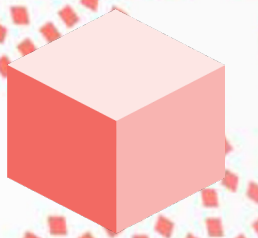# Dataversity Enterprise Data World

Denodo Data Architecture
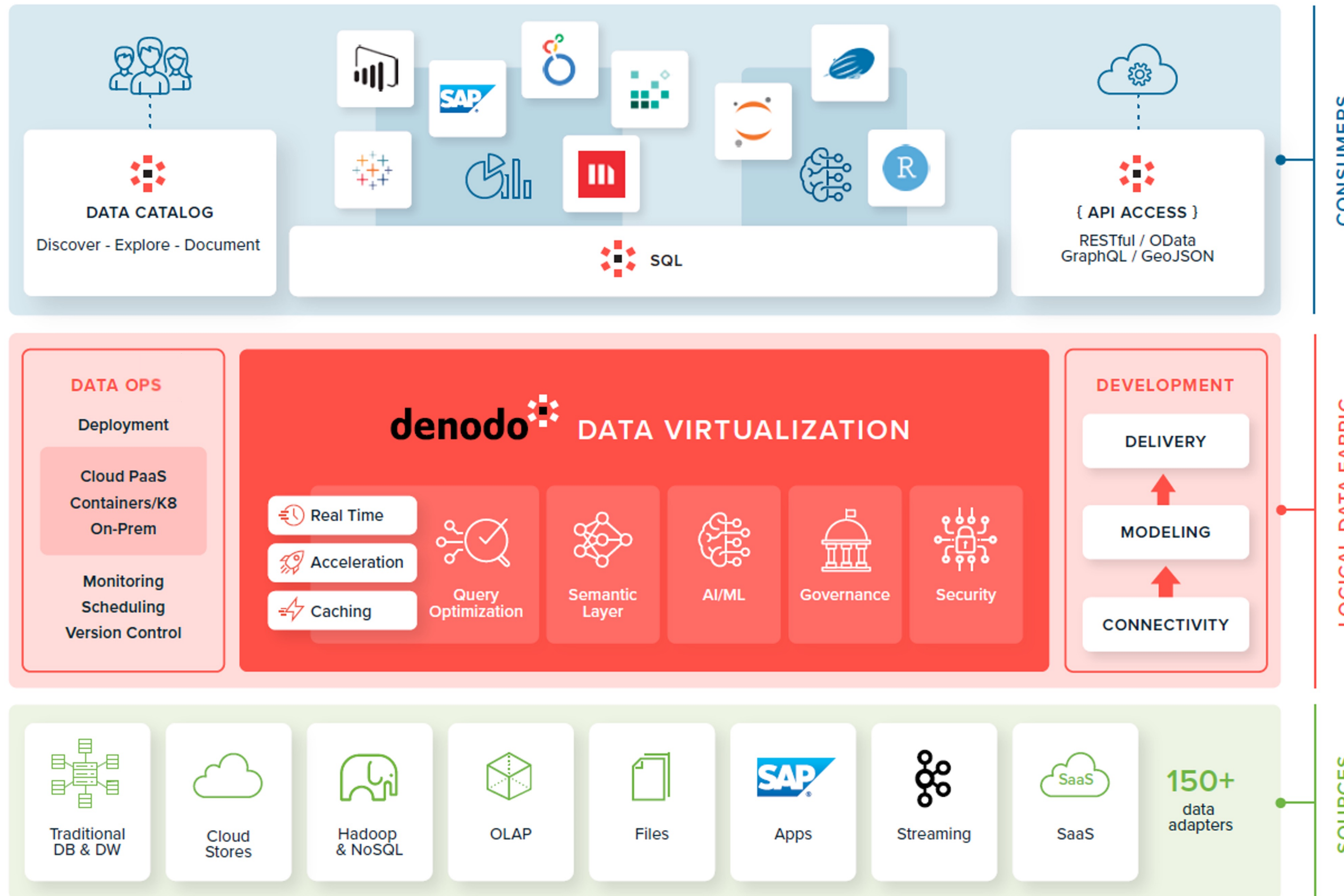
Sardor Davlenov

Sales Engineer – Partner and Channels

Denodo's Architecture - Overview

CONSUMERS

DATA CATALOG
Discover - Explore - Document

SQL

{ API ACCESS }
RESTful / OData
GraphQL / GeoJSON

LOGICAL DATA FABRIC

DATA OPS
Deployment

Cloud PaaS
Containers/K8
On-Prem

Monitoring
Scheduling
Version Control

denodo DATA VIRTUALIZATION

Real Time
Acceleration
Caching

Query Optimization
Semantic Layer
AI/ML
Governance
Security

DEVELOPMENT

DELIVERY
MODELING
CONNECTIVITY

SOURCES

Traditional DB & DW
Cloud Stores
Hadoop & NoSQL
OLAP
Files
Apps
Streaming
SaaS

150+ data adapters

Benefits for IT and Data Architects

Centralized Security

    Secure, controlled delivery of any data

    Integration with IdPs, SSO and AD

Higher Control

    Decouple processing and storage from data delivery

    Provides flexibility to modify the back-end without impacting the business

        Migration to cloud

Lower Cost

    Reduces the need to create and manage unnecessary copies of data

    Allows easily reusing transformation logic and policies across multiple use cases

denodo

Benefits for End Users

Increased agility and better TTM

    Single place to access any data

    Hides complexity of back-ends

        On-prem vs. Cloud

        Formats and protocols

Trust

    Consistency of results regardless of access method

    Documentation, lineage, governance, etc.

Reusability

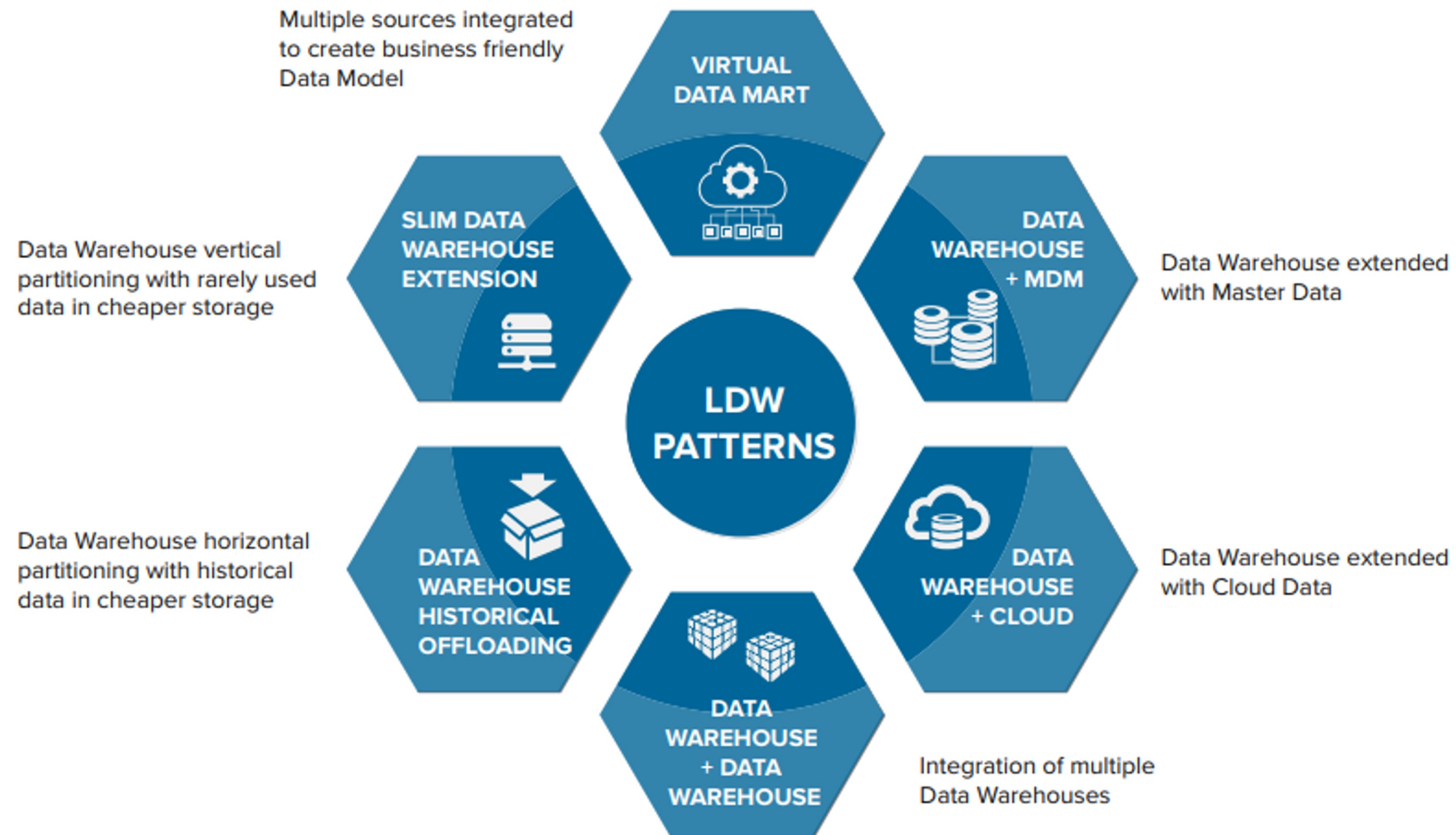   Access to same data and metrics from anywhere as logic is not coded in a BI tool

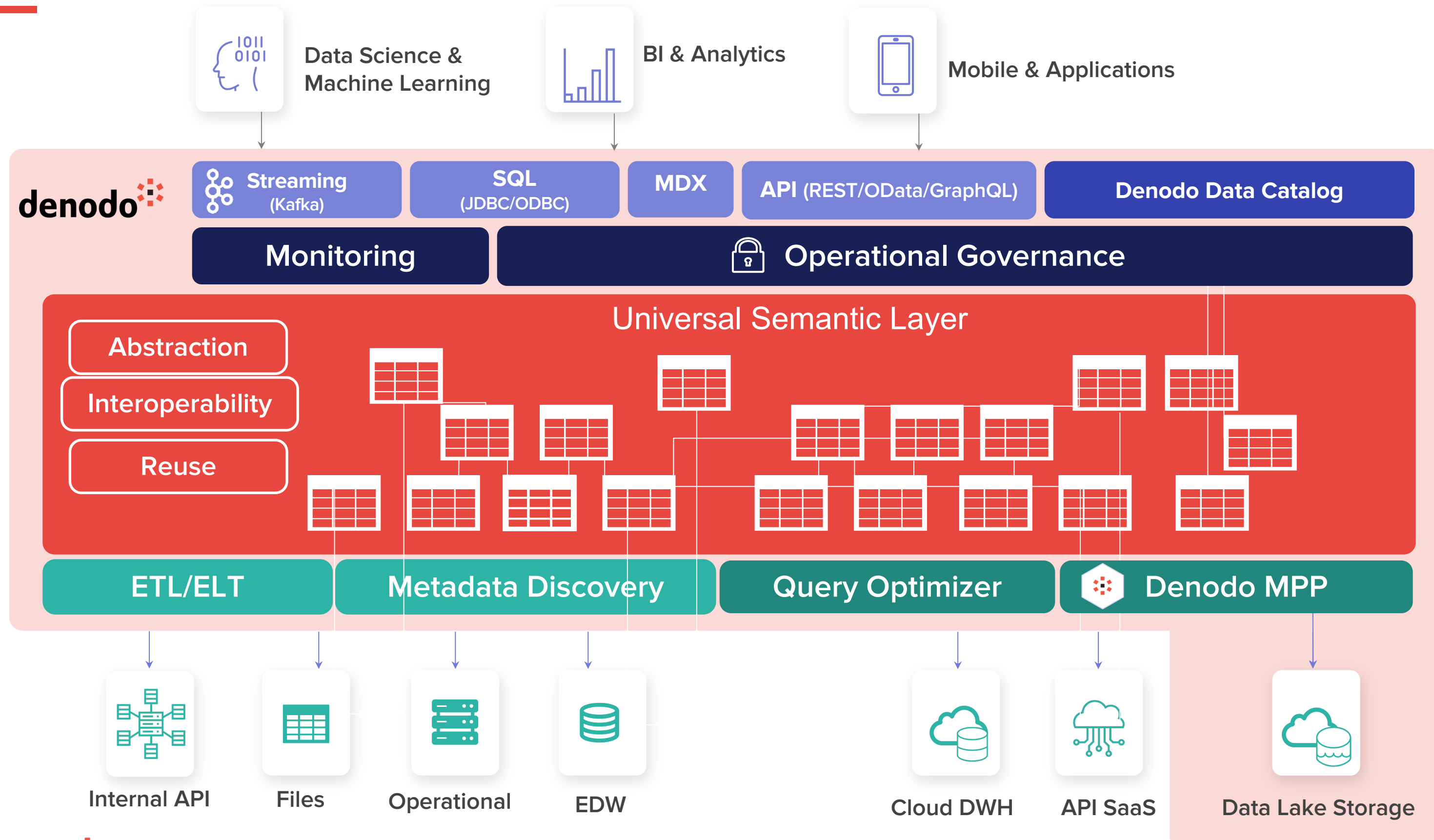   Any data available in multiple standardized formats: SQL, REST, OData, GraphQL, etc.

denodo

Logical Data Warehouse – Six Patterns

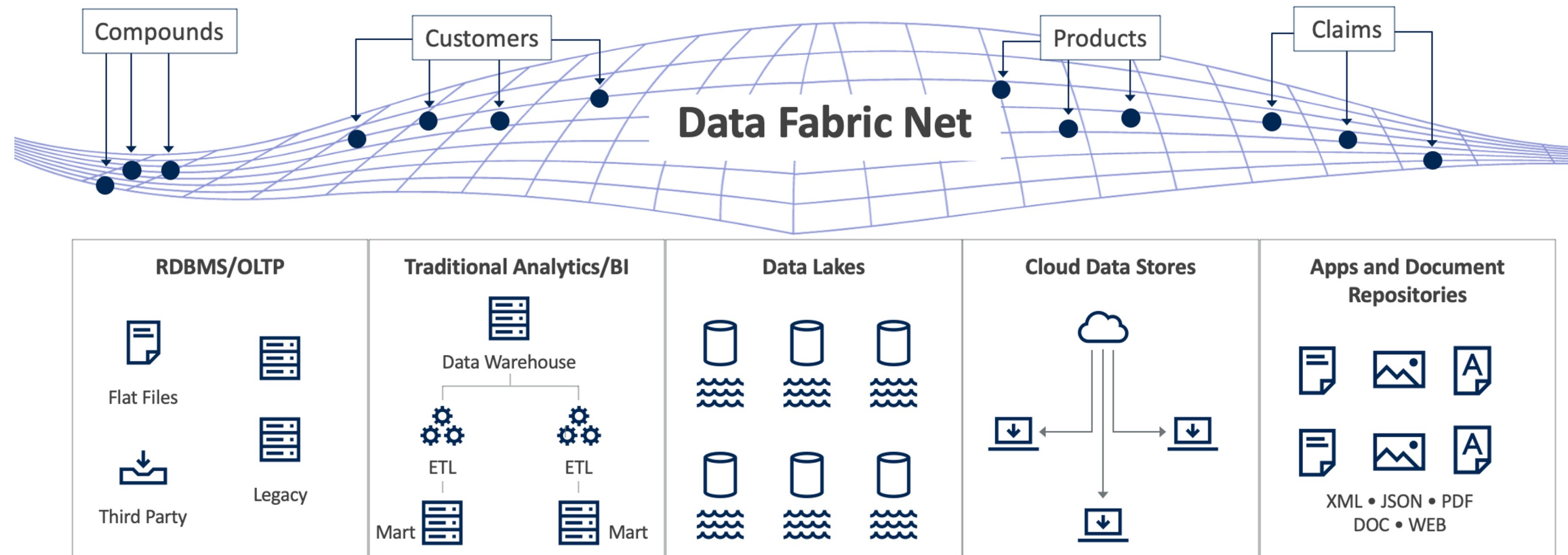Multiple sources integrated to create business friendly Data Model

VIRTUAL DATA MART

SLIM DATA WAREHOUSE EXTENSION

Data Warehouse vertical partitioning with rarely used data in cheaper storage

DATA WAREHOUSE + MDM

Data Warehouse extended with Master Data

LDW PATTERNS

Data Warehouse horizontal partitioning with historical data in cheaper storage

DATA WAREHOUSE HISTORICAL OFFLOADING

DATA WAREHOUSE + CLOUD

Data Warehouse extended with Cloud Data

DATA WAREHOUSE + DATA WAREHOUSE

Integration of multiple Data Warehouses

denodo

Logical Data Warehouse

Data Science & Machine Learning

BI & Analytics

Mobile & Applications

denodo

| Streaming (Kafka) | SQL (JDBC/ODBC) | MDX | API (REST/OData/GraphQL) | Denodo Data Catalog |

| Monitoring | Operational Governance |

Universal Semantic Layer

- Abstraction
- Interoperability
- Reuse

| ETL/ELT | Metadata Discovery | Query Optimizer | Denodo MPP |

Internal API

Files

Operational

EDW

Cloud DWH

API SaaS

Data Lake Storage

denodo

Data Fabric

Data Fabric for End-to-End Data Management

In Layman Terms

- "Integrate data" from disparate data sources

- Securely deliver an "integrated view" of the different data objects

- Consume the "integrated data" for analytics and operational purposes

- Automate the entire process using AI/ML

Logical Data Fabric with Denodo

Powered by Data Virtualization

This is where Data Virtualization can help, acting as the delivery layer in a Logical Data Fabric.

Logical Data Fabric Powered by Data Virtualization:

- Enables fast on-boarding of data, abstracting the complexity of the different technologies in which the data resides

- Enables fast creation of semantic data objects that are easily discoverable to support self-service

- Ensures consistent security policies are applied across all the data platforms integrated by the data fabric

- Enables universal access to the data objects to support any consumer, using common access methods
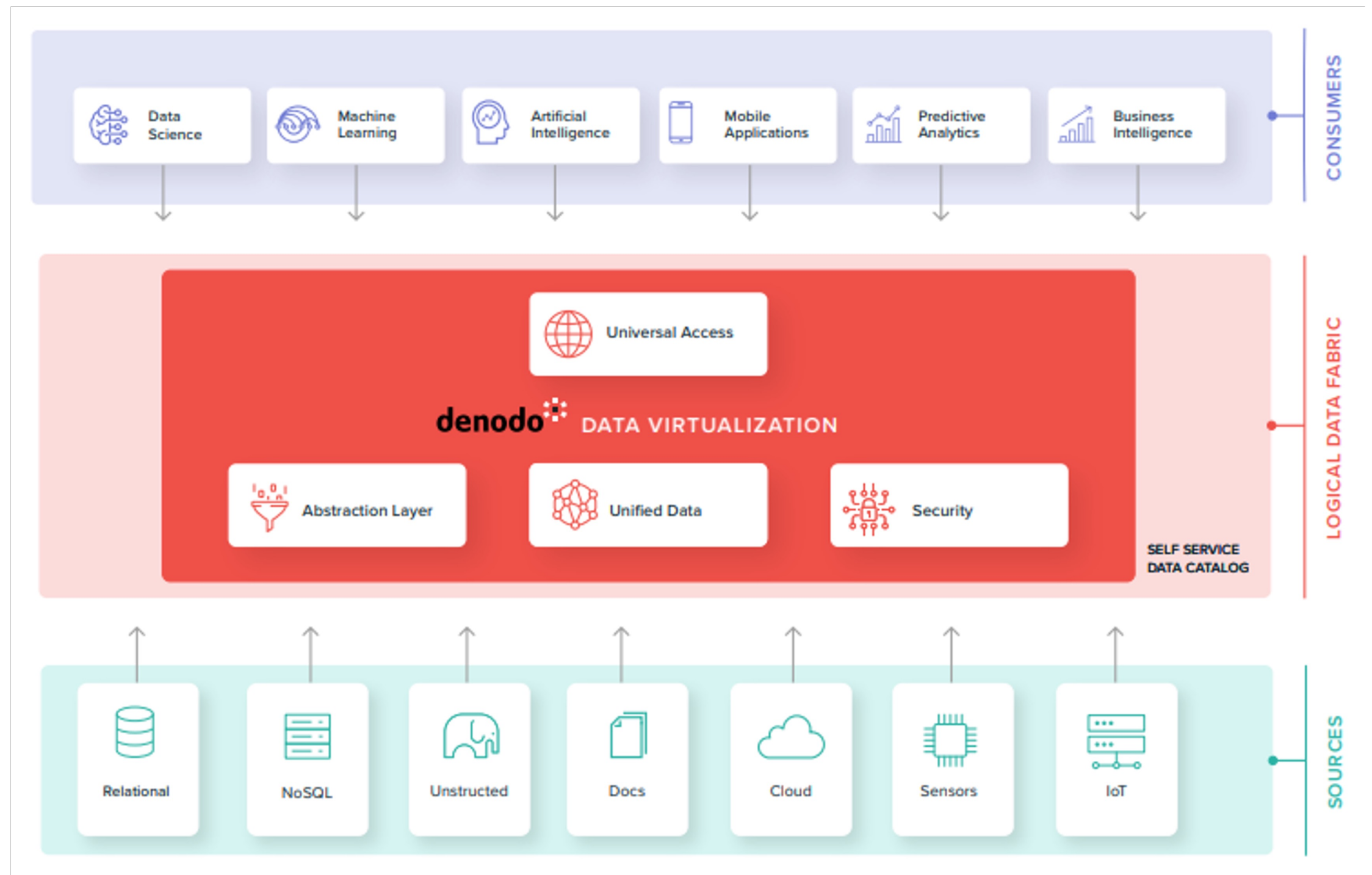
Multi-Region Multi-Cloud Architecture

Denodo Hybrid Deployment (On-premise & Cloud)

We have data sources with significant amount of data both on-prem and in the Cloud

- Recommended architecture whenever there are data sources with significant amount of data both on-prem and in the cloud
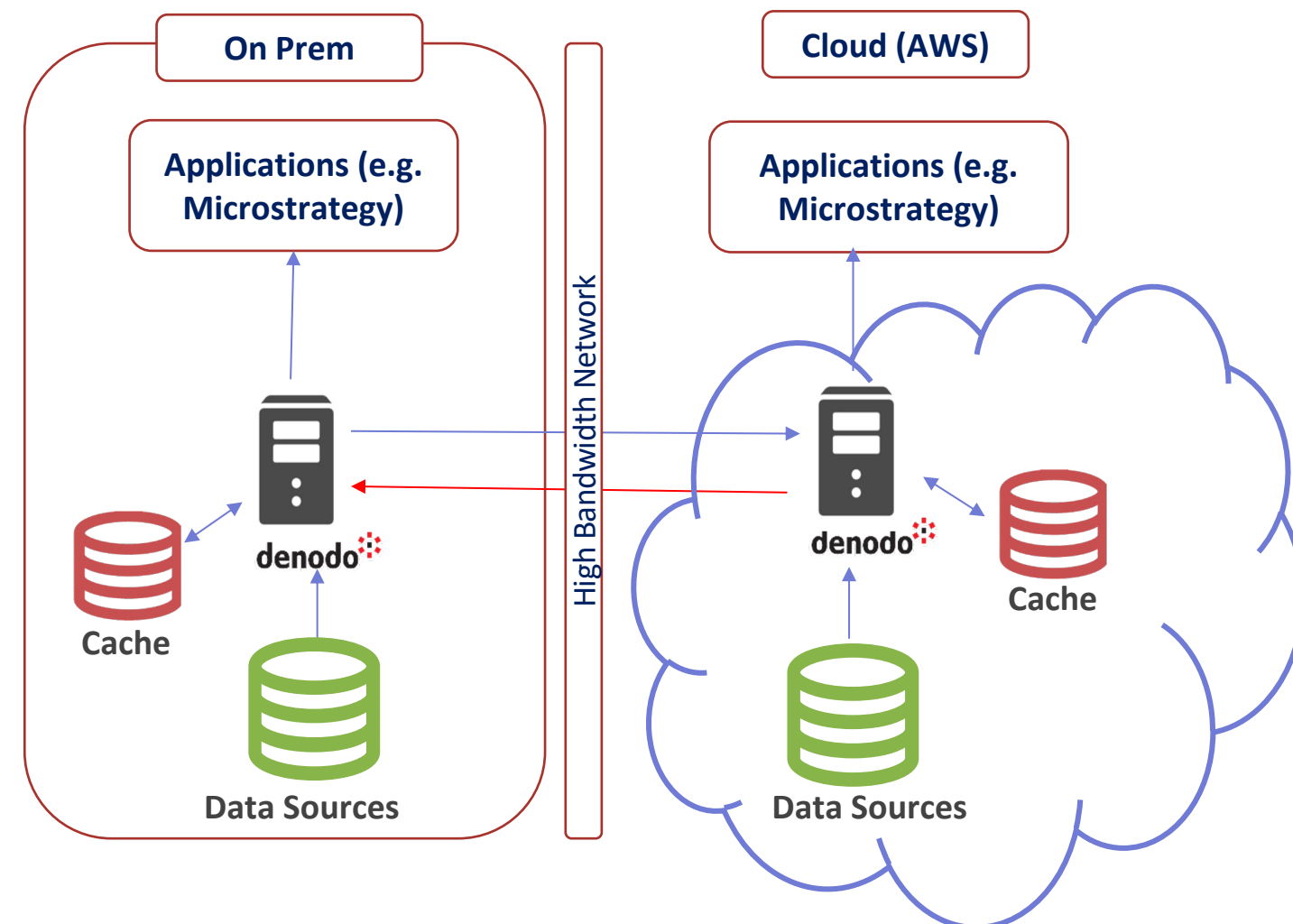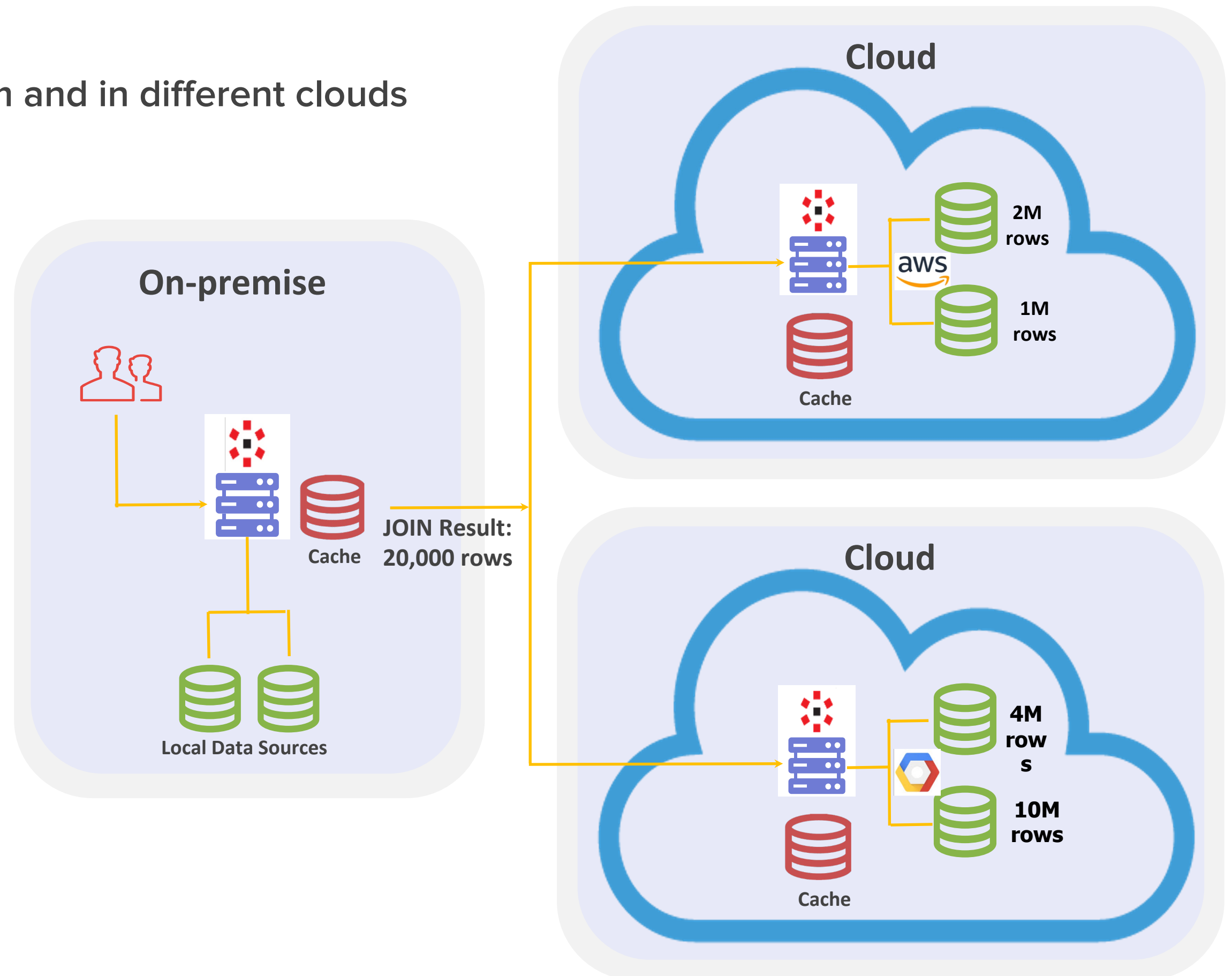

- Denodo can be configured in a layered architecture
- Applications in each zone (on-premise and cloud) access data from the other zone via Denodo.
- On-premise applications have access to external data via the local Denodo instance leveraging the performance optimisations provided by Denodo in both zones.
- Ensures only aggregated data from the external applications is transferred across the network in both directions, in response to requests from consumers.
- Assuming significant amounts of data and applications are hosted both on-premise and in the cloud, minimises the network overhead between on-premise and cloud installations
- Considerations of increased physical hardware/software maintenance costs and scalability in comparison to the cloud.

Denodo in a Multi-location Multi-Cloud Architecture

Significant amount of data on-prem and in different clouds
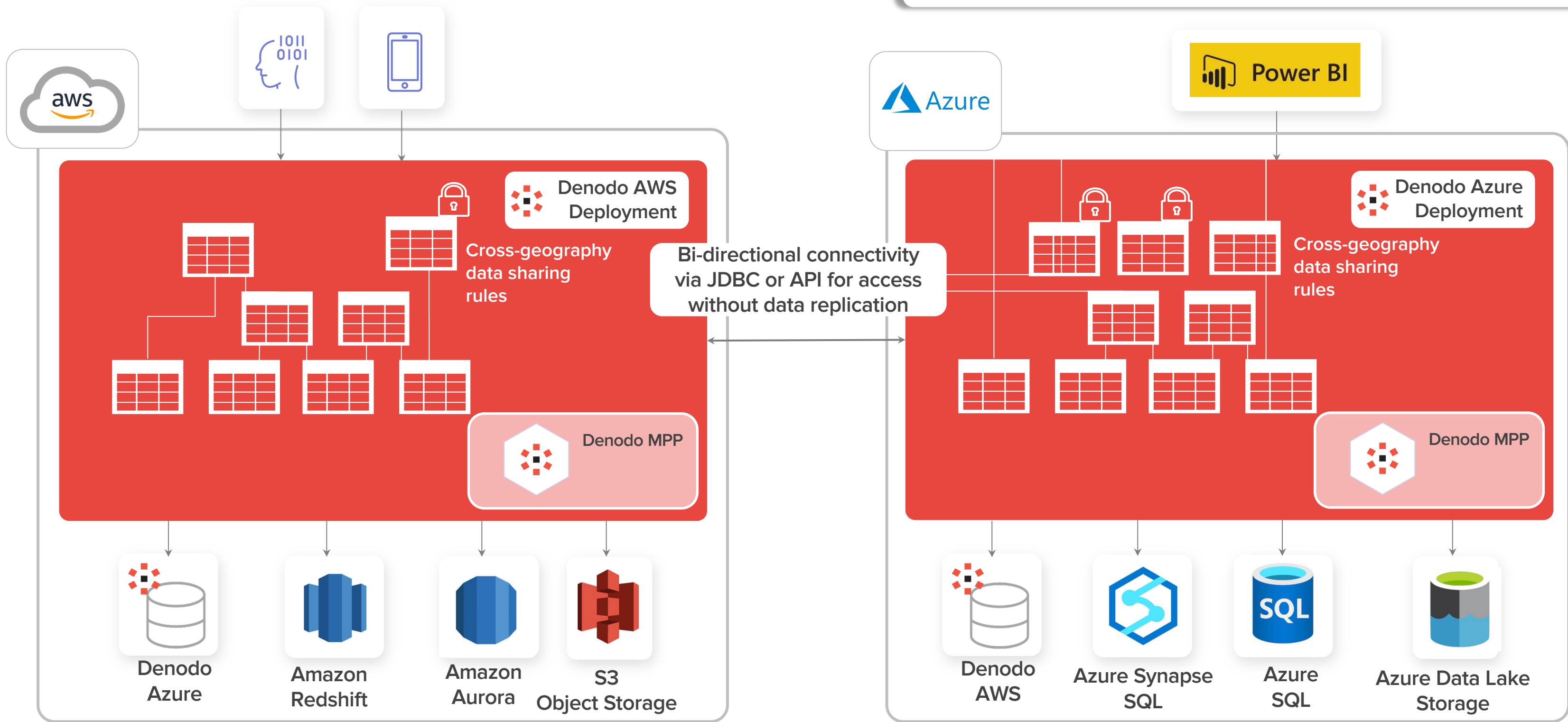
- Recommended architecture whenever we have on-prem and multi-cloud data sources with a significant amount of data at each location

- Denodo-to-Denodo connections are also possible, enabling convenient *layered architectures*

- Costly combinations can be *delegated to local instances*, avoiding unnecessary, expensive data movement.

Cloud

2M rows

1M rows

Cache

On-premise

Cache

JOIN Result: 20,000 rows

Local Data Sources

Cloud

4M rows

10M rows

Cache

Multi-Cloud/Multi-Region Pattern
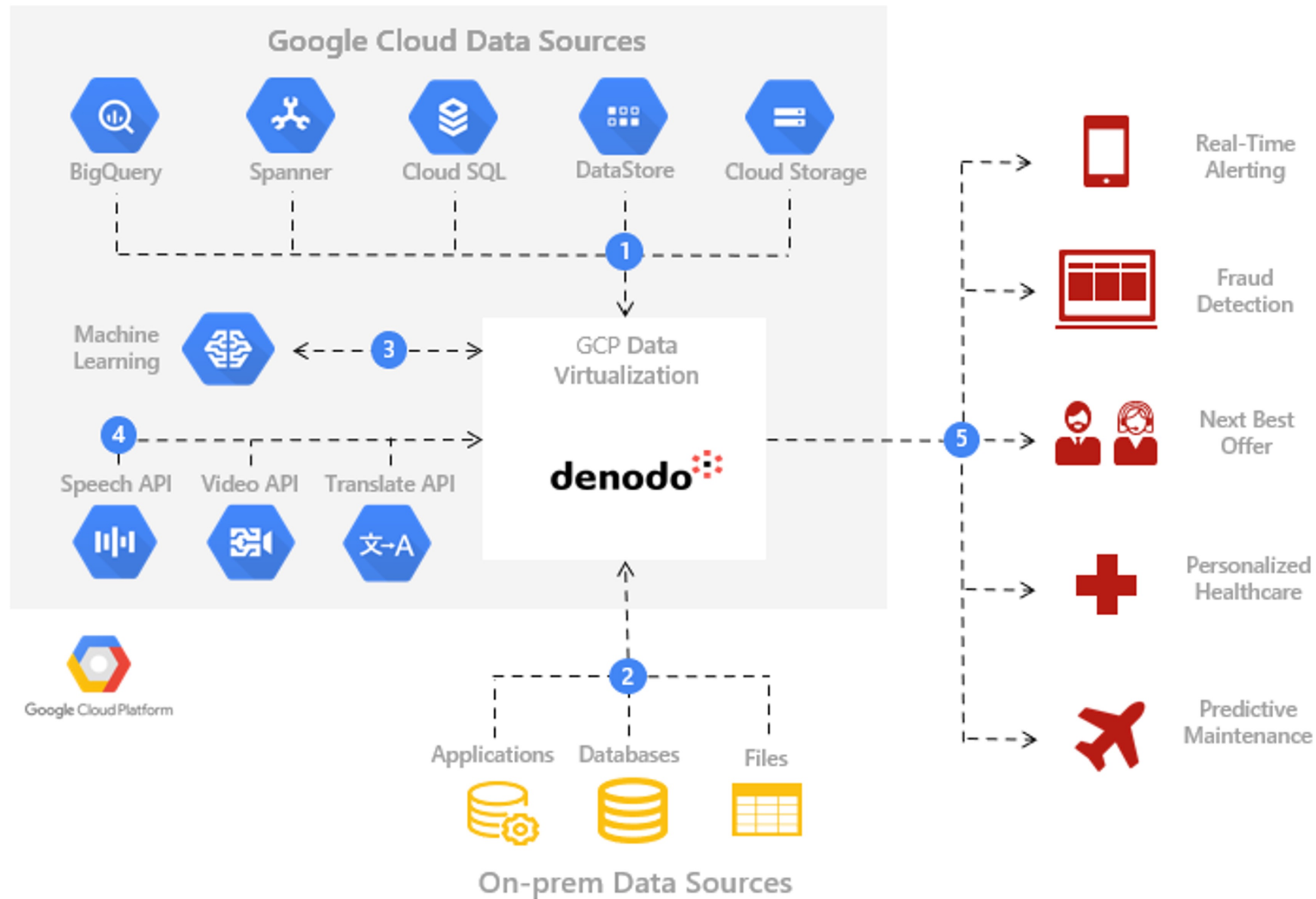
Other Architectures and Use Cases

Real-time Analytics on Big Data – Microsoft Azure
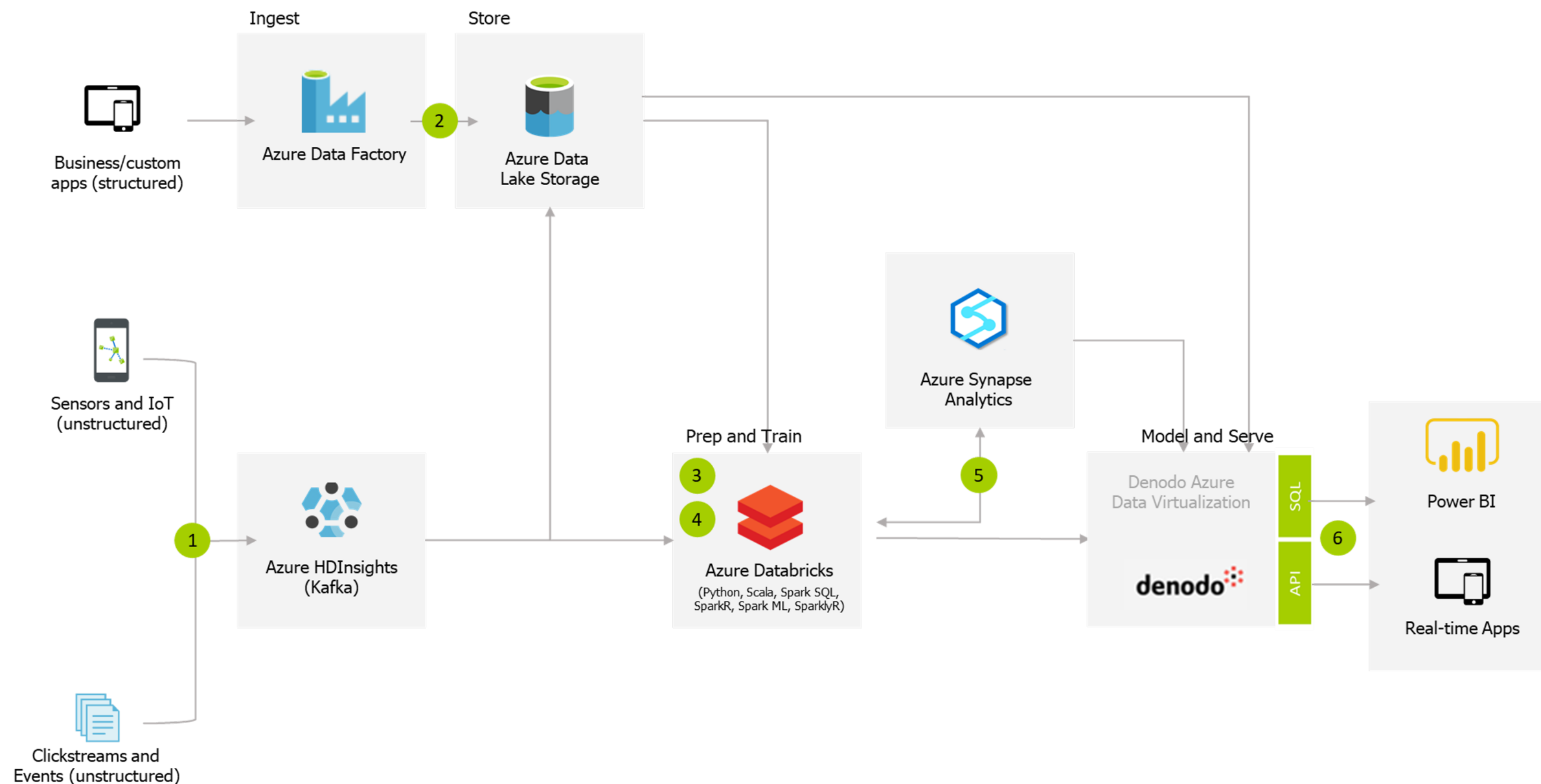
DATA FLOW

1. Easily ingest live streaming data for an application using Apache Kafka cluster in Azure HDInsight.

2. Bring together all your structured data using Azure Data Factory to Azure Blob Storage.

3. Take advantage of Azure Databricks to clean, transform, and analyze the streaming data, and combine it with structured data from operational databases or data warehouses.

4. Use scalable machine learning/deep learning techniques, to derive deeper insights from this data using Python, R or Scala, with inbuilt notebook experiences in Azure Databricks.

5. Leverage native connectors between Azure Databricks and Azure Synapse or use Denodo Platform bulk load APIs to move data between Azure Databricks and Azure Synapse Analytics
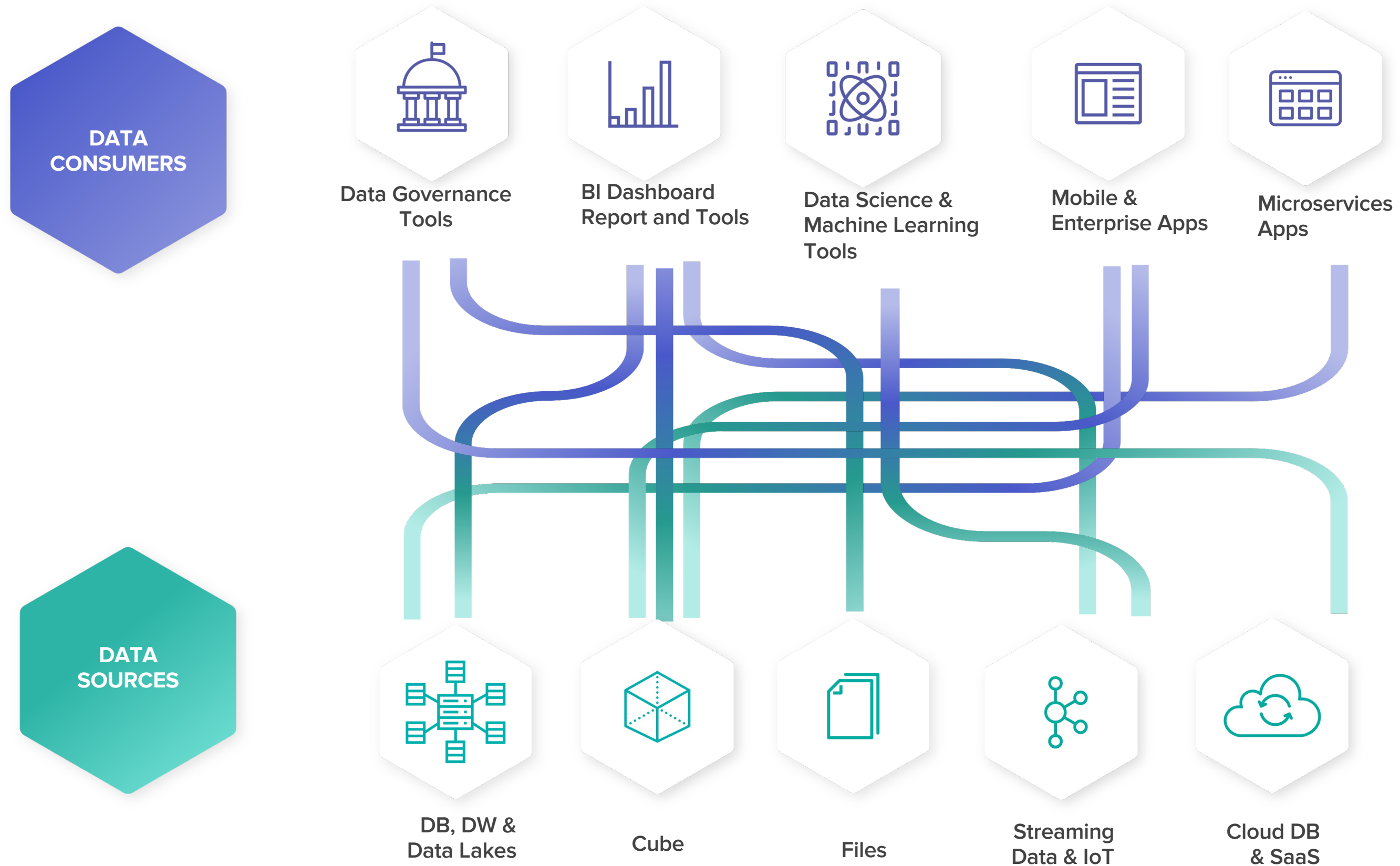
Operational AI – Google Cloud

Google Cloud Data Sources

BigQuery | Spanner | Cloud SQL | DataStore | Cloud Storage

Machine Learning

GCP Data Virtualization

Speech API | Video API | Translate API

denodo

Google Cloud Platform

Applications | Databases | Files

On-prem Data Sources

Real-Time Alerting

Fraud Detection

Next Best Offer
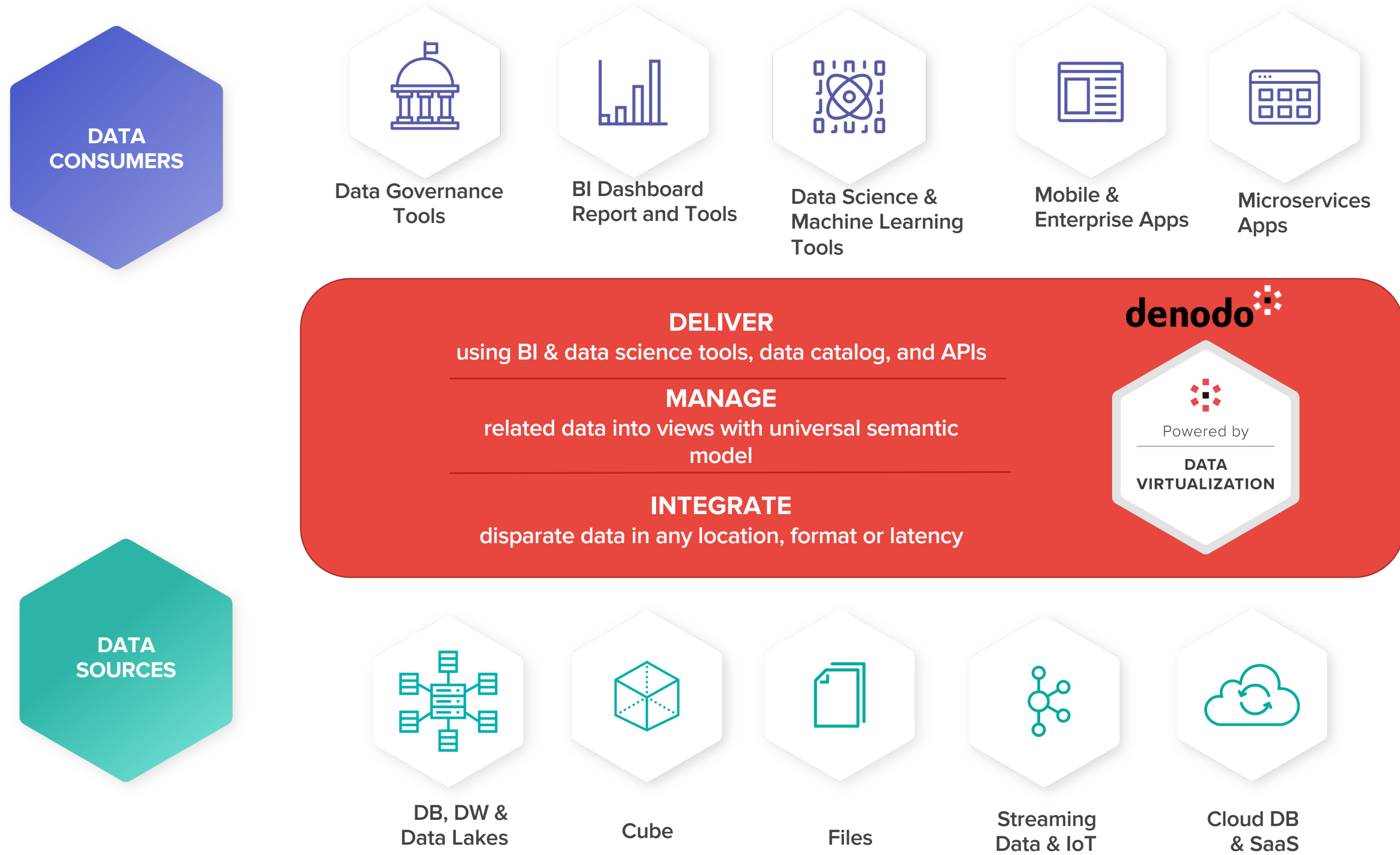
Personalized Healthcare

Predictive Maintenance

DATA FLOW

1. GCP-based data sources (BigQuery, Spanner, Cloud SQL, DataStore, Cloud Storage) connected to Denodo Data Virtualization platform providing unified hybrid abstraction layer.

2. Third-party on-premise (SAP, Oracle) and cloud data (Salesforce etc.) connected to Denodo Data Virtualization platform.

3. Google Machine Learning as well as Google Speech, Video, and Translate API can be utilized to get the valuable insights and provide real-time alerts, recommendations, scoring etc.

4.

5. All connected data sources are combined, secured, and exposed as Data Services over SQL and API interfaces. Exposed virtual datasets consumed by real-time dashboards, portals, Looker BI, analytical applications, any real-time and mobile applications.

denodo

Distributed Data Landscape

DATA CONSUMERS

Data Governance Tools

BI Dashboard Report and Tools

Data Science & Machine Learning Tools

Mobile & Enterprise Apps

Microservices Apps

DATA SOURCES

DB, DW & Data Lakes

Cube

Files

Streaming Data & IoT

Cloud DB & SaaS

denodo

Distributed Data Landscape - Simplified

DATA CONSUMERS

Data Governance Tools

BI Dashboard Report and Tools

Data Science & Machine Learning Tools

Mobile & Enterprise Apps

Microservices Apps

denodo

DELIVER
using BI & data science tools, data catalog, and APIs

MANAGE
related data into views with universal semantic model

INTEGRATE
disparate data in any location, format or latency

Powered by
DATA VIRTUALIZATION

DATA SOURCES

DB, DW & Data Lakes

Cube

Files

Streaming Data & IoT

Cloud DB & SaaS

denodo

Next Steps to Take:

1. Watch a Webinar to learn more - Building the Foundation of a Modern Data Architecure

Watch On-Demand & Learn:

● The need for a modern cloud data architecture
● Trends and Best Practices for a modern data stack
● The foundational capabilities required when building a Data Fabric and/or Data Mesh architecture
● How to simplify your data governance and security programs

1. Take a look at our Whitepaper - The Denodo Platform 8.0 Technical Architecture

This whitepaper introduces the Denodo Platform 8.0, covering all the features available not only in this release but also in previous ones, with special emphasis on the latest features:

● Enhanced, Unified, Web-Based User Experience
● Performance Acceleration for Complex Analytical Queries in Logical Data Warehouse / Data Fabric Scenarios
● PaaS support for Cloud and Hybrid Environments
● Enhanced Data Services APIs with Graph-Like Access to Denodo Views
● Going Beyond the Logical Data Warehouse: New Support for Data Science and ML
● Enhanced User Experience in the Data Catalog

1. Play with Denodo Platform on your own:
   Use Denodo Express for personal or prototype projects:

1. Register for a 30-day Free Trial (Denodo Professional)
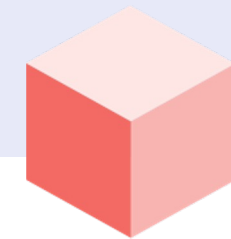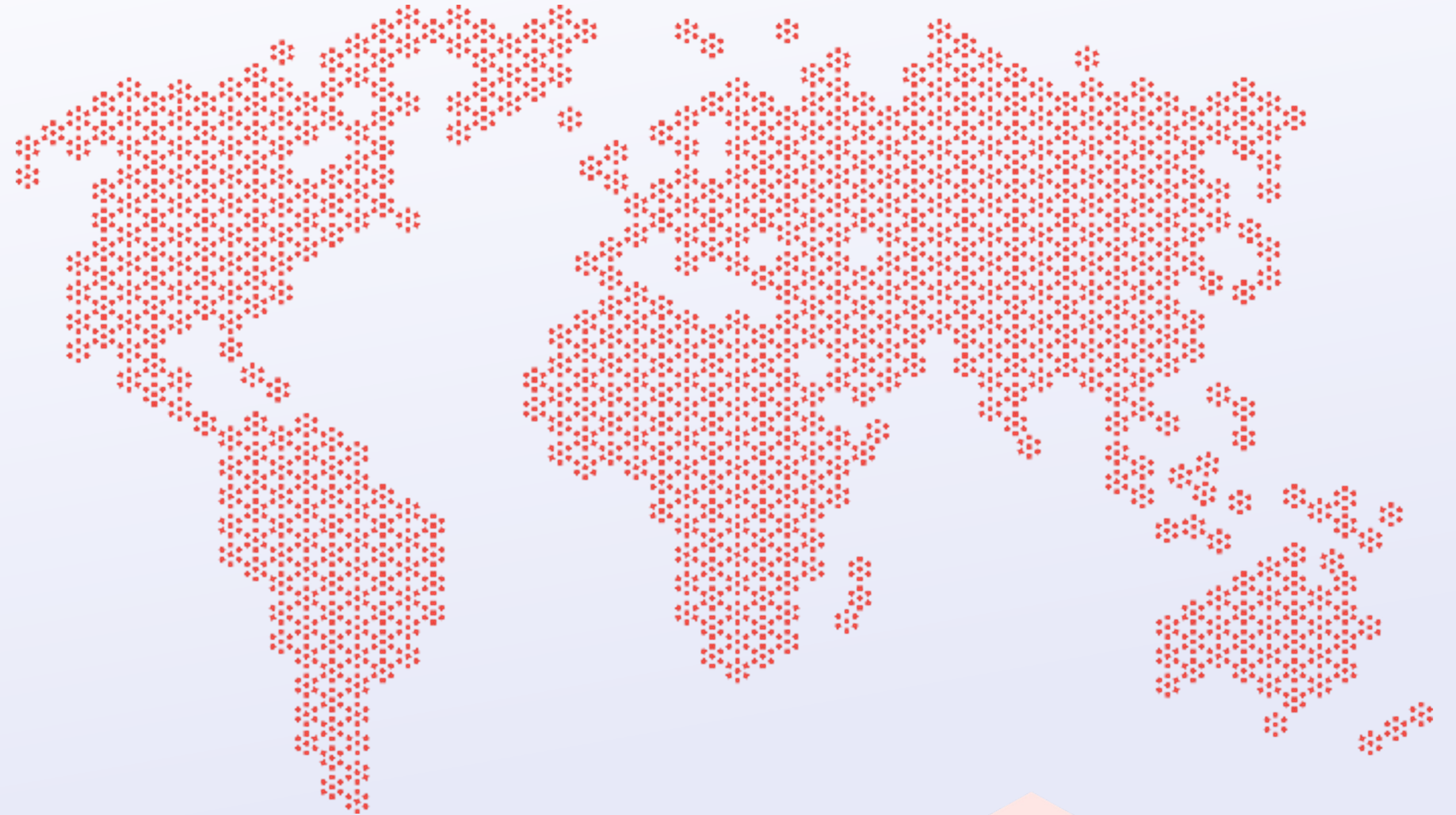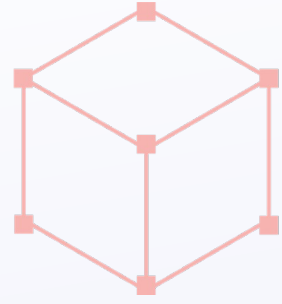
Thanks!

denodo

www.denodo.com | info@denodo.com