

 DENODO PROFESSIONAL

The smart way to leverage data virtualization for your data integration projects

The accelerating adoption of cloud platforms, cloud-based data stores, and software-as-a-service (SaaS) applications introduces new complexities and challenges for organizations that need to integrate their cloud data so that all systems work together seamlessly. The accelerating complexity of data in today's world means that organizations need to connect everything together in a way that is flexible and agile. However, it can be challenging and resource intensive to integrate all of the data and make it available in real time for analytics or data services. Traditional approaches, which use extract, transform, and load (ETL) processes to copy all of the data into a new repository, are time consuming and expensive, and they can result in data that is duplicated and out-of-sync with the source.

Data virtualization is the key technology to overcoming this challenge. As the only data integration style designed for distributed architectures, data virtualization provides a logical data access layer on top of multiple heterogeneous systems in distributed architectures. Today, the data management ecosystem is distributed in nature, so a logical data access layer such as data virtualization is the best fit.

Denodo Professional incorporates features that accelerate the delivery of governed data to business applications in the most appropriate format for each consumer, providing data integration across multiple heterogeneous systems.

Denodo Professional offers strong support for core data virtualization use cases: logical data warehouses and data services APIs. And it goes beyond these traditional data virtualization scenarios to better support new types of use cases, such as data science and machine learning (ML) initiatives, data lake management, and more.

In addition, it represents a big step forward in platform as a service (PaaS) cloud strategy, with capabilities for automatically managing the cloud infrastructure from a centralized web console.

What Are the Key Features of Denodo Professional?

Denodo Professional is built on the same core Denodo Platform as other Denodo subscription offerings, enabling organizations to quickly connect to data and make it immediately available for business analysts, data scientists, and consuming applications.

Denodo Professional has all the capabilities necessary to get data integration projects into production with minimal effort and resources, within the budgets of small teams or initial projects. With out-of-the-box connectivity to more than 150 data sources types and the web-based, visual Design Studio with which to combine, integrate, and normalize data, Denodo Professional enables data workers to connect, combine, and consume data quickly and efficiently. Developers can leverage the Design Studio to quickly build views and data services without writing any code. Denodo Professional supports single sign-on (SSO) using Kerberos, SAML, OpenID, and OAuth, enabling seamless connectivity to all Denodo components.

In addition to SQL access, Denodo Professional offers advanced support for data services with flexible web services delivery options (REST, SOAP, OData, and OpenAPI), the ability to expose data in multiple formats (XML, JSON, HTML, RSS), and unified support for the latest security protocols (OAuth, JSON Web Tokens, SAML, Kerberos, HTTPS, HTTP Basic Digest Authentication, or WS-Security). Denodo Professional supports GraphQL, simplifying the queries of multiple REST endpoints.



In terms of execution and performance, Denodo Professional offers state-of-the-art dynamic query optimization capabilities including AI/ML-powered, cost-based, and rules-based engines. For scenarios that require selective replication, Denodo Professional offers an advanced caching engine and flexible options for data ingestion and replication using ETL and ELT execution models.

Denodo Professional includes automated cloud infrastructure management features, which accelerate all of the tasks related to installing, configuring, deploying, and upgrading Denodo clusters.

Denodo Professional also includes the Apache Zeppelin-based Denodo Notebook for data scientists, with integrated security and SSO. This enables data scientists to find the right data more quickly and put it to work in machine learning and predictive analytics models.

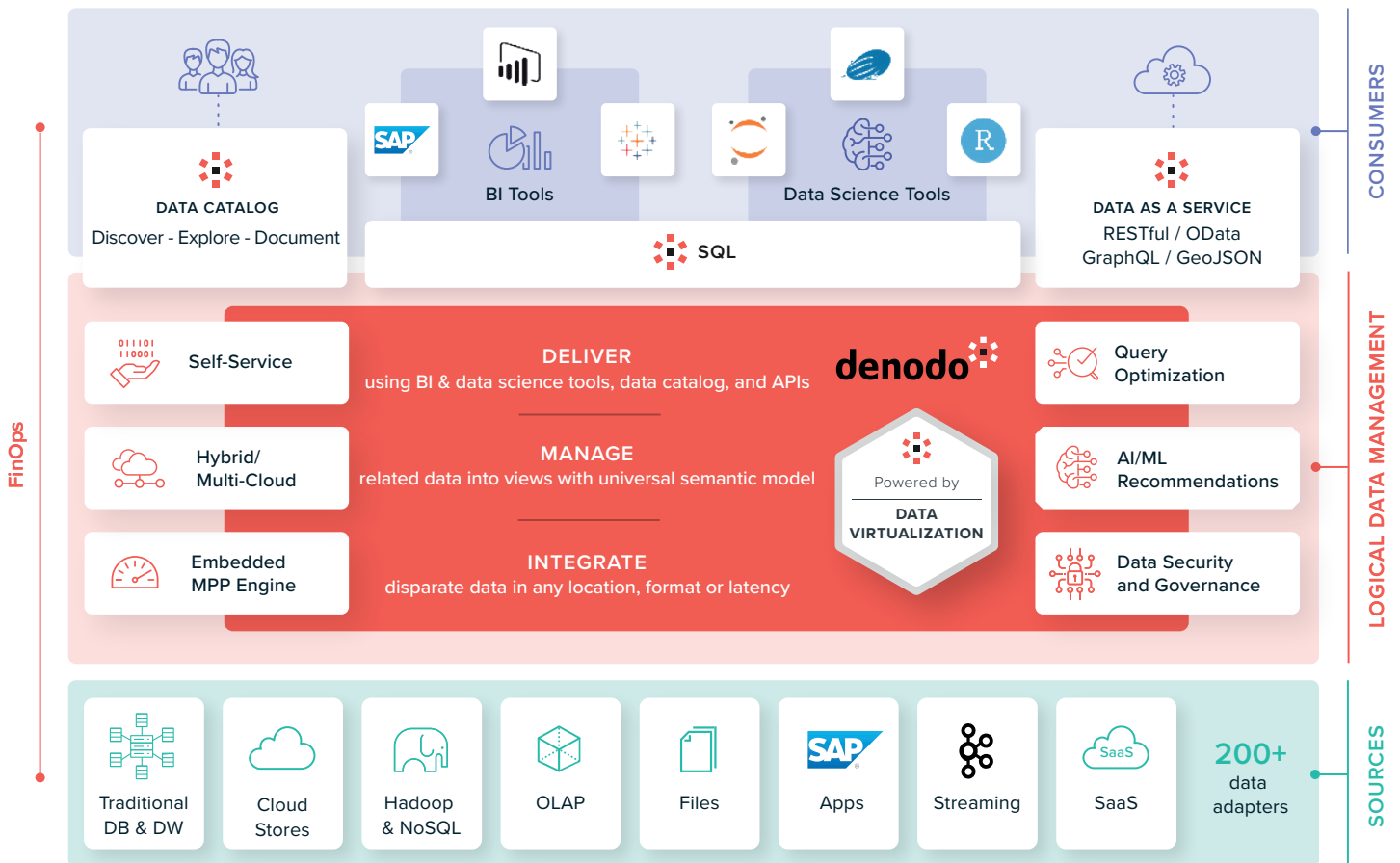


Fig. 1 Denodo Professional Architecture



KEY FEATURES OF DENODO PROFESSIONAL:



A full, web-based interface for all Denodo tools, with SSO support: an integrated, web-based experience across all tools.



200+ out-of-the-box connectors to quickly connect to up to 5 data sources and unlimited targets, in real time.



Web-based Design Studio for developers, providing ease-of-use across all of the steps in the data-services development process, including the development of different data views.



An Advanced Optimization and Caching engine that leverages 20 years of experience in complex multi-source queries.



Automated infrastructure management capabilities for the cloud including cluster configuration, start/stop controls, and automatic installation of updates.



GraphQL, REST, and OData support: zero-code creation of data APIs, with first-class performance leveraging the Denodo query optimizer.



“Apache Zeppelin for Denodo” Notebook, enabling data scientists to construct narratives that combine queries, code, and text, to aid in data analysis and to help them to explain their work and share it with colleagues.



Connect to up to 5 different data sources: Structured DB, unstructured data streams, web APIs, or even flat files



Business Benefits

HIGHLY ECONOMICAL

Integrate up to 5 data sources easily and reliably, in real time, at a fraction of the cost of traditional data integration approaches such as extract, transform, and load (ETL) processes, and with significantly greater agility.



FASTER PATH TO VALUE

Deliver contextual, reliable information, faster, for more actionable insights. Agile enterprises are industry leaders.



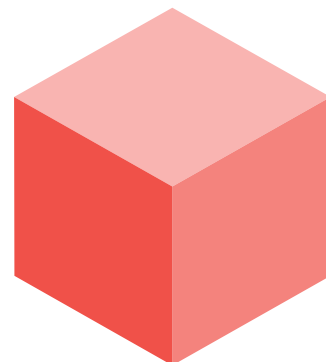
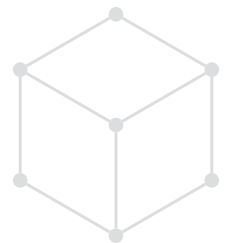
BUSINESS-FRIENDLY

Abstract the complexity of modern data ecosystems (myriad sources, multiple formats, distributed, heterogeneous, diverse) from business. Expose data in the right format and use the naming conventions required by every type of user and application at almost no cost. Rapidly adjust to changes in requirements.



PERFECT FOR PROFESSIONAL USE

State-of-the-art data virtualization engine with support for multiple users. Development and production environments included to separate modeling of new datasets from production usage.





COMPARING THE VARIOUS SUBSCRIPTION TIERS

Users looking for broader data management capabilities beyond data virtualization can leverage one of the other Denodo Platform subscription options. Designed to suit the needs of various projects, from small departmental projects to enterprise-wide digital transformation.

	 DENODO PROFESSIONAL Small, single-use-case projects within individual departments	 DENODO STANDARD Multiple use cases within individual departments	 DENODO ENTERPRISE Enterprise-wide deployment for multiple use cases and groups and large data volumes	 DENODO ENTERPRISE PLUS Comprehensive automation, collaboration, and advanced security for enterprise-wide deployments
Number of Data Sources Supported	5	Unlimited	Unlimited	Unlimited
FinOps Logging and Integrated Dashboard	✓	✓	✓	✓
Integration with Version Control Systems (VCS)		✓	✓	✓
Smart Query Acceleration using summaries			✓	✓
VQL procedures			✓	✓
Advanced Diagnostic & Monitoring Tool			✓	✓
Data Catalog			✓	✓
Integration with external Massive Parallel Processing (MPP) engines like Impala, Spark, and others			✓	✓
Integration with 3rd party data modeling tools (ER/Studio, Erwin, etc.)			✓	✓
Embedded Presto-based MPP engine				✓
Automatic recommendation of summaries				✓
Global security policies				✓
Import data governance tags from external catalogs				✓
Data Catalog: Natural language queries and AI-based recommendations				✓
Data Catalog: Dataset collaboration through endorsements, warnings, and deprecation notes				✓

Denodo Professional Capabilities Sheet

DATA SOURCES

Relational Databases

- Generic (JDBC)
- IBM DB2 (JDBC): 8, 9, 10, 11, 12 for LUW, 9,10 for z/OS, AS400
- Multi Layered Denodo deployments (JDBC): 5.5, 6.0, 7.0, 8.0
- Apache Derby (JDBC): 10
- Informix (JDBC): 7, 12
- MS SQL*Server (JDBC, ODBC): 2000, 2005, 2008, 2008R2, 2012, 2014, 2016, 2017
- MySQL (JDBC): 4, 5, 8
- Oracle (JDBC): 8i, 9i, 10g, 11g, 12c, 18c, 19c
- Oracle E-Business Suite (JDBC): 12
- PostgreSQL (JDBC): 8, 9, 10, 11, 12
- Sybase Adaptive Server Enterprise (JDBC): 12, 15
- MS Access (ODBC)

In-Memory Databases

- SAP HANA (JDBC): 1,2
- Oracle TimesTen (JDBC): 11g
- Oracle 12c In-Memory

Parallel Databases and Appliances

- Exasol (JDBC)
- GreenPlum (JDBC): 4.2
- HP Vertica (JDBC): 7, 9
- Netezza (JDBC): 4.6, 5.0, 6.0, 7.0
- Oracle Exadata (JDBC): X5-2
- ParAccel 8.0.2 (by using ParAccel 2.5.0.0 JDBC3g with SSL driver)
- SybaseIQ (JDBC) 12.x, 15.x
- Teradata (JDBC): 12, 13, 14, 15, 16, 17
- Yellowbrick

Cloud Data Warehouse / RDBMS

- Alibaba ApsaraDB for OceanBase MySQL (JDBC)
- Alibaba ApsaraDB for OceanBase Oracle (JDBC)
- Alibaba ApsaraDB RDS for MySQL (JDBC)
- Alibaba ApsaraDB RDS for PostgreSQL (JDBC)
- Alibaba ApsaraDB RDS for Microsoft SQL Server (JDBC)
- Alibaba ApsaraDB PolarDB for MySQL (JDBC)
- Alibaba ApsaraDB PolarDB for PostgreSQL (JDBC)
- Alibaba ApsaraDB AnalyticDB for MySQL (JDBC)
- Alibaba ApsaraDB AnalyticDB for PostgreSQL (JDBC)
- Amazon Redshift (JDBC)
- Amazon Athena (JDBC)
- Amazon Aurora (JDBC)
- Amazon DynamoDB
- Azure Cosmos DB
- Azure SQL Database
- Azure Synapse Analytics
- Delta Lake

- Google AlloyDB (JDBC)
- Google Big Query (JDBC)
- Google Cloud SQL (JDBC)
- Google Spanner DatBoost (JDBC)
- MongoDB Atlas
- Snowflake (JDBC)

Big Data

- Apache Hive (JDBC): 0.12, 1.1.0, 1.1.0 for Cloudera, 1.2.1 for Hortonworks, 2.0.0
- Impala (JDBC): 2.3
- Spark SQL (JDBC): 1.5, 1.6, 2.x, 3.x
- PrestoDB (JDBC)
- PrestoSQL / Trino (JDBC)
- Databricks Delta 2.x

NoSQL

- MongoDB
- Cassandra
- HBase

Multi-Dimensional Sources

- SAP BW (BAPI/XMLA): 3.x
- SAP BI 7.x (BAPI): 7.x
- Mondrian (XMLA): 3.x
- IBM Cognos TM1
- MS SQL Server Analysis Services 200x
- Essbase (XMLA): 9, 11

Data Lake Storage

- S3
- Azure Data Lake Storage
- Azure Data Lake Storage Gen 2
- Azure Blob Storage
- Google Cloud Storage
- Parquet
- Avro

Web Services

- SOAP
- REST (XML, RSS, ATOM, JSON)
- OData

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

- CMS, file systems, text
- ElasticSearch 6.4, 6.7

Cloud, SaaS, Web Sources with Simplified OAuth Security

- Adobe Analytics
- Amazon
- Google
- Google Sheets
- Facebook
- LinkedIn
- MS Azure Data Lake
- MS Sharepoint (via OData)
- MS Dynamics 365 Business Central / Customer Engagement
- Marketo
- ServiceNow
- Salesforce (SOQL)
- Twitter
- Workday
- many more through configurable JSON and XML adapters

Active Directory as Source or Leveraging Security

- LDAP v3
- Microsoft Active Directory 2003, 2008

Message Queues

- MQSeries
- SonicMQ
- ActiveMQ
- Tibco EMS
- Other JMS compatible services

Semantic Repositories

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

Packaged Applications

- SAP ERP/ECC (BAPIS and tables)
- Oracle E-Business Suite 12
- Siebel
- SAS (SAS JDBC Driver): 7 and higher

Mainframe

- IMS
 - IBM IMS native drivers: 8, 9
 - IMS Universal Drivers: 11

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
- OAuth, OAuth 2.0 (JWT)
- SAML
- SSL
- WS-Security
- JMS listeners for message queues
- Denodo Scheduler for batch process and lite ETL

PERFORMANCE OPTIMIZATIONS

- Full and partial aggregation and join pushdown, even in federated views
- Support for alternative data sources
- On-the-fly data movement for optimization
- 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

CACHE AND DATA MOVEMENT OPTIONS

- Multi-mode caching: full, partial, incremental, or total refresh, event-based or scheduled, configured at the view level, incremental queries for SaaS sources
- Amazon Athena
- Azure SQL
- Azure SQL Data Warehouse
- Azure Synapse Analytics
- Amazon Redshift
- Databricks 2.x
- Delta Lake
- IBM DB2 (8, 9, 10, 11 for LUW, 9,10,11 for z/OS)
- Hive 2.0.0
- Impala
- MS SQL Server (2000, 2005, 2008, 2008R2, 2012, 2014, 2016, 2017)
- MySQL (4 and 5)
- Netezza (6 and 7)
- Oracle (8i, 9i, 10g, 11g, 12c, 12c in-memory, 18c, 19c)
- Oracle TimesTen 11g
- PostgreSQL (9 and 10)
- Presto
- SAP HANA
- Snowflake
- Spark (1.5,1.6 and 2.x)

- Teradata (12, 13, 14, 15, and 16)
- 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

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

Authentication

- Native and LDAP/Active Directory based Support for Kerberos and Windows SSO
- Base64
- 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

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

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 DQ tools

MONITORING

- 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)

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.
 - Installation and launch of the Denodo servers.
 - Update of Denodo version
 - Enable SSL in the Denodo servers.
- Resource Manager to limit and allocate resources to each session, role or user in a way that optimizes resource utilization for each application
 - Change resource 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

- Production and Development servers included
- Public cloud options
 - Denodo Platform for AWS
 - Denodo Platform for Azure
 - Denodo Platform for GCP

USER INTERFACES

- Central Web Console: integrated access to all Denodo UIs with SSO (Kerberos, SAML, OpenID and OAuth)
- Solution Manager: Centralized UI for administrators to manage deployments and promotions, including automatic management of cloud infrastructure (AWS)
- Design Studio: Web-based Development Studio, drag-and-drop and low-code developer studio geared

to data-oriented developers such as data engineers, power users, and citizen integrators; publish data services with a few clicks.

- Desktop Dev. Studio (VDP Admin tool)

OPERATING SYSTEMS

- Microsoft Windows (32-bit and 64-bit platforms): Windows Server 2019, Windows Server 2016, Windows Server 2012, Windows Server 2008, Windows 10, Windows 8.1, and Windows 7
- Linux (32-bit and 64-bit platforms): Red Hat Enterprise Linux (RHEL) 6 and 7, Oracle Linux 6 and 7, Ubuntu 12.04 LTS and 14.04 LTS, CentOS 6 and 7
- Any Java 11 or greater compatible OS

Available through cloud market places only

* 4 core or 8 core, please refer to AWS, Azure or Google Cloud for full details on available images, memory requirements and disk space

