APRIL 2025

Data in Motion Acceleration With Cloudera and IBM

Stephen Catanzano, Senior Analyst

Abstract: Real-time data, analytics and AI have accelerated in importance as organizations look to bring AI-powered solutions to market at scale. Managing data in real time opens the door for new opportunities and creates competitive advantages. Cloudera and IBM have made significant moves to help enterprises process and utilize real-time data effectively to power advanced analytics and AI. Unlike single-purpose streaming platforms, Cloudera delivers a unified data platform that spans the full data lifecycle—from ingestion and streaming to analytics and AI— all with integrated governance and security. Cloudera's data, analytics, and AI platform works together with IBM watsonx AI and governance features to deliver a trusted, scalable hybrid cloud solution that businesses can use to accelerate their development and go-to-market plans, utilizing the capabilities of two industry leaders.

Data for AI Is Growing Quickly

The amount of data organizations are using for AI continues to grow rapidly. Research from Enterprise Strategy Group, now part of Omdia, showed that 64% of organizations collect data from 100 to 499 unique sources on a daily basis, many of which are streaming and managing constantly updated and refreshed data (see Figure 1).

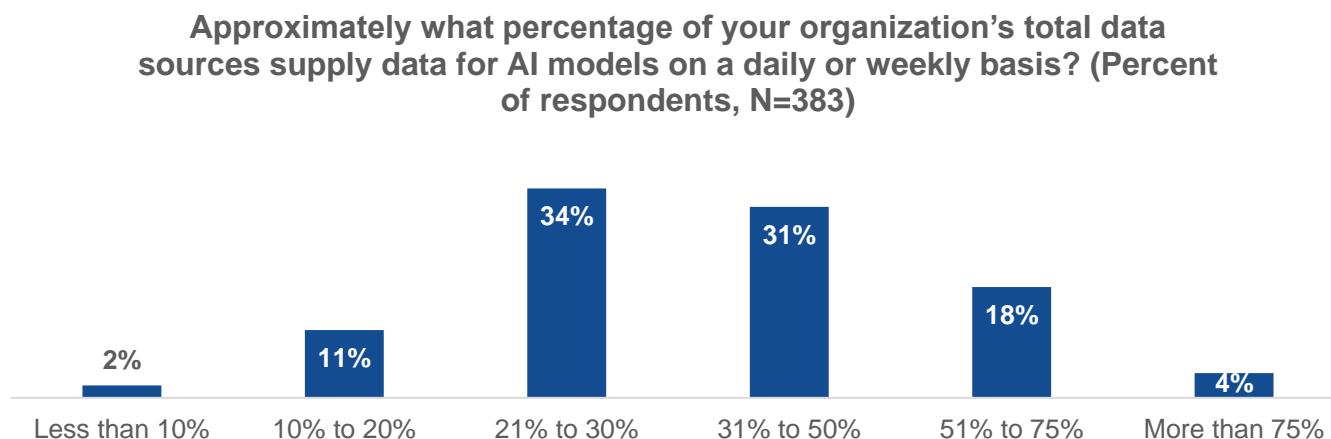**Figure 1.** Many Data Sources Being Managed by Organizations

From approximately how many different sources does your organization collect data on a daily basis? (Percent of respondents, N=385)

- Fewer than 50: 4%
- 50 to 99: 24%
- 100 to 249: 37%
- 250 to 499: 27%
- 500 or more: 8%
- Don't know: 1%
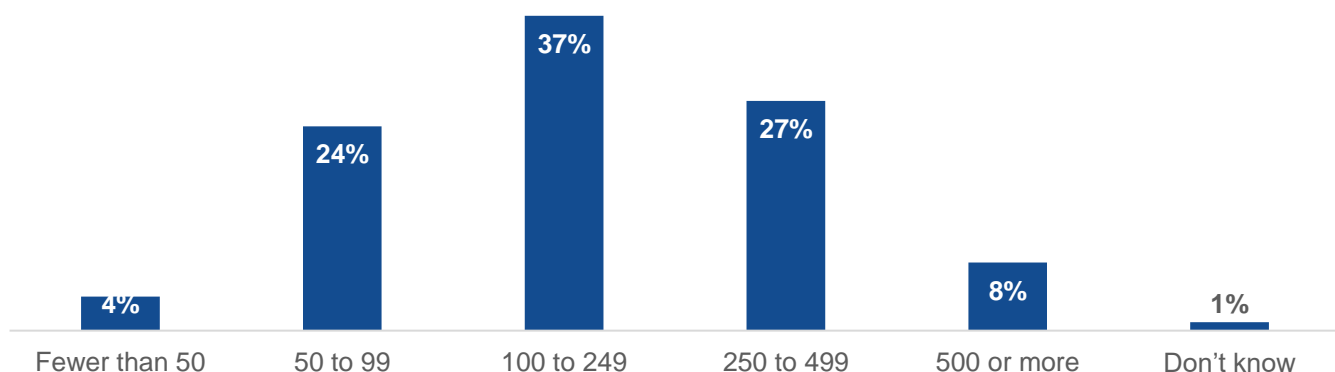
Source: Enterprise Strategy Group, now part of Omdia

AI is also driving the expanded use of data by organizations. Enterprise Strategy Group research found that 65% of organizations use 21% to 50% of their data for AI on a daily or weekly basis (see Figure 2).[1]

---

[1] Source: Enterprise Strategy Group Research Report, *Data Readiness for Impactful Generative AI*, April 2025. All Enterprise Strategy Group research references and charts in this Showcase are from this report.

Figure 2. A High Percentage of Enterprise Data Is Being Used for AI

Approximately what percentage of your organization's total data sources supply data for AI models on a daily or weekly basis? (Percent of respondents, N=383)

Less than 10%	10% to 20%	21% to 30%	31% to 50%	51% to 75%	More than 75%
2%	11%	34%	31%	18%	4%

Source: Enterprise Strategy Group, now part of Omdia

This underscores the pivotal role of a true hybrid data platform in driving business success, especially in an age where data, analytics, and AI are essential for both survival and growth. A scalable and robust data platform becomes essential to gain the full potential of data-driven strategies as organizations continue to strive for faster decision-making and enhanced employee and customer engagement.

Every Organization Needs to Manage Data in Motion

Data is no longer something that is collected and analyzed at a later time. Real-time insights are a driving factor for organizations to stay competitive, adapt quickly, and make immediate decisions. Data in motion is the data that is actively moving through networks, systems, or devices, providing up-to-the-second intelligence for smarter, faster decision-making.

Real-time data analysis makes data in motion extremely valuable across numerous industries because it maximizes the data's effectiveness. For example, insurance companies can leverage real-time insights to expedite claims processing by instantly validating reported incidents against external data sources, such as traffic reports or weather feeds, reducing manual intervention and improving customer satisfaction. Patient monitoring systems can facilitate swift medical interventions. Financial institutions can use real-time data to detect and prevent fraud through transaction analysis during its occurrence. Data in motion delivers real-time analytics, which helps businesses make quicker and more intelligent decisions, resulting in better efficiency and improved results. Some of the key elements of managing data in motion include:

- **Hybrid environment flexibility** for easily processing data, analytics and AI workloads in the public cloud and on premises seamlessly.

- **Performance optimization** to accelerate the crunching and parsing of large data sets on an infrastructure engineered to be as efficient and reliable as possible.

- **Support for diverse workloads** to provide AI, data processing, and analytics on structured and unstructured data for a wide range of enterprise workloads.

- **Automation and self-service functionality** to enable customers to independently analyze, interpret, and extract data from AI and analytics.

Market Insight

54% of organizations believed that real-time data availability is a major factor in their data-readiness strategy.

- **Real-time processing** to provide real-time insights and drive dynamic decisions and operational efficiency.
- **Scalable infrastructure** to meet growing demands by scaling easily and reliably for even the most demanding data, analytic, and AI workloads.

Overall, the right platform should empower organizations to take full advantage of data, AI, and analytics to drive performance and innovation at all levels.

Cloudera

Getting useful insights from data, analytics, and AI requires a high-performing, scalable, and flexible platform. Cloudera is the only true hybrid platform for data, analytics, and AI. Cloudera empowers global enterprises to transform diverse data across public and private clouds into actionable insights. A few of the core features to enable organizations to drive actionable insights include:

- **An open data lakehouse.** Cloudera's open data lakehouse design bridges the gap between data lakes and data warehouses and provides scalable and secure data storage for both structured and unstructured data. It enables businesses to easily collect and visualize data in hybrid environments without compromising scalability and vendor lock-in.
- **Enterprise AI acceleration.** Organizations consider data a key asset and differentiator to competition, and AI is the magic wand that extracts maximum value out of this data. When you have trusted data running on a high-performance platform, an organization can accelerate its path to using enterprise data to create unique AI-powered experiences.
- **Data transformation and pipelines.** Through the powerful data transformation and pipeline tools it offers, Cloudera can automate and simplify data preparation for organizations' analysis. From ingest to analysis, the platform facilitates many different data flow scenarios, delivering high-quality data fit for big data and AI use cases.
- **Scale and performance.** Cloudera takes enterprise data management to new heights, enabling organizations to analyze and consume data in real time to aid mission-critical processes and rapid decision-making.

Why Cloudera Excels at Data in Motion

Cloudera delivers a comprehensive, enterprise-grade solution for data in motion—powering real-time analytics, AI, and decision-making at scale. Built on proven open source technologies like Apache NiFi, Kafka, and Flink, Cloudera's platform goes beyond simple data movement to enable secure, scalable, and intelligent streaming pipelines across hybrid and multi-cloud environments. Key capabilities include:

- End-to-end stream processing with tools like Kafka, Flink, and Streams Messaging Manager, enabling real-time ingestion, enrichment, and delivery of high-velocity data.
- 450+ universal connectors for seamless integration across diverse data sources and destinations—on-prem, in the cloud, or at the edge.
- Cloud-native, no-code pipeline development that empowers teams with self-service data flow creation, reducing reliance on engineering resources.
- Security and governance built-in, with enterprise-grade encryption, access controls, and monitoring to ensure data in motion is always protected.
- Unified architecture that eliminates data silos and tool sprawl, streamlining operations and reducing total cost of ownership compared to single-purpose platforms.

With Cloudera, organizations gain the agility to act on live data as it happens—accelerating insights, improving outcomes, and unlocking the full potential of real-time decisioning.

The Power of Cloudera and IBM

Cloudera and IBM enable organizations to harness real-time streaming data, including AI, to derive the most value for real-time business needs.

The Cloudera and IBM partnership elevates organizations through powerful capabilities while delivering smooth data flows via NiFi and Kafka. The partnership provides exceptional price-performance, a secure AI foundation model, and built-in governance benefits, while supporting flexible hybrid deployment options. Advanced retrieval-augmented generation capabilities deliver operational agility and regulatory compliance with prompt lab operations. These future-ready solutions optimize IT infrastructure efficiency, security, and scalability while establishing a benchmark for partner collaboration and customer satisfaction.

Cloudera and IBM work together to deliver an end-to-end solution for real-time data streaming and analytics. Cloudera provides the foundation—managing data collection, real-time streaming, and secure data lake operations—while IBM's watsonx platform adds powerful AI and analytics capabilities to turn that data into actionable insights. Together, the companies help organizations seamlessly collect, manage, and analyze streaming data at scale. An example of this is shown in the fraud detection use case in Figure 3, which illustrates how Cloudera and IBM watsonx work together.

Figure 3. Cloudera and IBM watsonx Fraud Detection Use Case
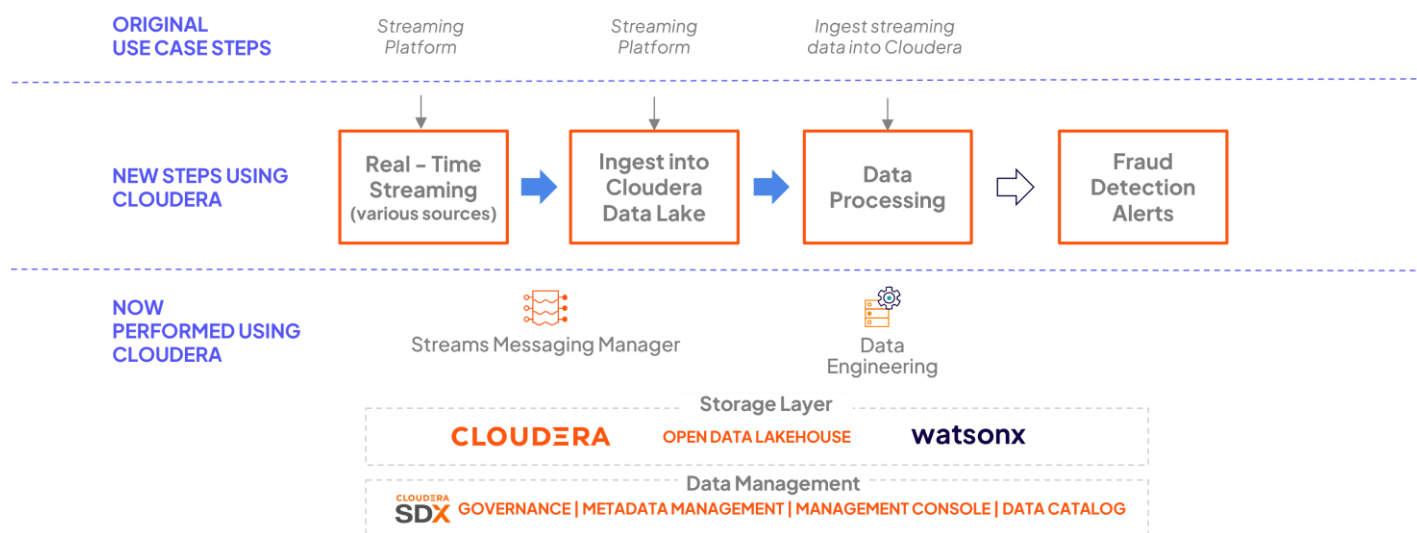
Source: Cloudera

This illustration shows how Cloudera simplifies and enhances the process of collecting, managing, and analyzing real-time data for fraud detection. Instead of using separate systems for streaming, ingesting, and processing data, Cloudera brings everything together into a unified platform. Data from various sources is streamed in real time, ingested into the Cloudera Data Lake, processed efficiently, and then used to generate fraud detection alerts. IBM's watsonx platform works alongside Cloudera to help organizations analyze and gain insights from this data to alert and inform the organization.

By integrating real-time streaming, data ingestion, and advanced processing into a single, secure platform, Cloudera enables customers to simplify complex data pipelines and accelerate decision-making. This streamlined approach reduces operational overhead while ensuring faster, more accurate fraud detection. Combined with IBM's watsonx platform, customers can further unlock the power of their data, using AI-driven analytics and machine

learning models to uncover hidden patterns, drive actionable insights, and improve outcomes. Together, Cloudera and IBM deliver an open, governed, and scalable data foundation that empowers businesses to make smarter, faster, and more confident decisions.

Conclusion

Cloudera and IBM complement each other in providing the perfect combination of scalability, speed, and innovation to process and utilize real-time data in combination with historical data to power advanced analytics and AI effectively. Cloudera's data, analytics, and AI platform works together with IBM's watsonx AI and governance features to deliver a trusted, scalable hybrid platform solution that businesses can use to accelerate their AI application development and deployments, capitalizing on the broad capabilities of two industry leaders. If your organization is looking to turn its data into actionable insights, Enterprise Strategy Group strongly recommends considering the Cloudera and IBM solution.

To learn more about the joint Cloudera and IBM solution, visit https://www.cloudera.com/partners/solutions/ibm.html.