

Achieving High-Quality Data on Azure with Cloud Data Preparation from Trifacta

Background

For companies looking to advance their analytics in the cloud, Microsoft Azure offers a suite of cloud services to support modern BI/Reporting and AI/ML initiatives. Organizations can leverage Azure Data Lake, SQL data warehouse, or Azure Storage Blob to store large volumes of data in its native form—structured, semi-structured and unstructured format to enable broader analytics use cases through analytics services such as PowerBI, or AzureML.

However, to achieve analytics excellence, first, you need great data at the foundation. A cloud-native data preparation solution is a critical component of the cloud ecosystem to ensure the success of any analytics project in the cloud. [Trifacta](#) offers an [industry-leading](#), machine learning-powered cloud data preparation solution natively integrated with a rich set of Azure services to deliver clean, trusted, and well-prepared data to support all your analytics needs.

Challenges

While migrating to the cloud provides tremendous agility and cost savings to the end-user organizations, wrangling data on the cloud is no easy feat due to the varying sizes and shapes of data stored in the cloud environment such as Azure Data Lake or Storage Blob. The existing data management solutions fail to deliver clean and well-prepared data in Azure cloud due to the following limitations:

Rigid architectural design: Legacy tools such as ETL/ ELT tools were designed to process structured data with predefined schema, they are unable to refine and prepare raw data in a complex form in Azure Data Lake, or Storage Blob, therefore limiting the analytics use cases companies can explore.

Lack of self-service: The existing tools were primarily designed for IT/technical users as opposed to business users who understand data best, and the same users also rely on data to drive everyday decisions.

Poor integration with cloud services: Existing solutions lack native integration with many of the services offered by the cloud provider, leading to increased management complexity and cost, as well as delayed time to analytics insights.

“

Clean data is critical to any information-based organization, but the process of cleaning and preparing data for use is time-consuming and challenging. Trifacta, by leveraging Microsoft Azure big data and advanced analytics services, arms our shared customers with the ability to simplify those processes in order to more efficiently analyze the data and seek out meaningful insights.”

Charlotte Yarkoni
Corporate Vice President,
Azure Growth + Ecosystems

Supported Services



Azure Data Factory



Azure Data Lake



Azure Storage Blob



Azure SQL Data Warehouse



Azure Databricks



Azure HDInsights



Azure Active Directory



Azure ML



Power BI

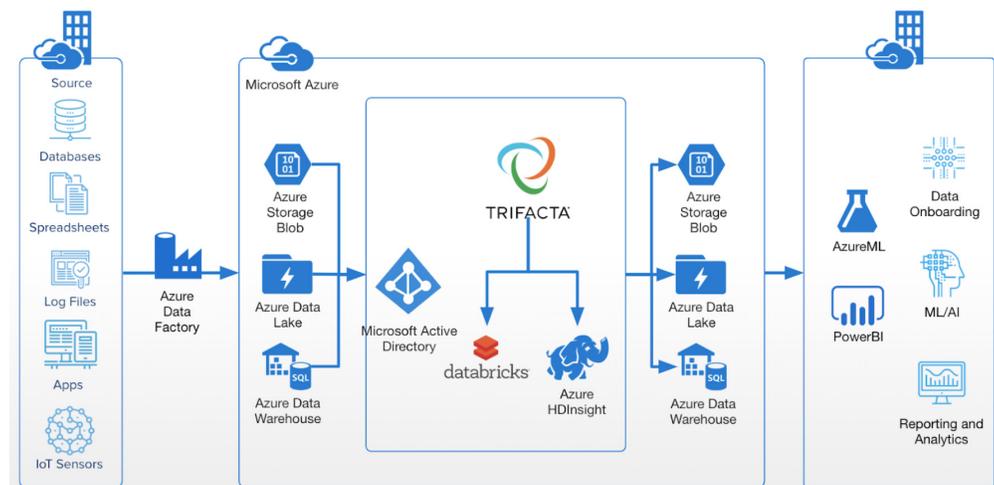
About Microsoft Azure

Microsoft Azure is an ever-expanding set of cloud services to help your organization meet your business challenges. It's the freedom to build, manage, and deploy applications on a massive, global network using your favorite tools and frameworks. Learn more at azure.microsoft.com

Why Trifacta for Azure

Trifacta Wrangler Enterprise is a serverless data preparation service tightly integrated with the Azure ecosystem. A leader in data preparation, Trifacta allows analysts, data scientists, data engineers, and business users to prepare data of any form and size and quickly and easily transform data from its raw format into a refined state for a broad range of analytics and machine learning initiatives. Trifacta Wrangler Enterprise is integrated with a rich set of Azure services including Azure Data Factory, Azure Data Lake and Storage Blob, Azure HDInsights and Azure Databricks to wrangle data at the scale offered through Microsoft Azure. Whether you are building a cloud data lake on Azure, migrating from your legacy data warehouse to Azure SQL data warehouse, or launching ML/AI project by leveraging AzureML, Trifacta accelerates your data preparation to enable faster time to insights and innovation with clean, connected, secure and timely data.

Reference Architecture



With Trifacta data preparation service on Azure, organizations gain the following advantages:

Seamless Integration with Azure Ecosystem

Native integration with Azure Data Lake, Storage Blob, and Azure SQL data warehouse, processing engines such as Azure HDInsight and Azure Databricks, analytics services such as PowerBI and AzureML, as well as Azure Active Directory security service for ease of management, greater agility, and enterprise-class security.

Accelerate Data Preparation on Azure

Automate data preparation process with a visual, interactive and AI-powered platform to ensure clean, connected and trusted data is immediately available on Azure to support data services, modern BI/Reporting, and AI/ML initiatives.

About Trifacta

Organizations that embrace data-driven decision making compete on differentiated data. Trifacta empowers data professionals of all levels of technical expertise to connect and wrangle data into its final state for reporting, analytics and machine learning, all in a tightly governed, cloud native environment. Trifacta blends visual guidance and machine learning to create an intuitive user experience built to accelerate time to value and automate repeatable workflows. Learn more at trifacta.com.

Centralized Data Governance and Access Control

Centralize data governance, security, lineage and access control to a single platform instead of disparate spreadsheets or desktops that are impossible to manage, reducing operational burden and cost.

Business Self-service, Intelligent Data Preparation

Empower business users who know the data best with simple, interactive, visual, and machine learning-powered platform to accelerate data preparation and increase productivity and time to insights.

Enable Quality Data Services with Azure Data Lake

Trifacta quickly transforms messy data from internal and external data sources into clean and well-prepared data in Azure Data Lake. With great data at the foundation, your analysts and data scientists can jump-start their BI reporting and advanced analytics, driving faster time to insights.

Accelerate Data Prep for BI Reporting with Azure SQL

Migrating your data from on-premise systems to Azure SQL data warehouse can be a daunting task. Trifacta expedites data preparation in Azure SQL, ensuring clean, well-prepared data is readily available in your Azure SQL data warehouse to support your BI reporting and other analytics projects.

Automate Data Prep for AzureML Service

High-quality data is critical to the success of AI/ML. For data scientists and developers working on AI/ML projects using AzureML service, Trifacta automates data preparation to reduce the time data scientists spend on data wrangling and allow more time on model building and testing.



Start Wrangling

trifacta.com/start-wrangling

[Get Trifacta on the Azure Marketplace >](#)



Trifacta is the industry pioneer and established leader of the global market for data preparation technology. The company draws on decades of academic research in machine learning and data visualisation to make the process of preparing data faster and more intuitive. More than 100,000 data wranglers in 10,000 companies worldwide use Trifacta solutions across cloud, hybrid and on-premises environments to support a variety of analytic and operational use cases. Leading organizations such as Deutsche Boerse, Google, Kaiser Permanente, New York Life and PepsiCo count on Trifacta to accelerate time-to-insight and discover opportunities that drive success. Learn more at trifacta.com.

For Additional Questions, Contact Trifacta
www.trifacta.com | team@trifacta.com

Experience the Power of Data Wrangling Today
www.trifacta.com/start-wrangling