

CASE STUDY | RETAIL PHARMACY

Agile Data Engineering for Data Warehouse Migration

Company Background

A Fortune 100 consumer retail pharmacy chain with stores located across the United States

The Challenge

Like all pharmacies in the United States, this Fortune 100 company processes insurance claims that pay for the prescription drugs prescribed by physicians. The claims data resides in an adjudication system where information goes back and forth between their systems and the insurers, which ultimately cover most of the cost of the prescription medications sold in the United States.

At a technical level, a subset of this data was moved from the claims system into a Teradata warehouse where the analytics were processed using BTEQ (Teradata's proprietary scripting language) and SQL to generate claims analytics reports.

The company set a business goal to reduce the need for costly confirmation calls to doctors' offices for pre-authorized prescription claims. Automating this process would require more advanced analytics than they had been previously performing and also meant more data would have to be loaded into an already overloaded Teradata warehouse. They simultaneously required an improvement in response time from 8 hours down to only 15 minutes. All of these changes necessitated a move to a big data platform like Hadoop. Unfortunately, the organization lacked sufficient in-house big data expertise to complete the project.

The Solution

Infoworks' agile data engineering (ADE) software platform simplified the project complexity and addressed these challenges by automating the company's data engineering and data operations processes. Hadoop would become the pre-processor of data, where more complicated machine learning analytics could be performed along with standard analytical queries, while Teradata would be focused purely on standard analytical reporting.

Challenges

- Teradata was too rigid, slow and expensive to process new machine learning algorithms
- Lack of skilled big-data engineers

Results

- 26 days from initial project start to completion versus a 9 month estimate from a systems integrator
- 2 data engineers completed the project from start to finish
- 300 tables, 70 SQLs and 5 BTEQ control flows were migrated in 26 days

Benefits

- ✓ Ability to adjudicate claims processing every 15 minutes eliminated the need for expensive validation call to doctor's office
- ✓ Existing employees now have the skills to implement big data-based analytics projects
- ✓ Poised to add more analytics use cases very quickly

The Solution (Continued)

The Infoworks ADE was then used to automate the ingestion of 300 tables from various data sources into Hadoop, and ensure they remained synchronized. Change data capture was implemented with a single mouse click. The Infoworks ADE was also used to automate the migration of 70 SQL queries from Teradata to Hadoop. Automating the SQL code migration ensured that the code was properly moved with no worries about human error.

Finally, all of the BTEQ control flows were translated to run in the Infoworks ADE Orchestrator, which also provides visualization of the workflows. Additionally, the Infoworks ADE delivers the ability to start, pause, stop and restart workflows from a visual dashboard while automatically eliminating data flow collisions. These are all capabilities that did not exist in the legacy Teradata system.

The entire project took 26 days from start to finish and required only two data engineers to complete. Performing this same migration using traditional system integration efforts was expected to take 9 months, a 9x improvement!

“ With Infoworks we were able to complete our project plan for the entire year, in just a few days!

Lead Enterprise Architect

The Results

This Fortune 100 retail pharmacy now has the ability to perform much more sophisticated analysis on their claims, meeting their service level agreement by running 32x faster, all at a lower cost. In the past, they were adding new analytical uses cases at the rate of one a year. By automating their enterprise data operations management environment on Infoworks, they are now adding a new use case every month and are using analytics to leverage data more effectively and as a competitive differentiator.

Infoworks' agile data engineering platform accelerates the pace of delivering business intelligence and machine learning analytics projects running on any big data fabric. Infoworks' code-free development environment allows organizations to develop and manage end-to-end data workflows without requiring an army of big data experts. Infoworks delivers capabilities to automate and simplify development of data ingestion, data preparation, query acceleration and ongoing operationalization of production data pipelines. Infoworks supports cloud, multi-cloud, and on premise-based deployments, enabling customers to deploy projects to production within days, dramatically increasing business agility and accelerating time to value.

Learn more at www.infoworks.io