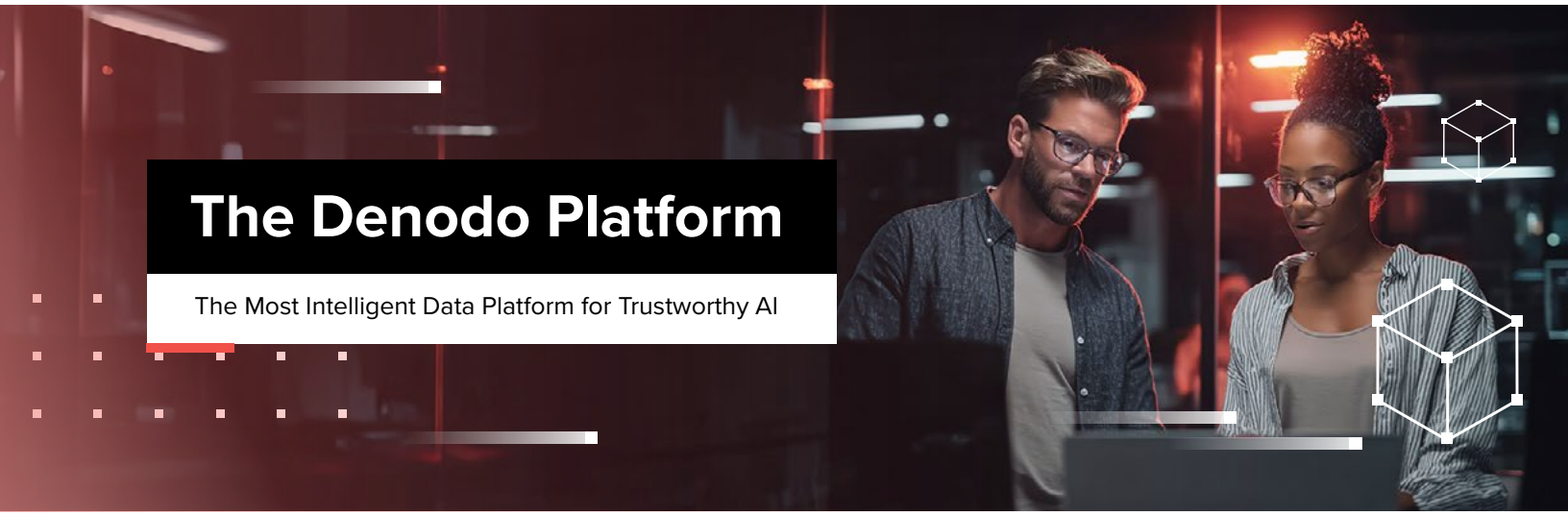


The Denodo Platform

The Most Intelligent Data Platform for Trustworthy AI

Organizations today operate across complex data environments that span cloud platforms, lakehouses, operational systems, SaaS applications, and legacy infrastructure. At the same time, demand for real-time analytics and AI-driven applications continues to grow.

Traditional approaches that rely solely on physically consolidating data into warehouses or lakehouses cannot support every use case. Many operational and AI workloads require **secure access to live data across distributed systems**.

The **Denodo Platform** provides an intelligent data layer that unifies distributed enterprise data in real time. Using a logical approach to data management, Denodo connects data across hybrid, multi-cloud, and on-premises environments while maintaining centralized governance and security.

By delivering **live, unified, and governed data**, the Denodo Platform enables organizations to accelerate analytics, build trusted AI applications, and scale data products across the enterprise.

Architecture for Modern Data and AI

To support this approach, the Denodo Platform focuses on the architectural capabilities required to deliver trusted data to the people and systems that use it.

For many organizations, the biggest challenge is no longer storing or processing data at scale—it is delivering the right data to the right consumer in a consistent, governed way. Data teams often spend significant effort building pipelines between systems, replicating datasets for different tools, and maintaining business logic across multiple environments. This increases operational complexity and makes it harder for analysts, developers, and AI systems to access reliable data.

The Denodo Platform addresses this challenge by providing a unified data access and delivery layer that simplifies how data is integrated, governed, and consumed across the enterprise.

This architecture is built on five core capabilities.

LIVE DATA DELIVERY

Many modern use cases require access to current operational data rather than periodically replicated datasets. Denodo enables real-time access to distributed data sources so analytics platforms, applications, and AI systems can work with up-to-date information.

Instead of relying solely on large-scale data movement, Denodo can query and combine data across systems at runtime while optimizing execution through pushdown processing and selective caching. This helps organizations reduce unnecessary replication, simplify pipeline management, and deliver fresher data for analytics and AI workloads.

SEMANTIC CONSISTENCY ACROSS SYSTEMS

As data environments grow, business concepts are often defined differently across systems. Analysts, developers, and AI models may interpret data inconsistently depending on the source.

Denodo provides a universal semantic layer that standardizes business definitions, metadata, and relationships across distributed sources. This allows teams to centralize transformation logic, maintain consistent definitions across tools, and improve data discovery and understanding.

CENTRALIZED GOVERNANCE AND SECURITY

Broader data access requires strong governance. Denodo enforces security and governance policies centrally while allowing data to remain in existing systems.

The platform supports fine-grained access controls, lineage tracking, and integration with enterprise identity and governance tools. This allows organizations to apply consistent security policies, track data usage across analytics and AI systems, and enable self-service access without compromising governance.

COST AND PERFORMANCE OPTIMIZATION

Accessing distributed data sources can introduce performance challenges if not optimized properly. Denodo includes advanced query optimization capabilities such as pushdown processing, parallel execution, and intelligent caching.

For lakehouse environments, Denodo also provides the Lakehouse Accelerator, powered by the Velox execution engine, which improves performance when accessing open data lake formats such as Parquet, Delta, and Iceberg.

These capabilities help organizations deliver responsive analytics while managing infrastructure and compute costs.

BROAD ACCESS ACROSS THE ENTERPRISE DATA LANDSCAPE

Critical data rarely resides in a single platform. Operational systems, SaaS applications, partner environments, and data lakes all contain valuable information.

Denodo connects across this distributed landscape and exposes data through standard interfaces such as SQL, APIs, and event streams. This allows BI tools, applications, data scientists, and AI agents to access governed data using the interfaces they already rely on.

Enabling the Last Mile of Data Delivery

Together, these capabilities address the final step in modern data architectures: delivering trusted, usable data to the people and systems that depend on it.

Denodo provides a consistent access layer that simplifies how data is discovered, integrated, and consumed across the enterprise. This enables organizations to deliver curated datasets for analytics, reusable data products for business domains, and reliable data access for applications and AI systems.

By simplifying the last mile of data delivery, Denodo helps organizations modernize their data architecture while continuing to leverage existing investments in lakehouses, warehouses, and operational systems.

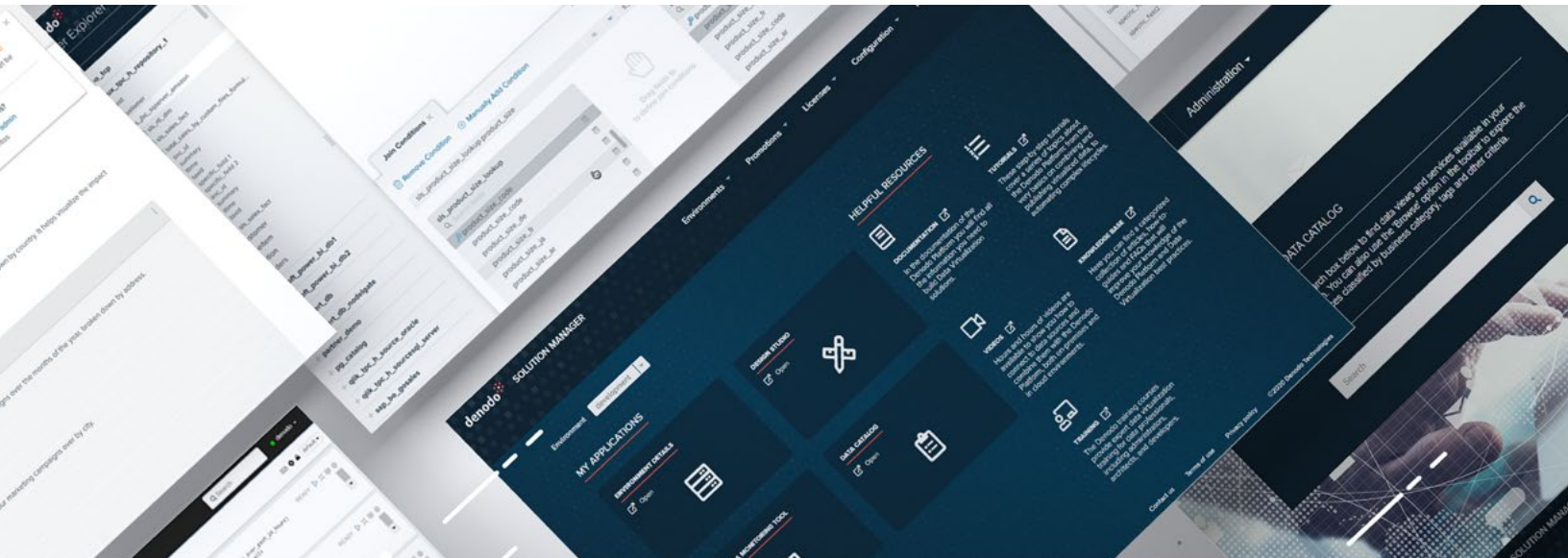
The result is an architecture where trusted data can be delivered consistently to business users, data teams, applications, and AI systems alike.

Flexible Deployment and Scalable Editions

The Denodo Platform supports multiple deployment models to align with enterprise infrastructure strategies. Organizations can deploy Denodo in their own environments—on-premises, in private cloud, or within their public cloud accounts—or choose Denodo's fully managed cloud service.

Denodo is also available in multiple platform editions designed to support different workloads and adoption stages. Teams can begin with a focused deployment and expand over time as additional data domains, users, and workloads are added.

The table below outlines the capabilities available across the different Denodo Platform editions.



	Team Empower small teams to deliver focused data projects with fast time-to-value	High Availability Support growing workloads with reliable performance across multiple projects and teams	Business Critical Unify data at full enterprise scale with resilient, mission-critical operations
Configurations			
Deployment	Single Server[1]	Clustering	Clustering
Non-production environments	1 Development and 1 Staging	Unlimited	Unlimited
Disaster Recovery environment	1	1	2
Solution Manager Instances (Included)	1	1	1
Maximum Cores	8	16	48
Data Products included	100	225	750
Data Volume included[2]	7.5 TB/year	25 TB/year	75 TB/year
Data Sources	Unlimited	Unlimited	Unlimited
Data consumers	Unlimited	Unlimited	Unlimited
Features			
Data Marketplace Users included	✔	✔	✔
Data Marketplace	✔	✔	✔
Dataset collaboration (endorsements, warnings, and deprecation notes)	✔	✔	✔
VQL procedures	✔	✔	✔
Denodo AI Assistant	✔	✔	✔
AI SDK	✔	✔	✔
Global security policies	✔	✔	✔
Smart Query Acceleration	✔	✔	✔
FinOps logging/integrated dashboard	✔	✔	✔
Advanced Diagnostic & Monitoring tool	✔	✔	✔
Data Lakehouse Accelerator (cores in MPP cluster)[3]	✔ (up to 512 cores)	✔ (up to 1024 cores)	✔ (up to 3072 cores)
Integration with VCS	-	✔	✔
Integration with 3rd party data modeling tools	-	✔	✔
Import governance tags from external catalogs	-	✔	✔
Clustering	-	✔	✔
Support			
Support Level (Included)	Standard	Premium	Global
Success Services (Included)	Enterprise Basic	Enterprise Advanced	Enterprise Advanced Plus
Onboarding Services (Not Included)	Basic	Advanced	Advanced

1. The Team tier is single server only, but it can be deployed in an Active-Passive failover configuration.
2. All data management units (KB, MB, GB, TB) used in this table and for billing purposes follow the Internation System of Units (SI) decimal standard where 1 unit = 1,000 of the preceding unit. Binary units (e.g., 1,024) do not apply.
3. The Data Lakehouse Accelerator will have an optional Support package for customers needing technical support and assistance with the Data Lakehouse Accelerator.

Capabilities

DATA SOURCES

Relational Databases

- Generic (JDBC)
- IBM DB2 8.2, 9, 10, 11, and higher; 9,10, and 11 for z/OS
- Denodo Virtual DataPort - 8.0, 9
- Apache Derby 10 and higher
- Informix 7, 12
- Microsoft SQL Server 2000, 2005, 2008, 2008 R2,
- 2012, 2014, 2016, 2017, 2019, 2025, and higher
- MySQL: 4, 5, 8, and higher
- Oracle: 8i, 9i, 10g, 11g, 12c, 12c In-Memory, E-Business Suite 12, 18c, 19c, 23ai, and higher
- PostgreSQL 11, 12, 13, 14, 15, 16, 17
- Sybase ASE / SAP ASE 12, 15
- MS Access
- Microsoft Dataverse
- Huawei GaussDB

In-Memory Databases

- SAP HANA 1.0, 2.0
- Oracle TimesTen 11g
- Oracle 12c In-Memory

Parallel Databases and Appliances

- Apache Doris
- ClickHouse
- Exasol 7.1
- Greenplum 4.2
- Vertica 7, 9
- Netezza 4.6, 5.0, 6.0, 7.0, and higher
- Oracle Exadata
- ParAccel 8.0.2 (by using ParAccel 2.5.0.0 JDBC3g with SSL driver)
- SybaseIQ
- SQreamDB
- Teradata 12, 13, 14, 15, 16, 17
- Yellowbrick

Cloud Data Warehouse / RDBMS

- Alibaba ApsaraDB for OceanBase MySQL
- Alibaba ApsaraDB for OceanBase Oracle
- Alibaba ApsaraDB RDS for MySQL
- Alibaba ApsaraDB RDS for PostgreSQL
- Alibaba ApsaraDB RDS for Microsoft SQL Server
- Alibaba ApsaraDB PolarDB for MySQL
- Alibaba ApsaraDB PolarDB for PostgreSQL
- Alibaba ApsaraDB AnalyticDB for MySQL
- Alibaba ApsaraDB AnalyticDB for PostgreSQL
- Alibaba MaxCompute
- Amazon Redshift
- Amazon Athena
- Amazon Aurora (MySQL and PostgreSQL)
- Amazon DynamoDB
- Azure Cosmos DB
- Azure SQL Database
- Azure Synapse SQL (aka Azure SQL Data Warehouse)
- Delta Lake
- Google AlloyDB
- Google BigQuery
- GCP (Google Cloud Platform) SQL for MySQL; GCP SQL for PostgreSQL
- Google Spanner (Includes support for the service Spanner Data Boost)
- MongoDB Atlas
- Snowflake

Big Data

- Hive 0.13.0 (Hive Server 2); 1.1.0 (Hive Server 2); 2.0.0 (Hive Server 2); 3.1.2 and higher (Hive Server 3); Hive for Cloudera 1.1.0; Hive for Hortonworks 1.2.1
- Impala 1.2.4., 2.3, 3.x kudu
- Spark SQL 1.5, 1.6, 2.x, 3.x and higher
- PrestoDB
- Trino
- Databricks

NoSQL

- Amazon DocumentDB
- MongoDB
- Cassandra 3.x

Multi-Dimensional Sources

- Multidimensional database (generic)
- Azure Analysis Services
- SAP BI 3.x and BW 7.x
- Mondrian: 3.x
- IBM Cognos TM1
- MS SQL Server Analysis Services
- Essbase

Data Lake Storage/Formats

- S3
- Azure Data Lake Storage Azure Blob Storage
- Google Cloud Storage
- Parquet/Avro
- Delta/Iceberg
- Huawei Object Storage Service (OBS)

Web Services

- SOAP
- REST (XML, RSS, ATOM, JSON)
- OData v2.0 and v4.0

Flat and Binary Files

- CSV, pipe-delimited, regular expression-parsed
- MS Excel xls 97-2003
- MS Excel xlsx 2007 or later
- MS Access
- XML
- JSON
- SAS Files (SAS7BDAT)
- All files can be local or in remote filesystems, through FTP/ SFTP/FTPS, and in clear, zipped and/or encrypted format.

Indexes and unstructured content

- Query vector databases and relational sources together while applying the same security, lineage, and semantics.
- CMS, file systems, text
- Elasticsearch

Cloud, SaaS, Web Sources

- Adobe Analytics

- AWS
- Google Analytics
- Google Sheets
- Facebook
- LinkedIn
- MS Azure Data Lake
- MS Sharepoint
- MS Dynamics 365 Business Central / Customer Engagement
- Marketo
- ServiceNow
- Salesforce
- Twitter
- Workday
- many more through configurable JSON and XML adapters

Active Directory as Source or Leveraging Security
- LDAP v3
- Microsoft Active Directory 2003, 2008

Streaming/Messaging systems
- Kafka
- MQSeries
- SonicMQ
- ActiveMQ
- Tibco EMS
- Other JMS compatible services

Semantic Repositories
- Semantic repositories in Triple Stores / RDF accessed through SPARQL endpoints.

Packaged Applications
- Native SAP integration through ABAP Core Data Services (CDS) views.
- SAP ERP/ECC (BAPIS and tables)
- Oracle E-Business Suite 12
- Siebel
- SAS 9.4

Hierarchical Databases
- Adabas (SOA Gateway and Denodo's SOAP connector): 5, 6

Denodo SDK for Custom Connectors

PUBLISHING OPTIONS
- SQL Based access via JDBC, ODBC and ADO.NET
- Web Services
  - REST
  - OData
  - Open API (a.k.a Swagger)
  - GraphQL
  - SOAP
- Model Context Protocol (MCP) support for governed access by AI agents
- OAuth, OAuth 2.0 (JWT)
- SAML
- SSL
- WS-Security
- JMS listeners for message queues
- Denodo Scheduler for batch process and lite ETL

DATA MARKETPLACE
Cataloging
- Web-UI for seamless data discovery and exploration for business users
- Description and documentation
- Customizable business properties for richer metadata management
- Business categories and tags Intelligent search with smart ranking of results

Governance
- Graphical lineage spanning Denodo and Consumer tools such as Tableau and PowerBI
- Extended visibility into reports and dashboards
- Integrated request management (access, changes, data quality issues, etc.)
- Endorsement of datasets, comments, warnings, etc.
- Usage statistics: who uses what data, when and how
- Data profiling information

Self-Service
- Automatically matches UI language to browser settings
- Last-mile data preparation wizards for customizing datasets by non-technical users
- Full-featured SQL shell facilitates the execution of complex queries
- Export to CSV, Excel, and Tableau Data Extracts
- Save personal queries for easy access
- Share queries or publish them as new data products

PERFORMANCE OPTIMIZATIONS
- Smart query acceleration for analytics
  - Aggregate Aware Summaries
- Massively parallel processing (MPP) integration for query acceleration and caching
- Support for Arrow Flight SQL, enabling high-performance data transfer
- Full and partial aggregation and join pushdown, even in federated views
- Support for alternative data sources
- On-the-fly data movement for optimization
  - Option to restricted data movement for sensitive data
- Cost-based optimization (data statistics, data source indexes, data source execution model and parameters, network transfer rates)
- Pushdown of selections/projections/joins/groupby operations also on federated views
- Multiple join strategies
- Simplifying partitioned unions (partition pruning)
- and many more

DATA MATERIALIZATION OPTIONS
- Multi-mode selective data materialization (i.e. cache): full, partial, incremental, or total refresh, event-based or scheduled, configured at the view level, incremental queries for SaaS sources
- Directed Acyclic Graph (DAG) cache management automatically analyzes table dependencies and sequences replication tasks for optimal performance
- AlloyDB for PostgreSQLAmazon Athena
- Supports schema evolution without data loss
- Amazon Aurora MySQL

- Amazon Aurora PostgreSQL
- Amazon Redshift
- Azure SQL
- Azure Synapse SQL (previously known as Azure SQL Data Warehouse)
- Clickhouse
- Databricks (with Apache Arrow support for accelerated pushdowns)
- Exasol
- Google BigQuery
- GCP Cloud SQL for MySQL
- GCP Cloud SQL for PostgreSQL
- Hive 2.0.0; Hive 3.1.2, and higher (HiveServer2)
- IBM DB2 (8, 9, 10, 11, and higher for LUW; 9,10,11 for z/OS)
- Impala 2.3; 3.x Kudu
- Microsoft Fabric Data Warehouse
- MS SQL Server (2000, 2005, 2008, 2008R2, 2012, 2014, 2016, 2017, 2019, and higher)
- MySQL 4, 5, 8, and higher
- Netezza 6, 7, and higher
- Oracle 8i, 9i, 10g, 11g, 12c, 12c In-Memory, 18c, 19c, and higher
- Oracle TimesTen 11g
- PostgreSQL 9, 10, 11, 12, and higher
- Presto
- SAP HANA 1 and 2
- Snowflake
- Spark SQL 2.x, 3.x, and higher
- Teradata 12, 13, 14, 15, 16, and 17
- Sybase
- Trino 4xx
- Vertica 7 and 9
- Yellowbrick
- Configurable “generic” adapter for other databases with JDBC drivers

DATA PIPELINES

- Remote Tables (created through UI or stored procedure)
- Denodo Scheduler
- VQL stored procedures

EMBEDDED MPP

- Presto-based MPP query engine leveraging the Velox C++ execution framework for high-performance access to data lake storage
- Graphical introspection of object storage (S3, ABFS, GFS, HDFS, etc.)
- Support for Parquet, Delta, and Iceberg
- New advanced optimization techniques to federate data lake content with any other data source

THIRD-PARTY MPP OPTIONS

- Impala
- Presto
- Spark 1.5, 1.6, 2.x
- Databricks 2.x

DATA GOVERNANCE

- Data source refresh, change impact analysis, dependency tree, full data lineage
- Denodo Governance Bridge: integration with IBM Information Governance Catalog
- API to publish metadata and lineage information to data governance tools like Informatica EDC, Collibra, etc.

SECURITY

Data in Motion – secure channels

- Using SSL/TLS
- Client-to-Denodo and Denodo-to-source
- Available for all protocols (JDBC, ODBC, ADO.NET and WS)

Data at Rest - secure storage

- Cache: third party database. Can leverage its own encryption mechanism
- Swapping to disk: serialized temporarily stored in a configurable folder that can be encrypted by the OS

Encryption/Decryption

- Support for custom decryption for files and web services
- Transparent integration with RDBMs encryption
- Encrypted metadata import/export

User and Role Based Including Integration with AD/LDAP

- Row and column level authorization
- Advanced customizable masking
- Custom policies for specific security constraints and integration with external policy servers

Global Policies

- Tag-based security policies
- Support for RBAC and ABAC
- Dynamic Access Control Polices
- Column and row restrictions, multiple masking options, deny execution

Authentication

- Native and LDAP/Active Directory based Support for Kerberos and Windows SSO
- Kerberos
- NTLM
- OAuth, OAuth 2.0 (JWT)
- SAML
- Two-factor authentication (through supported identity providers: Okta, Duo, etc.)
- SSL
- WS-Security
- Pass-through session credentials to leverage existing source privileges

ADVANCED SEMANTICS

- Global security policies
- View- and column-level classification tags
- Support for the importing of external data governance tags from Collibra and other external data catalogs

DATA MODELING

- Design Studio: Web-based development studio for data modeling
- Desktop version also available
- Bottom-up and top-down (through Interface Views)
- Integration with third-party modeling tool
  - ER/Studio Data Architect
  - ERwin Data Modeler

- IBM InfoSphere Data Architect
- SAP PowerDesigner
- Sparx Systems Enterprise Architect

DATA QUALITY

- Library of transformation, filter, and matching functions and quality rules for validating, cleansing, enriching, standardizing, matching, and merging data
- Extensible through custom functions
- Integration with external data quality tools

MONITORING

- Denodo Diagnostic and Monitoring Tool (DMT) integrated in the Solution Manager
- Extensible usage and metadata dashboards integrated in Apache Superset
- FinOps dashboard to monitor and understand key metrics associated with cloud costs like egress, query cost, etc.
- Detailed monitoring information is available in logs for integration with log management tools like Splunk, ELK, Cloudwatch, etc.
- Monitoring is also available via SNMP and JMX standards. Therefore interoperate with most leading systems management packages (e.g., HP OpenView, Nagios, Zenoss, Osmius, IBM Tivoli and Microsoft WinRM)

DENODO ASSISTANT

Data Marketplace

- A chatbot allowing users to discover data, explore views, and understand complex relationships
- Natural language queries with GenAI
- Support for OpenAI, Azure OpenAI, AWS Bedrock, and customer LLMs
- Personalized dataset recommendations
- Smart SQL-fragment autocomplete based on previous activity

Design Studio

- Summary recommendations for smart query acceleration
- View and column description suggestions
- Business friendly column name suggestions
- Diagnose incorrect queries and explain queries
- AI-generated recommendations for join conditions

VQL Function

- Automatically summarize, extract, analyze sentiment, remove sensitive data, and translate text using LLMs.

OPERATIONS

- Solutions Manager to automate operations and promotions tasks
  - Centralized management and distribution of updates to clients
  - Centralized management of license keys
  - Define promotion revisions and their dependencies and deploy them to a production cluster with zero downtime
  - Centralized management of data source properties and logs
  - REST API for automation of tasks from DevOps tools (e.g. jenkins)
- Integrated Infrastructure Management for Cloud (AWS)
  - Creation and management of clusters: define type of EC2 instances, number of EC2 instances, etc.
  - Creation of load balancers and Auto Scaling groups.
  - Installation and launch of Denodo servers.
  - Update the Denodo version
  - Enable SSL in the Denodo servers.
- Multi-user development with version control integration
  - Ingratiation with source management system such as Git
  - Workspaces for isolated development and testing
- Resource Manager to limit and allocate resources to each session, role, or user in a way that optimizes resources utilization for each application
  - Change resources priority
  - Enforce limited timeouts or limits on number of rows
  - Add daily quotas per minute/day/month: e.g. only 50 queries per day

DEPLOYMENT PATTERNS

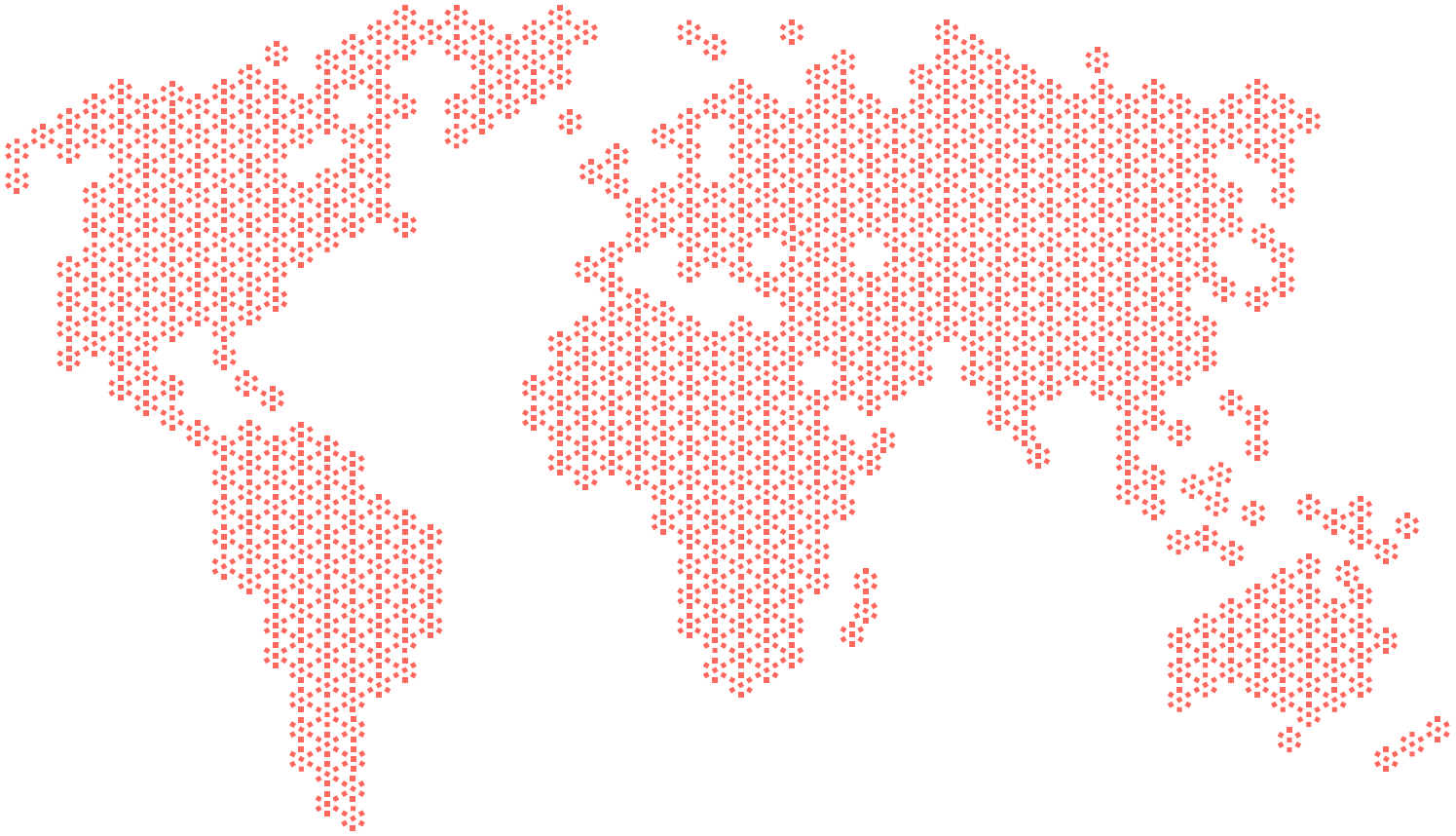
- On-premises, private cloud, public cloud
  - On-premises, private cloud, public cloud
  - Basic single server configuration
  - HA cluster with load balancing (Active-Passive and Active-Active)
  - Shared or distributed local cache
  - Geographically distributed server environments
  - Multiple Denodo instances, peer-to-peer or multilayered
  - Containerization support through Docker
- Public cloud
  - Denodo Platform for AWS
  - Denodo Platform for Azure
  - Denodo Platform for GCP
  - Denodo Platform for Alibaba Cloud
- Auto-scaling support both in AWS and Azure
- Agora - the Denodo Cloud Service

OPERATING SYSTEMS

- **Microsoft:** Windows Server 2022, Windows Server 2019, Windows Server 2016, Windows 11, and Windows 10
- **Linux:** Amazon Linux 2023, Ubuntu 20.04 LTS or later
- CentOS 9.x and 8.x, Red Hat Enterprise Linux (RHEL) 9.x and 8.x, Oracle Linux 9.x and 8.x, SUSE Linux Enterprise 15.x

MINIMUM HARDWARE REQUIREMENTS

- Processor: Intel Xeon quad-core or similar. High-load scenarios or cases with complex calculations may require 8 cores or more.
- Physical memory (RAM): 16 gigabytes of memory so the Denodo server can allocate a runtime heap space up to 8 gigabytes.
- Disk space: Minimum: 5 gigabytes, Recommended: 100 gigabytes. Denodo only needs around 1 GB of disk space. If the cache is installed on the same server, more disk space will be required.



denodo

Denodo is a leader in data management. The award-winning Denodo Platform is the leading logical data management platform for transforming data into trustworthy insights and outcomes for all data-related initiatives across the enterprise, including AI and self-service. Denodo's customers in all industries all over the world have delivered trusted AI-ready and business-ready data in a third of the time and with 10x better performance than with lakehouses and other mainstream data platforms alone.

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