

Denodo Virtual Data Integration



Data Sheet Version 4.5

IT Organizations are shifting their data integration strategies from tightly coupled hardwired development to loosely coupled dynamic configuration. The time has passed for data and application integration nightmares, heavy and costly architectures, long development projects and hard-coded solutions. Existing installations of monolithic, tightly coupled data integration tools are expected to be migrated to legacy system status in the next 2 years. Businesses are demanding data integration solutions that offer the capability to accommodate different data structures, flexibility in style of data integration (physical vs. virtual), flexibility in delivery of data (real-time vs. batch mode) and much more. With Denodo's data integration, each solution is architected to make data integration simple, flexible and scalable; by following standards, complying with service-oriented principles and, most importantly, using a single platform to include Web/Cloud and information in flat files into the radar of the corporate data centre.

Denodo offers a unified data services platform based on lighter integration architectures. Denodo's data integration platform provides superior data integration capabilities while effortlessly complementing your existing middleware. Based on Data Virtualization technology, Denodo's ability to harness all of the relevant information on a timely basis in order to present it as reusable data services is critical to addressing your organization's growing need to quickly and flexibly deploy next generation business applications. Denodo platform is based on three simple steps:

- **Connect:** Denodo's enterprise class capabilities allow you to connect to and access all your enterprise data. Denodo's extended capabilities allow you to connect to the web as well as to flat files, emails, PDFs etc to extract relevant content with the same enterprise class performance.
- **Combine:** This powerful layer combines disparate data quickly and effectively. The visual query builder allows you to transform (built-in transformation functions, fuzzy logic and custom functions), normalize, cleanse and relate data using common metadata, semantic tools to create composite data views.
- **Publish:** This composite data is fed to applications and processes using Java API, JDBC, ODBC, Web Services (SOAP, REST), RSS feeds, HTML pages etc.

Product

Denodo is an award winning virtual data integration platform for integrating data from a wide variety of sources, regardless of the structure of the data or its location. Denodo's powerful data integration engine enables you to quickly author new composite views through its visual interface. These views are published in real-time or batch to drive various business applications, dashboards, portals etc through standard interfaces such as JDBC, ODBC, Web Services, and Java API. Denodo also has the capability to automatically navigate the web to extract relevant content for your projects or to write back to the web. Denodo's web process extraction uses example-based learning and automatically maintains itself when the Web site changes.

The Denodo Platform presents a lightweight architecture for accessing existing data non-invasively, which means that the underlying sources do not need any modification. This guarantees short deployment times and a faster time-to-value. The platform offers enterprise-class performance including features such as advanced query optimization, asynchronous execution, caching, clustering, LDAP integration, strong security, read-write capability, reliability etc.

Advantages

Superior real-time data integration: The Denodo platform provides real time access to mission-critical information regardless of the complexity of the underlying data sources.

Ability to combine all data formats: Denodo's accesses heterogeneous data sources (including even web and less structured data such as email, MS Office docs etc) and combines the data meaningfully to give you unparalleled breadth and depth of information.

Complements existing middleware: Denodo leverages your investments in data integration as well as SOA middleware by virtualizing data services across these tools as well as enriching these sources with new/updated data.

Semantic relationships: Denodo normalizes the data and builds meaningful relationships between the data. This data can then be published to any application/portal/dashboard.

Flexible data integration: Denodo gives you the flexibility to choose the style of data integration appropriate for your project.

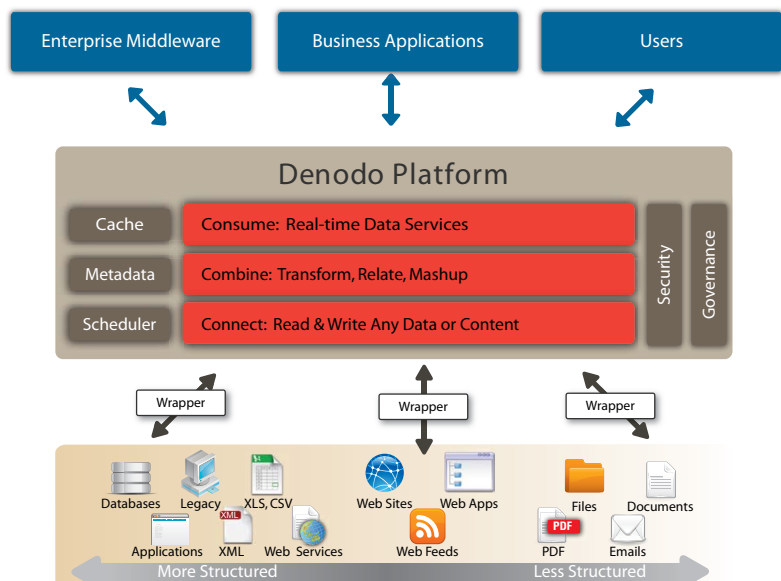
Physicality: Choose physical data integration or virtual depending on the requirements of your project.

Latency: Get data in real time. Or choose cached/scheduled delivery instead.

Granularity: Choose whether you want data delivered row-by-row or in bulk

Reusable data services: Denodo's data services created for one application can be reused for other applications and services making Denodo an ideal tool companies implementing SOA

Denodo Platform Architecture



Product

Access to Heterogeneous Data Sources:

- Access relational databases (JDBC, ODBC), XML, Web Services (SOAP and REST), delimited files, Microsoft Excel, Web 2.0 feeds (RSS, ATOM, etc.) and other structured data sources
- Navigate and extract content from any Web site application, including AJAX-based and dynamic Web 2.0 pages, blogs, reviews, wikis, competitor and password-protected sites
- Web Navigation provides access to hidden/protected data and precise control over what data is extracted and indexed
- Connect to file systems, content management systems, POP3/IMAP/MS Exchange email and federated search engines
- Native support for parsing of Microsoft Word and Adobe PDF formats
- Index unstructured data from Web sites, Word, PDF, flat file documents, email, relational databases, and RSS feeds
- Plug-in architecture allows development of custom connectors and legacy sources

Data Combination and Export:

- Integrate Web data with enterprise databases, applications, and unstructured information for presentation as either relational views or Web services.
- Use graphical point-and-click tool to author new unified views and insights by transforming and combining different data
- Manage hierarchical information natively, with no need for cumbersome transformation and flattening
- Export composite data views as RSS/Atom feeds and simple HTML interface as well as SOA Web services and JDBC
- Stored procedure in Java allow for total control of query pre and post processing
- Metadata exchange and visualization
- Supports Service Oriented Architecture (SOA)
- Integrates with any system via JDBC, Java API, Web Services (WS-* and REST), Search interface RSS feed, or other Web Mashup protocols
- Automatic query capability propagation for all formats
- Built-in data cleansing with option to extend with external plug-in tools
- Fully read/write capable, including full transaction support

Structure unstructured data to combine with structured data using text mining, taxonomy filters, and other semantic tools

Web Process Automation:

- Point-and-click technique enables access to any type of Web data.
- Workflow process modeling allows flexible automation of Web integration processes
- Library of pre-built components for browsing, extracting and structuring data from the Web
- Sophisticated data handling with iterations, loops, page and textual comparison, and data composition
- Native GUI support for building and sharing a library of user-defined JavaScript components
- Automatic processing of Web navigation complexities such as AJAX and JavaScript, authentication mechanisms, secure servers, cookies or sequences that involve popup windows
- Execution using Microsoft IE or Firefox, or basic HTTP + JavaScript client that optimizes for complexity and performance
- Ability to build self-maintaining Web extractions using the Automatic Maintenance Server

Data Mashup Task Scheduler:

- Orchestrate complex mashup processes by scheduling of tasks associated accessing and mashing data.
- Detailed reports of results of task execution including options to send the reports by e-mail
- Enables extraction of data from sources with limited query capabilities
- Support persistent tasks through continuation of query after restart and transparent retries in case of failures
- Support parallel execution of queries making up a task

Performance:

- A data integration engine that meets Enterprise Class requirements.
- Asynchronous delivery of data enables rapid service response time
- Intelligent query processing delegation pushes processing to the most capable data source
- Advanced, configurable cache system enables view-by-view tuning to answer queries directly from cache
- Seamless session transfer among IE/Firefox browsers and lightweight HTTP + JavaScript clients enables the best of both worlds-automatic advanced navigation and extraction efficiency

Security, Scalability and Reliability:

- Granular authentication and control of databases and data views using LDAP or built-in security
- Firewall support: All components can be distributed in different network segments
- Communication between modules can be encrypted and authenticated using SSL
- High Availability: Support for 3rd party load balancers to distribute the workload among servers
- Robustness: Automated swapping, support for XA-compliant endpoints, fully transactional catalog storage
- System Management via JMX standard
- Query Execution Plan Trace: Visually inspect all details of query, before and after execution, by view and by source

Specifications

Platform

- Hardware: Pentium IV 2.4 GHz
- Memory: 1 GB RAM
- Disk space: 400 MB
- Operating System: Windows (2000 Server, 2000 Advanced Server, 2003, XP, Vista), Linux Fedora Ubuntu, CentOS, and any OS with Java support

Data Sources

- Any JDBC-compliant repository, Oracle (8i, 9i, 10g), MS SQL Server (2000, 2005), Sybase 12.5, Postgres (7.x, 8.x), MySQL 4.x, IBM DB2 8.x, etc
- Any ODBC-compliant repository (e.g. MS Access, MS Excel))
- Web services (with SOAP/REST protocol)
- XML (with or without schema)
- JSON
- LDAP
- Delimited Files
- Content extraction from Web sites, PDF, Word, POP3/IMAP/MS Exchange Server e-mail, RSS feeds, Salesforce.com (and other SaaS)
- Web connectors
- Customer connector API

Access Interfaces

- Web Services using SOAP 1.1 or REST
- JDBC 3.0
- ODBC via ODBC/JDBC gateway
- Fully documented Denodo Java API

Directory Services

- LDAP v3
- Denodo embedded service

Standards

- Access interface JDBC, ODBC, WS-*, WS-REST, J2EE
- Query: WSDL, SQL, Xpath
- Security: HTTP, SSL, WS-Security
- Authentication: LDAP
- Combination: SQL (with unstructured, hierarchical, and constrained data commands)
- Data Retrieval/Update:
 - Protocols - HTTP, HTTPS, SSL, JDBC, ODBC, SOAP, REST, LDAP, FTP, File Systems, POP3, IMAP, etc.
 - Formats - SQL, WSDL, HTML+JavaScript, XML, RSS, ATOM, JSON, PDF, Office documents, CSV, etc.
- Management: JMX
- Transaction Management: X/Open XA



530 Lytton Avenue,
Suite 302
Palo Alto, CA 94301
USA

Phone (+1) 650 566 8833

Fax (+1) 650 566 8836

Email: sales.us@denodo.com &
sales.apac@denodo.com