



9 data products success stories with Starburst

The end-to-end analytics platform for both cloud and on-premises workloads



Introduction

The volume of data generated and collected by companies has grown exponentially over the decades, becoming both an asset and a challenge. Modern enterprises face the daunting task of managing diverse data sprawled across multiple systems, leading to complex data architectures and silos that hinder effective data analysis and decision-making. The critical issues at hand span from integrating disparate data sources, ensuring data quality, and enabling fast, reliable access to insights.

The rise of data products

Data products emerge as a strategic solution to these challenges by offering a structured approach to managing and utilizing data. They enable companies to package, curate, and share data across teams and systems, ensuring consistency, quality, and accessibility. However, developing and managing data products efficiently requires powerful tools and platforms capable of handling the complexities of modern data ecosystems.



Transforming data management with Starburst

Starburst accelerates the journey from data to insight, making it faster and easier to efficiently turn data silos into valuable, purpose-built data products. This is achieved through an open data lakehouse model, enhancing the exploration, building, governance, access, and sharing of data across more than 50 sources, on-premises, and across clouds. Starburst's solution increases data team productivity, ensuring data consumers and producers benefit from a streamlined, efficient process that enhances the business value of data across all lines of business.

In this eBook, learn how nine companies leverage Starburst data products to transform their data ecosystems, driving efficiencies, fostering innovation, and achieving a competitive edge in their respective industries.

Contents

Introduction	2
Table of Contents	4
Resilience streamlines biomanufacturing analytics with data products	5
Halliburton turns a data swamp into data products with real-time insights	6
Gilead Sciences turns a data swamp into data products with real-time insights	7
Genus improves animal genetic testing with faster data discovery and data product creation	8
Glovo advances its data mesh initiative and optimizes its marketplace with decentralization	9
El Toro leverages data products to build a data marketplace.....	10
Sky implements a data mesh architecture for decentralization and easy data product creation	11
Executive branch selects Starburst to improve Veterans' benefits	12
Leading pharmaceutical organization adopts Starburst data products to speed time-to-market for clinical trials	13
Conclusion	14

RESILIENCE

Resilience streamlines biomanufacturing analytics with data products

Overview

Resilience is a first-of-its-kind biomanufacturing innovation partner dedicated to broadening access to complex medications by revolutionizing how they're made, funded, scaled, and accessed.



“With data products in Starburst, we’re able to work with business partners to codify a single accurate representation of their data, and share numbers with the rest of the organization that we’re 100% confident in.”



ADAM MENDEZ
Head of Data Engineering
and Architecture

Challenge & solution

Resilience faced a complex data landscape due to rapid expansion and the acquisition of diverse legacy systems, leading to inefficient data management and reporting inconsistencies.

To address these challenges, Resilience adopted a data mesh architecture by choosing Starburst Enterprise hosted on AWS. This strategic move facilitated the creation and management of data products, enabling standardized, scalable, and efficient access to data. Starburst’s solution empowered Resilience with improved data analysis capabilities and consistent query performance, streamlining their data operations and fostering a more agile data environment.

What Starburst delivered

- 240 hours of data copying saved per month
- 20+ data products across 6 domains
- 10X increase in BI tool user adoption

HALLIBURTON

Halliburton turns a data swamp into data products with real-time insights

Overview

Halliburton is an American multinational corporation and the world's second largest oil service company which is responsible for most of the world's largest fracking operations.



“Previously it would take 2 to 3 weeks to get an answer to an ad hoc question. By embedding an LLM with Starburst’s data products architecture, data consumers can ask questions in plain language, have it converted to SQL, and get the answer back immediately.”



FAHAD AHMAD
Data Science Leader,
Halliburton

Challenge & solution

Halliburton had hundreds of disparate data sources, which caused data access and integration challenges. The central infrastructure team was overburdened by ad-hoc requests and the constant maintenance of their ‘data swamp.’

Halliburton’s journey with Starburst data products provided a solution by enabling decentralized data management, where each of their 13 Product Service Lines maintain their own high-quality data products. This ensures a single, consistent view of data across the organization.

The introduction of Large Language Models (LLM) further enhanced Halliburton’s approach, enabling users to formulate queries in natural language which are then translated into SQL for real-time responses.

What Starburst delivered

- Instant access to any data within the company
- No brittle data movement pipelines
- Fast data-driven decision making



Gilead Sciences builds 500+ federated data products with Starburst

Overview

Gilead Sciences is a biopharmaceutical company that focuses on advancing innovative medicines to prevent and treat life-threatening diseases, including HIV, viral hepatitis and cancer.



NGHI HO

Sr. Director, Head of Data & AI Platform and Enterprise Data Governance, Gilead Sciences

“Starburst fits well at Gilead, because it supports open standards and seamless integration with governance tools, data catalogs, data products, and it provides a secure data platform.”

Challenge & solution

Gilead has its analytical data stored across different database technologies such as SAP Enterprise HANA, AWS Redshift, Amazon S3, and Oracle Database. This made it challenging to achieve insights quickly and manage data access.

Starburst enables a federated semantic layer and distributed query engine as a global access layer (with fine grained access control), allowing Gilead staff to access data quickly and easily, no matter where it lives.

What Starburst delivered

- 500+ data products across nine business functions
- 4X data product development speed
- 5000+ direct and indirect data product users



Genus improves animal genetic testing with faster data discovery and data product creation

Overview

Genus PLC is an award-winning animal genetics company. The company researches and develops innovative animal breeding technologies that support a more sustainable food system for generations to come.



“With Starburst, we have accelerated data discovery, simplified data pipelines, and have a unified query layer across all data sources. These three points are critical to what we do.”



PATRICE LINEL
Sr Manager of Data Science &
Data Engineering, Genus

Challenge & solution

Genus faced a significant challenge in achieving dataset interconnectivity essential for innovation in animal breeding and genetics, with separate databases for different types of genetic information and a complex data storage infrastructure. This resulted in slow analytical response times, data quality issues, and lost productivity.

By adopting Starburst Enterprise and Starburst data products to support its data mesh architecture, Genus streamlined data access through a unified query layer, improving scalability, manageability, and performance.

What Starburst delivered

- 75% faster time-to-insight
- 150X faster data product validation & creation
- Hundreds of thousands of dollars in cost savings per year from reduced infrastructure costs



Glovo advances its data mesh initiative and optimizes its marketplace with decentralization

Overview

Glovo is a pioneering multi-category app founded in Barcelona in 2015. It connects users with businesses, and couriers, offering on-demand services from local restaurants, grocers and supermarkets, and high street retail stores.



SIMONE GRANDI
Engineering Manager in Data Platform,
Glovo

“Starburst offers a very nice user interface that allows not only admins, but also the data users to monitor the clusters and optimize the queries. In our self-service mindset, this is very useful.”

Challenge & solution

Glovo aimed to optimize its multi-category platform serving customers, couriers, advertisers, and brands by ensuring reliable access to high-quality data products for analytical and operational use cases.

To address the challenge of managing data at scale and facilitating data-driven decision-making, Glovo adopted a data mesh approach, selecting Starburst for its scalability, ease of deployment on Kubernetes, autoscaling capabilities, and enterprise-grade features.

What Starburst delivered

- 50% increase in query performance
- Reduced Total Cost of Ownership (TCO) from separation of compute from storage
- Self-service data management (via data products)



El Toro leverages data products to build a data marketplace

Overview

El Toro is an advertising technology company that allows advertisers to target individual homes based on IP addresses and the demographic information of the people living there with extreme precision.



“The needs we were struggling with were exactly why we ended up with Trino, and Starburst was a very natural next step for us as an enterprise company — it makes data access easier, better, more supported, more stable, and more developed without needing to put the resources in place.”



RICHARD TEACHOUT
CTO, EL Toro

Challenge & solution

On any given day, El Toro’s Machine Learning algorithms digest over 350 billion data points. Processing large amounts of complex data across multiple platforms was slow and time-consuming — El Toro relied on complex ETL pipelines to facilitate queries for ad-hoc analytics.

El Toro deployed Starburst Enterprise to accelerate data access with features such as enterprise-grade performance, connectivity, security, cluster management, and support. Additionally, Starburst enables data producers to create and maintain an array of data products, including a client-facing data marketplace for downstream users to query live data.

What Starburst delivered

- 300% improvement in query performance
- \$5 million in cost savings compared to a cloud data warehouse
- Thousands of campaigns launched per day



Sky implements a data mesh architecture for decentralization and easy data product creation

Overview

Founded in 1999 as the UK's first satellite TV service, Sky is Europe's leading media and entertainment company.



“We are in the process of creating data products, which Starburst is really helping with. Previously, without a single point of secure data access, creating a data product was not possible. Starburst has become our analytics engine for the Data Mesh.”



RITESH RANJAN
Lead Data Architect, Sky

Challenge & solution

Sky faced challenges in accessing and analyzing petabytes of data distributed across various cloud and on-premises storage systems, which previously took months to process via ETL pipelines.

With the implementation of a data mesh architecture through Starburst Enterprise, Sky achieved a decentralized approach, allowing for immediate access to all their data through a single point. Sky is also able to apply a product mindset to datasets and create data products. This allows Sky to give users access to the data they want and ultimately helps drive its business forward.

What Starburst delivered

- Drastic reduction in time-to-insight, from two months to minutes
- Integration with critical technologies (Google Cloud Platform, AWS, and Tableau), supporting Sky's data mesh architecture
- Enhanced data access and security through Starburst Stargate

ANONYMOUS

Executive branch selects Starburst to improve Veterans' benefits

Overview

This executive branch provides lifelong healthcare services to eligible military veterans. The department aims to accelerate analytics to identify veterans quicker, and inform patients about the veterans' benefits they're eligible to receive.



“Starburst is helping us unlock data silos and access data products without having to do a ridiculous amount of ETL and data movement across our platforms. I’m excited to start branching out with more of our end users as we get into production.”



ANONYMOUS
Director of Data & Analytics

Challenge & solution

This government agency had a requirement to enable a unified view of data across diverse analytics and data sources while improving data accessibility, governance, and query performance.

Starburst was selected as the analytics query accelerator to support the branch's R&D department in efficiently integrating and analyzing large volumes of data housed in disparate data platforms. Coming soon, the branch plans to leverage Starburst data products to create a Veteran 360 portal.

What Starburst delivered

- Empowers self-service analytics within business domains through data products
- Accelerates the creation of a centralized veteran experience for comprehensive health-related data, enhancing benefits and portfolio development
- Federated data access across Synapse, ADLS, and Databricks eliminates vendor lock-in

ANONYMOUS

Leading pharmaceutical organization adopts Starburst data products to speed time-to-market for clinical trials

Overview

This global pharmaceutical and biotechnology company provides innovative medicines to millions of patients throughout the world.



“Adopting Starburst data products has been a game-changer for us, reducing the time to insight from three to four months to just two days within our R&D department. This has not only expedited our clinical trial analysis but also ensured that data access is both secure and compliant, which is essential for the integrity of our trials.”



ANONYMOUS
Director of Data Platforms

Challenge & solution

This pharmaceutical company had long delays in getting access to clinical trial data, impacting their ability to get fast insights. They also faced security and compliance risks if someone gained unauthorized access to a clinical trial dataset. In rare cases, this could result in an entire clinical trial being scrapped after years of work.

Starburst provides a data analytics platform to accelerate and streamline access to clinical data and other data, enabling faster analysis on their AWS and Azure infrastructure. With Starburst data products, researchers can securely access curated data sets within days rather than months.

What Starburst delivered

- **<2 days to gain access to clinical trial data (versus 3 – 4 months previously)**
- **Improved security and compliance through granular access control to curated datasets**
- **Cost savings through faster time-to-market for clinical trials and reduced risk of data breaches**

Conclusion

The experiences chronicled in this eBook illustrate a clear trajectory towards a more interconnected, insightful, and innovative future in data analytics. The adoption of data products with Starburst not only resolves historical data dilemmas but also marks a shift towards a more agile, informed, and collaborative approach to data management.

READY TO UNLOCK THE POTENTIAL OF DATA PRODUCTS?

Get started today with a [free trial](#) on Starburst Galaxy.

