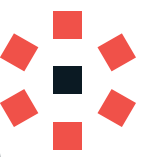


EBOOK

denodo  +  **databricks**

A MODERN DATA STRATEGY, ACCELERATED

Unlocking Data Agility with
Denodo and Databricks



INTRODUCTION
The Data Challenge



CHAPTER 1:
Semantic Unification of All Data



CHAPTER 2:
Data Self-Service



CHAPTER 3:
Federated Data Governance
with Centralized Oversight

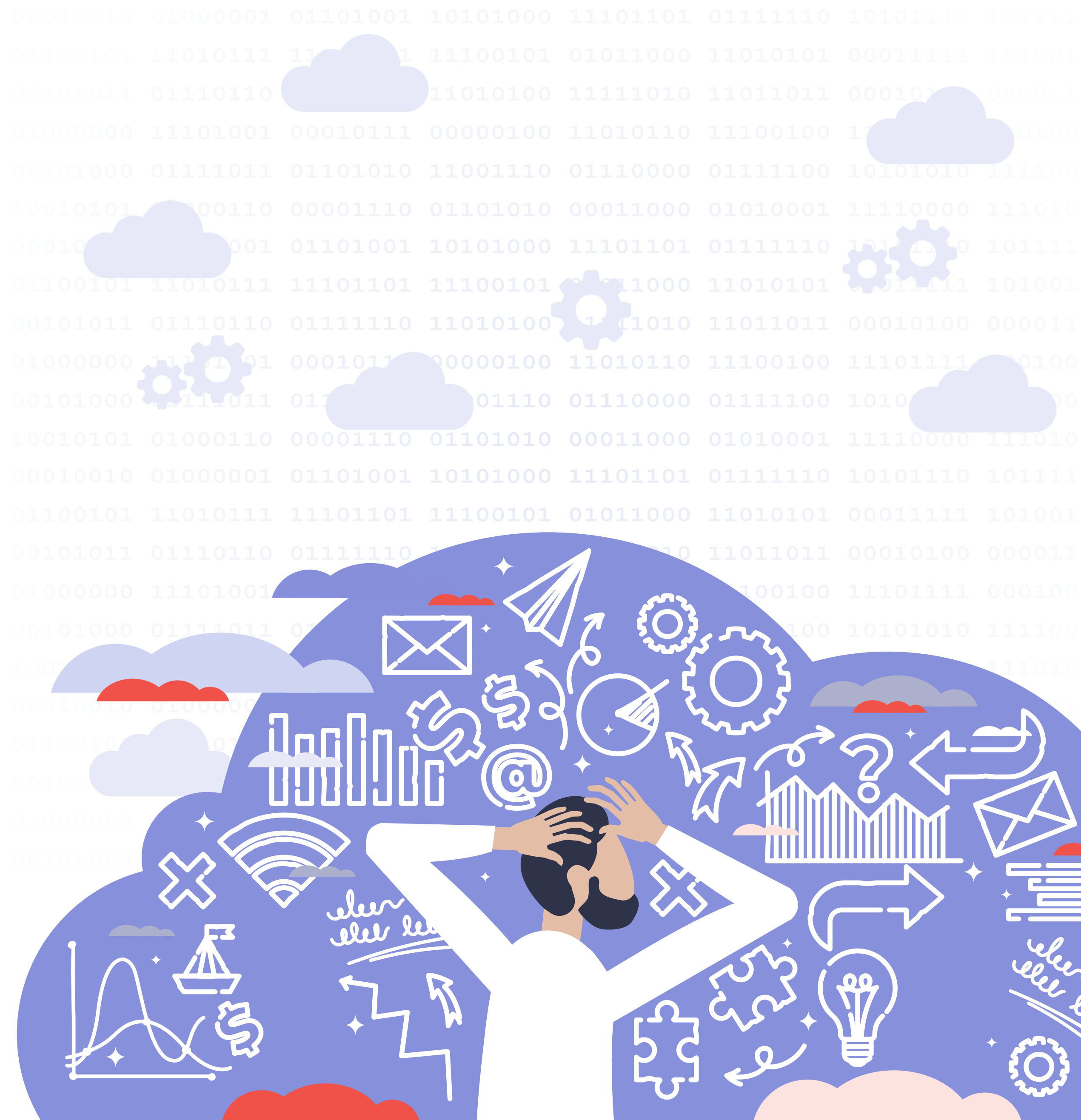
THE DATA CHALLENGE

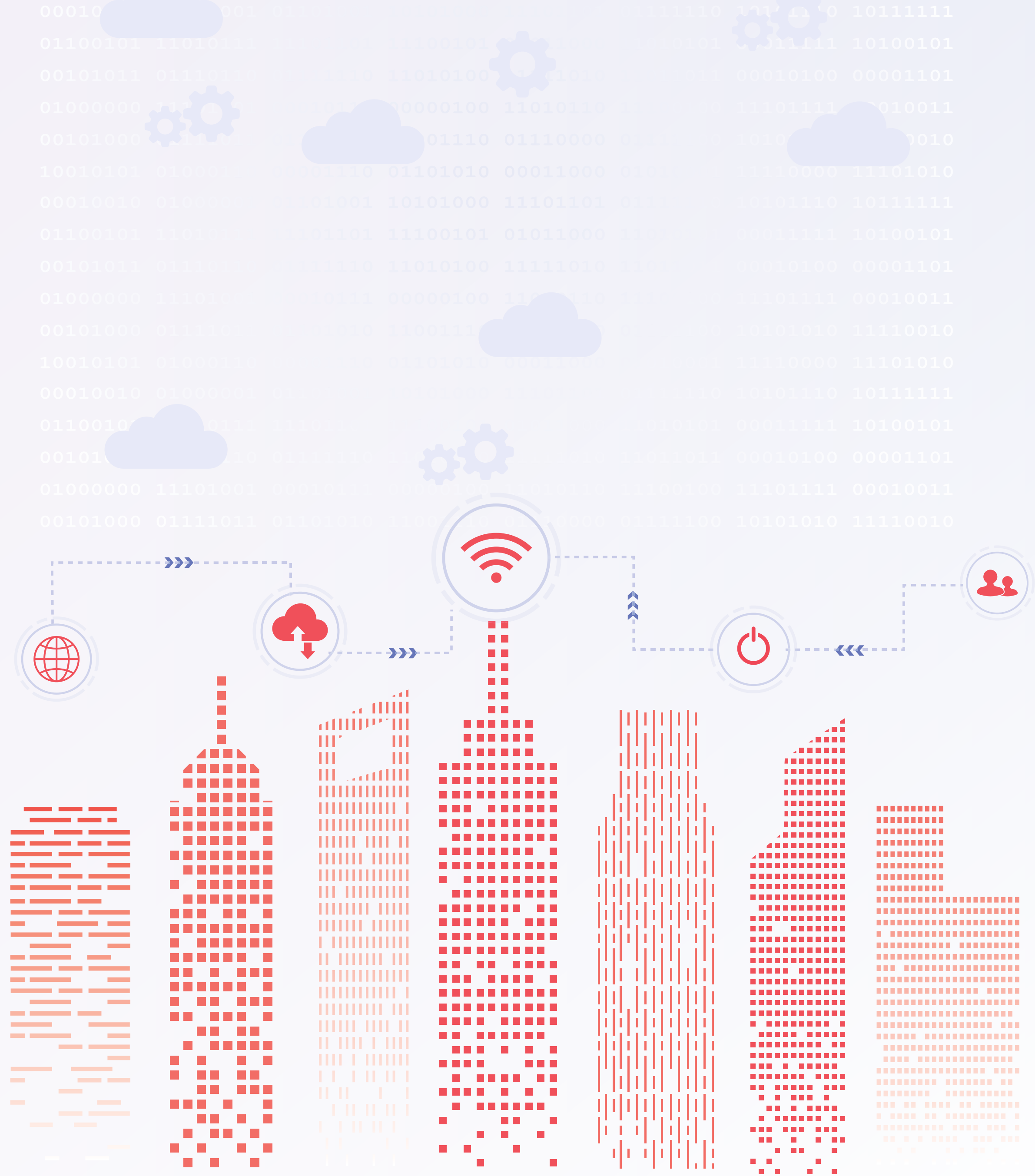
Organizations today face unprecedented challenges in managing and harnessing data to drive business success. Data is exploding in volume, complexity, and variety, spanning structured sources such as databases and unstructured inputs like IoT sensors, web logs, and customer interactions.

At the same time:

- Data remains **siloed across disparate systems**, including on-premises, cloud, and hybrid environments, creating barriers to seamless access and integration.
- Organizations must balance **cost efficiency** with the need for scalable, high-performance processing to meet growing analytics and AI demands.
- Businesses struggle with delivering **governed, accurate, and timely insights** without overwhelming IT teams.

The result? Teams are slowed down by disconnected tools, fragmented data landscapes, and rising infrastructure costs, hindering their ability to deliver real-time insights and innovation.





Scaling Databricks in a Distributed Data Environment

Databricks has emerged as a leading platform for advanced analytics, data engineering, and AI/ML innovation. Its powerful lakehouse architecture empowers organizations to build scalable AI models and derive insights from massive volumes of data.

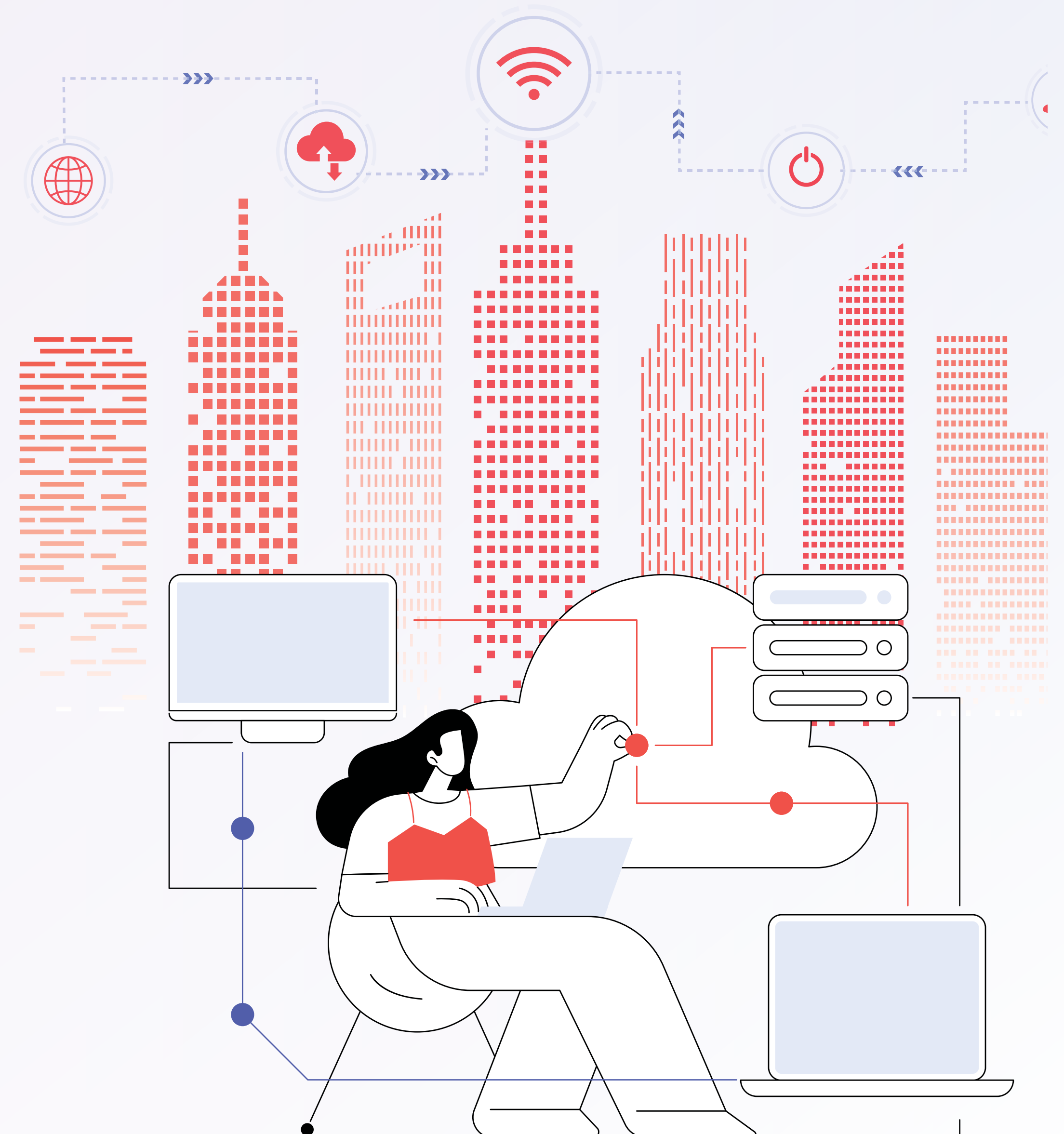
As enterprises expand their use of Databricks, many face new challenges: how to support real-time decision-making, how to extend governed data access to more users, and how to operate efficiently in a world where data is increasingly distributed across hybrid cloud and on-prem environments.

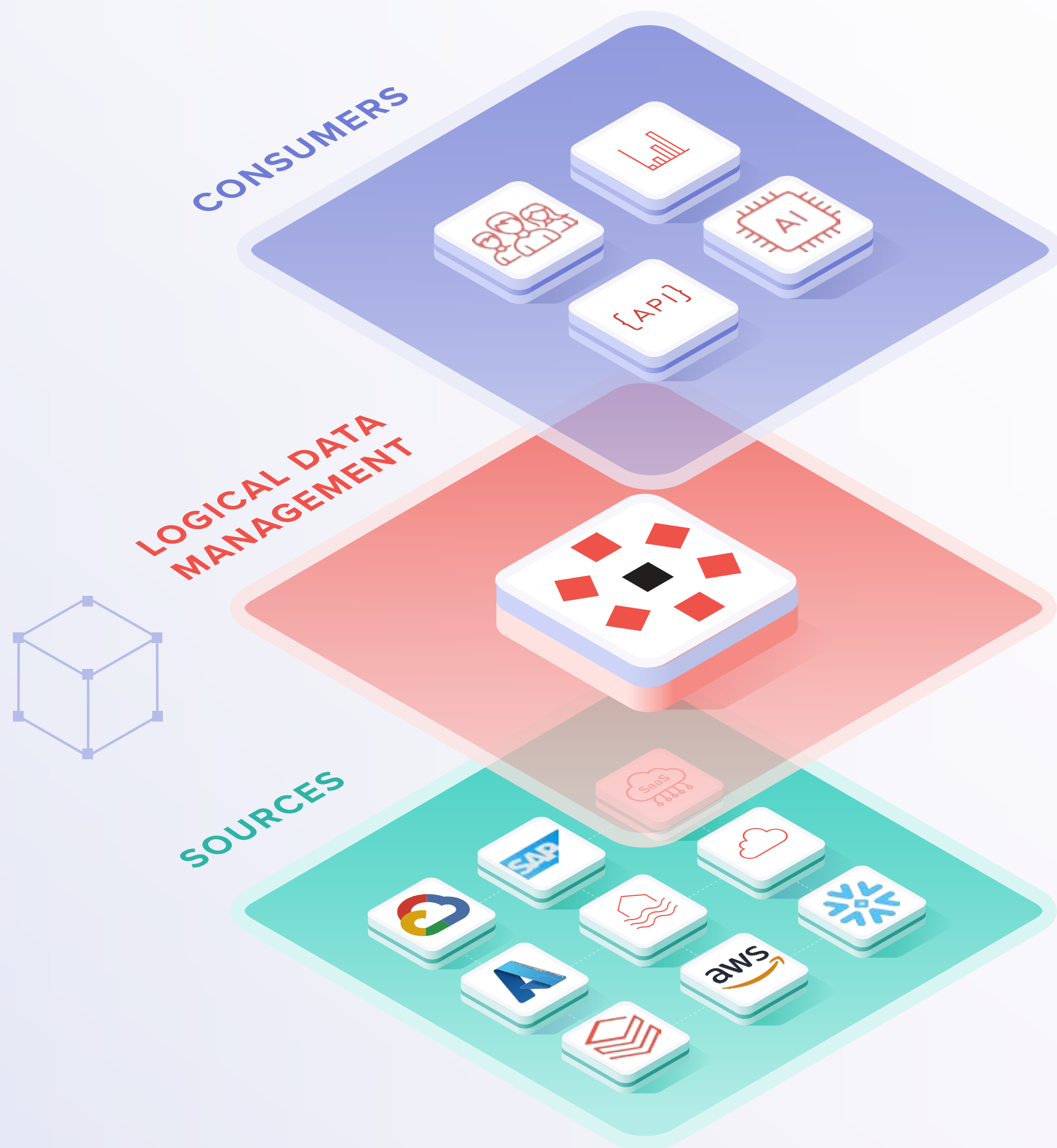
Not all data lives in Databricks—and not all use cases benefit from centralizing it.

Extending the Power of Databricks with Denodo

Real-time operational reporting, AI applications that depend on live, contextual data, and scenarios that require data to remain governed in place are often better served through a logical approach to data integration.

That's where Denodo adds value. The Denodo Platform complements Databricks by providing a unified data access layer that delivers real-time integration, semantic consistency, and secure, governed self-service access to data across the enterprise. Together, Databricks and Denodo provide the agility, scale, and governance organizations need to accelerate innovation and drive meaningful business outcomes.





denodo + databricks

A UNIFIED SOLUTION

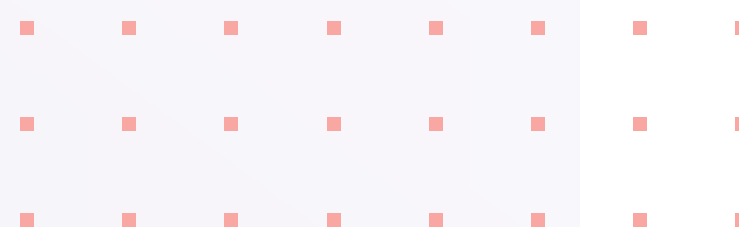
Denodo and Databricks work together to eliminate data silos, simplify operations, and accelerate insights. Together, they empower users to unlock the full potential of their data.

In this eBook, you'll see how Denodo enhances the value of your Databricks deployment by:

- Delivering **semantic unification of all data** to ensure seamless integration and collaboration across Databricks and additional data sources.
- Enabling **self-service data discovery** for business and technical users alike.
- Applying **federated data governance with centralized oversight** to build trust and ensure compliance.

Discover how Denodo and Databricks can **transform your data landscape—empowering your teams, enhancing efficiency, and driving innovation.**

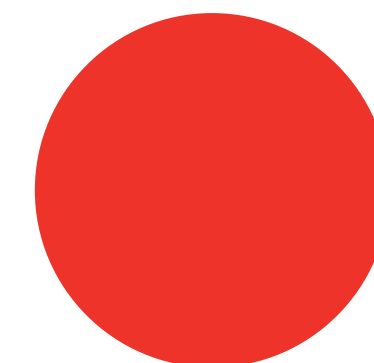
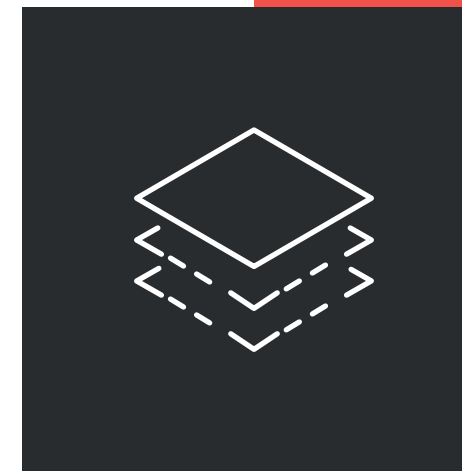
Let's get started!



CHAPTER 1:

SEMANTIC UNIFICATION OF ALL DATA

A universal data access layer enables seamless, secure access to data across all Databricks and non-Databricks sources.

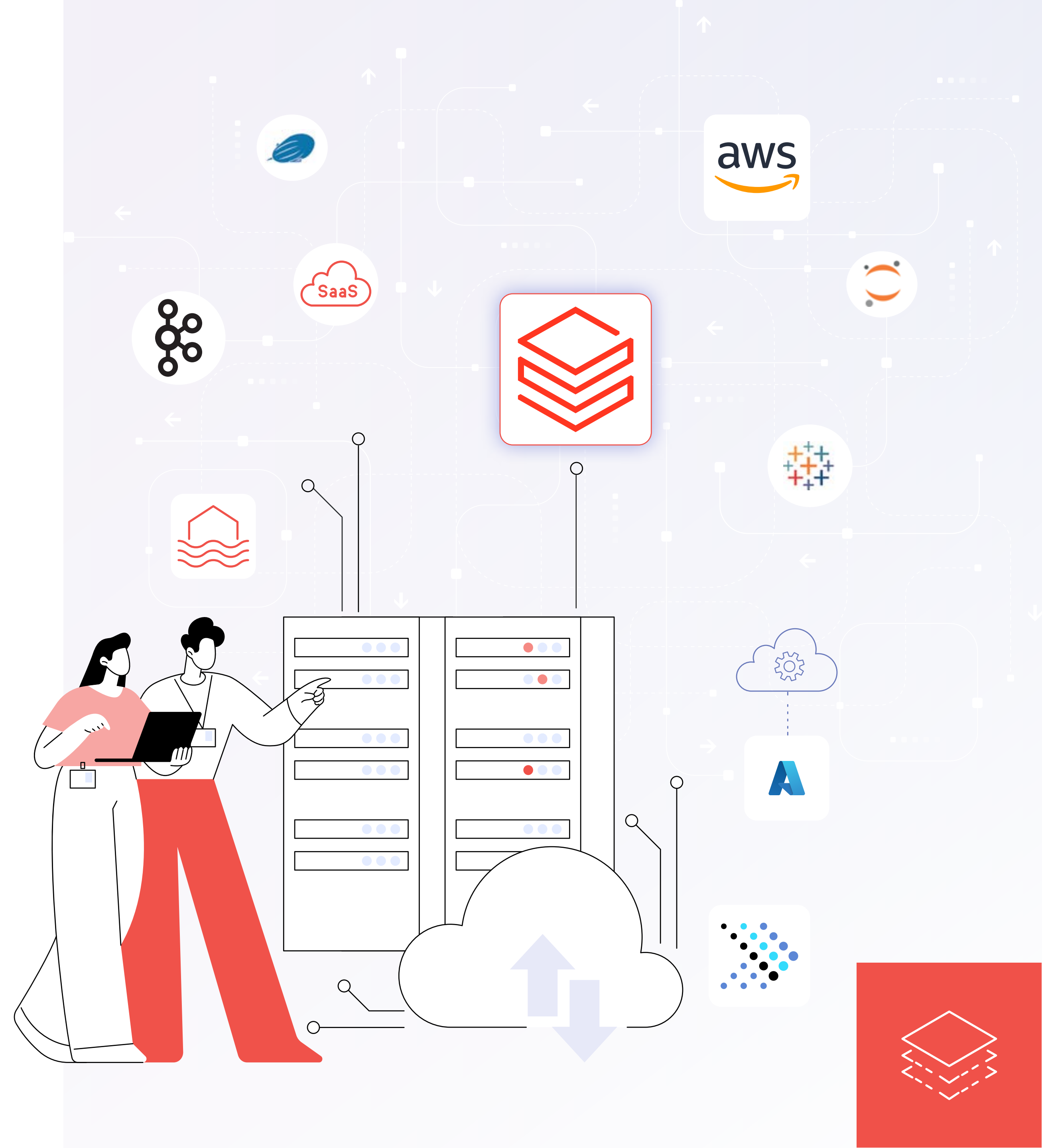


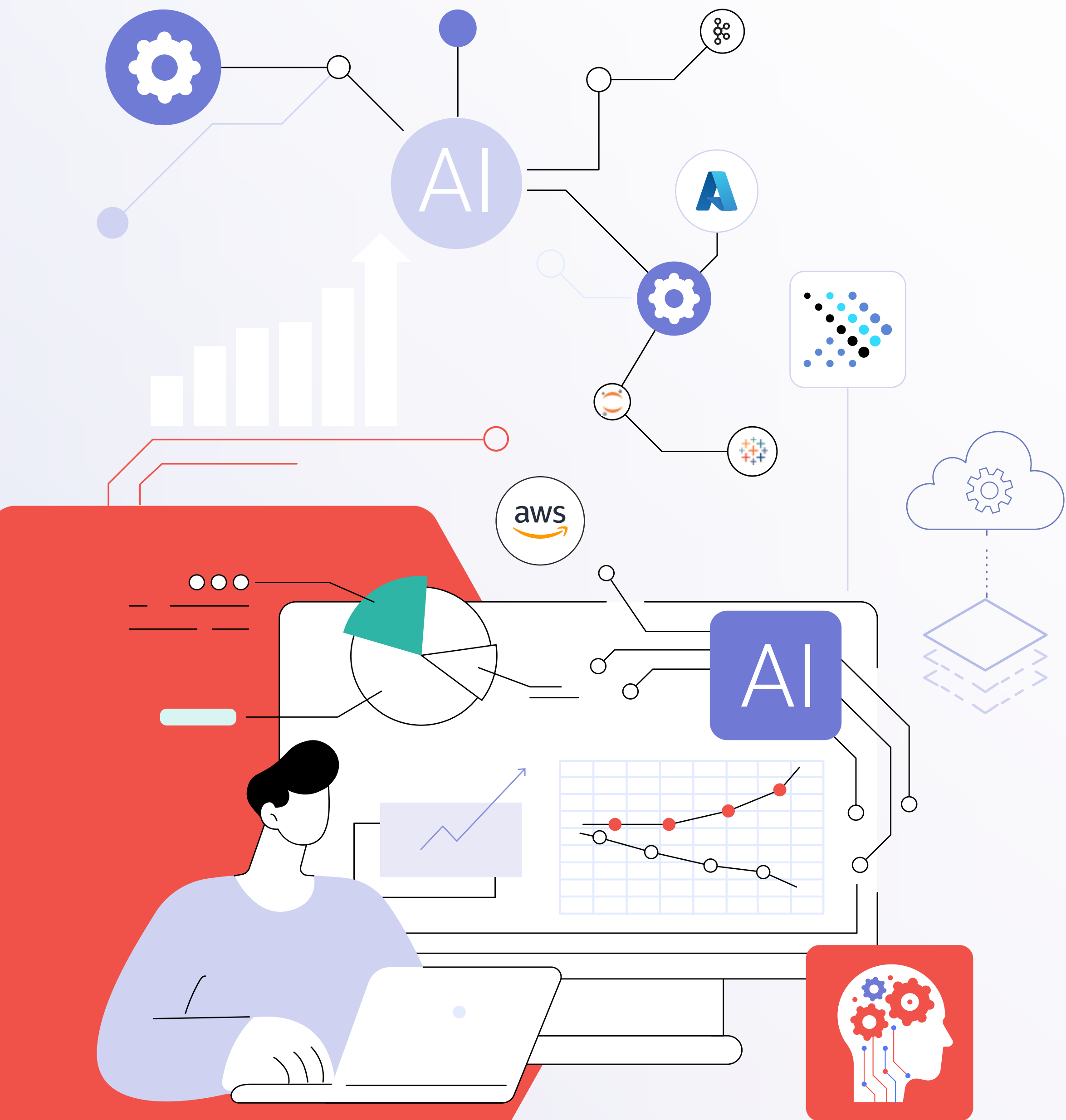
Semantic Unification Of All Data

Deep Integration with Databricks

The Denodo Platform's deep integration with Databricks leverages **SparkSQL** for seamless collaboration between the platforms. Additionally, the embedded Presto MPP engine supports a wide range of data formats, including **Delta Lake**, **Iceberg**, and **Parquet**. By combining Databricks' advanced data processing and storage with the Denodo Platform's data virtualization and access capabilities, organizations can efficiently process and analyze large datasets across various formats.

Beyond Databricks, the Denodo Platform connects to over 200 data sources, including relational databases, NoSQL databases, cloud platforms, and APIs. This extensive connectivity simplifies integrating Databricks with diverse systems, for streamlined access to critical data, whether it's on-premises or in the cloud.

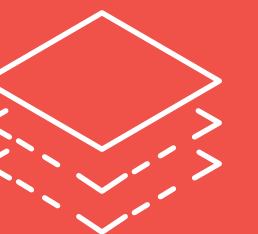




Semantic Unification Of All Data

A Broad Set of Data Delivery Options

The Denodo Platform offers a broad set of data delivery options beyond standard JDBC and ODBC connectors, including REST, OData, GeoJSON, and GraphQL APIs, as well as message queues like Kafka and JMS. These diverse options provide developers with the flexibility to choose the most appropriate method for delivering data to consumers, whether for building applications, real-time data streaming, or supporting multiple interaction models across different platforms



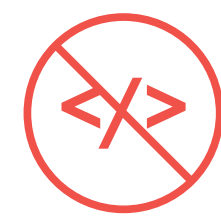
A Consistent View with a Unified Semantic Layer

The semantic layer within the Denodo Platform empowers organizations to operate efficiently by providing a unified, consistent view of business models and a common analytics vocabulary. It simplifies complex data, from a variety of sources, into actionable insights, enhancing data accessibility and governance across the enterprise. Here's how it adds value:



ENHANCED ACCESSIBILITY FOR ALL USERS:

The Denodo Platform's semantic layer translates complex data into business-friendly models, making data accessible to everyone—from business analysts to executives—who may lack technical expertise. This fosters a data-driven culture and puts insights within reach across the organization.



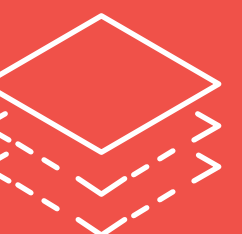
NO-CODE MODELING FOR QUICK ADAPTATION:

Users can build and refine models without SQL, enabling rapid iteration and quicker responses to changing business needs. This reduces time-to-insight, so teams can act faster.



CONSISTENT SEMANTICS ACROSS TOOLS:

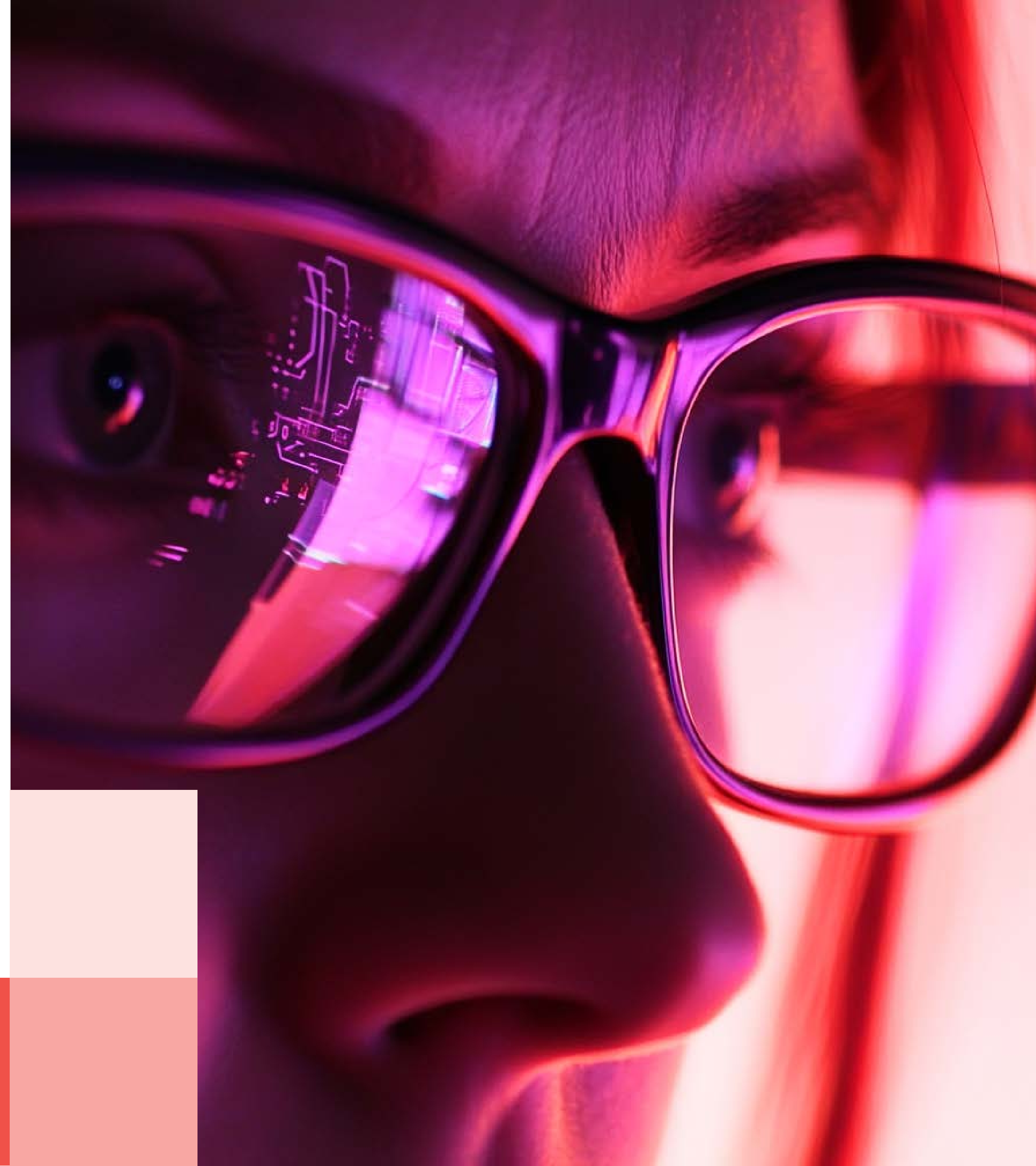
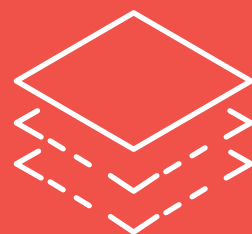
The Denodo Platform centralizes business definitions, eliminating the need to recreate them in each BI or analytics tool. This approach reduces data duplication, simplifies governance, and bases all analyses on a single, trusted source, streamlining data management and enhancing security.



Semantic Unification Of All Data

Smart Query Acceleration

The Denodo Platform analyzes data usage patterns and source information to generate recommendations, automatically creating pre-aggregated, persisted data sets that optimize the performance of analytical queries. This process is seamless and transparent to the end-user.





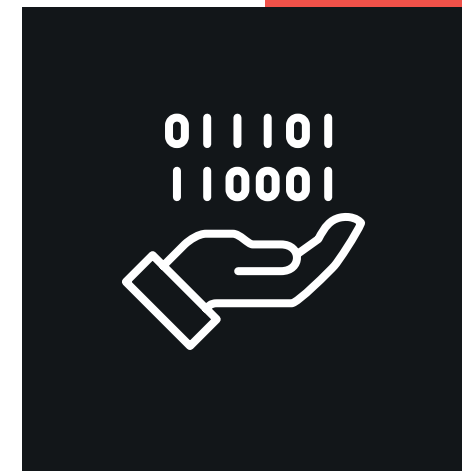
Semantic Unification Of All Data

Streamlined, Cost-Effective Data Access

A centralized semantic layer enhances the data infrastructure by optimizing resource usage and expanding access. While Databricks excels in advanced analytics and machine learning (ML), growing adoption can lead to unintended DBU consumption as other teams tap into the platform for operational or ad hoc queries. By integrating the Denodo Platform, organizations can provide secure, governed access to data in the lakehouse—often without consuming DBUs—reserving Databricks capacity for high-value analytics and ML workloads. Additionally, Denodo's FinOps dashboards offer visibility into data usage patterns and query behaviors to help manage costs and workloads more effectively. This ensures optimal use of Databricks, while enabling broader data access for business users and intelligent applications through Denodo, maximizing the value of both platforms.

CHAPTER 2:

DATA SELF-SERVICE

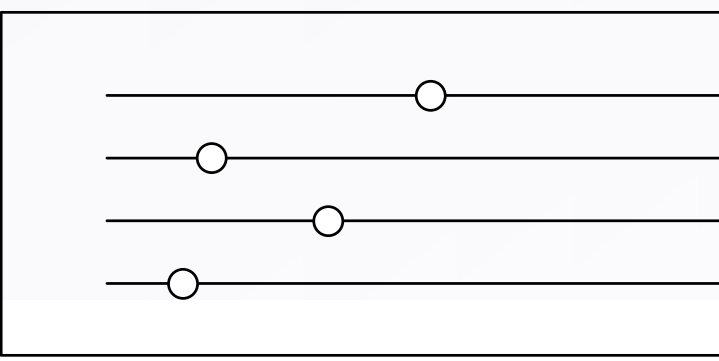
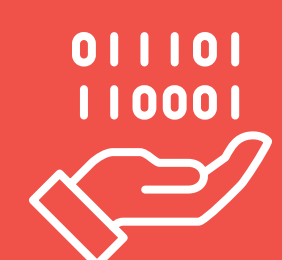
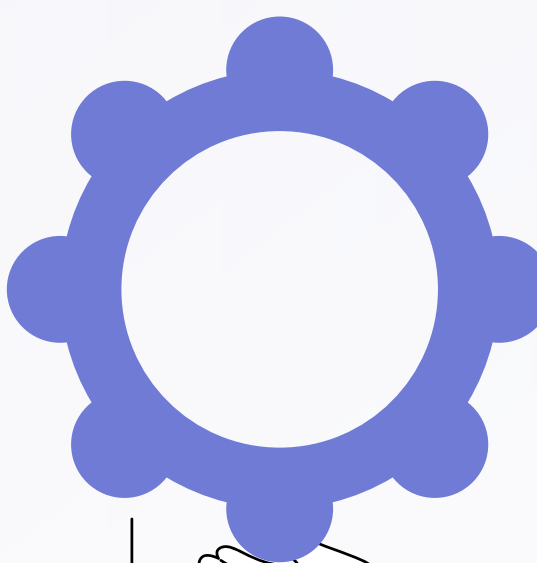
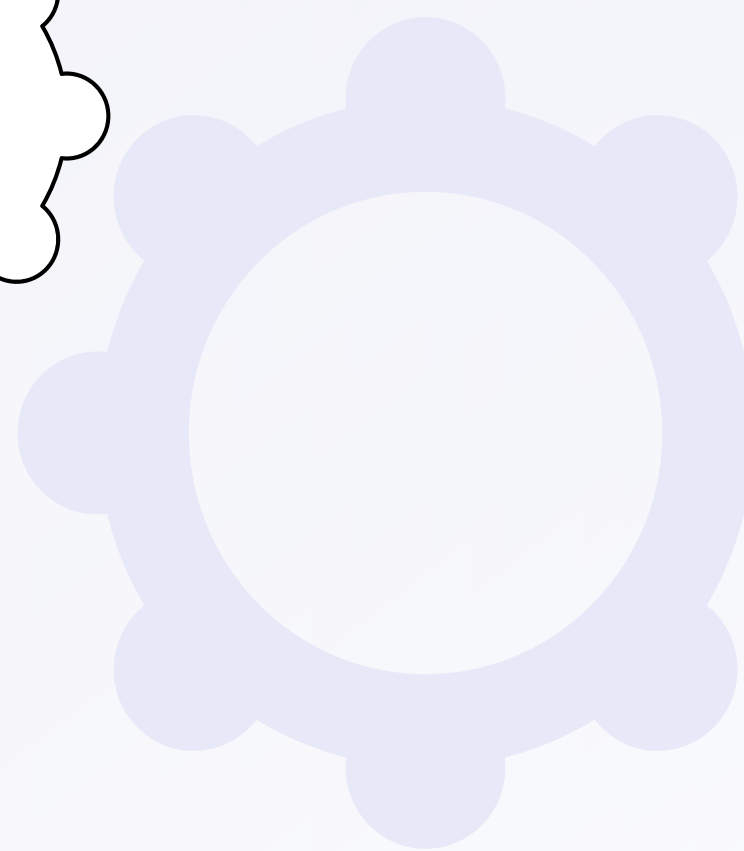
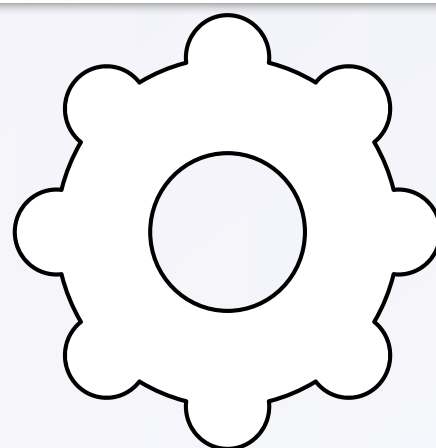
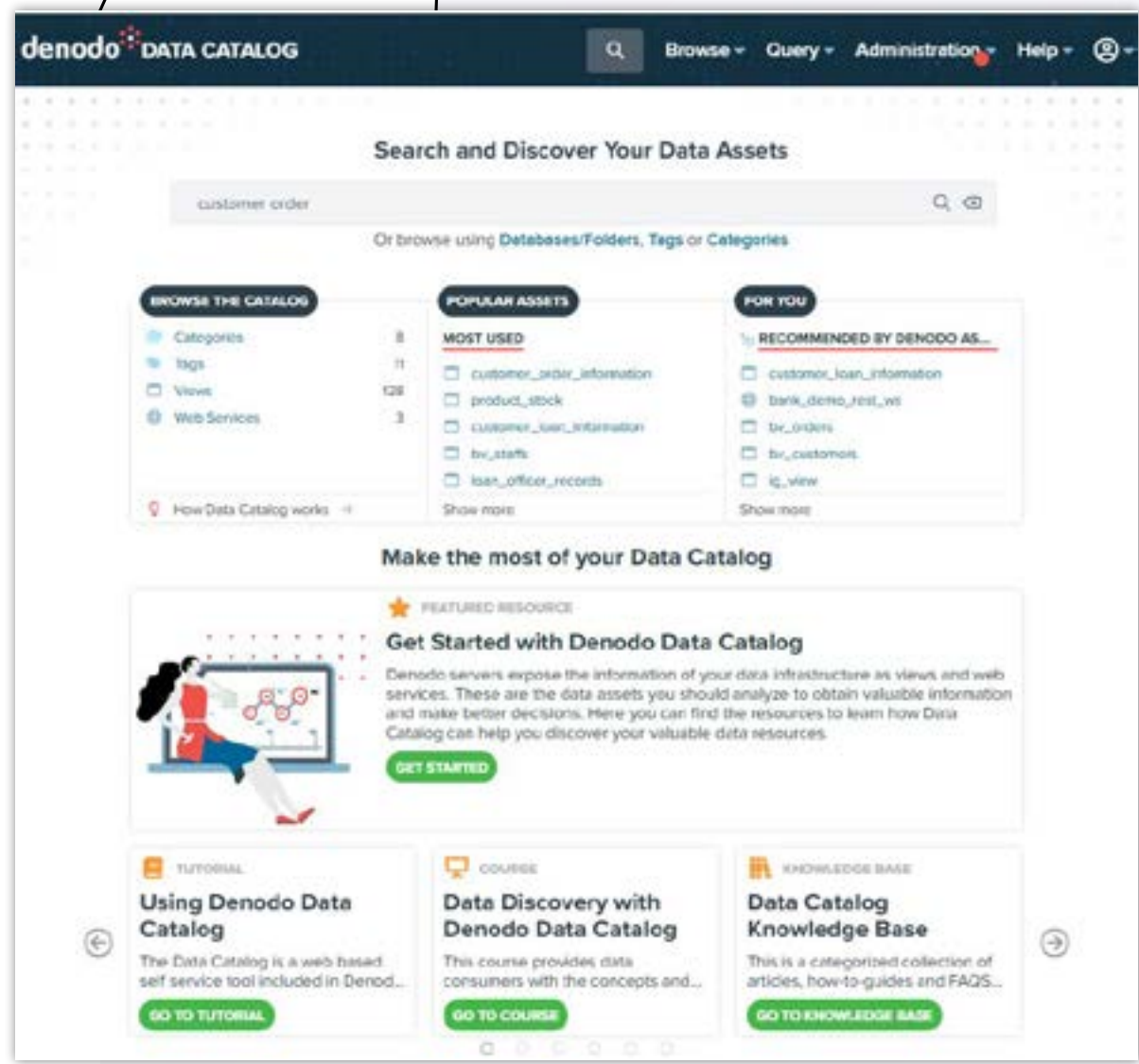
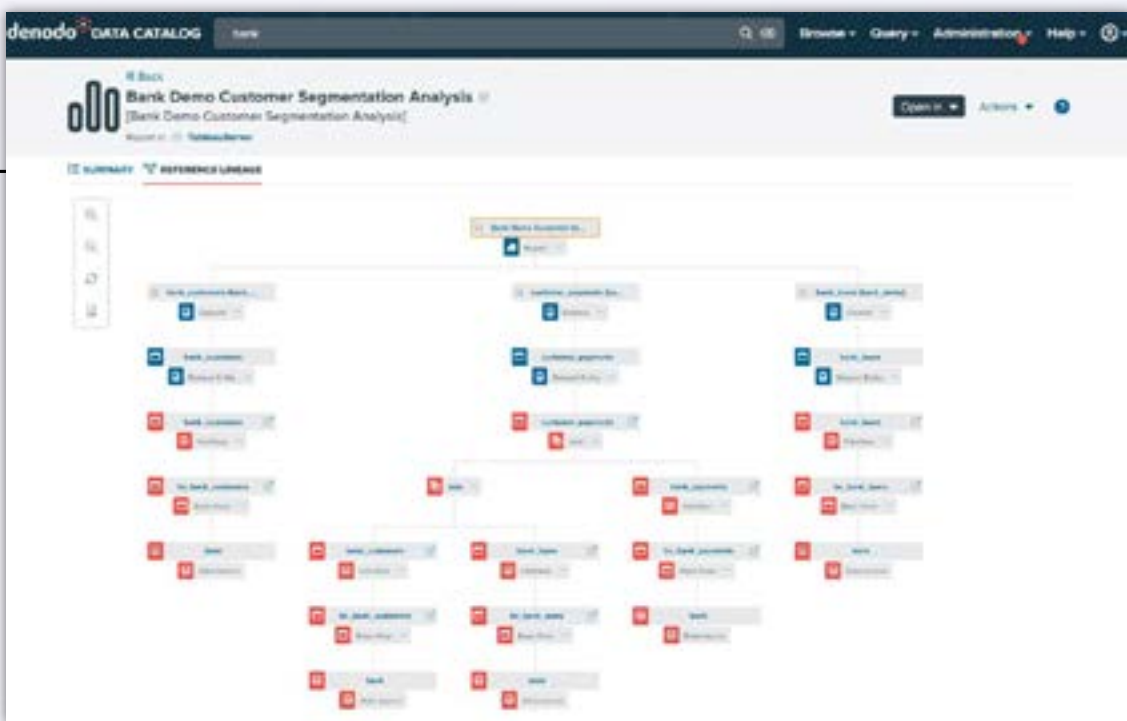
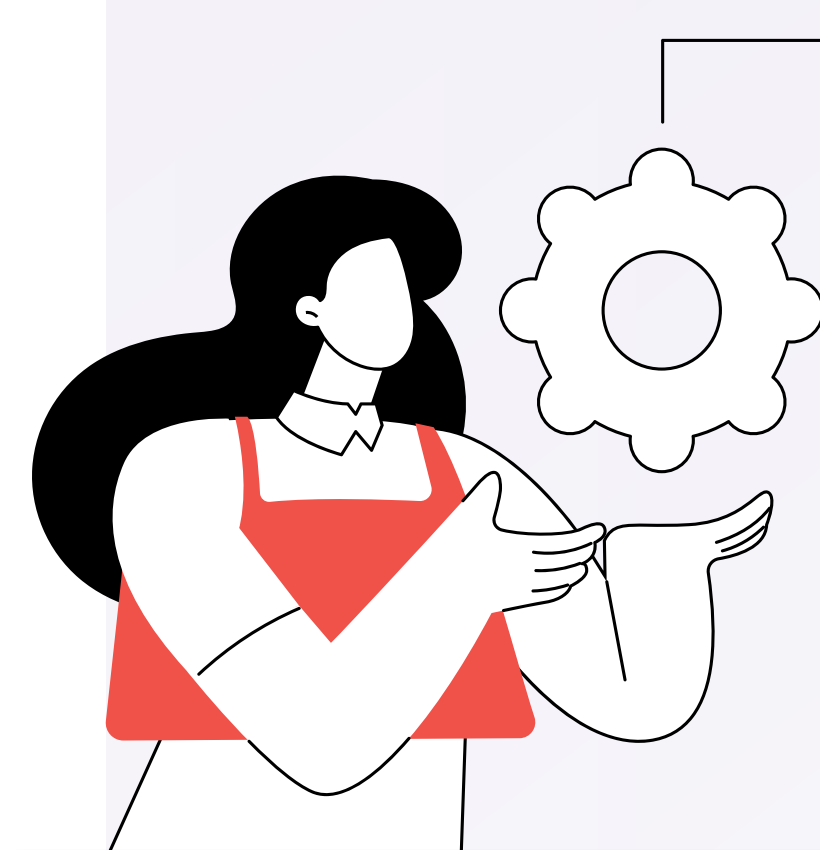


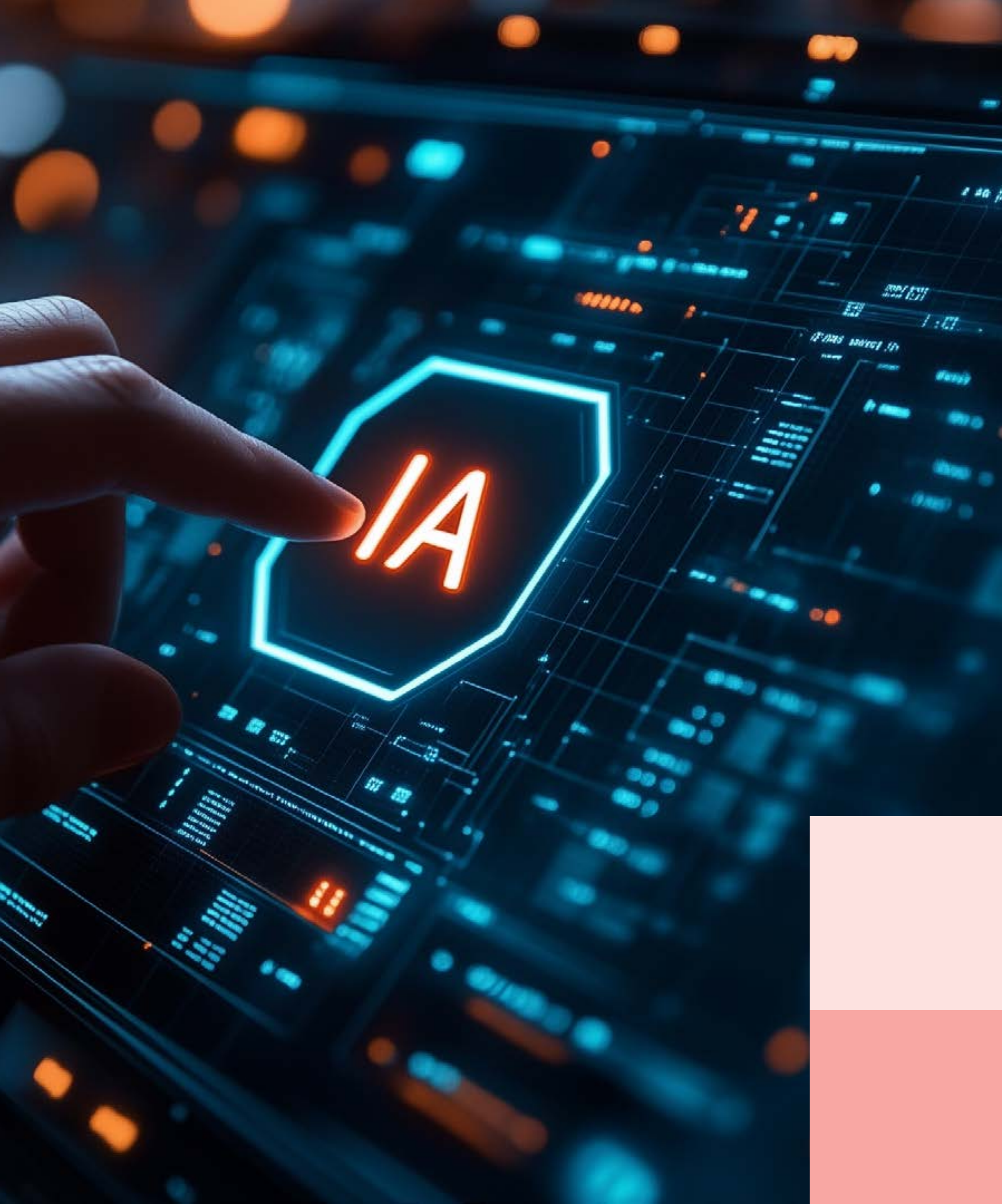
Data self-service capabilities enable users to easily explore, discover, and access the data they need without IT support.

Bridging the Gap Between Technical and Business Users

Databricks Unity Catalog provides a centralized repository for technical metadata and AI assets across workspaces, enabling data teams to define fine-grained access controls, track lineage, and streamline data discovery. However, its focus on technical governance makes it less ideal for business users seeking to explore and utilize data.

The Denodo Data Catalog complements Unity Catalog by offering a marketplace-like experience for business users, enabling them to explore, discover, and access data through an intuitive, codeless interface. By leveraging rich business semantics, active metadata, and detailed lineage, the Denodo Platform empowers anyone in the organization to find and use the data they need. Working together, the Denodo Catalog and Unity Catalog address the needs of both technical and business users, promoting data democratization while maintaining robust governance.





Data Self-Service

Simplifying Data Access with GenAI and Intelligent Data Discovery

The Denodo Platform simplifies data discovery by enabling business users to explore and query data using natural language through seamless integration with GenAI platforms like OpenAI, Azure OpenAI, and Amazon Bedrock. This eliminates complex query writing, while semantic models reduce hallucinations and improve query accuracy.

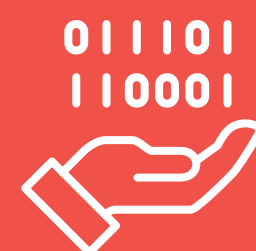
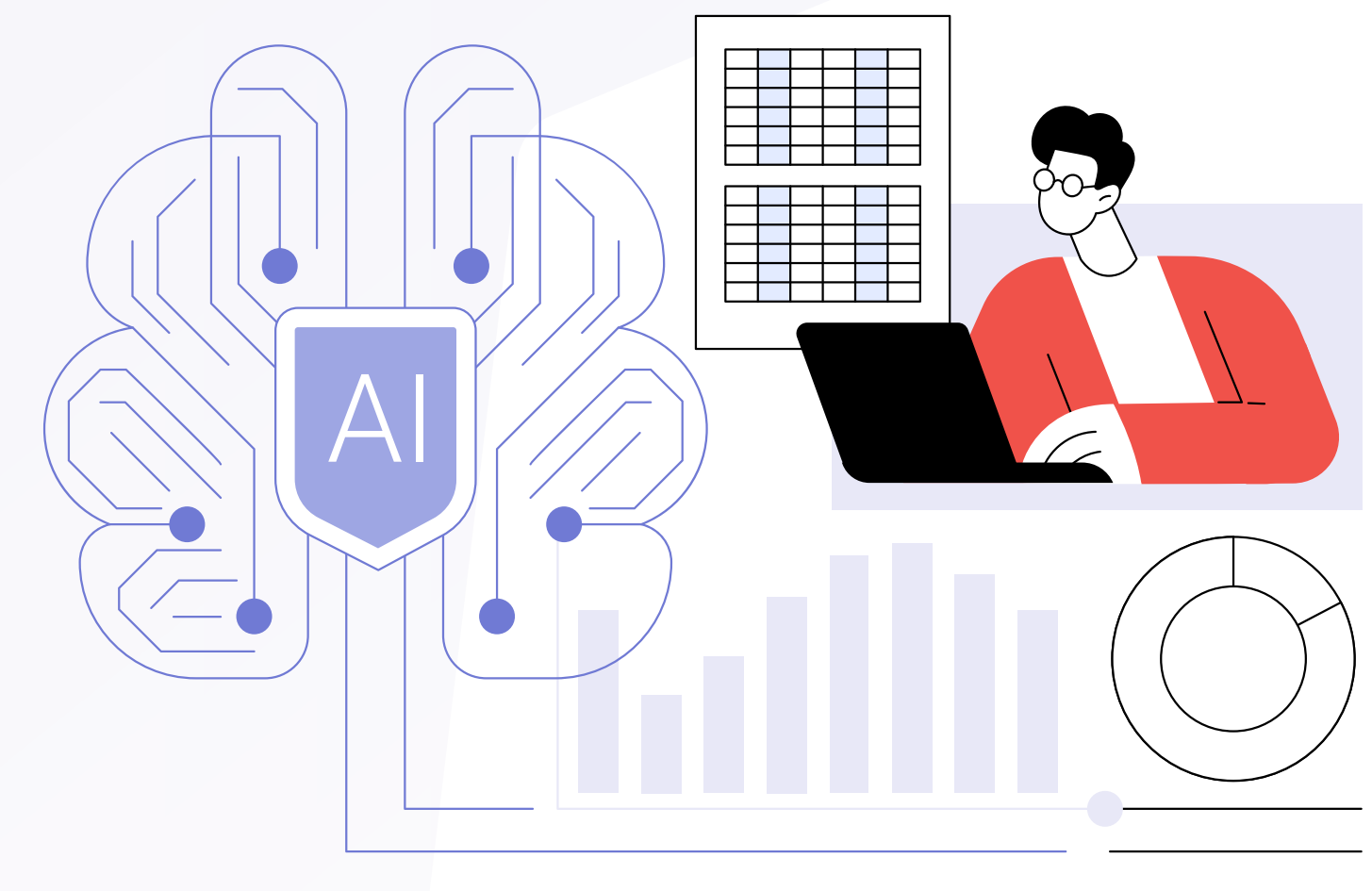
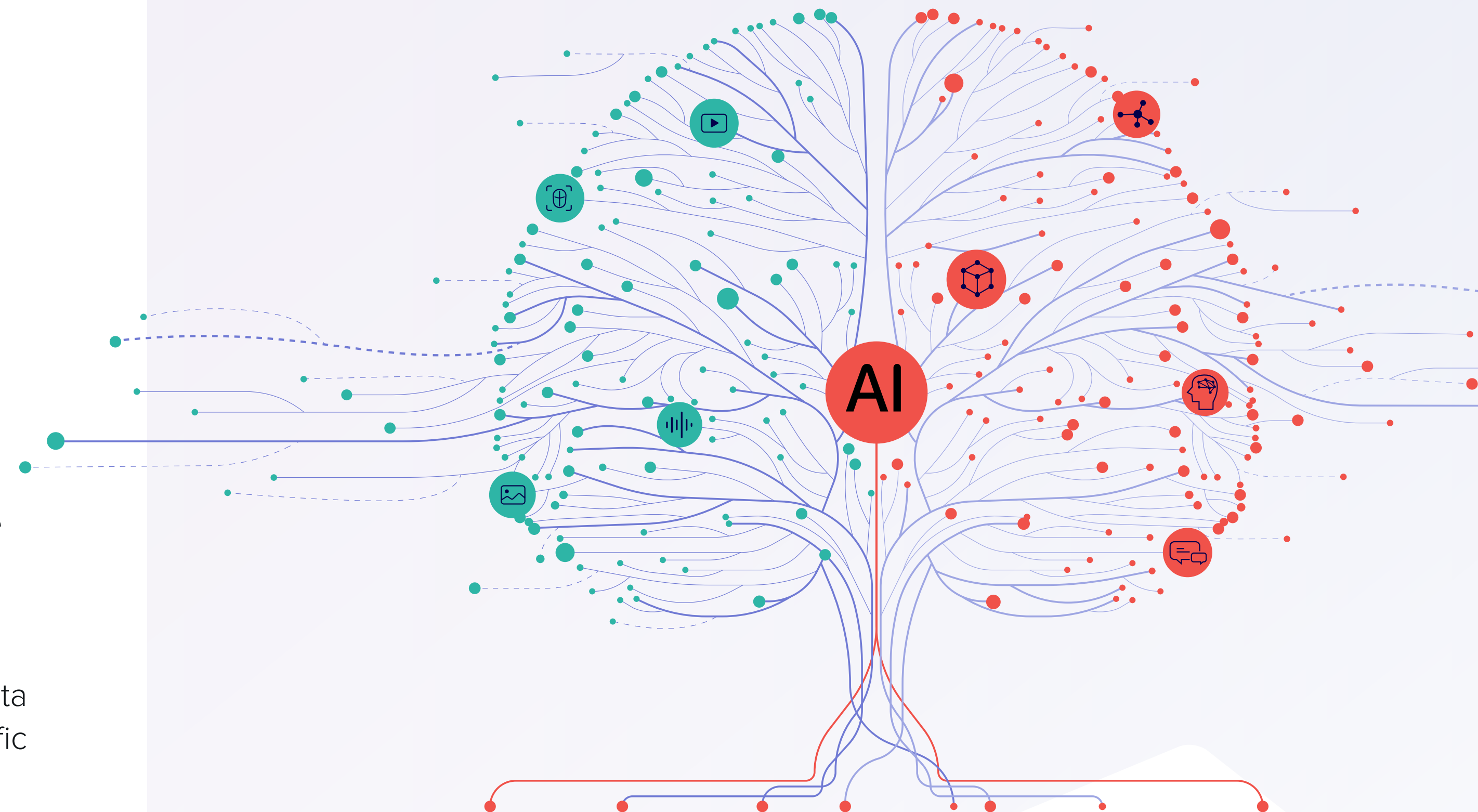
Combined with the Denodo Data Catalog, these features make it easier to find and leverage data, empowering non-technical users to gain insights and drive data-informed decisions.



AI-Ready Data for Foundational Models (LLMs)

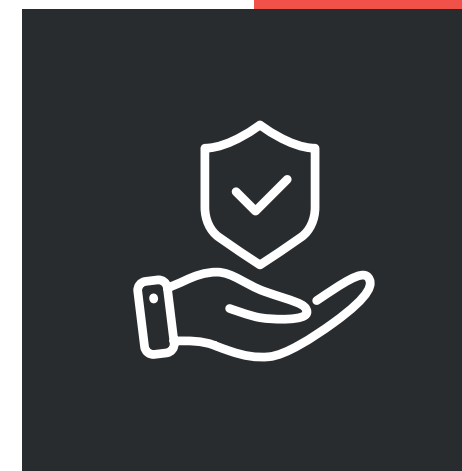
Foundational models/large language models (LLMs) benefit from the Denodo Platform's marketplace-like experience, in the same way as business users. By accessing rich metadata—such as data schemas, field descriptions, business-friendly field names, data relationships, and data profiles—LLMs can better understand the context of the data and pinpoint the most relevant and trustworthy information for specific queries.

The Denodo Platform streamlines data integration and the delivery of trusted data, enabling next generation AI applications to ground their responses in factual, reliable data. With the Denodo AI SDK, organizations can more easily build and deploy these applications, accelerating GenAI adoption while boosting performance and reliability.



CHAPTER 3:

FEDERATED DATA GOVERNANCE WITH CENTRALIZED OVERSIGHT



Logical data management solutions solve the data challenge. “Logical” means that it’s not tied to the restrictions of any one system or reporting tool.



Federated Data Governance with Centralized Oversight

A Unified Security Layer

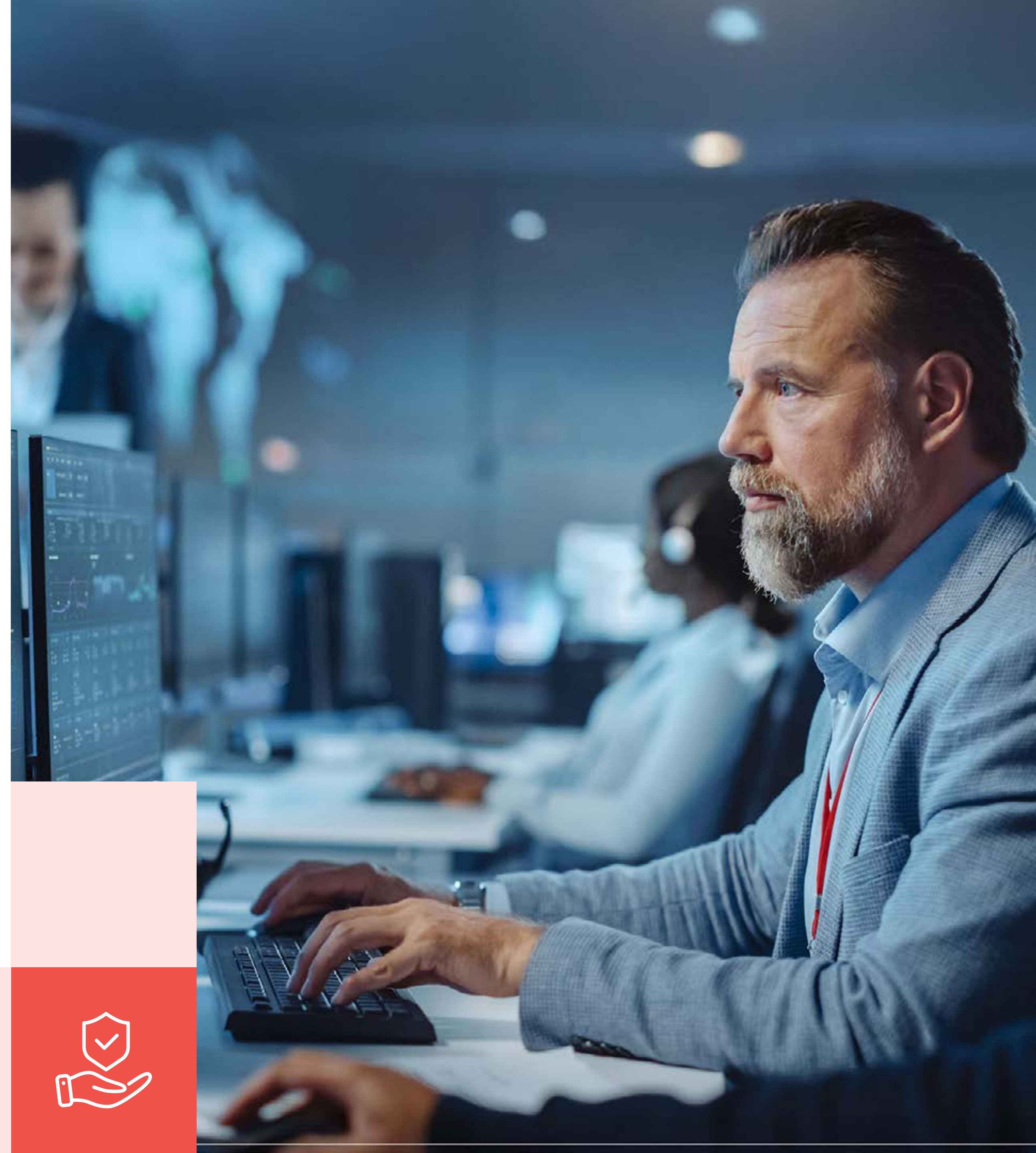
The unified security layer enabled by the Denodo Platform centralizes access control, securing all Databricks workspaces and distributed data sources through a single point. It supports fine-grained permissions—row, column, and cell-level access—with data masking and execution restrictions, and integrates with systems like LDAP, Active Directory, and Kerberos. This approach is widely used by organizations federating data product creation and ownership.

The Denodo Platform enables policy management at the domain level (including data products), empowering data owners to manage access and usage, while also supporting enterprise-wide controls to ensure consistent governance across all domains.

Federated Data Governance with Centralized Oversight

Advanced Semantics

The Denodo Platform's advanced semantics offer precise control over data access across sources, enabling organizations to implement fine-grained global policies without complex coding. By linking security policies to semantic tags and attributes, the Denodo Platform enables dynamic, context-aware permissions that account for user roles and data attributes. This approach also supports data masking, protecting sensitive information such as personally identifiable information (PII). With global policies and assigned tags, organizations can secure and manage access to sensitive data, regardless of format or origin.

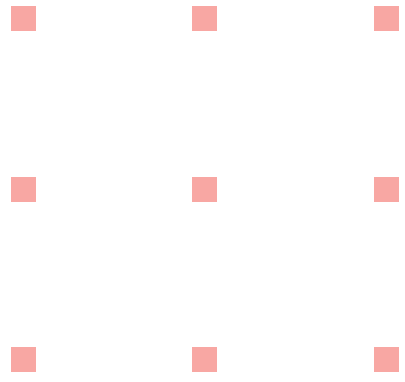




Federated Data Governance with Centralized Oversight

Robust Data Governance

The Denodo Platform offers comprehensive data governance features, including automated data source refreshes, change impact analysis, and full data lineage tracking. The Denodo Governance Bridge enables integration with leading governance tools, and its API enables the publishing of metadata and lineage information to platforms like Informatica EDC, Collibra, and others. Additionally, the Denodo Platform supports the importing of external data governance tags from Collibra and other data catalogs, to enable the consistent application of governance policies across the enterprise.



GETTING STARTED

Denodo and Databricks create a powerful combination, enhancing the strengths of each platform to help organizations maximize their data, analytics, and AI initiatives. While Databricks excels in advanced analytics and machine learning, the Denodo Platform complements this by providing a unified data access layer across Databricks and supporting data sources, seamless connectivity, and centralized data security and governance. This partnership enables all required data to be accessible, well-governed, and optimized for high-value tasks, reducing costs and enhancing overall efficiency.

Together, Denodo and Databricks enable organizations to break down data silos, streamline analytics workflows, and empower business users with self-service data access, driving innovation and accelerating time-to-insight. To discover how the Denodo Platform can further enhance your Databricks deployment, set up a time with one of our subject matter experts today and explore how we can help you unlock the full potential of your data landscape.

CONTACT US

