MASTECH INFOTRELLIS

Case Study

Modern Data Lake on Cloud that Reduced CAPEX and Improved Reporting

Overview

Mastech InfoTrellis helped a legal firm setup a modern Data lake on Cloud and build into it the ability to mine, store and report data with meaningful insights. The client was looking for a solution with the ability to configure modern ELT/ETL processes and integrate API for data consumption. We assisted the client in creating a single view of data, which they were struggling with, and modernized their data with 360-view of the customer.

At a Glance

Client: Legal

Geography: Americas

Offering: Azure Platform, IBM MDM, SSIS

The Challenge

- The client had multiple data entities with no central source of truth
- There was a delay in statutory data reporting and manual work was necessary to create a customer view across multiple systems
- The inability to extract insights from existing data was hindering the client from progress
- The client also faced difficulty managing semi-structured data and merging home data with social

The Solution

We designed and implemented a modern Data Lake solution on Azure Cloud that integrated data from various sources to create a customer master data with associated contents

We migrated legacy databases and files into Azure Data Lake, created a process for matching, de-duplication and enrichment of member information

We built an analytical solution to search contents, view customer data and links and predict new avenues for business development

The Outcome

- The solution enabled real-time and batch processing of crucial data for the organization
- The client could perform smart business analysis for all departments on multi-dimensional cubes of data
- We removed more than 20GB of redundant data from servers, leading to massive cost savings on data transfer on/off the cloud
- The solution acted as a mature and self-sufficient data platform with rich self-service capabilities in terms of reporting and integration
- The self-patching and upgrade features empowered the internal IT team to focus on priority areas than just performing housekeeping

Highlights

30% reports moved to self-service mode

Reduced CAPEX and improved data visualization capabilities

Streamlined integration with other agency and social data using API and data pipes

Predictive data analytics for non-IT uses groups with dynamic search capabilities