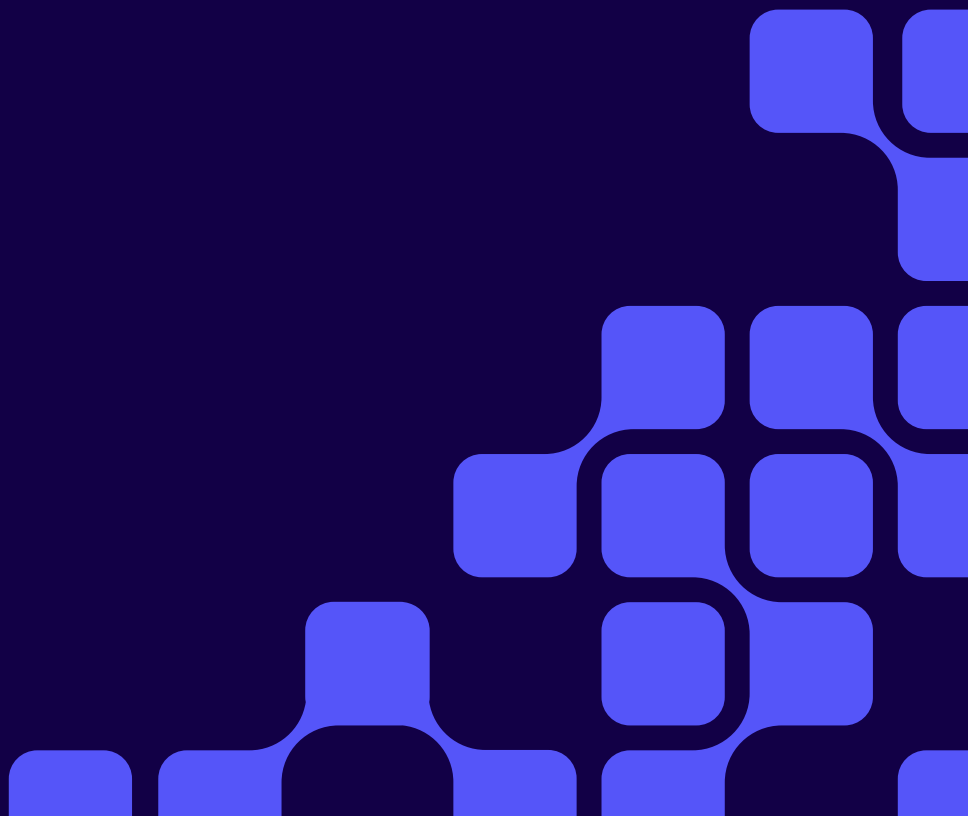
CLOUDERA | snowflake

A CLOUDERA PERSPECTIVE

Shape the future of Your Enterprise AI with Cloudera + Snowflake

Combine hybrid flexibility and advanced analytics to power tomorrow's AI innovations.



The Data Problem

AI has emerged as the transformative force shaping modern enterprises, unlocking faster insights, improving operational efficiency, and delivering exceptional customer experiences. Yet realizing the full potential of AI remains a challenge for many organizations. The root cause? Data. Or more specifically, the challenges of managing and leveraging data in today's complex IT environments.

Data is the foundation of AI. Whether contextualizing large language models (LLMs), deploying generative AI tools, or delivering predictive analytics, AI relies on vast quantities of diverse and high-quality data. But in most organizations, this data is anything but unified. Instead, it's scattered across on-premises systems, legacy infrastructure, and multiple cloud environments, creating significant roadblocks to scaling AI.

Data challenges fall into three critical areas:

1. **Data silos and fragmentation:**
   Enterprises today are grappling with fragmented architectures where critical datasets exist in isolated silos. Data resides across a mix of legacy systems, cloud platforms, and even departmental repositories. This disconnect slows down access and introduces inefficiencies in data preparation and analysis. AI models can't deliver meaningful results without a consistent, unified view of the organization's data.
2. **Hybrid complexities:**
   While many businesses have migrated workloads to the cloud, a significant proportion of applications and data remain on-premises, often in legacy environments like mainframes. These systems handle essential transactions but are notoriously difficult to integrate with cloud-based tools. This creates a complex hybrid environment where organizations must navigate between the old and the new, often at great expense.
3. **Scaling AI initiatives:**
   Generative AI applications—such as those powering conversational chatbots, content summarization, or predictive decision-making—need to consume data at scale. Without seamless access to high-quality data from all sources, enterprises face bottlenecks that slow innovation and limit the value of their AI investments. Scaling AI isn't just about technology; it's about making the data architecture fit for purpose.

These challenges leave organizations trapped in a cycle of high costs, inefficiencies, and limited AI scalability. For enterprises to truly unlock the value of AI and compete in a data-driven world, they must overcome these challenges by unifying their data landscapes and ensuring that all data—whether on-premises or in the cloud—is ready to fuel innovation.

The Cloudera + Snowflake Solution

Overcoming the challenges of siloed, fragmented, and hybrid data environments requires a transformative approach. That's where the partnership between Cloudera and Snowflake comes in. Together, these two industry leaders provide a seamless solution that empowers organizations to unify their data, streamline operations, and unlock the full potential of AI at scale.

Analysts agree: both Cloudera and Snowflake are recognized as leaders in their respective domains.

The *GigaOm Radar for Data Lakes and Lakehouses* highlights Cloudera's hybrid, open data lakehouse architecture, which excels in scalability, flexibility, and governance across on-premises and multi-cloud environments. Meanwhile, the *GigaOm Radar for Data Catalogs* praises Snowflake's robust, cloud-native analytics platform, designed for real-time insights and advanced analytics. Together, these platforms deliver a unified solution to address modern data challenges.

At the core of this partnership is Cloudera's open data lakehouse, powered by Apache Iceberg. Iceberg provides unified data access across environments, breaking down silos and enabling seamless interaction with structured, semi-structured, and unstructured data. Its multi-engine compatibility and ACID guarantees ensure scalability and reliability for enterprise workloads, as detailed in *Introducing Apache Iceberg*. Snowflake complements these capabilities with its advanced cloud-based analytics, empowering organizations to accelerate AI deployments and decision-making.

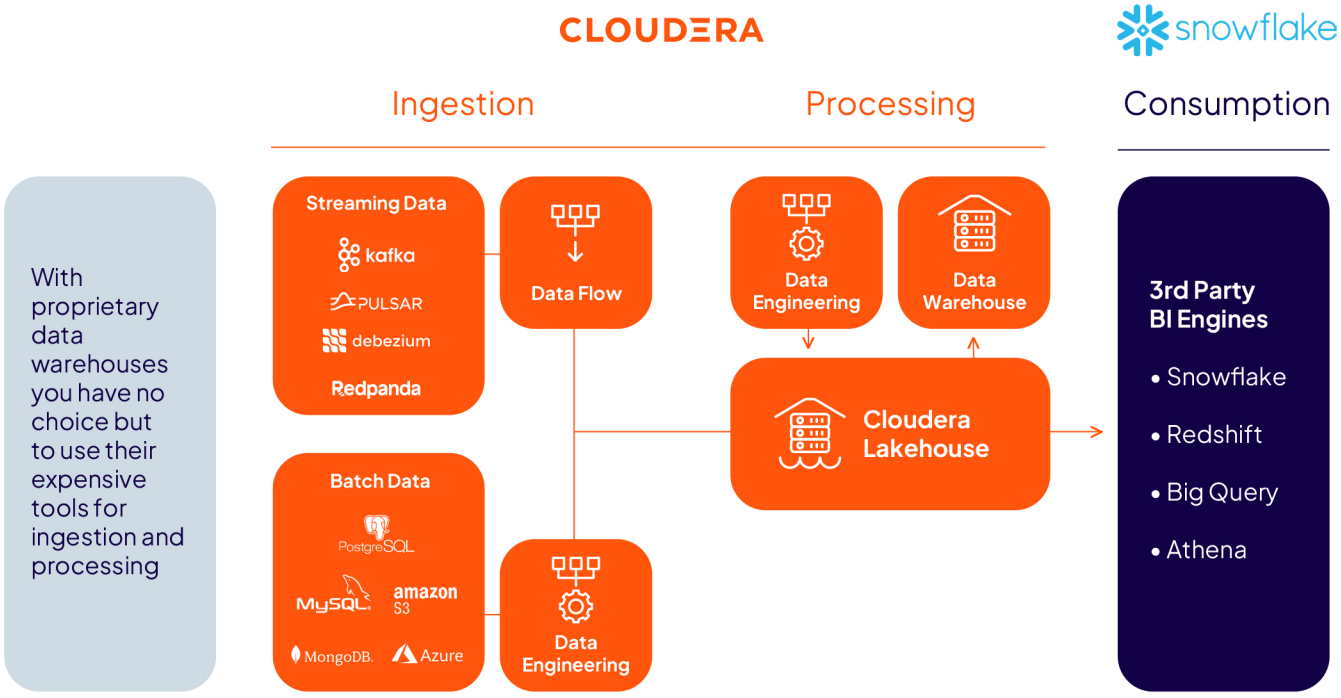
The Cloudera + Snowflake partnership offers three transformative benefits:

1. **Unified data access across environments:** Enterprises no longer need to struggle with fragmented data silos. Cloudera's open data lakehouse allows on-premises, private cloud, and public cloud data to be unified into a single source of truth. As noted in *Introducing Apache Iceberg*, this simplifies architectures, reduces costly duplication, and enables interoperability across diverse tools. Snowflake builds on this by providing real-time access to unified datasets, supporting advanced AI and machine learning applications without requiring extensive ETL processes. This combined approach is recognized in the *GigaOm Radar for Data Lakes and Lakehouses*, where Cloudera's hybrid capabilities stand out for unifying distributed data.
2. **Hybrid flexibility and scalability:** Many enterprises operate in complex hybrid environments, balancing the reliability of legacy on-premises systems with the agility of modern cloud platforms. Cloudera's hybrid control plane ensures consistent governance and management across all environments, while Snowflake's platform scales effortlessly to support even the most demanding analytics workloads. As the *State of Enterprise AI and Modern Data Architecture* report notes, 71% of enterprises cite hybrid architectures as critical for AI readiness. The Cloudera + Snowflake partnership is uniquely positioned to support enterprises in creating flexible and scalable hybrid environments.
3. **AI-ready data pipelines:** For AI to deliver actionable insights, it requires trusted, high-quality data. Apache Iceberg's advanced capabilities—such as time travel, schema evolution, and multi-engine compatibility—simplify pipeline creation and ensure AI applications have the data they need. According to the *GigaOm Radar for Data Catalogs*, Snowflake excels in delivering real-time data availability, a crucial factor in deploying generative AI and predictive models. Together, these platforms enable faster analytics, enhanced customer experiences, and operational efficiencies, as outlined in the *State of Enterprise AI and Modern Data Architecture* report.

By uniting Cloudera's hybrid flexibility and Snowflake's analytics excellence, this partnership empowers enterprises to transcend today's data challenges and embrace tomorrow's AI-driven opportunities. With a unified, scalable, and future-ready solution, organizations can innovate, adapt, and thrive in an increasingly competitive, data-driven world.

Cloudera Open Data Lakehouse

Reduce TCO: Use best-of-breed tools to ingest, prepare, and analyze data



Accelerating AI Together

AI is no longer a far-off vision of the future—it's here, transforming how businesses operate and innovate. From generative AI tools like chatbots and text summarization to predictive models driving operational efficiency, enterprises are racing to deploy AI capabilities that create real-world business value. However, scaling these initiatives requires a foundation that enables trusted, accessible, and AI-ready data.

The partnership between Cloudera and Snowflake accelerates enterprise AI by simplifying and streamlining the processes that underpin AI adoption. Together, these platforms create a seamless environment for building, training, and deploying generative AI and predictive models at scale.

Here's how:

1. **Simplifying data pipelines with Iceberg:** AI requires large volumes of high-quality data from diverse sources, but traditional architectures often rely on fragmented, complex pipelines that slow down innovation. Cloudera's adoption of Apache Iceberg changes the game. As noted in *Introducing Apache Iceberg*, Iceberg enables advanced features like time travel, schema evolution, and compaction, ensuring that data pipelines are both simplified and scalable. Organizations can ingest, process, and serve data faster, providing AI models with the trusted, up-to-date information they need to deliver accurate insights.
2. **Creating AI-ready architectures:** According to the *State of Enterprise AI and Modern Data Architecture*, 67% of enterprises are already deploying generative AI models, but many face roadblocks when attempting to scale these efforts. The combination of Cloudera's hybrid data lakehouse and Snowflake's cloud-native platform creates a modern architecture that is purpose-built for AI. Enterprises can expose data to AI models—whether small language models (SLMs) or LLMs—without complex migrations or costly duplication, ensuring seamless scalability.
3. **Accelerating time to insight:** Generative AI thrives on speed, but without fast, real-time access to data, businesses risk falling behind. Snowflake's analytics platform excels at delivering the rapid insights demanded by AI applications, as emphasized in the *GigaOm Radar for Data Catalogs*. Combined with Iceberg's ability to support multi-engine processing, enterprises gain the flexibility to train and operationalize AI models faster, driving innovation across use cases like customer experience, fraud detection, and predictive maintenance.
4. **Breaking down silos for AI-driven collaboration:** Many organizations are constrained by siloed data, limiting collaboration between departments and AI initiatives. Cloudera and Snowflake work together to eliminate these barriers, enabling data to flow freely across on-premises and cloud environments. This unified approach empowers teams to collaborate on AI projects without friction, ensuring AI models are trained on the most comprehensive datasets available.

By addressing the critical barriers to AI adoption—fragmented pipelines, scaling challenges, and slow time to insight—Cloudera and Snowflake empower enterprises to move faster and innovate further. Together, they provide the tools to operationalize AI, transforming data from a strategic asset into a competitive advantage.

Powering AI-Driven Innovation Together

AI is reshaping industries, unlocking possibilities, and driving competitive advantage. But behind every AI success story lies a foundation of unified, trusted, and scalable data. With Cloudera's hybrid open data lakehouse and Snowflake's cloud-native analytics platform, enterprises can overcome today's data challenges and build the AI-ready architectures needed to thrive in a rapidly evolving digital landscape.

Whether it's simplifying pipelines with Apache Iceberg, scaling generative AI deployments, or enabling hybrid data management, the Cloudera + Snowflake partnership empowers organizations to innovate, adapt, and lead.

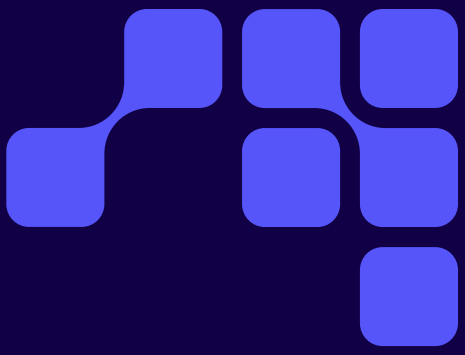
Ready to see the future of enterprise AI in action?

Request a demo today and discover how Cloudera and Snowflake can help you unlock the full potential of your data.

Offload Workloads

Cloudera + Snowflake for best-of-breed capabilities

About Cloudera



Cloudera is the only true hybrid platform for data, analytics, and AI. With 100× more data under management than other cloud-only vendors, Cloudera empowers global enterprises to transform data of all types, on any public or private cloud, into valuable, trusted insights. Our open data lakehouse delivers scalable and secure data management with portable cloud-native analytics, enabling customers to bring GenAI models to their data while maintaining privacy and ensuring responsible, reliable AI deployments. The world's largest brands in financial services, insurance, media, manufacturing, and government rely on Cloudera to be able to use their data to solve the impossible—today and in the future.

To learn more, visit Cloudera.com and follow us on LinkedIn and X. Cloudera and associated marks are trademarks or registered trademarks of Cloudera, Inc. All other company and product names may be trademarks of their respective owners.

CLOUDERA

Cloudera, Inc. | 5470 Great America Pkwy, Santa Clara, CA 95054 USA | cloudera.com