

Solving Modern Data Challenges with Data Hub



Structured and unstructured data volumes are exploding. [IDC predicts](#) that the “global Datasphere will grow from 33 zettabytes in 2018 to 175 ZB by 2025.” That is 175 trillion gigabytes. And much of it is unstructured data.

With extraordinary amounts of data come large problems. Unstructured data often needs to be combined with business knowledge and structured business data to make it meaningful. This can cause data silos, and the challenge is that you still need to move the data to utilize it. Compounding the concerns with moving data is that you still need to keep it inside compliance boundaries.

Microsoft [SQL Server 2019 Big Data Cluster](#) enables intelligence and cohesive computing over all your data. It helps remove data silos by combining structured and unstructured data across your entire data estate. This data hub integrates [Microsoft SQL Server](#) and big data open-source solutions. It’s deployed on scalable clusters using Apache Spark, HDFS containers with Kubernetes, and SQL Server. This data hub is an ideal hybrid transactional analytical processing (HTAP) system for online transactional processing (OLTP) and analytical processing, combined with AI predictions for business intelligence.

At the recent [Pure//Accelerate™](#) Digital event, Pure's Chris Adkin sat down with Microsoft's Buck Woody to explain how the [data hub](#) works at scale in Big Data Clusters. In their conversation, Buck provides examples of how the data hub is used to analyze hospital stay length in healthcare as well as fraud analysis in finance.