Étude de cas

Caterpillar, Inc.

Grâce au traitement du langage naturel à grande échelle par Neo4j, la réparation de matériel gagne en efficacité

SECTEUR
Industriel

CAS D'USAGE
Traitement du langage naturel / intelligence artificielle

OBJECTIF
Créer un outil de TLN permettant d'extraire du sens à partir de documents à grande échelle.

DÉFI
La base de données relationnelle renvoyait des chaînes de texte non analysées qui manquaient de sens.

SOLUTION
Création d'un outil de TLN évolutif avec Neo4j pour révéler les tendances et améliorer l'efficacité des réparations et de la maintenance.

RÉSULTATS
- Recherches en temps réel parmi des millions de documents
- Possibilité de relier la cause à l'effet et d'identifier des questions à un niveau plus approfondi

Caterpillar, Inc. détient plus de 27 millions de documents de suivi de réparation et de maintenance de véhicules. En utilisant Neo4j pour effectuer un traitement du langage naturel (TLN), l'entreprise peut désormais faire des recherches à grande échelle pour détecter les tendances et les problèmes issus des réparations, prescrire des solutions et faire de précieuses prédictions. Le tout afin d'améliorer l'efficacité des réparations et de la maintenance des véhicules dans toute l'entreprise.

L'entreprise

Caterpillar, Inc. est le premier fabricant mondial d'équipement de construction et d'exploitation minière, de moteurs diesel et à gaz naturel, de turbines à gaz industrielles et de locomotives diesel-électriques.

Les clients s'adressent à Caterpillar pour développer leurs actifs dans les infrastructures, l'énergie et les ressources naturelles. Son portefeuille de 20 marques offre des services et des solutions pour répondre aux besoins uniques d'une variété d'industries et de clients à travers le monde. En 2018, les ventes et le chiffre d'affaires de l'entreprise, qui existe depuis 90 ans, ont dépassé 54 milliards de dollars à l'échelle mondiale.

Le défi

À chaque fois qu'une machine de Caterpillar arrive en réparation ou en maintenance, le technicien crée un document de garantie qui enregistre la demande, l'analyse du problème et sa solution.

L'entreprise dispose ainsi d'un référentiel de documents techniques à grande échelle qui, en grande partie, présente une bonne qualité selon les standards de l'étiquetage et de la linguistique informatisée. Toutefois, il restait à connecter de nombreuses données hétérogènes.

Consciente que des données de valeur résidaient dans ses 27 millions de documents, l'entreprise a lancé la création d'un outil de TLN afin de mettre à jour les tendances et les relations non-visibles.

Depuis dix ans, Caterpillar explore le TLN à des fins de gestion de maintenance des véhicules et de chaîne d'approvisionnement. Une grande partie des données a pu être cartographiée correctement dans différents domaines, sans pour autant réussir à obtenir une représentation de ces connaissances, ni à les exploiter de façon pertinente.

« Nous voulions créer un système permettant de poser n'importe quelle question tant qu'elle restait dans le domaine » explique Ryan Chandler, Data Scientist en chef chez Caterpillar. « Il fallait mettre au point un système de dialogue pour tester l'utilisation des graphes, présenter une interface utilisateur ouverte capable de répondre aux questions et développer une fonctionnalité d'interface vocale homme-machine ».

Étude de cas

« L'enjeu du dialogue en langage naturel entre l'homme et la machine va sans doute porter sur l'analytique et le mécanisme pour y parvenir est le traitement du langage naturel - une adéquation idéale aux bases de données de graphes. »

–Ryan Chandler,
Data Scientist en chef
Caterpillar, Inc.

La solution

Comme les graphes constituent le premier niveau de structure d'une infrastructure et offrent une très grande flexibilité, ils conviennent naturellement au traitement du langage et au machine learning.

Les techniques de traitement du langage le plus souvent utilisées sont soit l'analyse des liens entre le verbe et les autres mots qui s'y réfèrent, soit par l'utilisation d'un arbre syntaxique. Ces deux structures sont des graphes.

Caterpillar a utilisé Neo4j pour les structures de données de graphes afin de créer un modèle logique de connaissances. Cette alternative NoSQL aux bases de données relationnelles a permis d'établir des ontologies et de faire des déductions.

Pour passer du langage naturel aux résultats d'une requête de graphes, l'équipe a mis au point une architecture de données qui intègre le texte via une boîte à outils open-source de TLN, qui utilise Python pour combiner les phrases en chaînes, ajuster les limites et ne pas tenir compte du texte inutile. Les données sont également importées depuis les systèmes ERP SAP et non SAP.

L'outil de classification de machine learning apprend à partir d'un ensemble de données déjà étiquetées avec des termes tels que cause ou demande qu'il applique ensuite au reste des données.

WordNet est utilisé comme dictionnaire lexicographique pour donner une définition aux mots, l'analyseur syntaxique de dépendances de Stanford analyse le texte et Neo4j identifie les motifs et les relations, établit les hiérarchies et ajoute les ontologies.

Le tout combiné permet aux utilisateurs de faire des recherches pertinentes à partir de simples requêtes Cypher.

Les résultats

La prescription d'une action découle naturellement du système, par exemple l'étape appropriée à suivre pour identifier le problème et son diagnostic si un moteur fait des heurts.

« Les solutions de ce genre vont beaucoup plus loin que ce que l'on peut faire avec une IA standard » déclare Morgan Vawter, directrice de l'analytique chez Caterpillar, Inc. « Elles incarnent l'état d'esprit de l'entreprise, son domaine de connaissances et sont ainsi le résultat d'une traduction minutieuse de l'humain à la machine. »

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

© 2021 Neo4j, Inc. Tous droits réservés.