous forme de métriques, plus de 6 000 classeurs et graphiques Tableau et plus de 1 500 articles de connaissance, les énormes quantités de données désordonnées nuisaient à l'avantage opérationnel de la société.

John Bodley a également remarqué que les employés comptaient sur des connaissances parcellaires informelles pour répondre à des questions ce qui, en fin de compte, entravait la productivité. « Nous menons régulièrement un sondage auprès de nos employés et l'affirmation « L'information dont j'ai besoin pour faire mon travail est facile à trouver » a toujours obtenu de très mauvaises évaluations » explique-t-il. Il savait qu'il fallait démocratiser les données afin que tout employé, quelle que soit sa fonction ou son aisance avec les données, ait les moyens de trouver les ressources et puisse compter sur des résultats de recherche pertinents et fiables.

La stratégie

« À très haut niveau, on veut simplement chercher quelque chose » note John Bodley, « du coup, comment structurons-nous nos données de manière judicieuse pour la recherche, le classement et la pertinence ? »

Étude de cas



« Pourquoi avons-nous choisi Neo4j comme base de données de graphes ? Nos données représentent un graphe, il était donc logique d'utiliser une base de données de graphes pour le stocker. Neo4j est agile. Nous voulions un système rapide et performant, reconnu. Il s'agit de la première base de données de graphes au monde. Et finalement, elle s'intègre très bien »

– John Bodley,
Ingénieur logiciel, Airbnb

Son équipe a commencé à développer Dataportal, un espace de données intégré en libre-service qui présente une vue contextuelle et holistique des données d'Airbnb. Cet espace permet aux employés de naviguer facilement et rapidement dans leur environnement de données lorsqu'ils ont besoin d'un accès ou de réponses dans leur travail quotidien.

John Bodley et son équipe ont déterminé que l'outil avait besoin de quatre caractéristiques principales : recherche, contexte et métadonnées, données axées sur les employés et données axées sur l'équipe. La connexion de relations entre chaque ressource de données - et leurs métatypes associés - allait être la clé pour fournir les liens nécessaires entre données et rendre l'outil entièrement fonctionnel et prêt à l'emploi par les employés.

La solution

Avec diverses ressources (tableaux de données, tableaux de bord, rapports, utilisateurs, équipes, résultats commerciaux, etc.), et différents niveaux de contexte et de relations pour chacune, John Bodley et son équipe ont rapidement compris qu'un graphe était le meilleur format pour représenter au mieux leur écosystème de données, ce qui les a conduits à la base de données de graphes Neo4j.

« Cela s'explique pour quatre raisons » poursuit John Bodley. « La première, nos données représentent un graphe, il était donc logique d'utiliser une base de données de graphes pour le stocker. Neo4j est agile. Nous voulions un système rapide et performant, reconnu. Il s'agit de la première base de données de graphes au monde. Et finalement, elle s'intègre très bien. »

En termes de vitesse, Dataportal se veut un moteur de recherche de ressources de données dans lequel des interactions rapides, détaillées et précises finissent par stimuler l'exploration. Neo4j offre le moyen le plus rapide de rechercher parmi des millions de relations entre données par seconde.

Airbnb disposait de sa propre pile technologique en place incluant Elasticsearch et Python. « Nous utilisons Flask comme framework web Python léger pour l'API, ce qui est cohérent avec un certain nombre d'outils de données open source d'Airbnb comme Airflow, The Knowledge Repo et Superset, » précise John Bodley. « L'application web à page unique exploite React et Redux. »

Neo4j s'intègre bien avec tous les langages de programmation préférés d'Airbnb, et ses équipes peuvent enrichir les classements de recherche par la topologie des graphes. Chaque jour, les données de Hive passent dans la base de données de graphes Neo4j - ce qui connecte les données en silo d'un point de vue relationnel - afin de favoriser des résultats de recherche contextuels rapides et très fiables.

Les résultats

Avec Neo4j, Airbnb a pu relier l'ensemble de son écosystème de données et le rendre consultable, pertinent et fiable, même pour les nouveaux employés les moins informés.

À la place d'un personnel lointain misant sur des connaissances tribales, ce qui donne des coups d'arrêt dans la production d'un travail de qualité, Dataportal est la ressource unique d'Airbnb pour trouver toutes les données pertinentes, particulièrement en termes d'informations centrées sur les employés et les équipes, essentielles à la performance quotidienne.

Et comme Neo4j offre une grande évolutivité, Dataportal est prêt à faciliter la croissance future de l'entreprise en connectant instantanément de nouvelles recrues et de nouveaux projets en temps réel.

Neo4j est le leader de la technologie des bases de données de graphes. Avec le plus grand nombre de déploiements au monde, Neo4j aide des entreprises mondiales - comme Airbus, [Michelin](#), [NASA](#), [Crédit Agricole](#) et [Volvo Cars](#) - à prédire et identifier la façon dont les personnes, les processus, les lieux et les systèmes sont interconnectés. Grâce à cette approche par les relations, les [applications](#) mises au point en utilisant Neo4j relèvent les défis associés aux données connectées, tels que [l'analytique et l'intelligence artificielle](#), la [détection de fraude](#), les [recommandations](#) en temps réel et les [graphes de connaissance](#). Pour en savoir plus, merci de consulter [Neo4j.com](#) et [@Neo4jFr](#).

Des questions sur Neo4j ?

Contactez-nous :
info@neo4j.com
neo4j.com/contact-us