



# Bureau

Bureau is a trust network that facilitates end-to-end identity verification, compliance, and fraud prevention for new-age businesses.

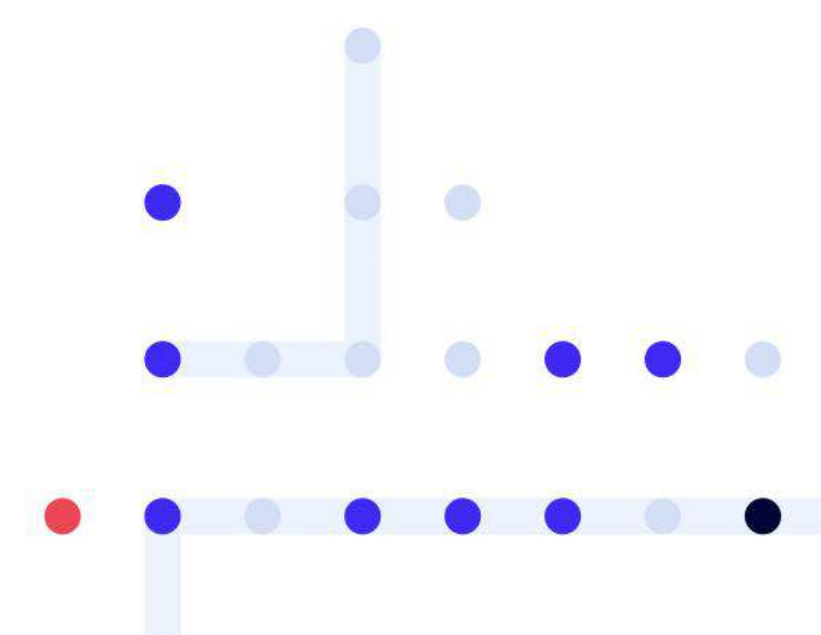
Report Type  
Case Study

Domain  
Security & Risk management

Headquarters  
 United States

*"KeyValue's passion, commitment, and technical expertise are truly remarkable and their ability for swift deliveries without compromising on quality has been vital for our progress."*

- Ranjan R Reddy Founder, CEO at Bureau Inc.



# Problem Statement

To build Bureau into a company with data science at its core – to understand risk, detect fraud & anomalies, and to make data-driven decisions in a few seconds. This requires extending the existing identity decision platform with a centralized data science stack that consolidates data from customer events, third-party integrations, and user behaviour.

Data Insights Stack Development

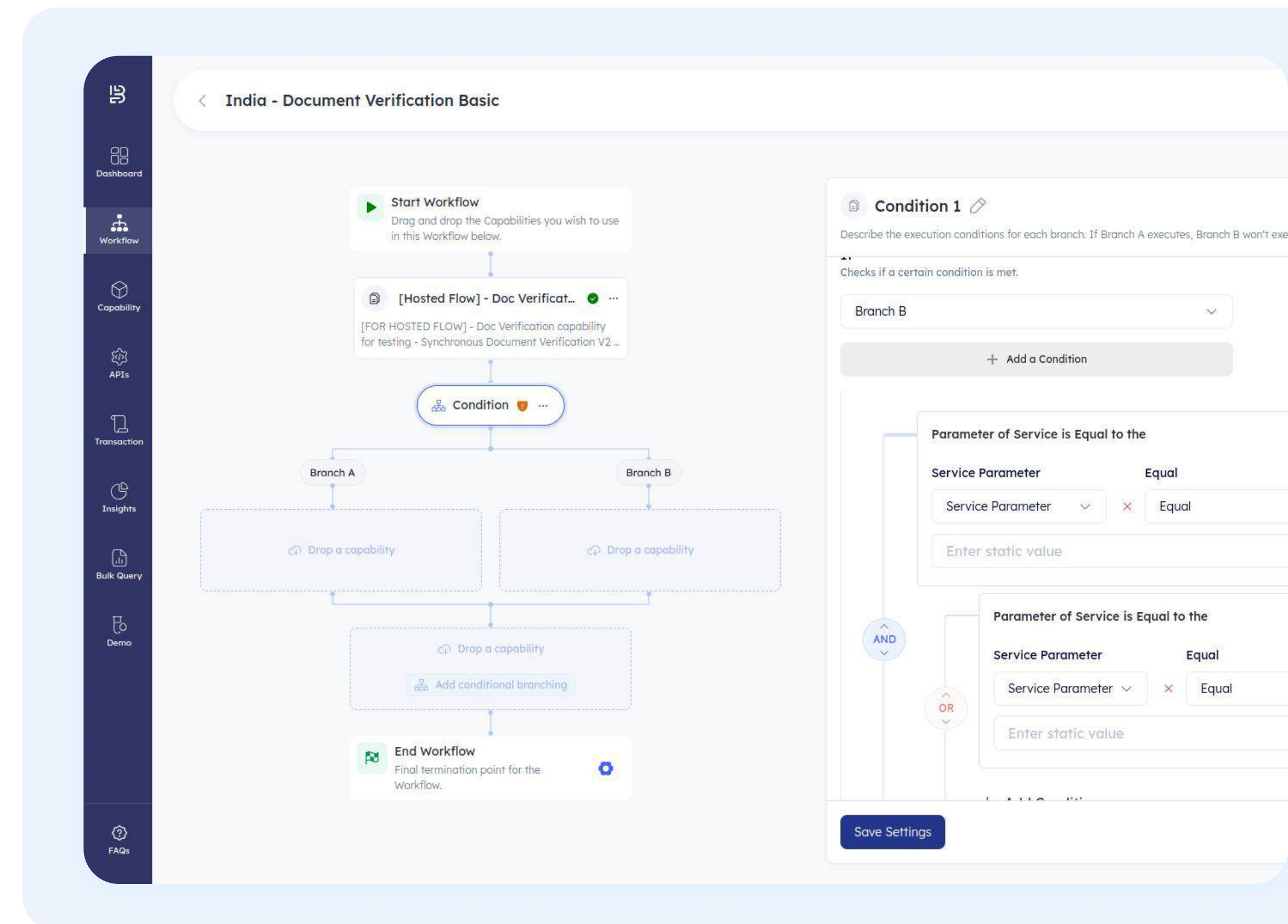
Data Centralization

Billing and Insights Platform

Customer Metrics Dashboard

Data Preparation for Machine Learning

Anomaly Identification



# Key Metrics

**7+**  
Data sources

**100GB**  
Data processed  
per day



**3M+**

Transactions processed daily

**10+**  
Data pipelines



**70+** Charts





# Business Impact

## Reduced Data Processing Time

Achieved a 40% decrease in data processing times, enabling faster decision-making across the organization.



## Real-Time Insights

Provided real-time insights through efficient data pipelines, enhancing the ability to respond promptly to market changes and customer needs



## Operational Cost Savings

Realized a 25% reduction in operational costs by streamlining data workflows and reducing manual interventions.



## Improved Fraud Detection

Strengthened fraud monitoring capabilities by integrating diverse data sources and enabling advanced analytics.



## Enhanced Decision Making

Empowered teams with comprehensive dashboards and accurate data, leading to more informed business strategies.



# Challenges





# Solution



## Data Unification

Centralized data from diverse sources into a unified data warehouse, eliminating silos and enabling seamless access across teams.



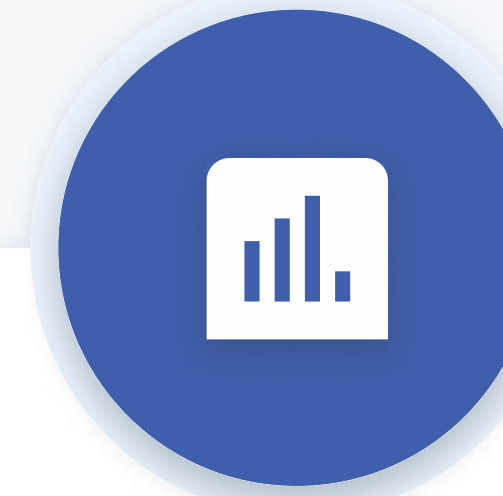
## Scalable Infrastructure

Developed a flexible, scalable platform using Airflow, Amazon Redshift, and AWS services for batch and streaming data processing.



## Integrated Dashboards

Created real-time insights dashboards with Metabase and Preset to monitor customer metrics, billing, and system anomalies.



## Enhanced Analytics

Prepared high-quality datasets for machine learning models, enabling advanced analytics and improving fraud detection.



# Implementation

## Flexible Schema Modeling

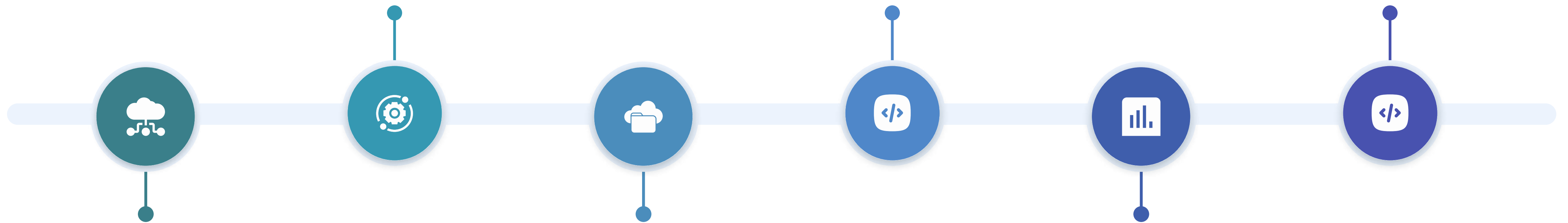
Designed adaptable schemas after analyzing the existing architecture to accommodate future changes and support various team needs.

## Third-Party Integration

Used Airbyte for integrating services like HubSpot and Google Sheets; implemented custom Python solutions when necessary.

## Anomaly Detection Mechanisms

Implemented monitoring and alerting systems to promptly identify and address anomalies across the platform.



## ELT Pipeline Development

Established a new data insights stack with an ELT pipeline from the ground up.

## Orchestration with Airflow

Employed Airflow for orchestrating extraction and loading processes in the ELT pipeline.

## Data Transformation with DBT

Utilized DBT for SQL-based transformations to prepare data for reporting and machine learning.



# Technology Stack

