

CASE STUDY

Michelin Transforms Data Access and Governance with Dremio's Open Data Lakehouse Platform

At a Glance

The Customer



Challenge

Michelin's traditional data lake architecture created significant barriers to data access and governance. Users could only access data through direct file connections without SQL capabilities, making it difficult to provide fine-grained access controls and support multiple data formats across their diverse user base of 100,000+ employees.

Solution

Michelin implemented Dremio as the foundation of their lakehouse architecture, providing unified SQL access, federation capabilities, and integrated governance tools. They developed an industrialized approach using infrastructure-as-code practices with DBT integration and connected Dremio to their existing governance tools like Collibra.

Results

Simplified data consumption experience for business users, enhanced security with row-level access controls, streamlined data product development through code-based deployment, and improved performance through intelligent caching and reflection capabilities.

The Customer

Michelin Group stands as a global leader in tire manufacturing and mobility solutions, employing over 100,000 people worldwide and generating more than 25 billion euros in annual revenue. The company sells approximately 200 million tires yearly while expanding beyond traditional tire manufacturing into tire-related services, mobility experiences including the famous Michelin Guide, and high-tech materials expertise. As Michelin increasingly leverages data to provide new services and features, their data strategy has become critical to supporting both traditional operations and innovative product offerings across their diverse business portfolio.

The Challenge

Four to five years ago, Michelin embarked on a comprehensive data transformation organized around three pillars: platforms, governance, and products. Initially, they implemented a traditional data lake architecture that could ingest, transform, and store data effectively. However, this approach quickly revealed significant limitations that hindered their data democratization goals.

The primary challenge was accessibility. Users could only connect directly to files stored in the data lake, with no SQL engine available for querying. This technical limitation severely restricted who could access and analyze data, essentially limiting usage to highly technical users comfortable with file-based data consumption. Additionally, Michelin's governance team required increasingly fine-grained access controls, but the data lake only supported directory-level permissions, making it impossible to implement sophisticated security policies.

Data format inconsistency created another major obstacle. The data lake contained information stored in various formats including CSV, Delta, Iceberg, and Parquet files with no standardization. This variety meant that different analytical tools couldn't consistently consume the data, creating fragmentation and limiting the value users could extract from their data investments. The lack of a unified interface forced users to navigate multiple systems and understand various technical formats, significantly reducing adoption across business teams.

The Solution

Michelin chose Dremio to transform their data lake into a modern lakehouse architecture that could address these fundamental challenges. Dremio provided the essential SQL engine capabilities they needed while offering federation across multiple storage systems, an integrated user portal, and advanced security features including row-level security management.

The implementation strategy focused on serving two distinct user personas. For data producers—primarily IT teams exposing data from source systems and data teams consolidating information from multiple sources—Michelin developed an industrialized approach using infrastructure-as-code principles. They integrated Dremio with DBT (Data Build Tool) to enable teams to define their data models as SQL code stored in Git repositories, with automated CI/CD pipelines deploying these models as virtual datasets in Dremio.

For data consumers, Michelin designed the Dremio interface to enable self-service analytics while maintaining governance standards. They organized their Dremio spaces into logical structures including source applications, 16 business domain directories reflecting their organizational structure, and application-specific folders. This organization ensured that users could easily discover relevant data while governance teams could manage access and quality standards effectively.

To enhance the user experience, Michelin connected Dremio with their existing Collibra governance platform. This integration automatically populated Dremio with metadata descriptions, ensuring users had comprehensive information about datasets directly within their analytical environment without switching between tools. They also implemented intelligent caching through Dremio's reflection capabilities, significantly improving query performance for frequently accessed data.

Results

The Dremio implementation delivered transformative results across Michelin's data landscape. The unified SQL interface eliminated technical barriers that previously limited data access to specialized users, democratizing analytics across their 100,000+ person organization. Business users gained the ability to perform complex data operations including joins, filters, and aggregations directly through Dremio's interface, enabling true self-service analytics.

Security and governance capabilities improved dramatically through row-level security implementation, allowing Michelin to provide fine-grained access controls based on user groups and business requirements. This granular security model supports their complex organizational structure while ensuring sensitive information remains properly protected.

The infrastructure-as-code approach using DBT integration enabled development teams to manage data products with software engineering best practices, including version control, automated testing, and deployment pipelines. This methodology significantly reduced deployment time and improved reliability while enabling teams to scale their data product development efficiently.

Performance enhancements through Dremio's reflection and caching capabilities reduced query response times and improved user experience, particularly for frequently accessed datasets. The integrated metadata management through Collibra connectivity ensured users have comprehensive context about data quality, lineage, and business definitions directly within their analytical workflows, reducing time spent searching for information and improving confidence in analytical results.

ABOUT DREMIO

Dremio is the intelligent lakehouse platform for the business, serving hundreds of global enterprises, including Maersk, Amazon, Regeneron, NetApp, and S&P Global. Based on open-source technologies like Apache Iceberg and Apache Arrow, Dremio provides an open lakehouse architecture enabling the fastest time to insight and platform flexibility at a fraction of the cost.

Learn more at www.dremio.com.

Dremio and the Narwhal logo are registered trademarks or trademarks of Dremio, Inc. in the United States and other countries. Other brand names mentioned herein are for identification purposes only and may be trademarks of their respective holder(s). © 2025 Dremio, Inc. All rights reserved.