

**As the volume and velocity of data continue to grow, streaming data represents a paradigm shift that introduces new data quality challenges.**

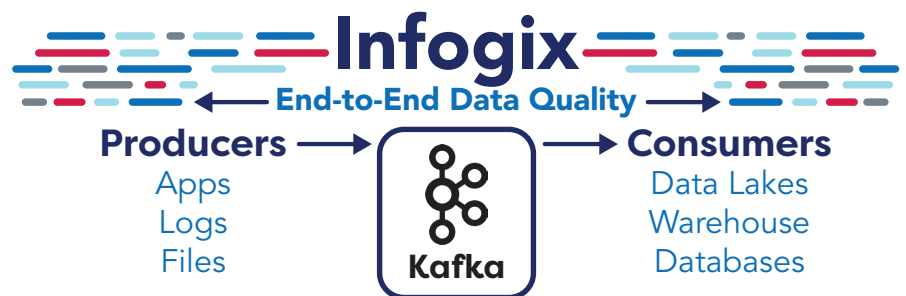
## The Challenge

The speed of business today demands that organizations enable access to real-time data. This has given rise to event-driven architectures. This fundamental shift from batch processing to streaming data provides new opportunities for increased data delivery reliability, more nimble reactivity and faster business insights. But it also presents new risks to data integrity as streaming sources, data volumes and architectural complexity continue to grow, affecting data reconciliation.

Among the many streaming data options, distributed streaming platform Apache Kafka has become the preferred software for real-time data communication. However, while Kafka is an agile, high-throughput, low latency option for managing data in motion, it cannot ensure the reliability and accuracy of real-time data streams which present many data quality challenges.

## The Solution

As reliance on streaming data continues to grow, so do demands from data consumers who expect real-time intelligence that's complete, consistent and trustworthy. That's why organizations need to ensure end-to-end data quality in streaming environments, as messages are sent from the sources that create them (producers) to the sources that consume them (consumers).



Businesses need a solution that provides a wide range of data quality capabilities, from in-line checks and validations to batch processing for robust reconciliations. Data360 DQ+ gives organizations confidence that the quality of streaming data won't degrade in motion and can immediately be acted upon in order to build trust throughout the business.

## Data360 for Kafka Delivers:

### Trust in Streaming Data with Producer-to-Consumer Validation

Reduce enterprise risk by ensuring that streaming data is validated, reconciled, and timely, to produce meaningful and reliable data insights.

### Accelerated Time-to-Value

Automatically connect and configure Kafka stream schema with zero-code setup based on integration with your platform schema registry. Easy setup combined with drag-and-drop rule development shortens time to value.

## Clarity and Insight across Your Data Pipeline

Validate data for streaming and non-streaming data producers to allow for validation across the entire data pipeline, from source system raw data, to Kafka message, to consumer (database, file, application, etc.).

## Single Solution Reduces Support and Costs

Provide an integrated toolset in a single solution, reducing the costs of multiple tools and the number of supported applications. Experience a breadth of functionality that helps consolidate required resources, skillsets and training with a single solution for IT, business users and audit.

## Key Solution Features

### Support for Real-Time or Batch Validation on Streaming Data

Applies user-configurable logic to real-time or batch streaming data to support inline or time-windowed validation, powered by the latest Spark Structured Streaming and SparkSQL. Easily build rules that match your latency requirements to support a variety of enterprise use cases.

### Ingest, Enrich, Transform and Output Multiple Data Types

A wide range of data type options can be captured as input and written as output, allowing users to input as one data type and write out as another. This flexibility allows traditional data sources like databases and files to be enriched or transformed and then output as Kafka messages, or vice versa. It also allows the blending of existing architectures without waiting for all data to fully transition to streaming.

### Apply Machine Learning to Identify Streaming Data Patterns or Anomalies

Machine learning models that identify patterns, highlight outliers and apply analytics across multiple facets of streaming data.

### Manage Exceptions and Visualize Results

Identify duplicate or failed messages and route them to a workflow to be assigned, managed and resolved. Visualize the results of your streaming data validations to gain deeper business insights while sharing and collaborating with teams through dynamic dashboards.

Develop, implement and optimize your data management and governance strategy with Infogix. Visit [www.infogix.com](http://www.infogix.com) or call +1.630.505.1800 (US, Canada and International), +44 1242 674 137 (UK and Europe).