



Aufbau eines Data Lakes auf Microsoft Azure

In enger Zusammenarbeit mit der IT-Abteilung der TX Logistik wurde die Systemlandschaft durch den Aufbau eines cloudbasierten Data Lakes zukunftssicher erweitert und erfolgreich „auf die Schiene gehoben“. Ziel des Data Lakes ist die schnelle, flexible und effiziente Bedienung der Businessbereiche mit qualitativ hochwertigen Daten für Analysen und Reports. Des Weiteren dient der Data Lake als Datenbasis für neue Use Cases im Advanced Analytics Bereich.

TX LOGISTIK AG

TX Logistik wurde 1999 als privates Eisenbahnverkehrsunternehmen gegründet und bietet mit Verbindungen in elf Ländern ein leistungsfähiges europäisches Netzwerk. Das Unternehmen verfügt über Tochtergesellschaften in der Schweiz, Österreich, Dänemark und Schweden sowie einer lokalen Präsenz in Italien. In den Geschäftsbereichen Intermodal und Rail Freight entwickelt TX Logistik umfassende Schienenlösungen für kontinentale und maritime Verkehre sowie maßgeschneiderte Konzepte für konventionelle Fracht. Mit rund 580 Mitarbeitern und 9,1 Mrd. gefahrenen Tonnenkilometer erwirtschaftete das Unternehmen 2018 einen Jahresumsatz von 250 Millionen Euro. Seit Januar 2017 gehört TX Logistik zu 100 Prozent der Mercitalia Logistics S.p.A., einer Tochtergesellschaft der Ferrovie dello Stato Italiane.



Kunde: TX Logistik AG, Troisdorf
Branche: Logistik
Technische Umgebung: Azure Cloud
Aufgabe: Aufbau eines cloudbasierten Data Lakes
Dauer: 3,5 Monate

„Die implementierte Data Lake Architektur ist eine sehr gute Basis, um die datenbasierte Steuerungsfähigkeit des Unternehmens zu erweitern.“

Mitarbeiter, TX Logistik

DIE AUSGANGSSITUATION

Die Vielzahl der Quellsysteme stellt die Fachbereiche bei TX Logistik zum einen vor die Herausforderung, geeignete Datenmodelle für die fachbereichsspezifischen Anforderungen aufzubauen. Zum anderen ist keine ausreichende Dokumentation vorhanden. Durch eine kontinuierliche Dokumentation von Datenbildungsregeln und eine Vereinheitlichung der Namenskonventionen sollen die Self-Service Möglichkeiten der Fachbereiche ausgebaut und einer fehlerhaften Datennutzung vorgebeugt werden.

DAS PROJEKT

Cologne Intelligence wurde im Rahmen des Projektes mit dem Aufbau eines cloudbasierten Data Lakes beauftragt, um die bisherige BI-Landschaft durch den Architekturchange für zukünftige Anforderungen vorzubereiten. Der Data Lake wird als zentrale Datenbasis positioniert, um unternehmensweite Use Cases mit strukturierten und unstrukturierten Daten zu versorgen.

Die auf Azure erstellte Infrastruktur wurde nach dem IaC-Prinzip (Infrastructure as Code) angelegt. Bei den verwendeten Diensten handelt es sich um Azure Data Lake Gen2, Data Factory v2, Databricks und Azure SQL Database. Bestehendes Reporting wurde auf PowerBI migriert und mit neuen Anforderungen erweitert. Ein wichtiger Erfolgsfaktor war die Schulung der TX Logistik Mitarbeiter, die zukünftig selbstständig Datenintegrationen durchführen können.

In enger Zusammenarbeit mit der IT-Abteilung der TX Logistik wurde die Systemlandschaft durch den Aufbau eines cloudbasierten Data Lakes zukunftssicher erweitert und erfolgreich „auf die Schiene gehoben“. Ziel des Data Lakes ist die schnelle, flexible und effiziente Bedienung der Businessbereiche mit qualitativ hochwertigen Daten für Analysen und Reports. Des Weiteren dient der Data Lake als Datenbasis für neue Use Cases im Advanced Analytics Bereich.

DER WEG ZUM ERFOLG

Gemeinsam mit der IT-Abteilung hat Cologne Intelligence in einem 3,5-monatigen Projekt einen cloudbasierten Data Lake aufgebaut. Die Umsetzung des Projektes erfolgte agil. Hierfür wurde das Projektteam in das SAFe Framework integriert und konnte die einzelnen Sprints detailliert planen, Abhängigkeiten im Vorfeld klären und am Ende jedes Sprints die Ergebnisse präsentieren. In einem der ersten Sprints wurde nach Aufbau der Infrastruktur die Einbindung derselben in die bestehende Netzwerk Infrastruktur der TX Logistik sichergestellt – die wichtigste Grundlage für den weiteren Projekterfolg.

Ein weiterer Erfolgsfaktor für das Projekt wird die Einführung eines Data Governance Boards, welches den Data Governance Prozess definiert und überwachen wird und somit eine einheitliche Dokumentation, eine kontinuierliche Verbesserung der Datenqualität und die Einhaltung von Sicherheitsanforderungen sicherstellt.

Sprechen Sie uns an

Cologne Intelligence (CI)
Johannes Schröter | Geschäftsführer
Phone: +49 221 99041 203
johannes.schroeter@cologne-intelligence.de
www.cologne-intelligence.de

