

**Solution Brief**

# Scale AI for all your data, anywhere

with the IBM® watsonx.data™ lakehouse



## Introducing the data lakehouse, a massively scalable, lower-cost alternative for large analytics workloads

AI starts with your data - how it is stored, how it is managed, and how it is governed. Without scalable, accessible data, you risk running into barriers that will impact your ability to run and scale AI workloads cost-effectively and at the speed of business.

Data lakes were supposed to provide these benefits, but properly updating them requires niche, expensive skills, which makes data lakes costly to create and maintain. Data warehouses do offer the high performance needed for processing of high volumes of structured data, but they can also become expensive, particularly for emerging workloads. When it comes to AI and analytics, both data lakes and data warehouses can create environments that are complex, inflexible, and cost-inefficient.

Unfortunately, as organizations grow, many end up with monolithic repositories complemented by multiple data lakes, data warehouses, and, as a result, data silos. In this complex environment, juggling costs, siloed data, and governance end up becoming a constant challenge that inhibit organizations' ability to innovate with AI.

Fortunately, a new paradigm has emerged that addresses these problems.

### The IBM Data Lakehouse

The IBM Data Lakehouse is an emerging architecture that offers the flexibility of a data lake with the performance and structure of a data warehouse. Data lakehouses provide the best of both worlds, while offering significant cost savings potential.

IBM watsonx.data, a modern data lakehouse solution designed for analytics and AI, combines high-performance query engines, low-cost storage, and a metadata governance layer.



Watsonx.data delivers **50%** reduction in data warehouse costs<sup>1</sup>

<sup>1</sup> [https://www.northdoor.co.uk/wp-content/uploads/2023/07/DD250\\_watsonx.data\\_Ebook-2.pdf](https://www.northdoor.co.uk/wp-content/uploads/2023/07/DD250_watsonx.data_Ebook-2.pdf)

IBM's Data Lakehouse solution, watsonx.data, provides the key components needed to scale and accelerate the impact of AI:



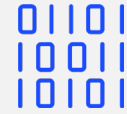
### **Infrastructure**

Where the lakehouse will be deployed and managed across any cloud or on-premises environment



### **Storage**

Physical storage in files and in open data formats, such as Apache Parquet and Avro



### **Open Table Formats**

Open table formats, such as Apache Iceberg, help provide structure, with the reliability and simplicity of SQL with big data

Governance is provided through metadata, which is also stored in open table formats:



### **Technical Metadata Service**

Required to understand what data is available in the storage layer



### **Data Catalogs**

Helps users find the correct data for the job and delivers semantic information for policies and rules



### **Policy Engine**

Enables users to define data protection policies and enables the engine to enforce those policies



### **Query Engine**

Accesses data in open table format and can be dynamically scaled to meet workload demands and concurrency

Data lakes such as watsonx.data can dramatically reduce storage and compute costs, while also freeing you from proprietary warehouse technology, reducing vendor lock-in, and lowering IT infrastructure overhead costs.

## Benefits of Using Watsonx.data for Analytics and AI

The IBM watsonx.data platform is an open, hybrid, and governed data lakehouse optimized for all data and AI workloads. It helps organizations derive value from their analytics ecosystem through three core benefits:



### Access all data through the hybrid cloud

Access any data source, wherever it resides, through a single point of entry and combine it using open data formats



### Get started in minutes

Accelerate time to trusted insights with built-in automation and governance that is unified across the entire ecosystem



### Reduce data warehouse costs

Workload optimization can help cut costs by up to 50% through workload optimization across multiple query engines and storage tiers

For decades, IBM has been handling the world's most demanding data workloads. With watsonx.data, they are maintaining their industry-leading position with a data lakehouse that enables organizations to get greater value from their analytics and AI.

## Putting AI to Work with Watsonx.data

Watsonx.data grants access to data across all cloud and on-premises environments. It lets you get started in minutes, leverage multiple query engines to run analytics and AI workloads, and cut data warehouse costs by up to 50%. If you are ready to take the next step in your data journey, consider putting AI to work in your business with watsonx.data.

## Take the Next Step with Logicalis and IBM

Logicalis, an award-winning IBM Platinum Partner with over 25 years of experience, helps organizations around the world drive transformation with IBM solutions. With a team of IBM-certified engineers and professionals, and offices in 27 countries, we can help you deploy, manage, and make full use of watsonx.data.

For more information on how we can support your data and AI initiatives, **contact your Logicalis account representative today.** Or, if you are not yet a Logicalis customer, get in touch to **schedule a free executive briefing on watsonx.data.**

Schedule a free executive briefing.  
[logicalis-hub.com/ibm](https://logicalis-hub.com/ibm)

 **LOGICALIS**  
Architects of Change

Platinum  
Business  
Partner

