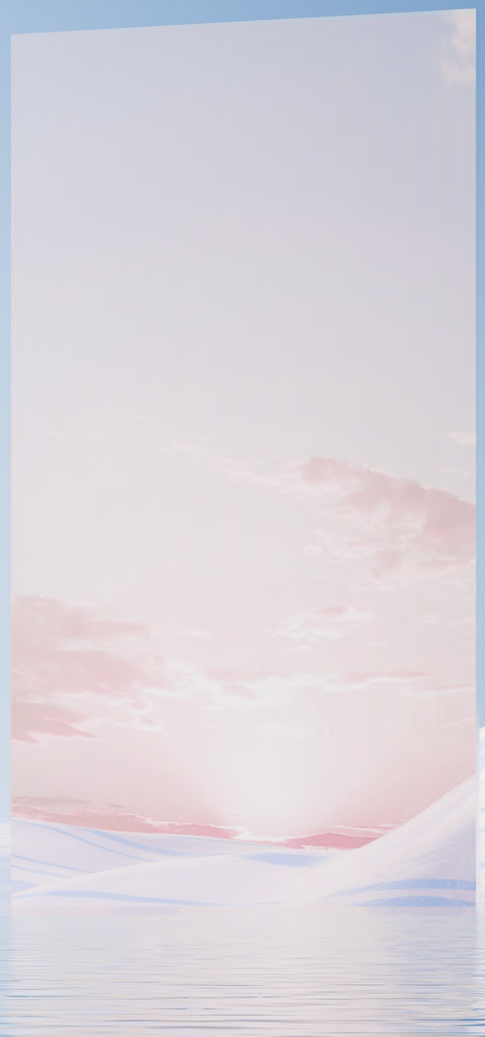
THE DATA LAKEHOUSE BLUEPRINT

The Missing Layer for Your Lakehouse

Connecting your lakehouse with operational data to deliver trusted insights for AI, analytics, and real-time decisions



CONTENTS

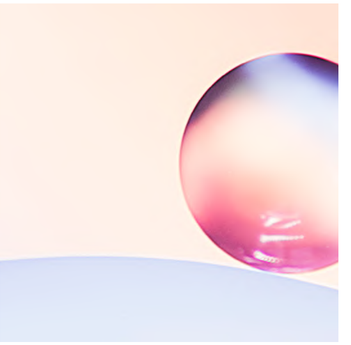
INTRO
The Assumption Nobody Questioned

SECTION 1
The Four Places Your Lakehouse Alone Is Not Enough

SECTION 2
What Changes When You Close the Gap

SECTION 3
What the Missing Layer Looks Like in Practice

SECTION 4
Organizations That Closed the Gap



THE DECISION
You Built the Foundation. This Is What Completes It.

THE ASSUMPTION NOBODY QUESTIONED

You made the right call. A modern data lakehouse, built right, running well, scaling as promised. It provides a strong foundation for data, analytics, and AI.

The platform scales. But the business outcomes don’t keep pace.

Analytics initiatives that should take days still take weeks. AI projects stall waiting on the data team. Business users raise tickets instead of opening dashboards. Your data engineers, the ones who built all of this, remain deeply involved in fulfilling requests, not because of how they built it, but because accessing and combining the right data is still too complex to do without them.

As new use cases emerge, the gap becomes more clear. Important context still lives across operational systems like SAP, Salesforce, and Oracle, as well as other distributed sources. Without unified access across both, every new question becomes another integration effort, slowing down insights and limiting what AI and analytics can deliver.

This is not a failure of execution. Somewhere along the way, a centralized storage and processing strategy got treated as a complete data delivery solution. **But it’s not.**

This situation is the natural outcome of centralizing data without a consistent way to connect, govern, and deliver it across all sources. It increases reliance on engineering and creates a gap between question and answer.

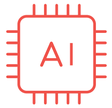
Closing it does not require a new lakehouse. It merely requires a layer that connects it to the rest of your data.

SECTION 1

THE FOUR PLACES YOUR LAKEHOUSE ALONE IS NOT ENOUGH

Today, many organizations are treating the lakehouse as the default architecture for modern analytics and AI. Most organizations adopting a lakehouse are facing the same set of challenges. Those that address them early are able to move faster with AI, self-service, and operational use cases. Those that don't continue to rely heavily on engineering effort to close the gaps.

Independent analyst firm Veqtor8 conducted research with enterprises **running Snowflake and Databricks at scale**, some with a logical data management layer, some without. A consistent pattern emerged: where that layer was missing, the same four challenges showed up:



Enabling AI and Agentic Applications

AI models and agents are dynamic. They determine what data they need based on the question or task at hand. The more data and context they can access across the enterprise, the more accurate, reliable, and ultimately trusted their outputs become. The lakehouse provides a strong foundation for analytical use cases, but important context still lives across operational systems and other distributed sources. Without a way to bring that together, AI outputs are limited, and projects struggle to move beyond early stages.

Supporting Business Self-Service

Your lakehouse catalog provides sufficient metadata, but it was designed for data engineers and other technical roles. But business users need something different. They need clear business definitions, relationships, and confidence in the data they are using. In practice, analysts still rely on data teams to combine data across systems, apply logic, and validate results. Self-service improves, but it rarely removes the dependency on engineering.

Providing Access to Real-Time Operations

Lakehouses are designed for large scale analytics. Many business use cases are not. Fraud detection, risk scoring, customer interactions, and operational APIs all depend on current data from source systems. These scenarios require connecting analytics to operational data as it exists, not just after it has been processed and loaded. Without that connection, teams introduce additional pipelines, which drive up latency, costs, and overall complexity.

Controlling Costs

Each new use case often brings new pipelines, new integrations, and more data movement. As data spans both the lakehouse and other enterprise systems, this effort adds up quickly. Over time, the cost is not just infrastructure. It is the ongoing engineering work required to keep everything aligned and usable.

SECTION 2

WHAT CHANGES WHEN YOU CLOSE THE GAP

To understand what actually changes, Veqtor8, an independent analyst firm focused on data and AI, spoke with enterprises running Snowflake and Databricks at scale. Some had added a logical data layer. Others had not.

The differences showed up quickly in how teams worked and how fast they delivered.

Organizations that introduced a logical layer were not replacing their lakehouse. They were extending it. The impact was less about the platform itself and more about how data could be accessed, combined, and trusted across the business.

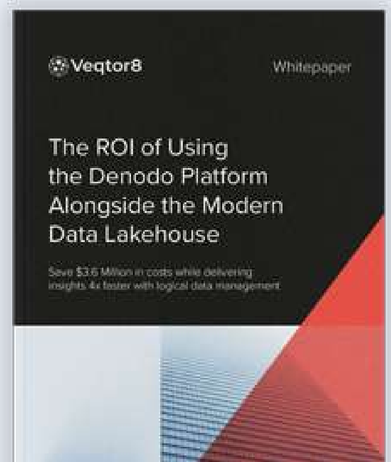
Across those organizations, five areas stood out.

BENEFIT AREA	IMPROVEMENT	HOW DENODO ENABLES IT
Time-to-Insight	**3–4x faster**	Virtualized access and semantic abstraction
Implementation Cost	**78% lower**	Reusable connectors, less custom coding
Data Engineering Productivity	**75% less effort**	No-code integration and reusable views
Governance & Consistency	**77% lower cost**	Centralized metadata and policy enforcement
Business Agility	**5–10x faster rollout**	Unified real-time and historical data access

These are not isolated improvements. They reflect a shift in how data is delivered across the organization. Instead of routing every request through engineering, teams gain direct access to trusted, connected data, regardless of where it lives.

The full Veqtor8 report goes deeper into how these changes show up across different industries and use cases.

Download the full Veqtor8 ROI Report →

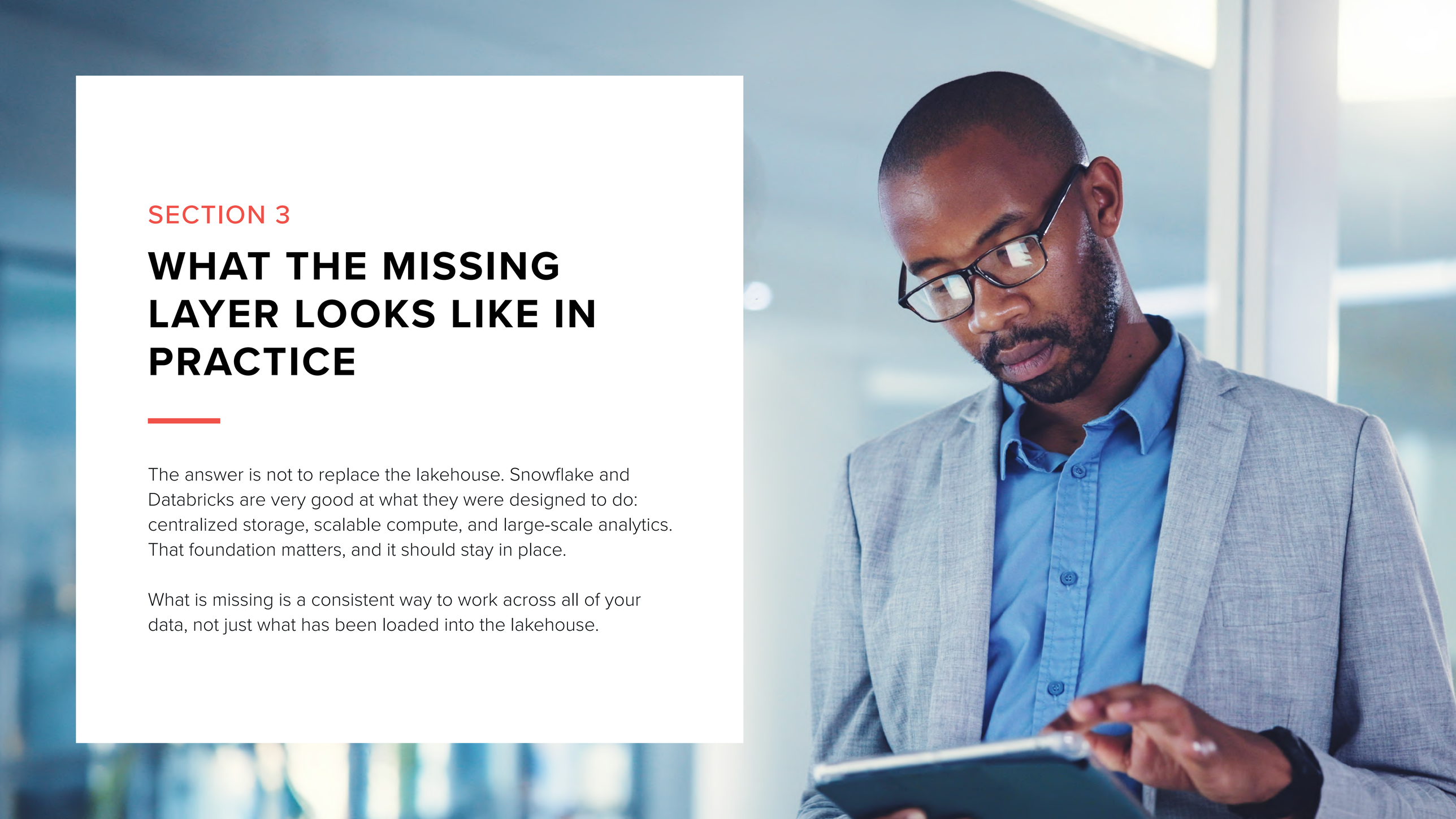


SECTION 3

WHAT THE MISSING LAYER LOOKS LIKE IN PRACTICE

The answer is not to replace the lakehouse. Snowflake and Databricks are very good at what they were designed to do: centralized storage, scalable compute, and large-scale analytics. That foundation matters, and it should stay in place.

What is missing is a consistent way to work across all of your data, not just what has been loaded into the lakehouse.



A logical layer sits above and alongside the lakehouse and connects it to the rest of your data landscape. It provides a way for users to discover, understand, and access data across systems, with business context built in so it can be used directly by analysts, applications, and AI. This includes data in the lakehouse as well as operational systems like SAP, Salesforce, and Oracle. Governance is defined centrally, but enforced across all data sources, so policies are applied consistently whether data lives in the lakehouse or in operational systems.

Denodo is designed to serve as that layer. It delivers governed, real-time access to data across sources, with consistent definitions and policies, without requiring data to be moved or restructured.

WHY EXISTING TOOLS DO NOT FULLY CLOSE THIS GAP

Most organizations already have strong tools in place. Each plays an important role, but they operate within a specific scope.

- Fivetran and Matillion ingest data into lakehouses.
- Unity Catalog governs data within Databricks.
- Dbt defines transformations after data has been ingested.
- Trino and Starburst can federate queries across systems.
- Collibra and Alation document and catalog metadata.

These tools are valuable, but they do not provide a single layer that combines access, business meaning, and governance across all sources, both inside and outside the lakehouse, in a way that can be used directly by all consumers.

That is why many organizations with mature data stacks still experience the same challenges described earlier.

The Veqtor8 report also highlighted that those organizations **operating a lakehouse without logical data management** ended up purchasing at least three additional tools to attempt to address these gaps. This adds significant additional cost, complexity and technical debt to their data architecture, all of which could be avoided through a single logical solution.

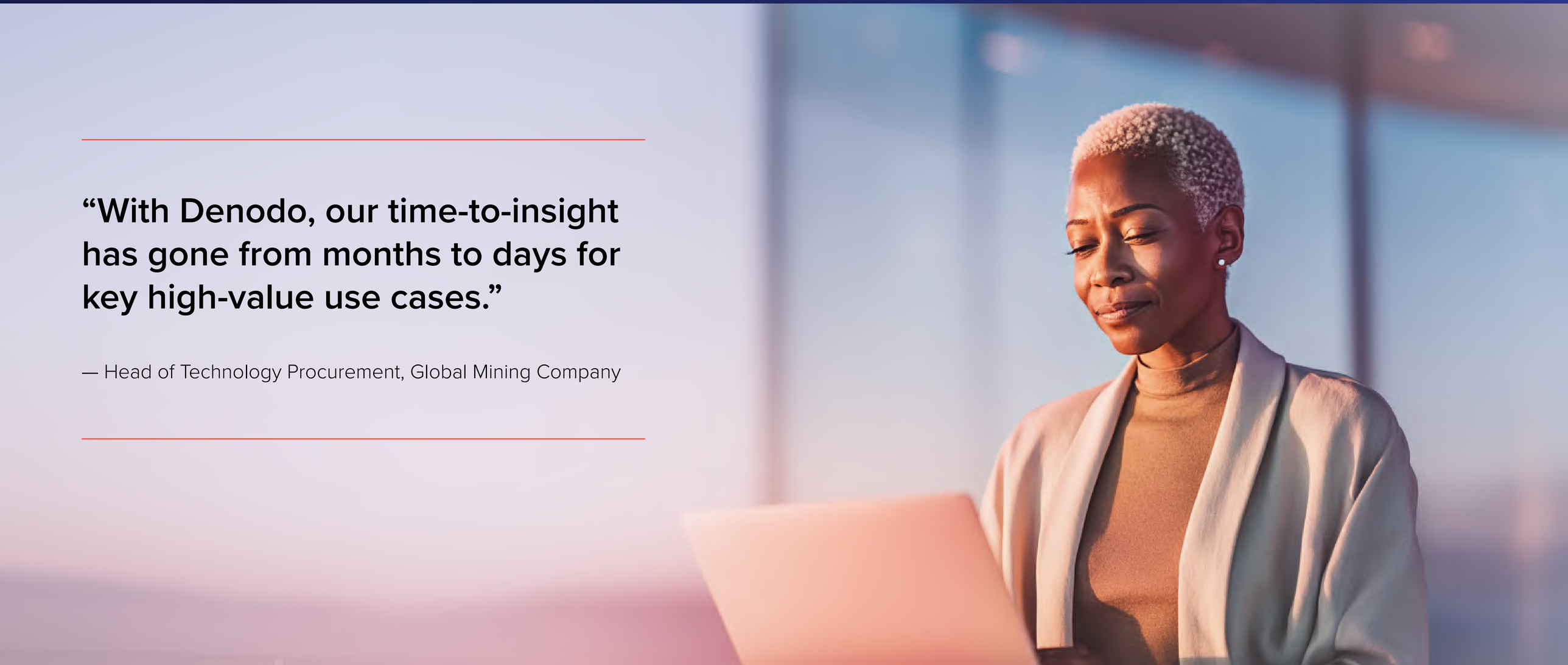
WHAT CHANGES IN PRACTICE

WITHOUT THE LOGICAL LAYER	WITH DENODO ABOVE THE LAKEHOUSE
An analyst needs a cross-source view of customer risk, combining data from Databricks, Salesforce, and a legacy CRM. A request is submitted. A data engineer builds a new pipeline. Data is ingested, transformed into a new table, tested, and deployed. This takes days, sometimes longer.	The analyst opens the Denodo data marketplace, searches for “customer risk,” reviews its definition and lineage, and queries it directly. The data is combined across systems in real time, with definitions and policies already applied.
An AI agent requests customer context. Only the data prepared in the lakehouse is available. Operational systems and SaaS applications are not included, limiting the completeness of the response.	The AI agent accesses data across both the lakehouse and operational systems. Context is broader, definitions are consistent, and governance is enforced at the point of access.
A new regulation requires masking sensitive data across multiple systems. Policies are implemented separately in each system, often on different timelines.	A single policy is defined once and applied consistently across all sources and consumers, including analytics, applications, and AI.
Time from request to trusted answer: several days.	Time from request to trusted answer: often within the same working session.

For AI builders, the **Denodo AI SDK** provides governed, real-time access to enterprise data through standard APIs. This enables developers to connect LLMs and agents to both lakehouse and non-lakehouse data without building custom integrations or bypassing governance controls.

“With Denodo, our time-to-insight has gone from months to days for key high-value use cases.”

— Head of Technology Procurement, Global Mining Company



SECTION 4

ORGANIZATIONS THAT CLOSED THE GAP

Leading organizations across industries are discovering that modern lakehouses alone do not solve the final challenge of enterprise data delivery: connecting fragmented systems, governing data consistently, and delivering trusted data at business speed.

The companies in this section closed that gap by adding the Denodo Platform as a logical data layer across enterprise systems. The result was faster analytics delivery, reduced engineering effort, stronger governance, and broader self-service access to trusted data for BI and AI initiatives.

Global Top-10 Insurer — From Manual Reconciliation to Real-Time Delivery

A global top-10 insurance company had built a modern data lakehouse. But key data still lived across cloud and on-premises systems. Analysts spent significant time reconciling sources, and each new use case required custom extract, transform, and load (ETL). Governance and masking were handled inconsistently across systems.

The insurer deployed Denodo alongside Databricks to introduce a unified semantic layer across both lakehouse and operational data. Manual reconciliation was reduced, and governed, real-time data products were made available to analysts across the business. The same layer now supports AI underwriting models with consistent, up-to-date data drawn from multiple systems.

Time to deliver a new analytics use case improved from months to weeks. Governance became more consistent, and reliance on custom engineering work was significantly reduced.

Energy Technology Company: Lakehouse Performance at Business Speed

A global energy technology company running Databricks at scale was still measuring data delivery in weeks. While the lakehouse handled processing well, delivering data for new use cases required repeated integration work.

Denodo was introduced as a logical layer above the lakehouse, connecting it with other enterprise data sources. Repetitive **ETL development was reduced, and business analysts gained more direct access to governed data across systems.** Data teams were able to shift focus from building pipelines to supporting higher-value initiatives.

Global Bank — Self-Service Without the Engineering Queue

A global bank deployed Denodo above its lakehouse to reduce reliance on ticket-driven data delivery. Business users gained direct, governed access to consistent data across both lakehouse and operational systems.

As a result, teams were able to respond more quickly to business needs, while data engineers spent less time managing requests and more time improving data and supporting new use cases. The outcome **was faster decision-making, lower operational overhead, and a more scalable approach to data delivery.**

“With Denodo, our business leaders can now access consistent data for decision-making without waiting for IT. This significantly accelerates how quickly we can respond to market needs.”

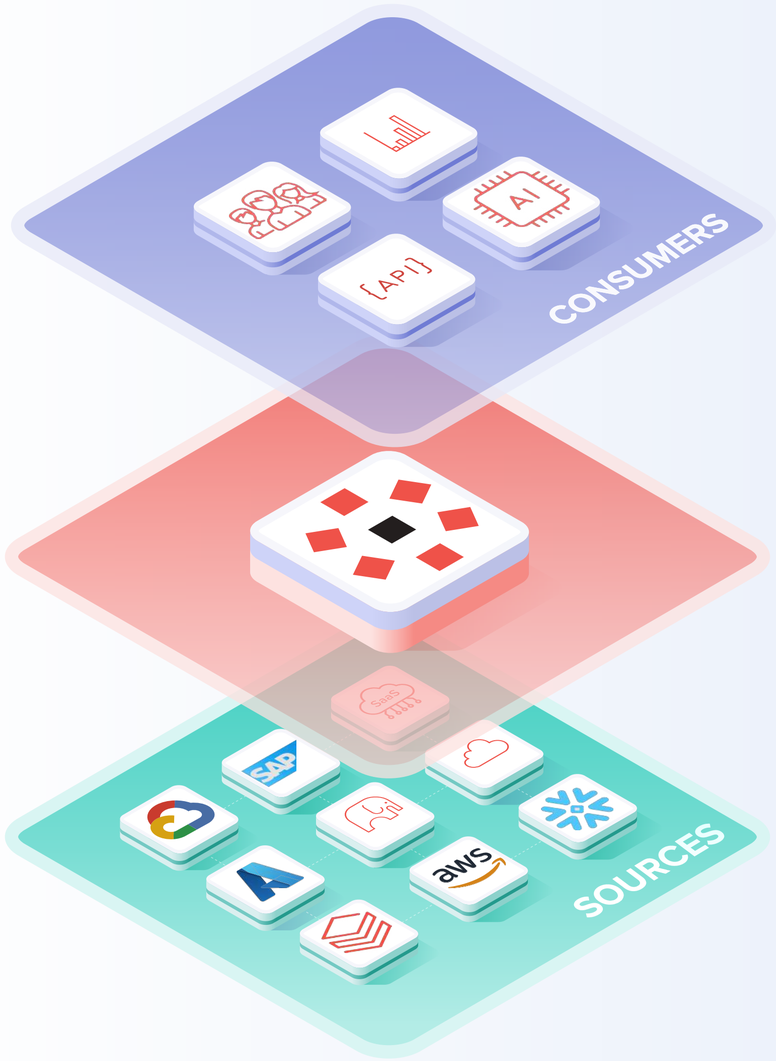
— Head of Technology, Global Bank

THE DECISION

YOU BUILT THE FOUNDATION. THIS IS WHAT COMPLETES IT.

Organizations getting the most from Snowflake and Databricks tend to approach them the same way. They treat the lakehouse as a foundation for storage and compute, not the complete solution for delivering data to the business.

What they add is a layer that connects that foundation to the rest of their data. It brings together lakehouse data with operational systems, applies consistent definitions and governance, and makes that data usable across analytics, applications, and AI.



Your lakehouse is the foundation. This is what completes it.

Without a logical layer, every new question is a pipeline project. With one, analysts and AI access the same trusted, governed data — without waiting on engineering.

For Data Architects and IT Leaders

See how a logical layer integrates with Unity Catalog, Delta Lake, and Snowflake in practice.

Book a Technical Session

For CIOs, CDOs and Data Leaders

See how organizations improved delivery speed, cut complexity, and unblocked AI initiatives, without replatforming.

Book a Strategy Session

Not ready for a conversation yet? Download the full Veqtor8 ROI Report.

denodo

Visit www.denodo.com | Email info@denodo.com | Discover community.denodo.com