INFOCEPTS

# Data Hub for a Leading Luxury Travel Retailer to Enable Data Discovery and Provide a Platform for Innovation



## Summary

Our client is a leading luxury travel retailer that owns a network of duty-free stores with prominent presence across several international airports and premium high-street stores. They wanted a Data Hub that would help democratize their data. We developed a Hadoop based Data Hub that effectively supports a variety of analytics applications.


## Industry

Luxury Travel Retail

## Users

C-Suite Executives, Marketing Managers, Sales Managers, Regional Managers, Category Managers

## Technologies

Hortonworks Data Platform, Apache Kafka, Apache Nifi

## Team Size

7 InfoCeptians, 3 Client Associates

## The Challenge

Our client is a leading luxury travel retailer that owns a network of duty-free stores across the globe. Owing to the rising scale of their business operations, they wanted to democratize data by enabling Self-service Business Intelligence (BI). As a result, they sought our expertise to improve their information management system's capacity and efficiency to support essential analytics applications.

They wanted a Data Hub to consolidate enterprise data from multiple sources into a common platform to ensure –

- Availability of data to fulfill complex business use cases which require real-time consolidated data sets across systems
- Greater access to data via a scalable Data Hub architecture
- Lower point to point data transfer between multiple applications

Our client wanted the Data Hub to serve as a central repository to facilitate operational analytics for improving customer engagement.

## The Solution

We studied the existing system landscape to assess their operational and real-time data needs. Our assessment identified a Hadoop based system to be the most effective option for addressing our client's use case.

Hence, we developed a Data Hub comprised of a Hadoop cluster that sources disparate (structured and unstructured) data from various sources. The Data Hub stores, processes and serves data through Hortonworks Data Platform, which is an Apache Hadoop Distribution. Effectively, it is a distributed, scalable, highly available open source software built on commodity hardware.

We used Apache Kafka to ingest data from various sources in multiple formats and varied velocities. The Data Hub uses Apache Nifi to ingest data from sources like CSV flat files and FTP.

Also, we developed an Operational Data Store (ODS) with the objective to bring data closer to consuming applications. The ability of the ODS to augment the Data Hub played a critical role in fulfilling our client's need for democratizing data.

## The Results

We were successful in establishing a Data Hub that can support a variety of BI applications. The Data Hub has been able to deliver the following benefits –

- *Quick response time* – The time it takes to bring data over to enterprise data warehouse has reduced
- *Rapid change implementation* – The Data Hub has eliminated the need to depend on multiple applications where changes need to occur before carrying out any analysis. Because of this, the development cycle time required to change data attributes within any application has reduced
- *Faster data availability* – Any new data added at source becomes available in the Data Hub with minimal or no code change. Consequently, the new data can be readily consumed by applications like Salesforce that are necessary for improving customer engagement
- *Platform for innovation* – It serves as a repository of clean and harmonized data that can be used by business analysts and data scientists for data discovery and to build analytical models
- *Reduce data redundancy* – The Data Hub provides on-demand access to data thereby eliminating redundant data


1750 Tysons Blvd, Suite 1500, McLean, VA 22102

**(703) 289 - 5117 | sales@infocepts.com | www.infocepts.com**