



# Infoworks

## THE AUTONOMOUS DATA ENGINE

Case Study:  
Retail Pharmacy

Automated Data Warehouse Migration and Offload

## Background

PharmaCorp ( company name with-held) is a Fortune 500 consumer retail pharmacy chain with stores located across the United States.

## The Situation

The current data warehouse environment used to process claims data is updated three times a day

Like all pharmacies in the United States, PharmacyCorp processes insurance claims that pay for the prescription drugs prescribed by physicians. In fact, they process millions of claims, if not tens of millions of claims, a year. Like other large pharmacies, they have claims data that 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 addition, a subset of this data is moved from the claims system into a Teradata warehouse where it is processed using BTEQ (Teradata's proprietary scripting language) and SQL to generate claims analytics reports. While claims are processed continually, the reporting data is updated only 3 times a day.

# 9x Improvement

## Offload a Legacy Data Warehouse to Hadoop in only 26 days

- Migrated 300 tables of data and metadata
  - In-house development: 2 mths
  - Infoworks: 7 days
- Migrated 70 SQLs
  - In-house development: 5 mths
  - Infoworks: 12 days
- Migrated 5 BTEQ control flows
  - In-house development: 2 mths
  - Infoworks: 7 days

## Challenge

The current environment cannot be cost effectively used for new machine learning based analytics

Claims data is the life blood of any pharmacy and is used by a wide variety of groups within the organization. The challenge for PharmaCorp was that as more users needed to access the claims analytics data, the query load would increase and the system would slow down. The company also had a desire to also perform more advanced analytics, combining claims data with other data, like CDC data, sales data, weather data, etc, and then use machine learning for near real-time analysis of consumer behavior. This would require yet more data to be loaded into an already overloaded Teradata warehouse while simultaneously requiring an improvement in response time from 8 hours down to only 15 minutes.

One solution would have been to increase the Teradata footprint. However, the practical challenge was that Teradata was already an expensive system and meeting the new SLAs, especially in a situation where the Teradata system would need to run machine learning algorithms, would be prohibitively expensive. In fact, PharmaCorp had already decided not to increase their investment in Teradata, knowing that it would organically grow even if they capped the addition of any new use cases. The only way to address this issue without adding capacity would be to free up capacity by moving data to Hadoop, which had a cost footprint 100x smaller than Teradata. The challenge with Hadoop however, was that PharmaCorp, like most similar organizations, lacked sufficient big data expertise in-house.

## The Solution

Automate the entire development and production process to move to a hadoop based architecture

Infoworks solved Pharmacorps challenges by working with them to automate the majority of the overall data flow process with the Infoworks Automated Data Engine. Infoworks first used automated crawling of the Teradata logs to identify workloads that could be moved to Hadoop. Workloads that required high levels of data integration, transformation and data management were targeted to be moved. Workloads that were more about performing queries and analytics would remain on Teradata. Hadoop would become both 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.

The Infoworks ADE was then used to automate the ingestion of 300 tables into Hadoop from various data sources and ensure they remained synchronized. Change data capture was implemented automatically with a single mouse click. The ADE then automatically migrated 5 BTEQ scripts and 70 SQL queries from Teradata to Hadoop. These workflows and SQLs had previously been blessed by legal to comply with all partner and government agreements. Automating the migration ensured that the code was properly migrated with no worries about human error.

In addition, because the new Infoworks platform provides a visual development environment, it was much easier to verify the workflow and transformation steps were correct. In the past, inspecting BTEQ and SQL code was always a difficult QA challenge requiring multiple reviews of hand written scripts.

Finally, all of the BTEQ control flows were automatically translated to run in the Infoworks ADE Orchestrator, which also provides visualization of the workflows. Additional, the Infoworks ADE provided the capability 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. Performing this same migration using traditional system integration efforts was expected to take 9 months. A 9x improvement!

## The Result

A 32x performance improvement

PharmaCorp now has the ability perform much more sophisticated analysis on their claims data while running 32x faster while also meeting their SLAs, all at a lower cost. In the past, PharmaCorp was adding new analytical uses cases at the rate of one a year. By automating their production data environment on Infoworks, they are now adding a new use case every month and are using analytics to make data a valuable asset and competitive differentiator.

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

Lead Enterprise Architect

## What Our Customers are Saying:

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

- Lead Enterprise Architect, Fortune 10 Retailer

“Infoworks reduces our time to introduce new end-to-end analytics models from 6 months to a week, without IT involvement. This allows our analytics teams to quickly meet fast-changing business requirements and directly enables growth”

- Director of Analytics, Leading CPG company

“With Infoworks we can quickly execute on our large backlog of data projects. If we were to custom build our ideal data platform it would be like Infoworks”

- Data Architect at a Fortune 100 technology company

“Prior to Infoworks we had long turn-arounds for data requests to IT. Now we have self-service reporting end-to-end”

- Business Analyst



# Infoworks



(650) 391-9306



sales@infoworks.io



www.infoworks.io