

CASE STUDY | MEDIA CONGLOMERATE

Agile Data Engineering for 360° Business Reporting

Company Background

A 100-year-old media conglomerate with cable TV, TV stations, magazines, newspapers and digital media needed to consolidate reporting for 60 business units

The Challenge

A major media conglomerate had grown through acquisition over the past one hundred-plus years. The CTO was interested in answering what he thought was a simple question, "What is the IT spend across the 60 divisions I am responsible for?"

They had tried multiple times to create an automated process to establish a consolidated 360-degree view of their IT spend but had repeatedly failed. Using a traditional enterprise data warehouse with ETL technology was estimated to take two data engineers at least 12 months to complete and required significant up-front planning to define a common schema. Just getting initial cooperation from each division was an exercise in herding cats. In addition, as part of the data consolidation, they also wanted to retire legacy mainframe systems as they moved to a more modern cloud-based platform.

The CTO had read about the potential of using a more modern approach based on the cloud and newer technologies like Hadoop and Spark. Unfortunately they lacked the skills in-house and also found it difficult to recruit the kind of data engineering talent required to be successful.

The Solution

With the emergence of big data solutions in the cloud, it was clear that moving to an Azure cloud-based solution would significantly simplify getting a big data infrastructure up and running. However, it still didn't address how they would develop the data pipelines that would be used to ingest and transform the data, prepare it for high speed querying, and then ultimately monitor and manage the pipelines on an ongoing basis.

Challenges

- Traditional Enterprise Data Warehouse approaches were too rigid and slow
- Lack of skilled big-data engineers

Results

- 4 weeks from initial project start to completion
- 1 data engineer completed the project from start to finish
- 35,000 tables ingested and processed

Benefits

- ✓ Visibility into technical spend analysis to support IT investment decisions completed
- ✓ Existing employees now have the skills to implement big data-based analytics projects
- ✓ Poised to add more analytics use cases very quickly

The client worked with the Infoworks software to build and deploy an end-to-end solution that automated their big data business intelligence analytics workflows. The software was installed in a matter of hours and after 4 weeks of implementation effort, the tech-spend analytics dashboard was completed. Because of the flexibility and agility provided by the combination of Infoworks and Azure HDInsight, they were able to add new data sources in an agile approach, and continuously enhance the data models as they went along. The up-front planning required in more traditional EDW implementations was eliminated. The complete solution included ingesting data with change data capture from 60 divisions and 12 different data sources including data from Oracle, DB2, Essbase, Salesforce.com, and flat files that were located both on premise and in the cloud.

The full configuration included building transformation pipelines to integrate the data and then generate in-memory models and cubes to ensure high-performance end user query performance. The entire environment also automatically tracks and captures changing data from the sources and provides a complete operational dashboard that monitors the status of the production environment, automatically orchestrating the production data flows and ensuring system processes uptime by starting, stopping and restarting system process... all automatically.

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

“ With Infoworks, we are able to implement data analytics projects without having to hire big data experts. These are people with skills that are hard for a 100-year-old media company to attract. Eliminating that need allows us to leverage our data just like our younger competitors.

The Results

The media company achieved a 24x improvement in time to deployment relative to their system integration estimates. More importantly, they have the ability to implement future projects without requiring an army of big data experts. The level of automation provided by the Infoworks software makes it possible for their in-house talent to quickly develop, deploy and manage new analytics use cases as their business and data analytics needs continue to evolve.