

# Customer Case Study: Discover Financial Services

Discover Accelerates Innovation with Faster Data Discovery and Active Metadata



- **Industry:** Financial Services
- **Headquarters:** Riverwoods, IL
- **Interesting Fact:** The Discover Global Network conducted \$550 billion in network volume in 2022.

## Challenge: Take “data-driven” to the next level

Since pioneering the first cash rewards credit card in 1986, Discover has become one of the most recognized brands in [financial services](#). The digital banking and payment services company is now one of the largest card issuers in the U.S. and operates a global payment network comprising millions of merchants and cash access locations. Today, 300 million cardholders use Discover’s global network at 70 million access points in over 200 countries.

“Discover is truly a data-driven company,” says Prakash Jaganathan, Senior Director, Enterprise Data Platforms at Discover. “Every decision, whether it’s made by an executive or an intern, is backed by data. Whether we offer credit to a customer, underwrite a loan, delight our customers with excellent service, or detect fraudulent transactions in seconds, our actions are supported by fresh, high-quality data.”

After more than 35 years in business, including multiple acquisitions, Discover has accumulated vast stores of data. Hundreds of petabytes of data are spread across on-prem systems, their Snowflake Data Cloud, and their AWS S3 data lake. As data accumulated, it became harder to strike a balance between speed to insight — what Discover calls

## Business Results



200,000 hours saved



1M datasets cataloged



2,500+ Discover users  
on Alation

data offense — and securing sensitive data and preserving [data integrity](#), known as data defense.

“We need to deploy new financial and behavioral models faster than ever to delight our customers and retain a competitive edge,” says Jaganathan. “But as a financial services company, we always need to balance that speed to insight with effective risk management.”

Getting data into the hands of data scientists and data analysts to create those new financial and behavioral models for innovation first required a rigorous but manual ETL pipeline creation process. The process, which took up to 30 days, required Discover’s data engineers to curate metadata, as well as to find, profile, and build quality control measures on data before they could create and run a pipeline to get the data into the hands of the people who needed it.

Furthermore, because Discover used multiple [metadata management](#) platforms, there was no single-pane-of-glass for analysts and engineers to quickly find the optimal data for decision-making. Even if they found the data they needed, understanding that data was difficult due to the lack of good up-to-date data documentation. The tribal knowledge about the data resided with the data stewards and subject matter experts. And there was no effective method to share this tribal knowledge across the organization. Dependency on such tribal knowledge did not scale for Discover’s large geographically distributed workforce.

***“We need to deploy new financial and behavioral models faster than ever to delight our customers and retain a competitive edge...But as a financial services company, we always need to balance that speed for insights with effective risk management.”***

Prakash Jaganathan  
Senior Director, Enterprise Data Platforms, Discover Financial Services

How could Discover get the right data into the right hands for faster decision-making while meeting all data security, data quality, and other risk management requirements?

“We wanted to cut the pipeline creation timeline from 30 days to one or two days,” says Jaganathan. “The only way we could do that was to automate and power those data pipelines using active metadata. In other words, we wanted to become truly metadata driven.”

## Objectives

Discover sought to deploy new financial and behavioral models faster to speed time-to-market for new customer products and experiences. To accomplish this goal, Discover first needed to:

- Leverage active metadata to speed up data pipeline creation
- Improve [data literacy](#) and speed up data insights
- Promote data self-service while protecting sensitive data
- Gain a deeper understanding of data migrated to the Snowflake Data Cloud

## Implementation: Enable faster data discovery and meta-driven pipelines in the Snowflake Data Cloud

To speed up analytics, Discover set out to modernize their analytical environment and complete the journey from on-premises to the cloud. As part of this data strategy, they deployed the [Alation Data Intelligence Platform](#). Alation captures and enriches metadata to enable data self-service throughout the company. With the addition of the [Alation Data Governance App](#) in 2022, Discover has moved to what Jaganathan calls “Governance 2.0,” providing decentralized governance that protects sensitive data and preserves data integrity by guiding compliance among empowered users.

[Alation connects to Snowflake](#) and other database platforms across Discover. These include Teradata, Oracle, SQL Server, DB2, Postgres, MySQL, and MariaDB. Alation was also used to scan reporting platforms such as Tableau. “Imagine the challenge of trying to find a fit-for-purpose dataset and understand its characteristics before we implemented Alation as our enterprise data catalog,” says Jaganathan. “In our data catalog, powered by Alation, we have about a million tables cataloged. Alation is used to capture active metadata, without which you cannot sift through millions of tables and find what you need.”

Discover captures four types of metadata in Alation:

- **Technical metadata** helps analysts and data engineers identify information such as schemas, tables, columns, data types, constraints and keys needed for insights and pipeline creation.
- **Business metadata** provides self-service governance by identifying names and descriptions, data ownership, privacy, and sensitivity tagging.

- **Operational metadata** such as top users and frequently used joins is captured via Alation's query log ingestion feature. These details help Discover leverage data in the Snowflake Data Cloud for faster insights creation and effective data management.
- Finally, Alation captures **social metadata** through articles and conversations. This allows knowledge sharing throughout the organization and reduces dependence on tribal knowledge.

"Metadata enables the automation we need to build data pipelines faster and monitor the quality of our data elements," says Jaganathan. "Alation has become our North Star, a single pane of glass in one system for our metadata."

## Result: Faster Snowflake data insights leads to faster innovation

Today, more than 2,500 users across Discover's global workforce employ Alation to find, use, and enrich the metadata in a million-plus tables already cataloged in Alation. With faster data discovery and 15x faster pipeline creation, Discover is well-positioned to speed up analytics. The productivity gains and faster time to insights lead to more time being spent on product innovation. Users across Discover can find information about datasets, including those migrated to Snowflake, in as little as 15 minutes and know they have quality data for decision-making. The productivity gains and faster time to insights lead to more time for innovation –and in Discover's case, that equates to 200,000 hours saved. Behind all these self-service capabilities lie Alation's governance features that provide the right people with faster access to the right data while keeping Discover compliant and audit-ready.

***"By improving the speed at which we are able to acquire and use data, we can turn that saved time into product innovations that help us remain No. 1 in customer satisfaction and delivering personalized customer experiences."***

Prakash Jaganathan  
Senior Director, Enterprise Data Platforms, Discover Financial Services

## Actionable metadata leads to faster pipeline creation

Prior to implementing Alation, data engineers at Discover spent a good amount of time on data discovery and metadata curation. Today, Alation provides the technical metadata needed to generate the pipelines and keeps the engineers apprised of changes to metadata.

“We have improved our data engineers’ developer experience,” says Jaganathan. “They have more time to understand the data that they are provisioning instead of just being the plumbers who are moving data, as in the past.” By understanding the data better, engineers build domain expertise, and can more quickly spot and solve data quality issues.

## Self-service boosts insights and organizational data literacy

Using Alation, the data discovery process takes as little as 15 minutes instead of up to 2 days and sometimes longer. “My goal on this journey is to ensure that people can come to Alation and find the best-fit data they need within 15 minutes, so they can spend the rest of their time analyzing the data and producing insights,” says Jaganathan.

“Data documentation used to be a laborious process,” says Jaganathan. “There was no excitement in documenting data, and we were not able to collect consistent quality metadata before Alation.” Today, Alation helps boost data literacy across the organization as users can easily document datasets and contribute quality metadata through articles and conversations. The result? Discover is no longer reliant on tribal knowledge. “Now, people are excited to use Alation’s friendly user interface to document their datasets.” Data stewards use Alation’s approval workflows to review and approve changes prior to publishing.

## AI and Alation: The future of data at Discover

Like every business, Discover intends to increase its use of AI and machine learning to support strategic decision-making. “The key to using AI and machine learning to make decisions is having context for the data,” states Jaganathan. “Alation can ensure that the datasets are labeled with accurate metadata, which tells the machine how to understand the context around the data it’s looking at. Only then can it answer our questions in a way that’s meaningful.”

## About Discover Financial Services

Since pioneering the first cash rewards credit card in 1986, [Discover](#) has become one of the most recognized brands in U.S. financial services. The digital banking and payment services company is now one of the largest card issuers in the United States. In addition to the widely known Discover® card, the company also offers private student loans, personal loans, home loans, checking and savings accounts, and certificates of deposit through their banking business. They operate the Discover Global Network® with millions of merchants and cash access locations; PULSE®, one of the nation's leading ATM/debit networks; and Diners Club International®, a global payments network with acceptance around the world. Discover's mission is to help people spend smarter, manage debt better, and save more to achieve a brighter financial future.

### Data Environment

- Snowflake
- Teradata
- Oracle
- SQL Server
- MySQL
- DB2
- Postgres
- MariaDB
- Kafka