NAVIGATE YOUR NEXT



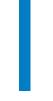
INFOSYS CLOUD DATA VALIDATION SOLUTION @ AZURE - Native Data Validation **Automation Solution**

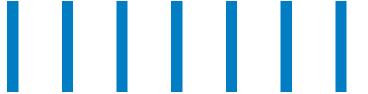














QA challenges in Cloud Data Modernization

01



Huge velocity and varied variety of data

02



Validation of legacy systems data while moving to cloud

03



Lack of data testing solutions to support new technology -Server less, Data Pipelines, etc. 04



Ensuring data integrity during movement across numerous technical layers

05



Ensuring data
accuracy
between
visualization
layer and
backend system

Infosys Cloud Data Validation Solution

A best in class cloud native data validation automation solution that helps streamline and accelerate testing of Data Integration / Data Analytics platforms. This is a UI-based automated data quality and data integrity approach for accelerated data validation.

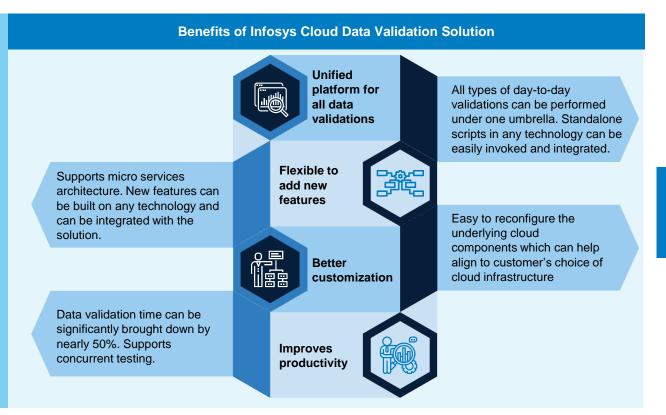


Infosys Cloud Data Validation Solution

Infosys solution helps in accelerated validation of Data Integration / Data Analytics platforms by bringing open source technologies and Azure Services together. It provides comprehensive technical approach for data validation between source and target databases / files like data quality checks, data reconciliation checks and data validation checks.

The solution is capable of validating data for use cases like:

- Modernization of Legacy Data Warehouse to Cloud Data Warehouse
- Application Modernization Data modernization to Cloud Storage
- New Data Lake hosted on Cloud
- Migration of on-premise Data Lake to Cloud and BI Reporting
- Migration of dashboards to Cloud based platform





Key solution components

images)

details

Infosys Cloud Data Validation Solution is a native solution catering to the industry needs of data validations. Its core components and features for Azure are depicted below:

Native Components Process Hadoop Discover, specific formats for understand the **Azure Kubernetes Service** early validation nature and quality (For container orchestration) Data compression of the data sources Data sanity and decompression **Databricks** (Spark Cluster to process Big Data) **Relational Data Store** Simulates Standardize (Application DB Server) positive and Test data generation **MODULES** Data conversion the format for negative test reconciliation **Azure Virtual Network** Data at attribute (Secure interaction between Azure level resources) Integration with JIRA **Azure Storage Account** Data comparison Reconcile the and ALM (Binary Large Object Storage) Establish two-way source and target data for straight and connection to **Azure Container Registry** derived moves. update defect (Secure Container to store Docker



Solution features



Columnar and Row-based results

 Results are available with a columnar view, row view and a summarized view.
 User can choose the mode of the view in the user specified path.



Batch mode operations

Comparison, conversion and compression can happen in a batch mode in any desired combination.

Status can be monitored at runtime.



Pattern retrieval

 Distinct patterns of a column in target can be extracted and reconciled against its counterpart in source.



Test data simulation

 Solution can simulate test data needed for validations, considering both negative and positive business rules



Eliminates the need for SQL syntactical knowledge

 User has provision to generate SQL query automatically



Nested file validations

 Solution can process complex files such as Nested JSON, Nested XML etc. The namespaces in the XML file would be automatically triggered.



THANK YOU



