snowflake®

THE ESSENTIAL GUIDE TO DATA ANALYTICS

Learn how to be AI-ready with a leading analytics engine, offering world-class price/performance across all data — whether in a data warehouse, lake or lakehouse

TABLE OF CONTENTS

INTRODUCTION

Organizations today have an unprecedented amount of data filled with business insights. Analytics is critical for organizationsto stay competitive, relevant and ready for the innovation that the age of AI is bringing.

Those who fail to embrace analytics risk falling behind competitors, missing crucial market opportunities, and making uninformed decisions that could prove costly. As AI continues to transform industries, analytics serves as the foundation for understanding and effectively leveraging these technological advancements.

In the age of data and AI, analytics is no longer a luxury but a necessity.

This guide will explore the benefits of analytics, the underlying architecture required to unlock its true value, data migrations, how to ensure your organization is ready for the future, and so much more. But first, let's define a few terms to ensure we're all on the same page.

DEFINITION

Analytics, at its core, is the systematic computational analysis of data or statistics that enables organizations to discover, interpret and communicate meaningful patterns in their data. While the specific applications are vast, analytical endeavors typically fall into these primary categories:

- **Descriptive analytics** explains what has already happened, providing insights into past events and performance

- **Predictive analytics** forecasts future trends and outcomes, enabling proactive decision-making

- **Prescriptive analytics** recommends specific actions to optimize results, guiding organizations toward desired objectives

These different types of analytics work together to provide a comprehensive view of an organization's operations, customer behavior and market position. However, to truly harness these benefits, organizations require an underlying architecture capable of efficiently storing and managing their data.

WHY ANALYTICS MATTER

The benefits of analytics are far-reaching and transformative for organizations. It helps them make more informed decisions based on data rather than intuition, leading to better outcomes and reduced risk.

Analytics enables deeper customer understanding, allowing companies to personalize their offerings and improve customer satisfaction. Operational efficiency improves as analytics identifies bottlenecks and opportunities for optimization, while real-time monitoring of key performance indicators ensures organizations can quickly respond to changing conditions.It also gives organizations a competitive advantage through market insights and trend identification. It supports innovation by uncovering new opportunities and patterns that might otherwise go unnoticed.

Cost savings emerge through optimized resource allocation and waste reduction, while compliance monitoring becomes more efficient and effective. Perhaps most importantly, analytics enables strategic planning by providing the insights necessary for setting and achieving long-term goals, ensuring organizations remain competitive and resilient in an increasingly data-driven world.

ARCHITECTING YOUR DATA FOR ANALYTICS

To truly unlock the benefits of analytics, all data must reside somewhere accessible. Organizations can choose from various architectural approaches. In this section, we'll briefly explore four different types of common architectures that have evolved over the years. We will then delve into data warehouses as a specific case study to show the evolution of a data architecture type into a modern cloud data platform. But first, the four types of architectures are:

1. Data Warehouses

A data warehouse is a centralized repository that stores data from various sources in a structured and organized manner, making it easily accessible for analysis and reporting. The primary purpose of a data warehouse is to provide a single, unified view of an organization's data, enabling business intelligence (BI) and analytics.

Analytics is a key driver for the creation and maintenance of a data warehouse, which provides the underlying foundation for analytics through:

- **Data integration:** Collecting and integrating data from multiple sources makes it possible to analyze data from different departments or systems.

- **Data standardization:** Standardizing data formats and structures enables consistent analysis and reporting.

- **Data quality:** Supports data accuracy, completeness and consistency, which is essential for reliable analytics.

- **Data accessibility:** Provides a single, unified view of data, making it easier for analysts to access and analyze data.

Analytics tools — such as SQL, online analytical processing (OLAP) and data visualization tools — are used to extract insights from a data warehouse. These insights are then used to inform business decisions, optimize operations and drive growth.

2. Data Lake

A data lake is a centralized repository that stores raw, unprocessed data in its native format, without any predefined schema or structure. A data lake is designed to store large amounts of data, including structured, semi-structured and unstructured data, from various sources.

Analytics plays a crucial role in extracting insights from the data lake, but the relationship is more complex than with a data warehouse:

- **Data exploration:** Analytics tools are used to explore and discover patterns, relationships and insights from the raw data in a data lake.

- **Data preparation:** Analytics tools help prepare the data for analysis by cleaning, transforming and structuring the data.

- **Data analysis:** Analytics tools, such as machine learning, natural language processing and data visualization, are used to analyze the data and extract insights.

- **Data refining:** Analytics helps refine the data by identifying data quality issues, handling missing values and creating data pipelines.

A data lake offers a flexible and scalable platform for analytics, allowing the analysis of large amounts of data and the discovery of new insights.

3. Lakehouse

A lakehouse is a relatively new concept that combines the benefits of a data warehouse and a data lake. It's a centralized repository that stores data in a structured and organized manner while also storing raw, unprocessed data. A lakehouse provides a flexible and scalable platform for analytics, enabling both batch and real-time data processing.

A lakehouse is designed to support a wide range of analytics use cases, including:

- **Real-time analytics:** A lakehouse enables real-time data processing and analysis, allowing for immediate insights and decision-making.

- **Batch analytics:** A lakehouse supports batch processing and analysis, enabling the analysis of large datasets and the creation of complex models.

- **Data science:** A lakehouse provides a platform for data scientists to explore, analyze and model data, using a variety of tools and techniques.

- **Business intelligence:** A lakehouse supports business intelligence and reporting, enabling the creation of dashboards, reports and visualizations.

A lakehouse provides a flexible and scalable platform for analytics, allowing organizations to adapt to changing business needs and data volumes.

4. Data mesh

A data mesh is a decentralized data architecture that treats data as a product, with each domain or department responsible for its data. A data mesh is designed to provide a scalable, flexible and self-service platform for data management and analytics.

A data mesh is designed to support a wide range of analytics use cases, including:

- **Domain-specific analytics:** Each domain or department is responsible for its analytics, enabling domain-specific insights and decision-making.

- **Self-service analytics:** A data mesh provides a self-service platform for analytics, allowing users to access and analyze data without relying on IT or centralized teams.

- **Data democratization:** A data mesh enables data democratization, providing access to data and analytics tools for a wide range of users, regardless of technical expertise

- **Federated analytics:** A data mesh supports federated analytics, enabling the analysis of data across multiple domains or departments while maintaining data ownership and governance.

A data mesh provides a flexible and scalable platform for analytics, enabling organizations to adapt to changing business needs and data volumes while maintaining data governance and ownership.

MIGRATING AWAY FROM TROUBLED LEGACIES

Advancing analytics for the modern AI era often means migrating from legacy systems that restrict data architectures. Organizations using legacy Spark clusters or traditional data warehouses often find themselves grappling with outdated systems that hinder their ability to extract meaningful value from their data.

These legacy data solutions frequently demand intricate management and time-intensive upkeep. Tasks such as manually scaling clusters, allocating resources and fine-tuning performance necessitate specialized expertise and impose a significant administrative burden — translating to rapidly escalating costs.

These systems often struggle to adapt to fluctuating workload patterns, leading to a costly dilemma: either over-provisioning resources, resulting in unnecessary expenses, or under-provisioning, causing performance bottlenecks. Cost management becomes a major hurdle as organizations are forced to maintain expensive infrastructure even during periods of low utilization, compounded by the hidden costs of maintenance, upgrades and specialized personnel.

The challenges don't end there. Data silos and integration complexities present another significant obstacle. Legacy systems typically lack seamless connectivity with modern data sources and tools, impeding the implementation of a unified data strategy.

Performance issues are commonplace, particularly when processing large datasets or handling concurrent queries, leading to prolonged time-to-insight and diminished business agility. Adding to these woes are the security and compliance requirements, which can be intricate for older systems not designed to meet contemporary standards for data protection and governance.

Furthermore, these legacy systems often lack support for modern data formats and processing demands, limiting an organization's ability to leverage cutting-edge technologies like machine learning and real-time analytics. The inability to easily scale resources up or down based on demand can result in either performance bottlenecks during peak times or wasteful resource allocation during low-usage periods, negatively impacting both operational efficiency and cost-effectiveness.

These challenges only intensify as data volumes and variety expand, making it increasingly difficult for organizations to maintain a competitive edge in today's data-driven business landscape, where AI capabilities are essential for staying ahead.

THE EVOLUTION OF DATA WAREHOUSES

The relationship between analytics and an underlying architecture, like data warehouses, is fundamental to understanding the evolution of analytics. A data warehouse traditionally served as the central repository where an organization stored its historical and current data from various sources.

This structured environment provided the foundation for analytical processing by ensuring the data was clean, consistent and readily available for analysis. Without a robust data warehouse, organizations struggled to perform meaningful analytics because they lacked the organized, integrated data necessary for comprehensive analysis.

Before data warehouses, companies relied on expensive and complex on-premises infrastructure for data storage and processing. These systems struggled with the growing volume of data and could not efficiently collect, store, process and analyze it. This limitation prevented businesses from gaining quick, valuable insights needed to stay competitive.

Organizations then shifted to building data warehouses to store and centralize large amounts of data generated across their organization. The goal was to perform fast analytics and reporting on this data so analysts and decision-makers could make informed decisions based on this data. They wanted to do this while controlling cost and maintenance.

However, some organizations didn't get the value they wanted out of their data warehouse. They wasted valuable time and money having to stitch together disparate services — including those for AI, BI, apps and collaboration. They were forced to create an analytics solution that often also suffered from disruptive downtime needed for cluster spin-up and spin-down, maintenance and upgrades, or code refactoring.

These solutions often failed to scale with the organizations as their volume of data increased. As organizations grow, so does the volume of data they collect and use; they must accelerate unlocking as much business value as possible by constantly evaluating, procuring and implementing the newest AI technologies to stay competitive.

Other solutions in the market keep up charging customers through a combination of increasing costs for newer and faster engines or demanding increasingly more tuning/manual optimizations to improve performance.

WHY MIGRATE TO THE CLOUD

Organizations relying on legacy Spark clusters or traditional data warehouses often find themselves seeking modern solutions to overcome inherent limitations. While the reasons for this shift — complexity, cost inefficiencies, performance bottlenecks and stifled innovation — are significant, the practical steps to modernize are equally crucial. For those looking to transition, the Snowflake AI Data Cloud offers a compelling path forward.

Migrating to Snowflake begins with embracing its cloud-native architecture, which immediately alleviates the burden of intricate infrastructure management and ongoing maintenance that plague on-premises solutions. This fundamental shift allows data teams to redirect their focus from system administration to core data analytics initiatives.

A key advantage of Snowflake lies in its disaggregated storage and compute model. This unique design enables independent scaling of resources, translating directly into cost savings as users pay only for the compute they actively consume. This contrasts sharply with the often fixed and potentially underutilized infrastructure costs associated with maintaining legacy Spark clusters or traditional data warehouses.

Snowflake streamlines data accessibility and collaboration through its built-in data sharing capabilities. Organizations can effortlessly share data, both internally and externally, without the complexities of creating multiple copies or managing cumbersome ETL pipelines. This fosters a more integrated and collaborative data environment.

Optimizing data storage and performance becomes significantly simpler with Snowflake. The platform's automatic features, including clustering and data compression, minimize the need for manual performance tuning — a common and time-consuming task in Spark environments. Furthermore, Snowflake's robust, built-in security features and comprehensive compliance capabilities address critical concerns in today's regulatory landscape.

Snowflake's versatility extends to handling diverse data formats. It natively supports semi-structured data like JSON, XML and Parquet alongside traditional structured data, eliminating the need for supplementary tools or intricate transformations. This simplifies data ingestion and analysis processes.

The platform's ability to instantly scale to accommodate fluctuating workloads optimizes performance without the need for manual intervention. Its support for multiple cloud providers offers deployment flexibility, while its standardized ANSI SQL interface minimizes the learning curve for existing SQL users. This combination of features makes Snowflake an attractive option for organizations aiming to modernize their data operations, reduce operational overhead and eliminate technical debt.

Finally, the Snowflake AI Data Cloud can support any data architecture — whether it be a data warehouse, data lake, data lakehouse or even data mesh. This flexibility is crucial for organizations to adapt to evolving data strategies and ensure their infrastructure can meet current and future analytical demands without costly re-platforming.

Facilitating this transition are Snowflake's AI-powered migration solutions. Solutions like SnowConvert AI, boasting over 96% accuracy in code conversion[1], and the Snowpark Migration Accelerator can significantly streamline the migration process. Developed by Snowflake experts, these solutions, coupled with their extensive partner network, reduce the risks of errors, reduce migration costs and accelerate timelines, allowing organizations to realize the benefits of a modern data platform sooner. Whether pursuing a full migration or a phased approach, these tools empower organizations to move forward at their own pace. Ultimately, migrating to Snowflake provides a pathway to unlock greater business value from existing data through advanced AI and ML capabilities, all within a robust and flexible cloud-native architecture designed to break down data silos and foster future innovation.

THE POWER OF
THE CLOUD DATA AND AI PLATFORM

A cloud data and AI platform is a suite of cloud-based services and tools that enable organizations to collect, store, process and analyze large amounts of data in a scalable, flexible and secure manner.

It provides a centralized platform for managing and integrating various data sources, applications and analytics tools, allowing businesses to gain insights, make data-driven decisions and drive innovation with analytics.

The relationship between analytics and a data platform is fundamental to modern data-driven organizations because the platform provides the foundation and capabilities needed to support effective analytics, while analytics drives the requirements and evolution of the platform.

A cloud data and AI platform typically includes a range of features and capabilities, such as:

- **Data ingestion:** Collecting and integrating data from various sources, including on-premises systems, cloud applications and IoT devices.

- **Data storage:** Storing and managing large amounts of structured and unstructured data in a scalable and secure manner.

- **Data processing:** Processing and transforming data in real time or batch mode, using technologies such as Apache Spark, Hadoop or SQL.

- **Data analytics:** Providing tools and services for data analysis, such as data warehousing, business intelligence and machine learning.

- **Data security:** Offering data security, governance and compliance capabilities, including encryption, access controls and data masking.

- **Data integration:** Integrating data from multiple sources, including on-premises systems, cloud applications and third-party services.

- **Data visualization:** Providing tools and services for data visualization, such as reporting, dashboards and data storytelling.

The advantages of a cloud data and AI platform for analytics

A cloud data and AI platform uses these features and capabilities to offer several key advantages for organizations. First, the platform provides scalable infrastructure and resources, enabling businesses to efficiently manage large amounts of data and handle variable workloads. Users also benefit from the flexibility to select from various data services and tools while maintaining seamless integration with their existing systems and applications.

From a financial perspective, cloud data and AI platforms are cost-effective, significantly reducing expenses compared to maintaining on-premises infrastructure. The platform also accelerates time-to-insight and decision-making processes through real-time data processing and analytics capabilities.

They also deliver improved collaboration and data sharing across teams and departments, breaking down organizational silos and promoting more efficient workflows. These are just some of the reasons organizations today are rapidly adopting cloud-based data platforms, which have evolved from cloud data warehouses, to gain scalable, flexible and cost-effective alternatives to legacy systems.

The transition from legacy systems to cloud solutions is critical for increasing revenue and operational efficiency. Companies can lose up to 30% in revenue annually due to inefficiencies from incorrect or siloed data, according to IDC Market Research. Traditional on-premises data warehouses, with their significant upfront investments and ongoing maintenance, often exacerbate these issues.

Understanding the role of cloud data and AI platforms

Data warehouses traditionally stored structured data from transactional systems, requiring extensive ETL (extract, transform, load) processes. Data lakes, on the other hand, accommodate a variety of data types, including structured, semi-structured and unstructured data, in their raw form.

Cloud data and AI platforms merge these concepts, offering a unified environment that integrates both data platforms and data lake capabilities. Modern cloud data platforms decouple storage and compute, allowing businesses to scale resources independently based on demand. This elasticity helps ensure efficient workload processing and cost optimization.

KEY TRENDS IN DATA MANAGEMENT

Several key trends are impacting data management, both today and in the future.

Trend	Impact
The rise of big data has overwhelmed traditional data warehouses.	Cloud data and AI platforms decouple storage and compute, scaling resources independently and handling diverse data types.
The increasing adoption of AI and ML requires seamless integration with data platforms.	Cloud data warehouses provide the necessary computing power, scalability and integration for ML workloads, enabling automated analytics.
The demand for real-time analytics is driven by the need for quick insights.	Cloud data and AI platforms enable faster data processing, facilitating near-instant analytics.
Unified data platforms support multiple workloads beyond traditional analytics.	Modern cloud warehouses integrate with data lakes, enabling storage of raw and processed data in a single environment.
Open table formats, such as Apache Iceberg, ensure businesses are not locked into proprietary systems.	These formats provide better performance, reliability and flexibility.
Data mesh architectures decentralize data management, giving teams ownership over their data while maintaining governance.	This enhances collaboration and maintains data usefulness across departments.

Architecting a powerful and secure cloud data and AI platform

A cloud data and AI platform is crucial for unlocking the full potential of data. Beyond just storage, a well-designed platform must be easy to use, connected across your data estate and trusted. Let's look at five principals of a robust AI and data platform to power analytics.

1. **Prioritize a fully managed, serverless architecture for ease of use and efficiency:** A robust AI and data platform should minimize operational overhead for users. This means embracing a fully managed, serverless model where the vendor handles infrastructure provisioning, deployment, maintenance and upgrades. This approach allows organizations to focus on data and AI innovation rather than platform management, leading to faster time-to-value, reduced total cost of ownership (TCO), and built-in cost optimization features. Automatic performance improvements and seamless, non-disruptive upgrades are key aspects of this principle.

2. **Foster seamless data connectivity and collaboration:** An effective platform must break down data silos and enable effortless internal and external data sharing. This principle emphasizes zero-ETL (extract, transform, load) sharing, allowing secure, privacy-preserving access to data across regions and clouds without physical movement. It also includes capabilities for secure collaboration through unified search, internal marketplaces, and external data exchanges, promoting a rich data ecosystem and maximizing insights from diverse data sources. Interoperability with open data formats is crucial to avoid vendor lock-in and support diverse architectural patterns.

3. **Embed comprehensive security, governance and business continuity by design:** Trust is paramount for an AI and data platform. This principle dictates that enterprise-grade security, robust data governance, and strong business continuity/disaster recovery capabilities must be integrated natively and operate uniformly across clouds and regions. This includes features like centralized access controls, automated data protection, and seamless replication for resiliency and failover, ensuring data integrity, compliance, and uninterrupted operations, especially for sensitive AI workloads.

4. **Enable built-in AI capabilities and an open, flexible development environment:** A robust platform should not only store and manage data but also bring AI directly to the data. This principle emphasizes integrating AI capabilities (like generative AI models and machine learning tools) directly into the platform, allowing users to leverage them without moving data. Furthermore, the platform should offer a flexible development environment with support for multiple coding interfaces (e.g., SQL, Python) and no-code options, empowering a wide range of users to build and deploy AI solutions efficiently.

5. **Provide granular cost visibility, control and optimization:** With increasing data volumes and AI workloads, cost management becomes critical. A robust platform must offer transparent cost visibility, enabling organizations to track and understand their spending with detailed drilldowns. Beyond visibility, it should provide intuitive cost control mechanisms, such as resource monitors and budgeting tools, to set limits, receive notifications, and automatically optimize resource utilization, ensuring economic efficiency and maximizing business value.

LEVERAGING CLOUD DATA AND AI PLATFORMS FOR ADVANCED ANALYTICS AND COLLABORATION

Cloud data and AI platforms enable quick analysis and collaboration on vast amounts of data using innovative technologies such as automation, AI and ML.

Enhancing data processing with AI, ML and geospatial analytics

Advanced analytics are transforming how businesses extract value from data. Geospatial analytics is becoming essential for industries that rely on location-based data. AI and ML integration drive automation across industries, enabling tasks like fraud detection, demand forecasting and predictive maintenance. Cloud data platforms must provide seamless integration with ML frameworks such as TensorFlow and PyTorch. Automation is also transforming the development and deployment of analytics applications, enabling real-time data processing.

Enabling seamless data sharing without duplication

Traditional methods of data sharing involved data movement or replication, which caused data redundancy and potential inconsistency — not to mention security risks. Cloud-native data-sharing solutions eliminate those issues by enabling real-time, governed access to shared data. Sharing data directly from its source without replication provides all users access to the most up-to-date, accurate version.

Data marketplaces and monetization: Turning data into revenue

Leading companies are monetizing their data in a way that safeguards sensitive data and complies with privacy laws. Companies can package and sell valuable datasets through data marketplaces, providing insights to external organizations and generating new revenue streams. For example, retail businesses can sell anonymized purchasing data to advertisers, while financial institutions can share economic indicators with investors. A well-designed cloud data-sharing architecture facilitates cross-cloud and multi-region collaboration, enabling seamless integration regardless of where the data is stored.

Automating development, deployment and data-driven insights

Automation is transforming the development and deployment of analytics applications. Organizations are increasingly using automated workflows to ingest, process and analyze data in real time. Features like automated query optimization, AI-driven insights, and no-code or low-code tools enable business users to leverage analytics without requiring deep technical expertise. Automation streamlines operations, reduces manual effort and accelerates the delivery of data-driven insights.

TRANSITIONING TO THE CLOUD AND SCALING
FOR THE FUTURE

A structured approach allows a smooth and successful implementation for organizations looking to transition to cloud data warehousing. Rather than jumping in without a clear plan, businesses should follow four key steps to evaluate needs, decide whether to migrate or start fresh, calculate the total cost of ownership (TCO), and set up a proof of concept (POC).

1. Evaluating business needs and building a cloud strategy

Before selecting a cloud data platform, businesses must assess their current and future data needs. This involves a comprehensive evaluation of various factors to ensure the chosen solution aligns with organizational objectives. Factors such as data volume, the variety of data sources, query performance requirements and compliance obligations should all be considered. Understanding the scale and complexity of the data is essential for choosing a suitable platform. Organizations should also identify key stakeholders, including data engineers, analysts and business users, to ensure the platform meets their diverse needs.

2. Migration vs. starting fresh: Choosing the right approach

Businesses must decide whether to migrate their existing data warehouse to the cloud or build a new system from scratch.

Migration is ideal for companies that have a structured data warehouse with valuable historical data. Migrating leverages existing investments and preserves historical context, but it requires careful planning to ensure compatibility with the cloud environment. This includes addressing data format differences, schema changes and potential performance bottlenecks.

Starting fresh allows organizations to design a modern data architecture without legacy constraints. This approach offers the opportunity to implement best practices and optimize the platform for current needs but may require additional time to implement. Building a new system from scratch also involves data migration from legacy systems, but with a focus on clean and streamlined data transfer.

Both approaches have their merits, and the best choice depends on the organization's specific circumstances, existing infrastructure and long-term goals.

3. Calculating TCO and budgeting for the future

Understanding the costs associated with cloud data warehousing is essential for budgeting and long-term planning. TCO includes not only storage and compute costs but also operational expenses, data transfer fees and licensing costs. Businesses should carefully analyze each component to get a comprehensive view of the financial commitment. Consumption-based pricing models — a key advantage of cloud platforms — allow businesses to optimize costs by scaling resources up or down based on demand. However, it's important to monitor usage and set budgets to avoid unexpected expenses. Calculating TCO also involves considering the long-term benefits, such as reduced maintenance costs and increased efficiency.

4. Running a POC and ensuring a smooth implementation

A proof of concept is a critical step in evaluating the performance and feasibility of a cloud data platform. Organizations should start with a small dataset and test core functionalities such as query performance, security controls and integration with existing tools. This allows them to see the platform in action and identify potential issues early on. Running a POC helps businesses identify potential challenges and refine their deployment strategy before fully committing to the cloud. It also provides an opportunity to assess the platform's usability and gather feedback from key stakeholders. A successful POC validates the chosen solution and builds confidence in the transition process. Ensuring a smooth implementation also involves careful planning for data migration, user training and ongoing support.

FASTER ANALYTICS AND A MORE INTELLIGENT FUTURE WITH THE SNOWFLAKE AI DATA CLOUD

Organizations today must be AI-ready, needing a data-leading analytics engine that handles diverse data types and architectures while offering world-class price and performance. All too often, however, organizations waste valuable time and money piecing together disparate services. This results in analytics solutions that frequently suffer from disruptive downtime, extensive maintenance and upgrades, or complex code refactoring. As businesses grow, so does their data volume, demanding agility to evaluate, procure, and implement the newest AI technologies to stay competitive.

This is precisely why Snowflake offers a fully-managed, AI-ready platform. It reallocates valuable time and resources, allowing you to launch data products faster with its built-in visualizations, application development tools, and seamless sharing capabilities. Snowflake integrates out-of-the-box AI into your everyday analytics workflows through LLM-powered assistants, data agents, text-to-SQL features, and machine learning, empowering intuitive experiences for faster insights. Furthermore, it continuously reduces compute costs with a leading engine that delivers world-class price/performance and continuously improves speed through AI-powered optimizations. Simply put, Snowflake makes data analytics easy, intelligent, and fast.

AI-powered BI: Snowflake's AI-powered BI empowers users to effortlessly query text, images and audio files alongside structured data for deep intelligence using intuitive AI-powered SQL. It also provides intelligent insights by empowering smarter decisions through intuitive, conversational data exploration and enables users to ask questions in natural language, effortlessly following up for deeper insights. It is fast, scaling batch inference with industry-leading LLMs for forecasting, audio transcription, sentiment analysis, anomaly detection, and more, with zero downtime and reduced data extracts.

Lakehouse analytics: With Snowflake's lakehouse analytics, users can readily unlock the power of their open and connected data ecosystem and access data and powerful insights with Snowflake's integrated platform and engine. It also provides intelligent data management, allowing users to govern and discover data across their entire architecture with a connected catalog experience and implement precise security with table-, row-, and column-level access controls for compliance. It's also fast, operating 2.4x[2] faster for Externally Managed Iceberg Tables and 2.1x[3] faster for Snowflake-Managed Iceberg Tables to extending Snowflake's leading analytics engine to open table formats.

Applied analytics: Snowflake's applied analytics offers an integrated platform for advanced analytics on any architecture, enabling analysts to gain richer insights with reduced data silos and access to a robust ecosystem of apps and data. It also provides intelligent insights, offering pre-built models and LLM functions that can be applied to both structured and unstructured data, including text or images associated with locations or time series data. It's also fast, providing a scalable and performant data layer through its multitenant architecture and enabling the building f data products without bottlenecks.

Interactive analytics: Organizations can scale with ease and security using Snowflake's interactive analytics, reducing their operational burden with a fully managed service. It also provides intelligent insights, enabling users to talk to their data and achieve high performance. It's also intelligent, enabling organizations to separate compute and storage, auto-scaling to support apps at scale, and the ability to run multiple compute clusters simultaneously on the same datasets without constraint. Organizations can also use AI-assisted app development using Copilot, AISQL, Agent API, Document

AI, and even third-party solutions and coprocessors from Snowflake Marketplace. It's also fast, enabling organizations to integrate data and apps in client environments across clouds and globally and commercialize their products directly on Snowflake to close deals faster and grow their business.

Migrations: Modernizing on the Snowflake AI Data Cloud is easy, connecting your data ecosystem on a trusted platform with automated cluster management, continuous performance improvements, zero-ETL collaboration, and built-in security and governance. Snowflake is a proven leader in data ecosystem migrations. We help you save costs on your migration with free, cutting-edge, AI-powered solutions, dedicated migration experts, and a robust network of highly trained system integrator partners. By migrating to Snowflake, you'll drive innovation and unlock business value from your existing data using AI/ML on our cloud-native platform that connects data silos. Plus, Snowflake's open standards support provides the flexibility to choose the best price/performance engine for each workload, unifying data, AI, apps and models for a seamless analytics experience.

Snowflake is easy

Snowflake is a fully managed, AI-ready platform that is user-friendly with a simple setup process, requiring minimal maintenance and no complex infrastructure management since it's fully cloud-based. The Snowflake AI Data Cloud reallocates valuable time and resources to help your organization launch data products faster with a fully managed, AI-ready platform that has built-in visualizations, application development tools and sharing.

Snowflake is intelligent

Snowflake is intelligent — it's the platform's smart features like automatic optimization, data sharing capabilities and built-in analytics tools that help you make better data-driven decisions.

Snowflake offers integrated LLM-powered agents, workflows and SQL functions. The AI Data Cloud enables you to achieve intuitive AI experiences for faster insights. Snowflake integrates out-of-the-box AI into your everyday analytics workflows through LLM-powered assistants, data agents, text-to-SQL and ML to empower intuitive experiences for faster insights.

Snowflake is fast

Snowflake enables high-performance computing capabilities right out of the box. The AI Data Cloud enables your organization to enjoy quick query processing and concurrent access, enabled by the platform's unique multi-cluster architecture that separates compute from storage, supporting rapid data access and analysis regardless of data volume or user load.
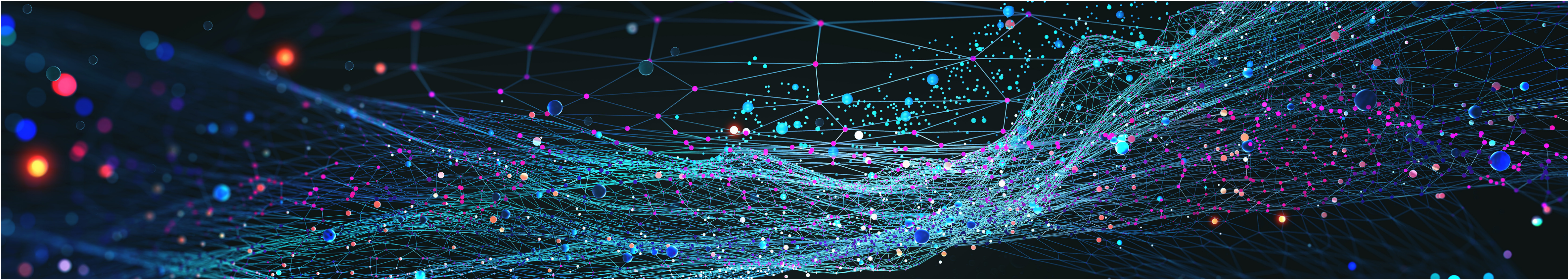
CUSTOMER STORY:

MARRIOTT SIMPLIFIES ITS DATA PLATFORM AND ACHIEVES LOWER TCO

It's hard to believe Marriott International started as a root beer stand in 1927. Today, the company comprises 32 global brands across 139 countries, with 8,300+ hotels offering 15 million hotel rooms, and 100,000 home and villa properties. As Marriott's business has grown over the past century, its data infrastructure has become more complex. In 2019, the company embarked on a mission to modernize and simplify its data platform.

Prior to 2019, Marriott was an early adopter of Netezza and Hadoop, leveraging the IBM BigInsights platform. Its technology team then moved Marriott Data Platform to AWS, which integrated batch and streaming data into cloud object storage, and processed data with a variety of technologies including Spark, Terraform and Snowflake as a data warehouse for users to consume data through Tableau. Many of those technologies made the stack complex, costly due to expensive upgrades, and difficult to operate. Data engineers spent 20% of their time on infrastructure issues such as tuning Spark jobs.

Now, the team is on an ongoing mission to use Snowflake's data platform to simplify the complexity of its tech stack. New functionality arrives automatically, eliminating the need for testing or infrastructure fixes. Snowflake simplifies data ingestion by consolidating batch and streaming, increasing Marriott's speed to market — as soon as a customer transaction occurs, the data is available for consumption. With Snowflake's Kafka connector, the technology team can ingest tokenized data as JSON into tables as VARIANT. Then, they parse and flatten that data using Snowpark.

Simplifying with Snowflake also allows the team to better support different people in the company at a lower cost. Business users have better access to data with less IT oversight needed. Analysts and data scientists can query data directly with their preferred language and tools. And third-party data is easily accessible from Snowflake Marketplace.

Today, simplifying its data platform on Snowflake has enabled Marriott to achieve transparency and control of its data, faster speed to market, improved collaboration and data sharing, a better user experience, and lower TCO. With Snowflake, Marriott has seen a dramatic improvement in performance and cost savings in comparison to Spark and Hive-based workloads. Many users from Marriott have commented on their improved experience with Snowflake, mentioning queries that used to take five hours or time out on Netezza now take one hour on Snowflake. Data that previously took 48 hours to one week in Hadoop is now available near-instantly in Snowflake.

Read about Marriott's full journey to the AI Data Cloud here.

CUSTOMER STORY:

ADVISOR360 BUILT AN AUTOMATED CUSTOMER SATISFACTION PIPELINE TO GAUGE CUSTOMER SATISFACTION IN JUST A MATTER OF DAYS

For Advisor360°, customer feedback is as essential as it is constant. Supporting financial advisors and wealth managers in nearly all aspects of their jobs, the company's integrated and configurable SaaS platform informs business decisions, improves client service and even drives innovation for firms. But getting a handle on all the emails, calls and support tickets had historically been a tedious and largely manual process.

To address that, the Advisor360° analytics and insights team built a sentiment model from scratch, using highly specialized, Python-heavy code that would extract data and push it out to a file, then incorporate it into a dashboard. But, of course, the model required constant maintenance and updating. When senior executives asked for insights using the most up-to-date data, they'd often have to wait the better part of a day to get answers. After migrating all of its historical feedback data to Snowflake, however, Advisor360° created an automated pipeline using Cortex AI to cover that end-to-end process

of gauging customer sentiment. Advisor360°'s Director of Analytics and Advanced Insights, Mark Coleman, hoped the team could establish the pipeline within a matter of weeks.

They had it finished by day 2. "Our customer sentiment pipeline is working extremely well," Coleman says. "It's integrated directly into Snowflake so we can call it as part of any of our data flows, which makes life really easy. Cortex is doing a great job for us."

Now, processing feedback occurs automatically every morning, and Coleman estimates that it saves at least a day's worth of work per month for a senior data scientist. What's more, the company found that using Cortex AI's built-in sentiment analysis functionality was about 1/25th the cost of the alternative, which would require using specialized virtual machines and calling in a much larger and pricier LLM to do essentially the same thing.

"When our data scientists tried to create this pipeline using native Azure tools, it probably took a solid month to figure out how to get permissions right and get through our firewall. We had to get the plumbing to work before we turned the water on," Coleman says. "But Cortex AI worked out of the box, integrating into our system seamlessly and translating into huge productivity gains for the team."

Now Advisor360° has instant access to the most up-to-date customer insights — allowing the firm to provide the enterprise-class customer care it is known for.

Read about Advisor360's full journey to the AI Data Cloud here.

CUSTOMER STORY:

AT&T PROVIDES FASTER INSIGHTS WHILE LOWERING ESTIMATED ANNUAL COSTS BY 84%

In today's connected world, you can't accomplish much without cell service or Wi-Fi. That's where AT&T comes in. Founded in 1885, this telecom giant has evolved throughout the decades to continually deliver on its vision to connect people to their world through the best technology and service.

And its efforts paid off. AT&T now provides smart solutions and mobile and broadband services to more than 100 million U.S. consumers and almost every Fortune 1000 company. To meet high demand and deliver a stellar experience across this sprawling customer base, AT&T needs a powerful, easy-to-use data management system that efficiently processes hundreds of petabytes of data every day. But its complex on-premises systems, including Hadoop, were slowing down business — and increasing costs. With the Snowflake Telecom Data Cloud, AT&T now derives more value from its data, boosting performance while lowering costs.

A golden source of truth for greater efficiency and better decisions

Hundreds of petabytes of data course through AT&T's networks every day. But this massive volume of information is only valuable if users can quickly access and act on it. "Prior to Snowflake, we had a very complex data environment on-premises," says Chief Data Officer Andy Markus. "That led to a more ineffective operating environment for our business partners, both from a speed and cost perspective."

Using Hadoop for storage and processing, AT&T's monolithic on-premises data warehouse hampered the team from collecting, storing, sharing and processing its vast stores of data. By moving to the Snowflake Telecom Data Cloud, Markus and his team achieved their goal of democratizing data across the business.

Achieving higher performance at lower cost

"Data democratization for AT&T is really important because we have so much data," says Markus. Delivering on this goal means the data team must keep up with business teams' surging demand for simple yet powerful business intelligence tools.

With Snowflake's separation of storage and compute, AT&T deploys new features or products for different business units without affecting each other, enabling faster and more agile responses to user needs. While launching new solutions for internal teams previously took weeks or months, now it takes just days.

AT&T Connects the world in the Data Cloud

In addition to releasing new products, the data team also evaluates how to optimize existing tools. Cue myRESULTS, a highly tuned, interactive reporting application that AT&T built on its previous solution 13 years ago. What started as a single dashboard tracking around 10 retail metrics has since transformed into dozens of screens that thousands of internal users rely on daily to track metrics across the business.

But as usage surged, so did associated costs, complexity and concurrency challenges — causing AT&T to turn to Snowflake. AT&T was initially skeptical that Snowflake could match the performance of its original application. However, during the migration process, the data team used stored procedures for API calls from their Java-based frontend. This allowed them to "pin" stored procedures to specific virtual warehouses, matching individual procedure workloads to corresponding warehouses — providing an advantage over the previous platform's macro API calls.

By migrating its reporting application to Snowflake, AT&T now has a cost-effective alternative with faster performance, which has benefited the company's bottom line. "Because of results caching and our Snowflake migration, we're now running at one-sixth of the cost that we thought we'd be operating at," says Chris Goldston, Lead Solution Architect at AT&T. "With the additional optimizations we've made in the last several months, we have shaved another 35% off of that lower number."

Faster reporting for better decision-making across teams

The newly transformed myRESULTS, built on Snowflake, is a self-service analytics application that unifies data from headquarters, call centers, retail locations, vendor partners and executive management. Users, which have grown to about 115,000, now access more than 80 custom screens to see real-time sales reporting, historical data trends, order lifecycle data and customer surveys. The app features a precalculated library of more than 2,700 metrics and customizable data sets, as well as more than 230 separate APIs.

Every day, about 35,000 unique users make more than 1 million API calls — with 90% of data requests coming back in less than one second. These quick responses help teams everywhere make faster, more informed decisions that can ultimately drive revenue and deliver a better experience to customers.

Read about AT&T's full journey to the AI Data Cloud here.

Snowflake makes data analytics easy, intelligent and fast with its user-friendly and high-performance platform. The platform's zero-maintenance and automated performance optimizations make it easy for your organization to focus on deriving insights from your data, rather than worrying about the underlying infrastructure.

Our commitment to security and governance helps ensure that sensitive data is protected, while the AI Data Cloud's ecosystem partners and AI-readiness enable seamless integration with a wide range of tools and models.

The platform's built-in machine learning and language models, as well as its selection of application development tools, makes it easy for your organization to build data-driven products and applications.

Snowflake's data analytics and products, including conversational assistants, interactive applications, unstructured data insights and lakehouse analytics, provide businesses with a wealth of options for deriving value from their data. With Snowflake, data analytics has never been easier, more intelligent or faster.
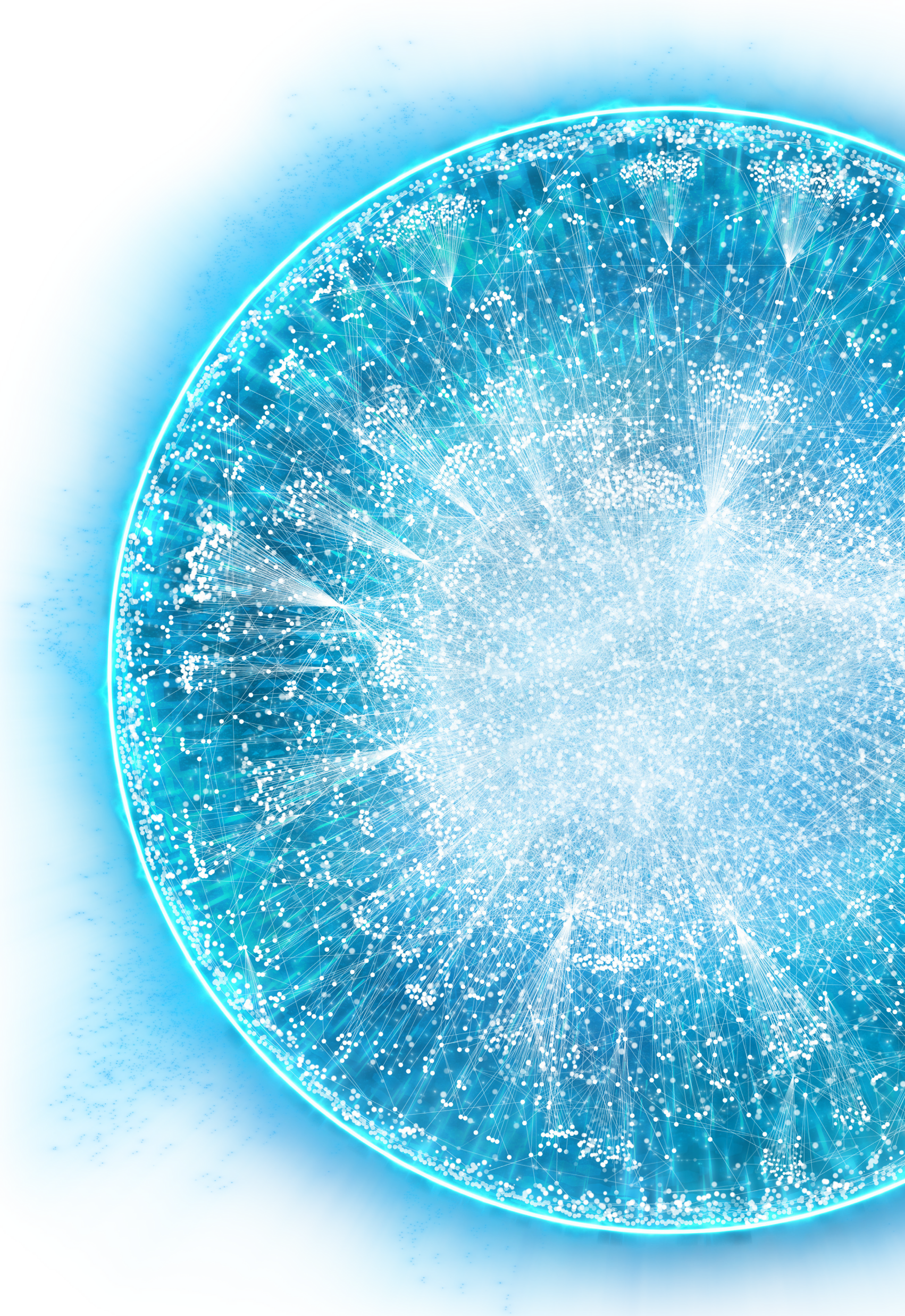
![snowflake®]

Snowflake is the platform for the AI era, making it easy for enterprises to innovate faster and get more value from data. More than 11,000 companies around the globe, including hundreds of the world's largest, use Snowflake's AI Data Cloud to build, use and share data, applications and AI. With Snowflake, data and AI are transformative for everyone.

Learn more at **snowflake.com**

(NYSE: SNOW)

1. Based on total lines of code for professional services engagements with SnowConvert for entire workloads and not individual objects, specifically for Oracle, SQL Server, and Teradata migrations. Amazon Redshift is excluded. The numbers include internal usage. Data from March 2020 - May 2025.

2. Snowflake performance improvements based on core analytics workloads on 2XL warehouse measured as of October 4, 2024 using Standard Warehouse and May 2, 2025 using Gen2.

3. Snowflake performance improvements based on core analytics workloads on 2XL warehouse measured as of May 3, 2024 using Standard Warehouse and May 2, 2025 using Gen2.