

Cybba Adds Comprehensive Data Lake House and AI Tools to Power Modern Analytics



Executive Summary

Cybba is a global digital software and services company that helps businesses accelerate online sales growth through an integrated suite of marketing and advertising solutions. The company reaches new customers and increases sales conversions, and works with major brands worldwide in a variety of sectors, including retail, travel, and education.

Recently, Cybba recognized it had an opportunity to revamp its data processing infrastructure so that clients could extract even more value from historical and future data. The company decided to bring in an experienced cloud solutions expert for a large-scale data engineering project that would set Cybba up for long-term success. Cybba relied on ClearScale's experience to integrate multiple AWS components into its data processing solution.

"ClearScale helped us make the critical shift from reactive to proactive data processing and analysis. With ClearScale, we can now harness and analyze the high volume of incoming data in real time to provide our clients with predictions and actionable insights about consumer behavior."

Christopher Peterson, Full Stack Developer, Cybba

The Challenge

Cybba's clients generate vast amounts of behavioral and demographic data every day from numerous websites and applications. Cybba needed to upgrade their on-premise MS SQL Server database to get more flexibility to process data for real-time campaigns on customer websites.

As Cybba grew, it needed a more efficient solution to keep up with the volume and velocity of incoming data. While the company's MS SQL Server environment could support necessary operations, data processing had been done manually instead of automatically. This process consumed a significant amount of time and monopolized company resources. Cybba sought out a way to aggregate and analyze customer data proactively, while minimizing the need for manual data analysis.

Cybba decided it was time to enhance the reliability and performance of its data infrastructure without significantly driving up their data costs.

As an [Amazon Web Services \(AWS\) Premier Partner](#) with a Data & Analytics Competency, ClearScale was well-positioned to help Cybba achieve its goals and extend the value of its digital marketing suite.

The ClearScale Solution

After understanding Cybba's current ecosystem and business needs, ClearScale suggested consolidating information from various sources into a single location and format with the deployment of a Data Lake House, using a plethora of sophisticated AWS technologies. To achieve this, ClearScale needed to aggregate Cybba's historical data, while simultaneously designing for scalability and data security, all while having minimal impact on daily operations.

ClearScale's first order of business was to collect and aggregate Cybba's on-premise historical data into a new data lake house, without disruption to Cybba's normal business operations. Once the data was loaded, ClearScale needed to crawl the historical data, determine its schema, validate inputs, and perform any necessary transformations. Following this, ClearScale designed a new data pipeline for ingesting future data directly into the lake house, ensuring data reliability, security, and resiliency during transit.

Today, all of Cybba's data is stored in a semi-structured manner and can be accessed through cost-efficient serverless or powerful managed services solutions. Information is processed more rapidly and consistently, opening up the door for advanced analytics. In this project, ClearScale provided the means to have a single source of truth for all of their customer data, and methods for amalgamating first party data with disparate data sources in order to provide a comprehensive picture of each user's experience.

In addition to the Data Lake House, ClearScale suggested leveraging AWS machine learning tools to further extract value from the aggregated data. Using tools with years of Amazon's retail and real-world data training, ClearScale applied predictive analysis to Cybba's existing data, enabling best in class personalization and forecasting for Cybba clients. Due to the serverless nature of these tools, AWS now takes care of burdensome backend administrative tasks, like server provisioning, patching, and maintenance, freeing Cybba's engineers to tune specific use-cases while preserving all the underlying model's traits like accuracy and precision. The solutions also enable Cybba to study customer segments more closely and pivot strategies accordingly.

The Benefits

Today, Cybba's data processing is more robust than ever. Cybba's new Data Lake House can now ingest and store data in a cost-efficient manner that can be used for both ad-hoc queries and ongoing analyses. Cybba's clients can load data in batches or in real-time and make informed decisions about their digital marketing campaigns.

This updated infrastructure will augment Cybba's MarTech and AdTech product offerings, which are designed to help brands engage and convert customers effectively. Cybba can now analyze disparate data sources and adjust targeting tactics for onsite and email campaigns based on all available data. Cybba also has access to deeper insight into how brands use its digital marketing tools. The team can use this information to identify and monitor brand user needs more closely.

The marriage of AI and ML capabilities allows Cybba's platform to handle large volumes of behavioral and demographic data without running up costs, which is essential for staying competitive in today's fast-moving digital landscape. There is no doubt about Cybba's ability to scale to meet demand or the needs of their customers.

Looking ahead, there are many ways that ClearScale can enhance Cybba's data infrastructure further. The partner can implement real-time data analytics with [Amazon Kinesis](#), build outlier-detection algorithms with [Amazon SageMaker](#), and even automate data deduplication with AWS Glue FindMatches, with the foundation ClearScale helped to build.