snowflake®

ACCELERATE TIME-TO-VALUE WITH A MODERN DATA STACK

Bring speed, scale and simplicity to Snowflake with Fivetran, dbt Labs and Sigma

INTRODUCTION

In modern business, data is central to almost all activities – and the scale of the data deployed to make key decisions grows by the second. In fact, more than 400 million terabytes[1] of data are produced every day. But it's not how much data you have that matters – it's what you do with it.

Putting data to use is integral to the fast, accurate decision-making at the heart of successful organizations. Data-driven practices are now table stakes for success, key to gaining competitive advantage and vital for future-proofing operations. But to turn masses of unstructured data into actionable insights, you need the right data infrastructure.

For many organizations, outdated, piecemeal data architectures and legacy processes restrict time to insight. They contribute to escalating costs, limit the amount of data that can be explored and prevent organizational agility due to their inherent inflexibility. As AI, machine learning and data processing advancements demand even greater performance from these infrastructures, there's extra pressure on organizations to reassess and upgrade their data strategies.

For those that want to get ahead, it's vital to simplify the pipeline from raw data to analytics, bringing new levels of scalability, agility and speed to data practices. This is the principle at the heart of a modern data stack (MDS). In this eBook, we'll explore the four key elements of an MDS and how it can help you embed advanced data analytics at the heart of your operations.

More than 400 million

terabytes of data are produced every day

1 https://www.statista.com/statistics/871513/worldwide-data-created/

WHAT IS A MODERN DATA STACK?

Your data stack dictates the speed of your teams' operations and serves as the backbone of impactful, data-driven analysis. But to meet modern demands, it must have certain attributes.

Unlike the on-premises monoliths of the past, the MDS is cloud-based, modular and flexible, providing immediate benefits and long-term advantages. By uniting cloud-native tools and technologies, it streamlines the process of collecting, storing, transforming, analyzing and visualizing data. And its emphasis on scalability, flexibility and ease of use makes it the antithesis of the rigid IT infrastructures of the past.

Done correctly, this configuration of tools should enable you to efficiently make data-driven decisions without your IT teams having to dedicate resources to infrastructure management or worry about expanding capabilities. Instead, they can focus on what really matters – deriving insight.

Vitally, each component of an MDS is designed for interchangeability and optimized for its specific role. While the perfect configuration will vary depending on your business requirements, one guiding principle remains a constant: an MDS should combine best-of-breed technologies that evolve to meet your changing needs – so you can work at the speed of data.

THE CORE COMPONENTS OF A MODERN DATA STACK

An MDS has four key components, including tools for **data ingestion**, **storage**, **transformation** and **visualization**.

1. DATA INGESTION TOOLS (ETL/ELT)

Efficient data ingestion is vital for gathering data from different sources (such as apps, databases and APIs) into a centralized repository. Popular examples of data ingestion tools include Fivetran, Airbyte, Stitch and Kafka.

2. DATA WAREHOUSES OR DATA LAKES

A central store for all data, in a scalable, queryable format, is essential for providing analysts with timely access and ensuring everyone has a single version of the truth. Examples of data stores include Snowflake, BigQuery, Amazon Redshift and Microsoft Azure.

3. DATA TRANSFORMATION TOOLS

Before raw data can be used by analysts, it must be cleaned, standardized and transformed into a usable format. This is where data transformation tools such as dbt and Astronomer come in.

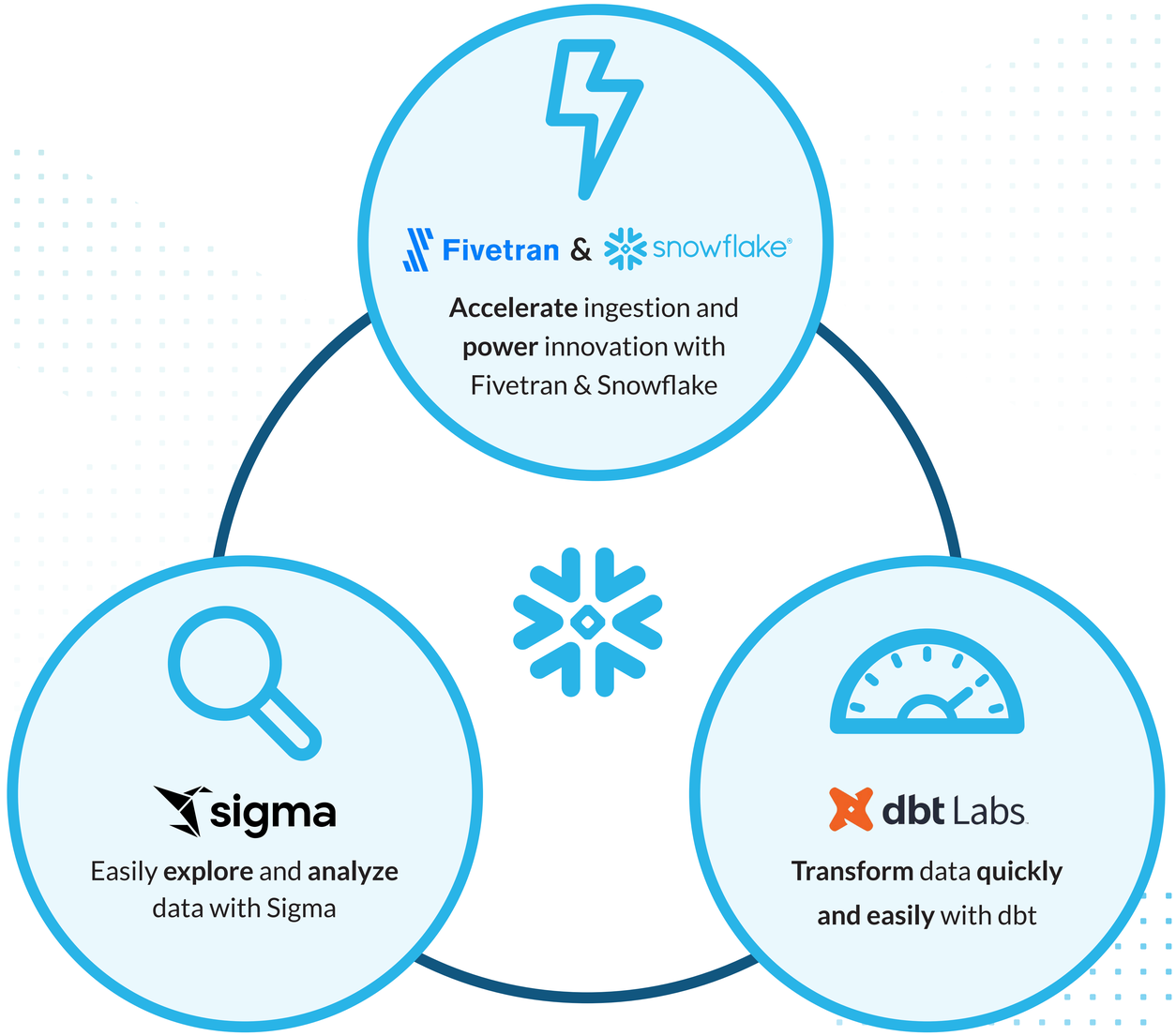
4. DATA VISUALIZATION AND BUSINESS INTELLIGENCE (BI) TOOLS

Finally, the insights from data must be shareable with non-technical business users who need to explore that data at will. This is the role of data visualization and BI tools. Popular examples include Tableau, Looker, Power BI, Mode and Sigma.

GET MORE FROM YOUR DATA WITH THE MODERN DATA STACK

A MODERN DATA STACK BUILT ON SNOWFLAKE'S AI DATA CLOUD

In this guide, we're exploring an MDS that consists of four industry-leading tools, all working in harmony while delivering their own best-of-breed capabilities across the four key areas outlined above.



ACCELERATE INGESTION AND POWER INNOVATION WITH FIVETRAN AND SNOWFLAKE

The first step to becoming a truly data-driven organization is ensuring data is available when you need it. But that's not always easy.

Data integration can be a complex process. You have to scope and scale infrastructure, ensure availability, make sure you can quickly recover from failures, and maintain and adapt systems in response to changing data sources and evolving business needs.

Many common data integration tools offer frameworks for these tasks, but they can still demand a considerable amount of engineering work, which means slower time to insight, higher costs and a greater strain on resources.

Fivetran is the global leader in data movement, helping thousands of customers worldwide use their data to power everything from AI applications and ML models to predictive analytics and operational workloads.

Its automated, reliable and secure data pipelines seamlessly integrate hundreds of SaaS applications and databases into Snowflake's AI Data Cloud. With 99.9% pipeline uptime and automated schema drift management, Fivetran ensures availability and reduces the risk of errors. This allows you to focus on using your data effectively without worrying about the complexities of data integration.

99.9%

pipeline uptime and automated schema drift management, Fivetran ensures availability and reduces the risk of errors

BY COMBINING FIVETRAN AND SNOWFLAKE, YOU CAN GAIN ACCESS TO A FULLY MANAGED, HIGH-PERFORMANCE DATA SOLUTION

AUTOMATED, ZERO-MAINTENANCE DATA PIPELINES

Fivetran's automated schema drift management, secure log-based change data capture (CDC) replication, and hybrid deployment capabilities deliver high volumes of data to users in real time without impacting source systems. With Fivetran, you can seamlessly extract and integrate business-critical data for a comprehensive, 360-degree view, while ensuring compliance and adhering to industry standards.

SECURE DATA REPLICATION FOR THE SNOWFLAKE AI DATA CLOUD

Fivetran's automated schema drift management and secure, log-based change data capture (CDC) replication deliver high volumes of data to users in real time without impacting source systems. That means you can easily extract and combine business-critical data for a complete, 360-degree view, all while maintaining compliance and adhering to industry standards.

SIMPLIFIED DATA MANAGEMENT FOR ALL ANALYTICS AND AI USE CASES

Fivetran enables you to break down data silos and unify teams on a single data platform. With Fivetran and Snowflake's AI Data Cloud, you can centralize data for analytics and AI, driving insights at scale and enabling new use cases through seamless data access and collaboration.

COMPANIES USING FIVETRAN HAVE SEEN:

400%+
average three-year ROI

$1.5m
average annual benefit

99.9%
uptime across 3+million daily syncs

TRANSFORM DATA QUICKLY AND EASILY WITH DBT LABS AND SNOWFLAKE

If ingesting data is the first step, ensuring the quality of that data is next. After all, the results of data activities are only as good as the data they're based on – and that's where data transformation comes in.

dbt Cloud combines SQL-based transformation with best practices from software engineering to make data transformation reliable and fast. dbt Cloud provides a fully automated transformation workflow that ensures consistent data transformation without the need for manual intervention. It also allows you to produce reliable data sets for reporting from directly within your data warehouse.

All of this means you can ship trusted data faster, so you can put it to use faster, building reliable, governed data and AI products on Snowflake's AI Data Cloud.

dbt Cloud also plays a crucial role in removing silos by centralizing workflows like observability, cataloguing and orchestration in a single data control pane to provide holistic context. Thousands of organizations rely on the dbt Data Control Plane to accelerate pipelines, improve collaboration and conduct data activities at scale.

Those looking to build an MDS on Snowflake's AI Data Cloud can use dbt Cloud to transform data in place, increase data quality with rapid testing and accelerate data engineering and development activities. With impressive capabilities like these, it's no surprise that dbt Labs has been named Snowflake's Data Integration Partner of the Year for the last two years running.

WITH DBT CLOUD AND SNOWFLAKE

- **Build trust**

  Help users understand transformations deeply, with automatic logging and alerting that uses column-level lineage.

- **Build data products**

  Reuse assets and get full support for a robust development process, so more people can build more quickly.

- **Ensure security and governance**

  Stay in control of your dbt Cloud and Snowflake data transformations with single sign-on, role-based access control, audit logging and 99.95% uptime.

- **Scale with ease**

  Use existing dbt models across all teams with dbt Explorer and maximize the value of your consumption.

- **Protect sensitive data**

  Combine dbt Cloud's fine-grained access controls with Snowflake's dynamic data masking for secure data pipeline development.

- **Develop and deploy more efficiently**

  Orchestrate jobs, reuse assets and automate version control with browser-based development.

- **Build trust in AI**

  With dbt's native application, you can build trusted AI applications and use Snowflake Cortex AI to implement the large language model (LLM) of your choice. Users can now issue Cortex requests directly from dbt and use the results in data pipelines.

DBT AND SNOWFLAKE USERS HAVE EXPERIENCED:

Upto
99.9%
uptime with dbt Cloud

18 hours
less query time per week

80%
faster time to insight

EASILY EXPLORE AND ANALYZE DATA WITH SIGMA AND SNOWFLAKE

Delivering insights to your business users is what data analytics is all about. Today's analytics platforms no longer need to host and process data themselves to present visualizations – and this can have a major impact on speed to insight.

Sigma is a next-generation analytics and business intelligence platform that scales billions of records using spreadsheets, SQL, Python or AI — all without compromising speed or security.

It's a SaaS application that provides a simple interface for users to query, profile, visualize and explore data stored and shared in Snowflake's AI Data Cloud. The app uses a direct, secure connection to Snowflake alongside a SQL generation engine to translate user interactions in a familiar spreadsheet interface into machine-optimized SQL. So, if you know how to use a spreadsheet, you're already an expert.

This workflow can help democratize data access across your organization, providing business users with limitless charts, graphs and pivot tables with a level of flexibility not found in other tools.

With Sigma, users can query, pivot, transform and visualize data sets containing billions of rows in real time and with no limitations. Teams can add context to the data to tell a complete story and securely write new data back to their cloud data platform alongside the original dataset. Above all, it's easy for both data teams and business users to use. And since it doesn't have to store any data, there's no need to maintain extracts from source systems or aggregate data in intermediate tables to optimize performance — a common issue with legacy BI solutions.

USING SIGMA ALONGSIDE SNOWFLAKE'S AI DATA CLOUD

Analyze trillions of rows of data in seconds

Gain endless scalability without compromising performance or security, with Sigma's unique architecture.

Simplify collaboration

Interact with data, and your colleagues, in real time with a single, collaborative data workspace.

Securely write back and automate

Add context to your work and automate actions, with the only BI platform that lets you input data directly into your analysis.

Explore the potential of AI

Seamlessly integrate your trusted LLMs into your teams' workflows on a secure and auditable platform and enable every business user to unlock data science capabilities.

Build powerful apps on your cloud data

Build scalable, code-free apps on your Snowflake data to meet both internal team and customer needs.

COMPANIES USING SIGMA HAVE SEEN:

90%
cost savings

20%
sales increase through predictive analytics

30%
more queries for the same cost with Snowflake

THE MODERN DATA STACK IN ACTION

An MDS can help your organization bring speed, agility and flexibility to its data processes, democratizing access to insight across both technical and non-technical users and enabling you to operate at the speed of change.

Users of this data stack, and the tools within it, have seen lower costs, faster time to insight, increased sales, greater revenue, and significant returns on investment.

95%

TROVATRIP SPEEDS TIME TO INSIGHT BY 95%

By modernizing its data stack with Snowflake, Fivetran and dbt, and integrating Sigma for business intelligence, TrovaTrip automated dashboards and enabled self-service analytics. This transformation reduced report delivery time from days to just an hour, empowering business users to create intricate data models without SQL. The results were 95% faster delivery of insights and significant time savings for the data engineering team.

PREPARE FOR YOUR DATA-DRIVEN FUTURE

Data will continue to play a central role in shaping business decisions, and the rapid rise of AI will accelerate the need to use it effectively – which means agility and flexibility are essential for future success.

Those who invest in the right foundations for data exploration now will experience better decision-making and improved operational efficiency, and will be able to deliver enriched customer experiences. But it all hinges on having a modern data stack at the heart of those data activities.

The stack you choose should be an evolving ecosystem capable of delivering scalability, ensuring data quality, enabling real-time analytics and adjusting to the changing needs of your business.

LET'S GET STARTED

- To learn more about building your own MDS, or to discuss your data needs in general, get in touch with one of our experts below.

ABOUT SNOWFLAKE

Snowflake makes enterprise AI easy, efficient and trusted. More than 11,000 companies around the globe, including hundreds of the world's largest, use Snowflake's AI Data Cloud to share data, build applications, and power their business with AI. The era of enterprise AI is here.

Learn more at **snowflake.com** (NYSE: SNOW).