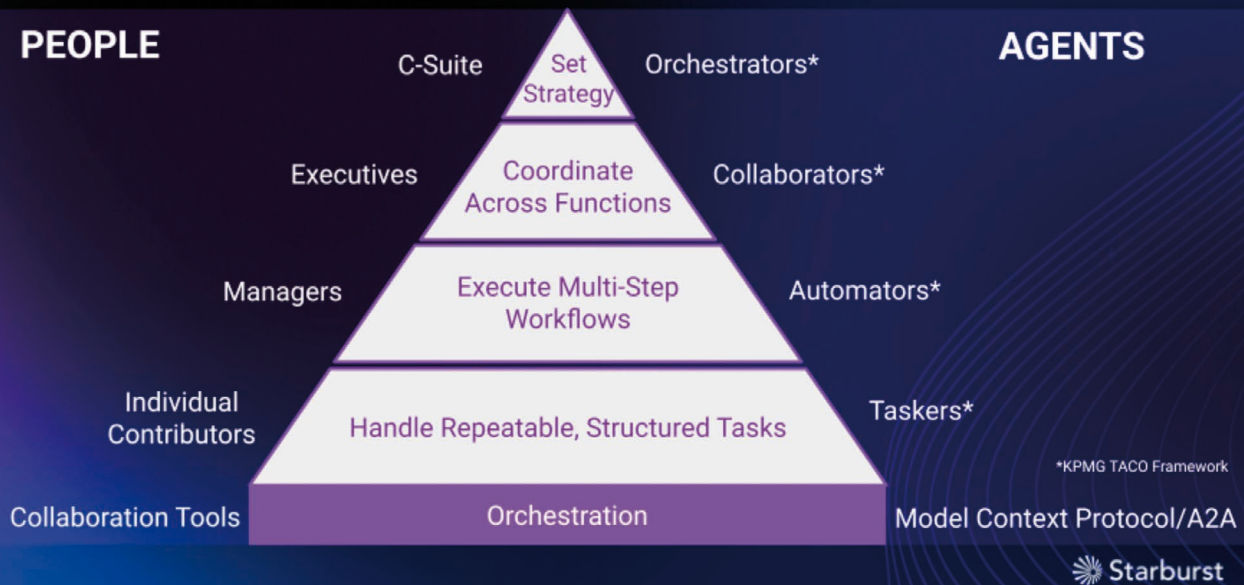


Erfolgsrezept für Datenanalytik und KI

Bei KI-Initiativen sollten Unternehmen beachten, dass die richtige Datenstrategie die Grundlage der KI-Strategie bildet

From People to Agents: A Shared Structure of Work



Unternehmen sind immer an Technologien interessiert, die echte Geschäftsprobleme lösen. Angesichts der stark zunehmenden geschäftlichen Bedeutung von Daten haben sich diese Möglichkeiten erweitert. Zu Beginn des Big-Data-Zeitalters haben Unternehmen schnell den Wert für das Tagesgeschäft erkannt, einfacher auf große Datenmengen zugreifen zu können. Trino hat seinen Ursprung in der Lösung dieses Problems. Damals hatten jedoch nur Unternehmen von der Größe von Facebook oder Netflix die Ressourcen, um solche Lösungen zu entwickeln.

Heute sollte jedes Unternehmen von Daten profitieren. Lösungen wie Starburst Galaxy machen die Nutzung von Trino einfach und zugänglich. Unternehmen können eine Grundlage für ihre Datenarchitektur schaffen, unabhängig davon, wo sich die Daten befinden und wie sie strukturiert sind, und diese durch eine strenge Governance schützen. Das Engagement von Starburst rund um Apache Iceberg avancierte mittlerweile zur Icehouse-Datenarchitektur, die Trino und Iceberg kombiniert. Diese Architektur hilft Unternehmen dabei, auf neue Weise Nutzen für das Business aus ihren Daten zu ziehen.

KI erschließt den Wert von Unternehmensdaten wie nie zuvor

Daten stehen niemals still. Es gibt neue Wertschöpfungspotenziale, insbesondere im Bereich der KI. Unternehmen nutzen mittlerweile Enterprise-Intelligence-Plattformen, um die Grundlage für ihren Datenstack zu schaffen.

Diese Grundlage kommt häufig für Analyse-zwecke zum Einsatz, aber zunehmend auch für KI. Im Kern sind sowohl Analytik als auch KI nur zwei Seiten desselben Datenproblems. Beide Ansätze versuchen, Rohdaten in Erkenntnisse umzuwandeln, die reale Probleme aus dem Business lösen. Die Antwort von Big Data auf dieses Problem war die Schaffung groß angelegter „Systems of Record“. Dies Lösung der KI hingegen ist die Entstehung von Systemen, die

auf Intelligenz und Erkenntnis basieren, sogenannten „Systems of Intelligence“. Erforderlich ist hierbei eine zugrundeliegende Schicht für die KI-Datenarchitektur.

Für Unternehmen, die ihre gesamte Datenarchitektur KI-fähig machen, verändert KI die Wahrnehmung über den Wert von Daten fundamental. Diese Unternehmen arbeiten in der Cloud, vor Ort und zunehmend auch in einer Mischung aus beidem, indem sie das Hybrid-Lakehouse-Modell nutzen. Entscheidend ist jedoch, dass sie diese Daten unabhängig von ihrer Datenstruktur erschließen und mit einer einfachen All-in-One-Lösung sichern.

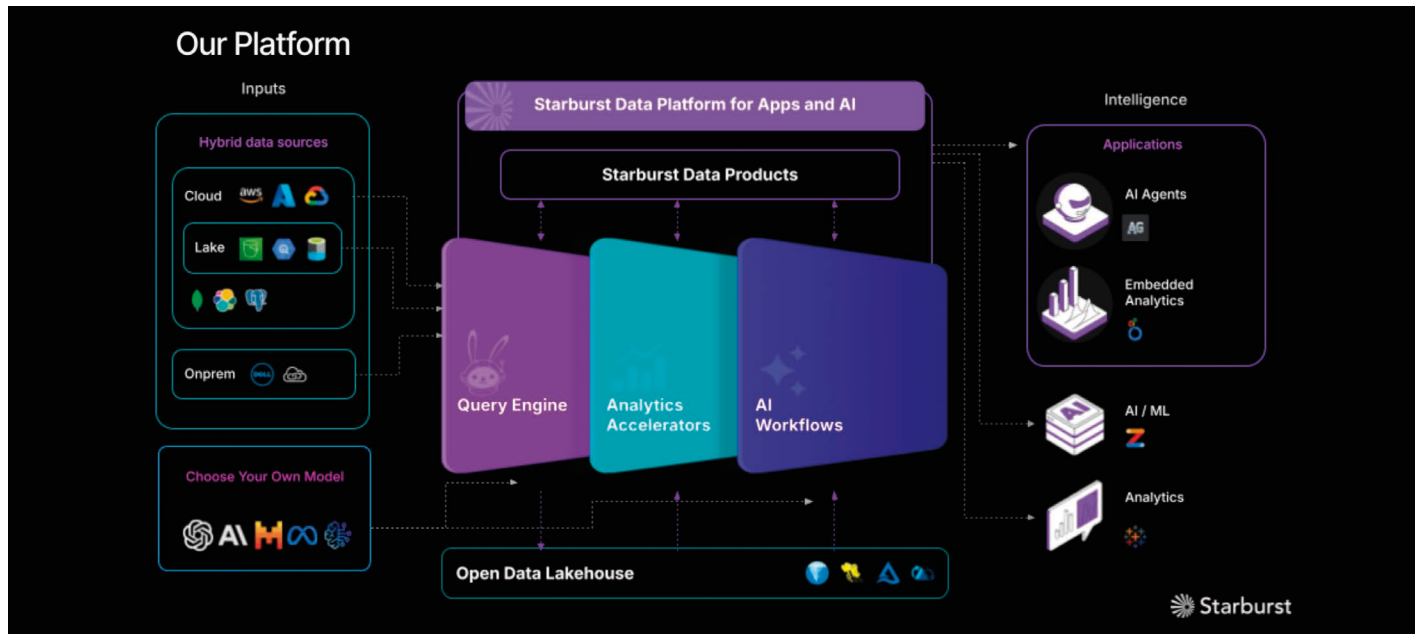
Daten und KI sind untrennbar miteinander verbunden

Dieser Trend wird sich noch verstärken, denn KI ist bereits untrennbar mit Daten verbunden. Das liegt daran, dass moderne KI auf Architekturen basiert, die auf Wahrscheinlichkeiten aufbauen und bei denen mehr Daten bessere Algorithmen bedeuten.

Damit KI Fortschritte machen kann, muss die Datenbasis, auf der sie aufbaut, breit gefächert sein – je größer, desto besser. Dies wiederum bedeutet, dass KI eine skalierbare Datenarchitektur benötigt, die auf Daten aus mehreren Quellen in verschiedenen Formaten zugreift und diese gleichzeitig sicher verwaltet. Eine solche Datenarchitektur trägt dazu bei, Engpässe in KI-Daten-



Autor:
Korbinian Zollner
Leiter Solutions Architecture, EMEA
Starburst Data, Inc.
www.starburst.io



Workflows zu beseitigen, genau wie es bei der Datenanalyse geholfen hat, diese zu beseitigen.

Gleichzeitig reicht Skalierbarkeit allein nicht aus. Einschränkungen bei universellen LLMs bedeuten, dass in vielen Fällen bessere Daten wichtiger sind als mehr Daten. Der Kontext ist entscheidend. Methoden wie Retrieval-Augmented Generation (RAG) helfen, diese Lücke zu schließen. Wichtig ist, dass dies jetzt geschieht – nicht morgen oder in ferner Zukunft. Im Wesentlichen ist eine kritische Masse erreicht. Dieser epochale Wandel hat Daten und KI mittlerweile untrennbar miteinander verbunden.

Über Analytik hinaus

Das Zeitalter der KI ist nicht nur eine Modeerscheinung, sondern der Weg, den Unternehmen aktiv beschreiten, um ihre Business-Probleme zu lösen. Den Wert von Daten zu erschließen, geht heute zunehmend über die Analytik hinaus – hin zur KI. Die Datenarchitektur erreicht damit eine neue Phase, in der dieselbe Data-Lakehouse-Architektur, die Analytik besser macht, auch generative KI- und Machine-Learning-Modelle speist. Um dieser neuen Landschaft gerecht zu werden, gilt es die wachsende Nachfrage nach Datenarchitekturen zu bedienen, die KI-Workloads unterstützen können.

Die Bemühungen im Enterprise-Intelligence-Markt bauen im Falle von Starburst auf der zentralen Icehouse-Architektur auf und umfassen eine Reihe von KI-Beschleunigern. Diese ermöglichen Unternehmen einen schnelleren Übergang von der Entwicklungsphase zum produktiven Deployment – vom KI-Proof-of-Concept (POC) zur vollwertigen produktiven Intelligenz. Hierzu zählen Workload-Optimierungen, Streaming-Ingestion, dateibasierte Datenerfassung, verbesserte Unterstützung für KI-Governance und Funktionen zur Unterstützung der RAG-Architektur.

Ohne Daten keine KI

Kein einziges KI-Modell funktioniert ohne Daten, mit denen es trainiert werden kann, und der kontinuierliche Datenfluss in die Modelle ermöglicht deren Wachstum. In diesem Sinne lassen sich Daten als Grundlage betrachten, nicht nur für Analytik, sondern auch für KI. Was Unternehmen darauf aufbauen, hängt von ihrer Fähigkeit ab, diese Grundlage auf sichere, zuverlässige und vorhersehbare Weise zu schaffen.

Starburst – hier als Beispiel – schafft bereits die Grundlage für die Datenanalyse und zunehmend auch für KI. In beiden Fällen ist die Rolle ähnlich, denn im Kern ist Starburst eine Engine. Diese Engine kann den Analytik-Datenstack antreiben, aber auch den KI- oder ML-Datenstack. Sie kann den Datenstack antreiben, der ein BI-Dashboard, eine Datenanwendung oder ein KI-Modell steuert. In diesem Sinne ist Starburst sowohl eine SQL-Abfrage-Engine für Analysen als auch eine KI-Abfrage-Engine.

Die gleichen Faktoren, die einen Analytik-Datenstack erfolgreich machen, machen auch einen KI-Datenstack erfolgreich. In beiden Fällen müssen die Daten zugänglich, organisiert und gut verwaltet sein. Angesichts des rasanten Wachstums von KI besteht in jedem Unternehmen ein enormer und wachsender Bedarf, das richtige Datenfundament aufzubauen, unabhängig davon, ob diese Unternehmen KI bereits heute einsetzen oder erst morgen einführen wollen.

Was Datenanalytik und KI gemeinsam haben

Tatsächlich haben Datenanalytik und KI viel gemeinsam. Beide basieren auf Daten und benötigen eine solide Datenbasis, um erfolgreich zu sein. Da Unternehmen ihre Analyseprojekte

weiter vorantreiben und gleichzeitig neue KI-Projekte einführen, ist der Bedarf an einer soliden Grundlage größer denn je. Beide Vorhaben leben von hochwertigen, kollaborativen und gut verwalteten Daten, sodass geeignete Datenplattformen eine grundlegende Voraussetzung für sichere und wirkungsvolle KI-gesteuerte Innovationen sind.

Wenn KI der Jet-Motor ist, der Innovationen antreibt, dann sind Daten der sorgfältig raffinierte Treibstoff, der den Motor antreibt. Ohne zugängliche, kollaborative und gut verwaltete Daten stottert der Motor und versagt. Die Datenarchitektur hatte schon immer einen enormen Einfluss auf den Erfolg von Datenanalyseprojekten, weshalb die richtige Umsetzung dieser Technologie in der Geschichte von Big Data den Unterschied zwischen Erfolg und Misserfolg ausgemacht hat.

Mit der Verlagerung hin zur KI rückt die Datenarchitektur erneut in den Fokus. Wie zuvor ist auch jetzt eine gute Grundlage für diese Datenarchitektur unerlässlich. Während die KI die Möglichkeiten für den geschäftlichen Nutzen auf neue und spannende Weise erweitert, bleibt die Erreichung dieser Ziele weiterhin im geschäftlichen Nutzen verwurzelt.

Fundament für Analytik und KI-Architektur

Heute befinden sich Unternehmen an einem entscheidenden Schnittpunkt zwischen Rohdaten als Ressource und Daten, die zur Wertschöpfung beitragen. Jedes Unternehmen steht vor dieser Herausforderung und möchte seine Daten in echte Erkenntnisse umwandeln. Hierfür ist ein Fundament nötig, auf dem die Analytik- und KI-Infrastruktur aufbaut. Entscheidend sind jetzt Lösungen, um die Voraussetzungen zu schaffen, damit ein Unternehmen dieses Fundament erstellen kann. ◀