On-Demand Delivery Service Drives AI Success by Boosting Data Trust with Alation & Monte Carlo

Industry: Technology
Headquarters: United States
Interesting Fact: The company's services reach over 60 million US households annually

Challenge: More data, more data quality issues

Data is the lifeblood of this same-day shopping and delivery service that serves major US retailers. Their customers expect a seamless experience on their app, including efficient search and personalized product recommendations based on their input. It's also crucial that the company's shoppers — the folks who shop in-store and deliver the goods — have a smooth experience. Additionally, the company must demonstrate to their retail partners that they are delivering value and enhancing their customers' experiences. All these processes require well-governed access to high-quality data for reporting, analytics, and AI modeling to generate insights that improve the overall customer experience.'

As the company's data volume and number of data sources grew, the data quality and metrics team faced several challenges. People needed trustworthy data for accurate metrics, but different stakeholders would see the same metric defined differently by various teams, which led to different interpretations of the data. The team began uncovering problems with data quality as they traced data to its source. "At some point, after we'd traced that lineage enough times and found those underlying quality issues, we understood that implementing a data quality program was critical," their senior manager says. Compounding these challenges, they experienced issues with concurrent processing on their backend Postgres databases, significantly slowing time to data insight.

The company's data teams needed to provide fast, enterprise-wide access to trusted data. "Data trust means that we have an understanding of how data is coming into our system from the source," notes the team's senior manager. "And when the data gets in front of the user, they know that what they're seeing is accurate, and how to use it." This required data observability from the source and a way to provide context around the data, especially for business users. The business needed a front end where it could document data definitions, lineage, and PII policies, and that would update automatically.

> By using Alation to certify data assets and Monte Carlo to demonstrate that those assets are being continuously monitored, my team can ensure that we're providing trustworthy data to the data science team for their AI models.
>
> **Senior Manager, Data Quality and Metrics**
> On-Demand Delivery Service Company

Objectives:

To ensure well-governed access to trusted data, data leadership sought solutions to:

- Provide a single platform for enterprise data to speed up processing
- Improve data quality and reliability by identifying issues at the source
- Deliver an intuitive interface so users can easily find and understand trusted data

Implementation: Enterprise-wide data modernization

The company's data leadership addressed their distributed data landscape and concurrent processing issues by migrating their data to the Snowflake AI Data Cloud. "The promise of Snowflake — unlimited concurrency and the separation of compute and storage — has been crucial for meeting our needs," says the senior manager of data quality and metrics.

The data platform team chose Alation as their data intelligence platform to provide a single access point to their Snowflake data. Alation provides an intuitive user interface (UI) that business users can easily navigate so they can search for, understand, and trust the data they find. "We compared it to other, more technical data catalog tools that were built more for an IT persona, but Alation was built more to be used by the users that we care about," says the senior manager. "With Alation, everything was focused on the simplicity of the user interface. It was something that we knew we wouldn't have a hard time getting adoption for across the enterprise."

The data quality and metrics team tackled their data quality challenges by implementing the Monte Carlo data observability platform. Monte Carlo provides an understanding of data lineage from source to destination, ensuring that downstream users can quickly identify and address data discrepancies. The platform began providing value during the proof of concept when it revealed discrepancies in data uploaded in an hourly ETL process. During the proof of concept, the team immediately saw discrepancies in the hourly update data that they had been reporting as consistent.

> "
>
> With Alation, everything was focused on the simplicity of the user interface. It was something that we knew we wouldn't have a hard time getting adoption for across the enterprise.
>
> **Senior Manager, Data Quality and Metrics**
> On-Demand Delivery Service Company

Results: Delivering trustworthy data for AI models

"Alation is now the front end for our data," says the senior manager. "It's where we want people to go to search for data and understand how to use it." By surfacing PII policies and usage rules directly in the Alation UI — along with documentation, definitions, and a business glossary — the company's data teams can provide enterprise-wide, governed access to data for metrics and reporting. Thanks to the tight integration between Alation and Monte Carlo, even non-technical users can quickly see data quality issues without leaving Alation.

For this fast-paced delivery company, data observability is the key to high-quality reporting data. With data observability from Monte Carlo, they can see and alert their users in real-time when there's an issue. Monte Carlo is a critical tool for data observability and trustworthiness. In the last quarter of 2022 alone, the company estimates that they saved an estimated $500,000 by proactively identifying outliers in critical machine learning (ML) models with Monte Carlo alerts.

With the rise of GenAI initiatives that support the continuous improvement of the customer and shopper experience, the data quality and metrics team is laser-focused on providing high-quality data to the data science team. "There's an ever-increasing usage of the data we're creating," concludes their senior manager. "By using Alation to certify data assets and Monte Carlo to demonstrate that those assets are being continuously monitored, my team can ensure that we're providing trustworthy data to the data science team for their AI models."

About the Customer

This retail shopping and delivery service connects people to the retail brands they already know and trust. The service is available in more than 5,000 US cities.

Data Environment: snowflake⬡ +tableau