



# Infoworks

## THE AUTONOMOUS DATA ENGINE

Case Study:  
High Tech Manufacturer

Automated Data Lake Creation and Management

## Background

A-Corporation ( company name with-held) is a Fortune 100 multinational technology company headquartered in the United States that develops, manufactures and sells high technology product and services.

# 13x Improvement

**Infoworks: 10 Weeks**

**In-house development: 130 weeks**

- Deploy Ingestion Framework
  - In-house development: 18 mths
  - Infoworks: 1 mth
- Implement change data capture and high speed replication
  - In-house development: 18 mths
  - Infoworks: 1 mth
- Deliver production ready solution with resource optimization
  - In-house development: 6 mths
  - Infoworks: 2 weeks

## The Situation

Early adopter of hadoop technology heavily invests to make it work in a production environment

Several years ago, A-Corporation saw the potential to use big data analytics for competitive advantage. However, because they were an early adopter, they began using the technology at a time when the underlying Hadoop software infrastructure was still too immature for production deployment.

Even seemingly simple problems turned out to be a challenge. Because of the high volumes of data they were dealing with, loading the data lake required a significant amount of work. Hadoop data ingestion infrastructure at that time wasn't able to handle the petabytes of data they needed to reliably load and there were no 3rd party commercial solutions available. The solution had to manage over 1500 data sources on a daily basis including one schema that contained over 1500 tables.

As a result, they recruited a Silicon Valley team of Hadoop experts who could augment the existing commercial platforms to create a data lake environment that could be made production-ready. A-Corporation allocated 5 software engineers and spent over 3 years to develop and successfully deploy their solution. However, once the solution was in place, because data sources were constantly changing they continued to need the 5 engineers to maintain the environment.

## The Challenge

Find a commercial replacement for the in-house data ingestion software

After several years had passed, A-Corporation had expected the commercial market to have caught up with their needs, so they set out to find a software vendor that could replace, or at least augment their in-house code. The ultimate goal was to reduce the number of engineers required to support the existing data ingest infrastructure. The idea was to focus the same engineering effort on new projects that would further simplify their Hadoop environment, so less sophisticated users could access the data lake for advanced analytics and decision making.

They evaluated 11 big data ingestion solutions over 8 months, considering a variety of vendors, ranging from small venture capital backed startups to well-known public data integration vendors. None of the 11 were able to meeting the needs for scale and for delivering a production ready solution.

## The Solution

Automation of the big data development and production environments

The team was about to give up and continue their in-house development, when Infoworks introduced the Infoworks Autonomous Data Engine into the market and was evaluated as vendor number 12. Over a 4 week POC, Infoworks proved out the ability to not only augment the existing code, but completely replace the in-house ingestion and synchronization software.

Infoworks also demonstrated the ability to handle changing source data, automatically adjusting to new rows & columns that were constantly being added to the source schemas, without requiring human intervention. In addition, Infoworks provided an operational dashboard to make it possible to visually monitor and manage all 1500+ data pipelines, identifying bottlenecks and automatically adapting the data ingestion process.

The ultimate solution required only 2.5 months to fully deploy into production

## The Result

13x Improvement

What had previously required 5 highly skilled big data engineers to maintain, now requires only 1 person to manage and update. The high level of automation that Infoworks delivers allowed A-Corporation to reallocate 4 of the 5 engineers onto a "data shopping cart" that uses the Infoworks APIs to execute data ingestion jobs. The shopping cart allows A-Corporation to expose the data in their data lake to a broader set of users, expanding the value of their big data environment for their overall business.

" 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, Fortune 100  
Technology Company**



(408) 899-4687



sales@infoworks.io



www.infoworks.io

## 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