

Mitigating Default Risk

Case Study | **Data Engineering**

LARGE RETAIL SERVICES COMPANY

THE PROBLEM

The company's services were based on the ability of their customers to obtain and pay for goods delivered to their home or office.

Because there was not good way to score a customer's credit real time, defaults and loss of revenue was occurring.

There was no common repository for data – departments had built their own custom solutions. Some kept data in the cloud, in Dropbox, or AWS RedShift databases.

THE SOLUTION

Mastech InfoTrellis worked with the customer to create a new data architecture and data lake of company-owned as well as third-party data for their customers. This included both real-time web services and batch; as well as structured, unstructured, and semi-structured data. The architecture included maturing the company's third-party integration, data transport, data bus, and analytics engines.

All of the data was cleaned and put into a format that could be easily analyzed. Analytics engines were developed that could perform real-time analytics on any given customer. Real-time Interfaces between the PoS and website were created allowing for quick determination of the risk score and if the customer was a high or low risk.

THE OUTCOME

Once the risk scoring infrastructure and tools were in place, defaults where reduced by over 30% and revenue increased.

Additional use cases have been enabled, particularly for advanced analytics. A Machine Learning (ML) algorithm governs certain risk score scenarios, eliminating entirely manual interactions in those cases.

The data lake has continued to grow and now handles archiving of historical transactional data – this has improved the risk scoring process as data analysts have been able to derive deep analytical insights from the trends in their data.