denodo

Whitepaper

The New Era of Gen AI: Enabled by Logical Data Management

If 2024 was the year that generative artificial intelligence (GenAI) started to become recognized as transformative, 2025 onwards will be the era when experimentation will give way to the realization of business and financial value. Organizations are starting to realize and measure this value, in the form of cost reduction, process efficiency, revenue generation, and better services across several industries.

However, GenAI-based applications can only be as reliable as the data they are based on. According to McKinsey, 72% of organizations identify data management as a key requirement in implementing AI use cases, and Boston Consulting Group found that "Only 22% of companies have advanced beyond the proof-of-concept stage [of GenAI] to generate some value, and only 4% are creating substantial value".

Unfortunately, many organizations are still struggling to implement a data foundation that can scale across the increasing number of data sources and increasing volumes of data required by GenAI, one that can enable all GenAI applications and other data consumers to access the data they need in real time, in a safe and privacy-compliant manner. Yet despite these data management challenges, GenAI yields considerable benefits. Industry-leading data management vendors have built GenAI-based functionality into their platforms to accelerate data preparation and consumption, automatically optimize for performance and cost-effectiveness, and keep up with ever-increasing data demands from the business, without sacrificing quality or security.

For example, the Denodo Platform can deliver data to all consumers, including GenAI, in the form they need, when they need it. To achieve this, the Denodo Platform leverages a logical approach to data management and integration. This approach abstracts access to multiple data systems for GenAI and other consumers, hiding the underlying complexity while delivering data according to predefined semantics and data governance rules.

Denodo recognizes the transformative potential of GenAI in fostering data democratization. The Denodo Platform not only automates tasks but also empowers users across a variety of roles, from end-users, to developers, to administrators, and data stewards.

In this position paper, we will cover each of these points in depth:

1. The use cases in which GenAI is enjoying early adoption and success in the enterprise
2. Some of the challenges of using GenAI in the enterprise
3. How the Denodo Platform addresses these challenges
4. How GenAI is being leveraged within the Denodo Platform

GenAI Use Cases in the Enterprise

Gartner predicts that by 2026, over 80% of businesses will utilize application programming interfaces (APIs) or GenAI models and/or implement GenAI-enabled applications in operational settings, a significant increase from the less than 5% reported in 2023.

The Deloitte AI Institute compiled 60 of the most compelling use cases for GenAI across six major industries, and the most common examples include:

Know your customer: GenAI can summarize information and create insights for processes such as know-your-customer (KYC), and create targeted personalized messages based on this information.

Marketing content assistants: Marketing content management is a challenge for organizations dealing with multiple websites in various languages across brand portfolios. Allocating time and resources for personalized content creation, such as product descriptions, images, videos, and audio, becomes difficult. GenAI provides a faster, more consistent solution compared to traditional tools, facilitating efficient content generation for enterprises.

Product design assistants: Product design traditionally takes time, with only one idea out of many making it to market. Overcoming human limitations in generating diverse ideas, fostering cross-industry inspiration, and simplifying concept testing are challenges. This process can be enhanced by integrating GenAI with CAD and other software aids in designing prototypes, enhancing creative thinking, brainstorming, and promoting out-of-the-box ideas.

Code assist for developers: Developers and other high-skilled professionals are in high demand and short supply. To overcome the talent gap, GenAI can be used to supplement a developer's effort by automating aspects of code creation and maintenance so the developer can focus on more complex code writing.

Customer support: GenAI has the potential to revolutionize customer interactions by employing speech-to-text and natural language inputs to produce empathetic and personalized conversations, particularly in after sales support and addressing customer complaints.

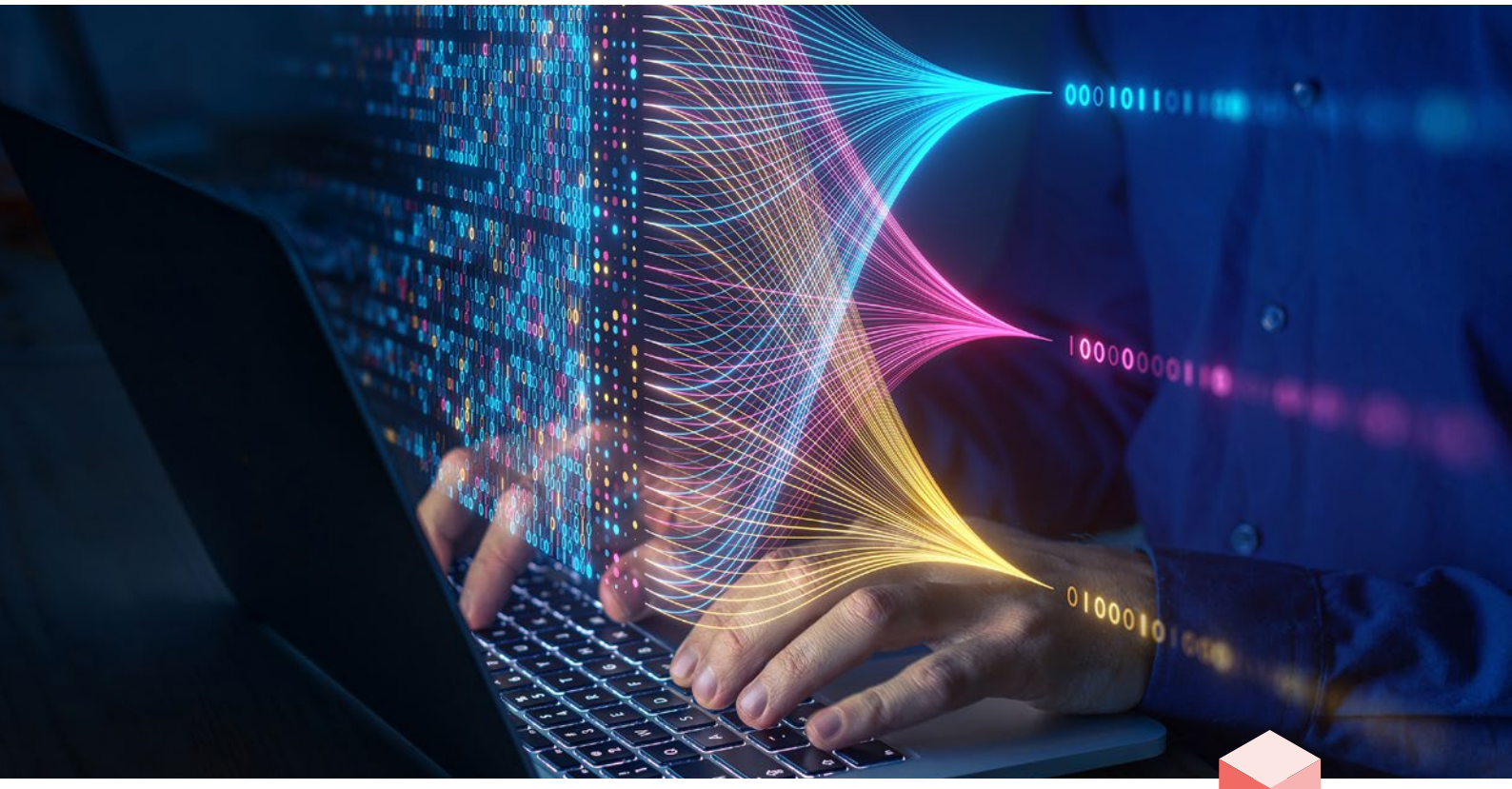
Asset maintenance planning: In the industrial sector, maintenance planning is crucial for preventing equipment failure and costly repairs, extending asset life. GenAI optimizes schedules based on operational factors, recommending efficient, cost-effective plans, while analyzing equipment data to minimize downtime and maximize availability.

Virtual field assistants for engineers: A virtual field assistant empowered by GenAI has the capability to function as a reference tool, offering swift access to an extensive array of technical information. In addition to furnishing pertinent details and guiding engineers to suitable resources, virtual assistants can aid in troubleshooting by addressing inquiries related to specific engineering concepts, principles, or calculations.

Resilient logistics and planning: GenAI can help identify and simulate potential disruptions or risks in the supply chain. By assessing port congestion, shipment routes, and tier-N supplier mapping, GenAI can be used to predict risks and their corresponding impact on operations and can also recommend actions to mitigate those risks.

Digital citizen services: Data about government and public services is often spread across multiple formats and locations (e.g., on-premises, cloud). A GenAI-enabled virtual assistant can serve as the interface between citizens and government information, answering questions in natural language.

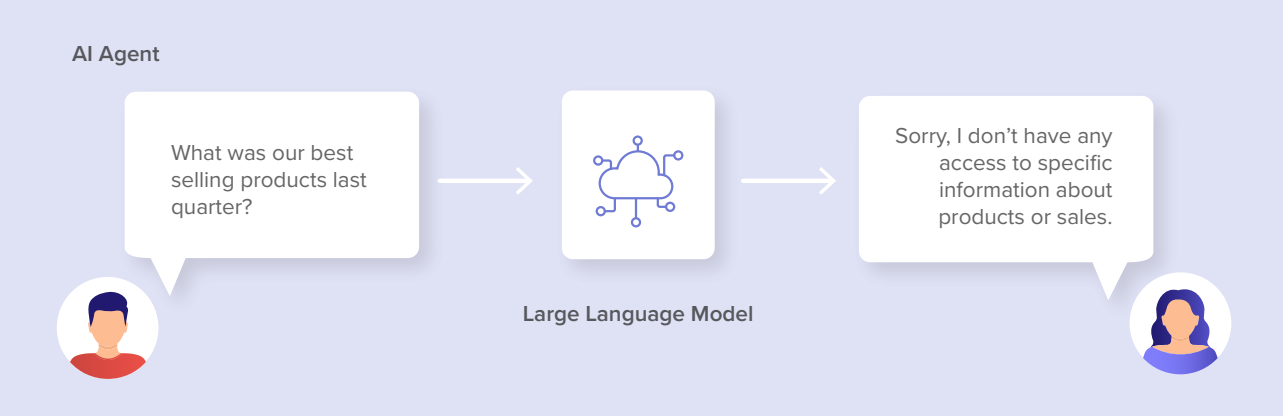
These are just some examples of how GenAI is transforming how organizations operate and do business. The next several years promise even more creativity and innovation, rivaling the Internet itself in the potential to transform how we live and work.



Challenges of Using Generative AI in the Enterprise

As organizations recognize GenAI's transformational potential, they begin to invest in relevant technologies and competencies and start experimenting with initial use cases. However, they quickly run into a variety of data-related issues. According to **McKinsey**, 72% of organizations already identify data management as a key challenge in scaling AI use cases. For example, for GenAI-enabled customer personalization to be safe and effective, the underlying customer data needs to be accurate and up-to-date, and near-real-time data access is required across multiple diverse data sources.

> "At least 30% of GenAI projects will be abandoned after proof of concept by the end of 2025, due to poor data quality, inadequate risk controls, escalating costs or unclear business value."
>
> — **Gartner**



GenAI Agent Struggling to Answer an Organization-Specific Question

The problem is, GenAI relies on large language models (LLMs), and these have inherent limitations. LLMs are only as smart as the data they were trained on. While LLMs have encyclopedic knowledge of historical events and all the literature ever digitally recorded, they know nothing about individual organizations, such as details about their customers, products, or operations, and the LLM's dataset is not updated in real time. Without up-to-date customer or business context, GenAI simply isn't useful for any operational use case.

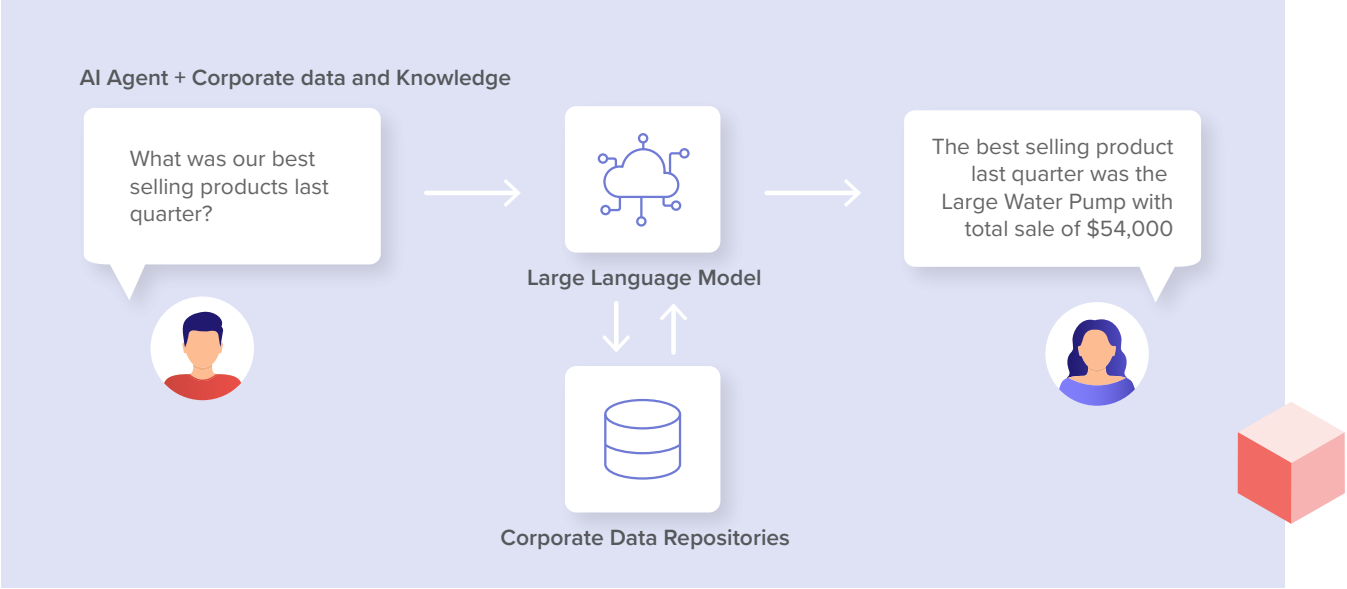
While it is possible to train and fine-tune existing foundational models with additional information and make them aware of corporate data and information, this path often leads to more complexity and challenges. Not only are the cost and skills required to re-train LLMs prohibitive for most organizations, but the ongoing need to make LLMs aware of the latest data and information makes the iterative process of constantly re-training LLMs simply impractical.

> "What won't matter as much for AI strategy is your choice of large language model (LLM). There will be many good options. Everyone will be using them. A shrewd strategy will instead emphasize what can set you apart — how you leverage AI with your institutional knowledge and proprietary data."
>
> — **PWC: 2025 AI Business Predictions**

In addition, training LLMs with corporate information (which is potentially sensitive), to embed that knowledge into the model itself, is also fraught with danger, as there is always the risk of data leaks. A savvy user can engineer prompts that elicit disclosure of such information, even if the model was trained to recognize sensitivity; such training is far from perfect.

Fortunately, an emerging implementation pattern promises to overcome the limitations of LLMs and deliver the knowledge that they need within an enterprise context in a safe and effective way. Retrieval augmented generation (RAG) architecture was first discussed in a **paper by Meta** in 2020, but it is quickly becoming the preferred method to augment LLMs with additional data and information in a cost-effective, secure manner.

RAG enables the incorporation of real-time-updated data into GenAI's results, but this still requires a comprehensive data management solution to provide relevant, high-quality, privacy-compliant data in real time.



GenAI Agent Augmented with Corporate Data and Knowledge

While data in traditional machine learning use cases have played a critical role in the upfront training process, GenAI and RAG are changing that paradigm and require data to play the role of the knowledge augmentation layer during the inference process. Instead of infusing knowledge into the LLMs by re-training the model, RAG architecture involves adding knowledge through additional context windows during the prompting process. The LLM can then use the additional context provided to generate the necessary response without that knowledge being embedded in the LLM itself.

The benefits of RAG include reduced hallucinations, more up-to-date and real-time information (overcoming the limitations of training the model on 'point-in-time' data); domain-specific knowledge (such as the product sales example depicted above); elimination of costly retraining, and visibility into the source of the data retrieved.

This simple approach can be applied to structured and unstructured information and is a more agile, cost-effective, and safer way to provide LLMs with additional knowledge and information. Coupled with the powerful code generation capabilities of LLMs (such as SQL to query the underlying enterprise data), next-generation AI applications using RAG can open up new types of powerful user interactions and new ways for organizations to unleash the value of their data.

However, one needs the necessary data management foundation to implement RAG within an enterprise context. While one can now find plenty of simple GitHub projects demonstrating RAG's power against a single table in a desktop setup, implementing the RAG architecture in an enterprise context against real production data across an often siloed, complex enterprise data landscape can be a nightmare. As organizations explore tighter integration between LLMs and enterprise data sources, they will inevitably encounter the same data management challenges in terms of overcoming data silos, dealing with diverse data source types, and managing complex and lengthy data delivery pipelines.

ORGANIZATIONS SHOULD CONSIDER THE FOLLOWING DATA MANAGEMENT CHALLENGES WHEN ADOPTING GENAI:

Regulations: New regulations, guidelines, and frameworks are rapidly emerging. These address a potential lack of transparency in the functioning of GenAI systems, the data used to train them, issues of bias and fairness, potential intellectual property infringements, possible privacy violations, third-party risk, and security concerns. The EU AI Act aims to establish comprehensive regulations for high-risk AI systems, with specific requirements for transparency, data governance, human oversight, and risk assessment. The Executive Order issued by the U.S. President establishes new standards for AI safety and security, protects Americans' privacy, advances equity and civil rights, stands up for consumers and workers, promotes innovation and competition, and more. In 2023, China enacted the world's first law specifically regulating GenAI, imposing restrictions on training data and outputs for companies offering such services. South Korea's AI Act is aimed at protecting users of AI-based services by ensuring trustworthiness of the AI systems through, for example, more stringent notice requirements for high-risk AI services and certification systems for AI trustworthiness. Other regional regulations and guidelines have advocated or even mandated the use of privacy-enhancing technologies for data collection, for example: Singapore's Model AI Governance Framework for Generative AI.

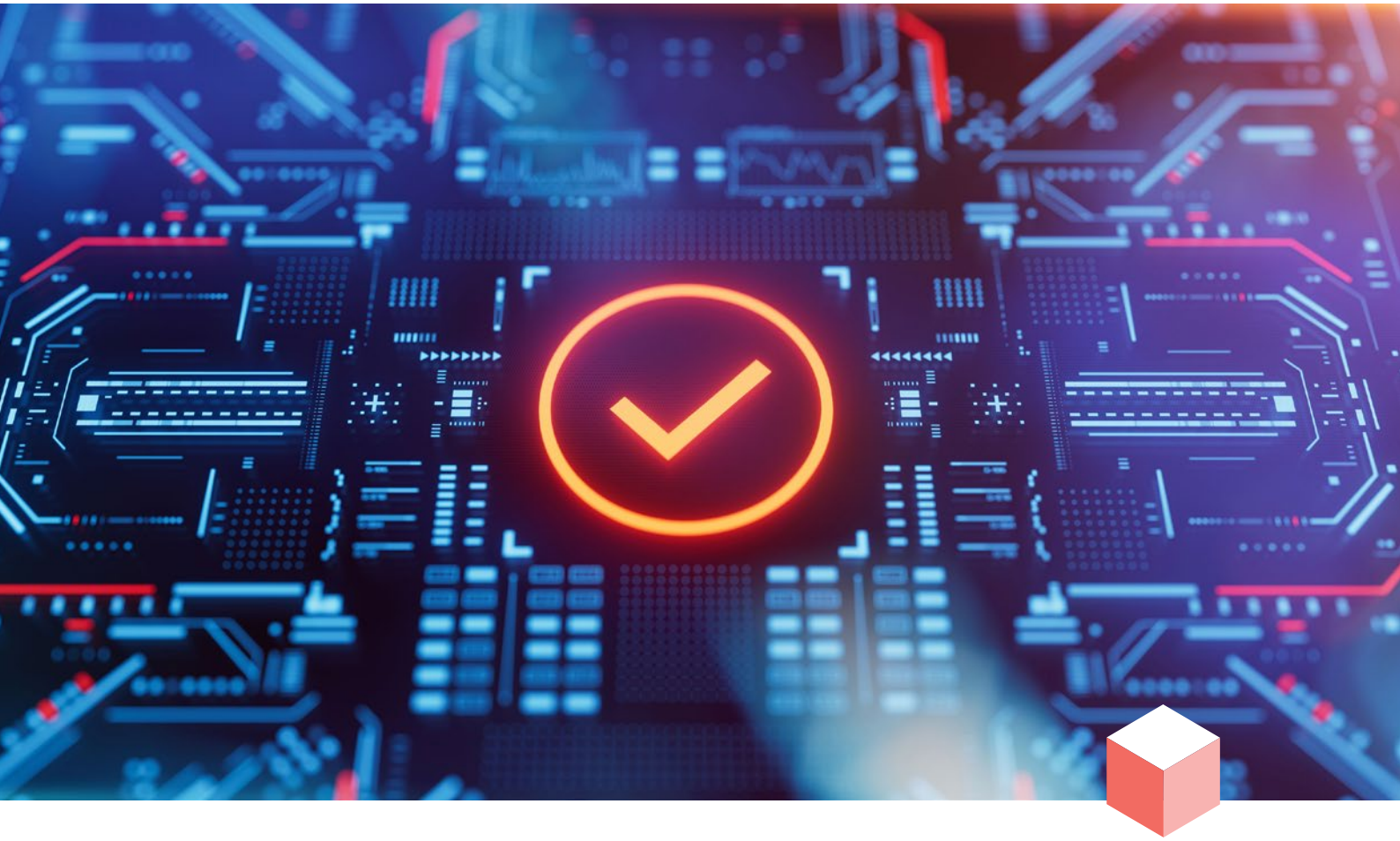
Additionally, these regulations require auditability and explainability of GenAI results. For example, the **EU AI Act** requires auditability that GenAI is free of bias and is not causing harm to consumers, and transparency to consumers that GenAI was being used during particular interactions (e.g. disclosing that the chatbot they are speaking with, or the recommendations being generated for them, are powered by GenAI). Responding to such audits requires the ability to see what data was used by GenAI when producing a given result and ensuring controls are in place to monitor what data is being used by which GenAI use cases, and when.

Data Quality and Explainability: Poor-quality data can result in incorrect or inconsistent behaviors – "garbage in-garbage out." This is true both during model training of the LLM, as well as subsequent real-time enterprise data access via RAG. The resulting "hallucinations" lead to end-user mistrust and, in the context of enterprises subject to regulations governing the ethical use of AI, potential compliance violations.

As a result, rigorous data quality is necessary, but not sufficient. GenAI also requires "data explainability," which is the ability to specify what data was accessed in order to produce a result. This way, GenAI users can readily ascertain that the data used was correct and accurate, and easily act to remediate, if it was not. Data explainability provides transparency, which then leads to greater end-user trust. Indeed, some of the above-mentioned regulations, notably the EU AI Act, requires this level of transparency on-demand.

Data privacy and security: Privacy regulations such as the General Data Protection Regulation (GDPR) in Europe still apply, but the challenge is not limited to ensuring that correct access controls and de-identification measures are in place. GenAI poses additional risks.

Let us look at an example of how a GenAI model may pose a privacy risk. If "Jane Doe's" private data was used in training a model, then a prompt such as "my name is Jane Doe and my phone number is..." could lead to the model revealing this information (Seth & Chang, 2024). Models have been known to "memorize" training data, compounding the problem. Several types of attacks, such as training data extraction, have been shown to leak sensitive data from LLMs (Seth & Chang, 2024). All of the above-mentioned AI regulations require organizations to have controls in place for mitigating the risk of such privacy violations occurring.



How Denodo Addresses These Challenges - Logical Data Management

Denodo believes that organizations must evolve and re-imagine data management, as the stakes have never been higher. Organizations must act now to build a data management foundation that can prepare them for a new AI-driven competitive landscape that is rapidly approaching us all.

Having considered the data management capabilities needed to power the next-generation AI applications, we believe that a logical approach to data management is needed to power next-generation AI applications.

The Denodo Platform, leveraging **data virtualization** technology, is a logical data management solution that eliminates the need for data movement or consolidation before augmenting an AI application. It provides AI applications with a single, consolidated gateway to integrated data. The Denodo Platform offers a number of key benefits, including:

- A unified, secure access point for LLMs to interact with and query all enterprise data (ERPs, operational data marts, EDWs, application APIs)
- A rich semantic layer, providing LLMs with the needed business context and knowledge (such as table descriptions, business definitions, categories/tags, and sample values)
- The quick delivery of logical data views that are de-coupled and abstracted from the underlying technical data views (which can be difficult to use by LLMs)
- The delivery of LLM-friendly logical table views, without first needing to physically combine multiple datasets
- Built-in query optimization relieves LLMs from dealing with specific data source constraints or optimized join strategies.

With these benefits, the Denodo Platform is a perfect enabler for RAG. The semantic layer makes data stored in the canonical data platforms accessible with one consistent, secure interface for GenAI applications. The Denodo Platform also has the necessary metadata to provide GenAI apps with the information they need, including data schemas, field descriptions with contextual information, and business names of the field.

INTRODUCING: QUERY RAG

Most RAG documentation and examples today focus on augmenting LLMs with unstructured data. As a result, the techniques for implementing RAG using data scraped from websites and PDFs have matured rapidly and are proving to be robust and scalable. However, augmenting LLMs with structured data presents a different set of challenges, such as:

- Tabular data not only needs to be retrieved, it must be accessed and queried via correct SQL.
- The access mechanism and optimal SQL needed to retrieve data can differ from vendor to vendor and platform to platform.
- Vector-embedding strategies must cover both metadata and actual data.
- Structured data (and its associated metadata) changes constantly.
- Existing data security and permissions controls need to be respected and integrated.

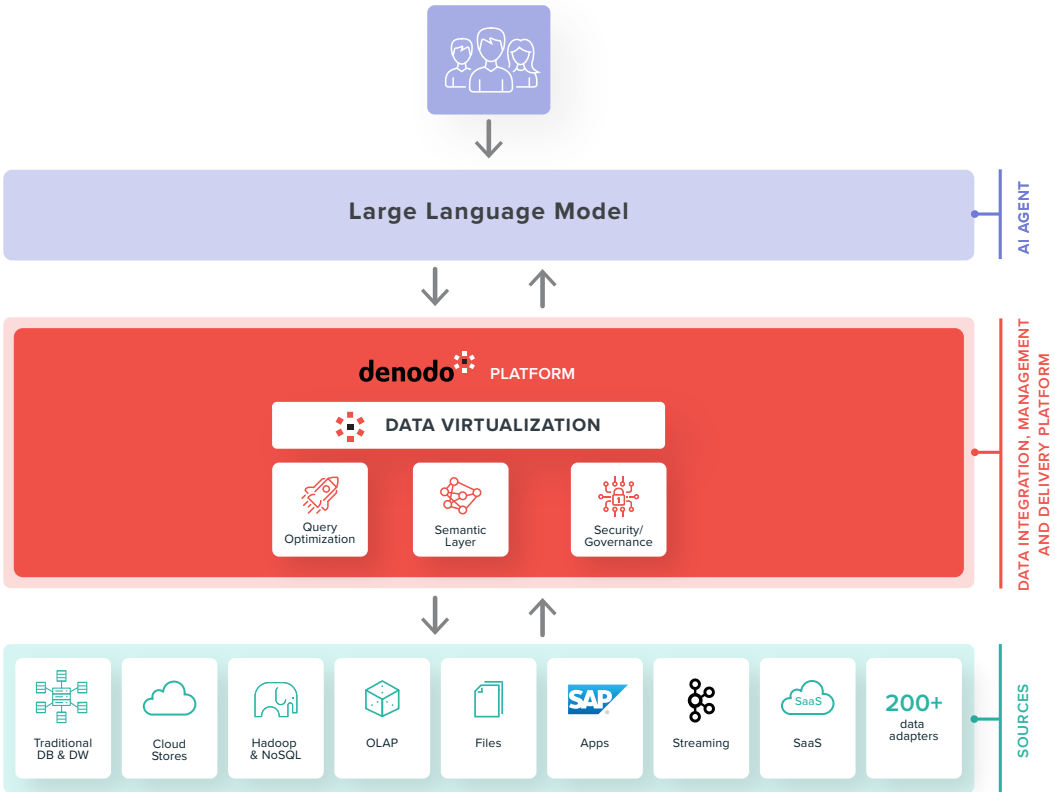
These unique challenges highlight the need for a distinct approach to handling structured data. The Denodo Platform supports the seamless implementation of Query RAG, a variation of RAG designed to meet these challenges. Query RAG provides these benefits:

- A robust and scalable way to create vector embeddings across all metadata, data sources, and platforms
- A unified SQL access engine that abstracts away the complexity of the underlying data sources and systems
- A powerful query optimizer that enables LLM-generated SQL to be highly performant across multiple data architecture patterns and use cases.
- A flexible, robust semantic layer that enables you to easily develop new data views and extend metadata
- A simple process for enforcing user permissions and data security at every step.

One of the more unique aspects of Query RAG is that the goal of the embedding search is not to find the text for augmenting the prompt but to find the right data view to query. Once the right data view and correct SQL have been identified, the tabular data is then retrieved from multiple data sources and fed to the LLM.

With Query RAG, enterprise data repositories can be utilized in a robust, scalable way. By leveraging embedding and exposing real-time data to AI applications, this approach can help organizations extend their GenAI capabilities beyond unstructured data, delivering real-time insights and improving decision-making in new, innovative ways.

The benefits of RAG include reduced hallucinations, more up-to-date and real-time information (overcoming the limitations of training the model on 'point-in-time' data); domain-specific knowledge (such as the product sales example depicted above); elimination of costly retraining, and visibility into the source of the data retrieved.



The Denodo Platform Augmenting and Supporting an AI Interaction

The Denodo Platform can become the single serving layer for one of the most critical components of a GenAI application - your data. The combination of LLMs and the Denodo Platform can significantly accelerate the development of powerful AI agents.

THE DENODO PLATFORM OFFERS THE FOLLOWING ADDITIONAL CAPABILITIES TO PROVIDE A SOLID DATA FOUNDATION FOR GEN AI:

DATA QUALITY AND USAGE MONITORING:

The Denodo Platform offers a vast library of transformation, filter, and matching functions, and quality rules for validating, cleansing, enriching, standardizing, matching, and merging data, such as conditional processing, partitioning, fuzzy match algorithms for deduplication and cleansing, as well as syntax based, thesaurus like, or semantic mappings. Additionally, the Denodo Platform is continuously monitoring what data was accessed when, and by whom. For every GenAI query initiated through RAG, the Denodo Platform logs the query as well as what data was returned, providing always-on explainability and transparency.

DATA SECURITY AND PRIVACY COMPLIANCE:

As a single point of data access for applications, it provides a central layer for enforcing access security restrictions that can be defined in terms of the canonical model with field-level granularity. The Denodo Platform supports user and role-based authentication and authorization mechanisms with both schema-wide permissions (e.g., to access Denodo databases and views) and data-specific permissions (e.g., to access the specific rows or columns in a virtual view). The Denodo Platform offers both row-based and column-based security, including the possibility of masking specific fields (e.g., managers are not allowed to view the “salary” column of higher-level management, so those cells would appear masked in the results).

These features make the Denodo Platform ideal for enforcing data security for RAG-enabled AI applications and helping organizations to meet a wide variety of regulatory requirements as described above.

The journey to harnessing the full potential of LLM-powered AI agents is ongoing and requires continued technological advancements and innovations by the industry as a whole. Denodo is at the forefront, evolving our offerings to meet the demands of an AI-enabled future. Our talented product management team continually monitors and assesses the newest advances in AI and other modern technologies to provide an enhanced foundation for data management. We are devoted to empowering our customers to effectively embrace these innovative technologies.

ENABLING THE NEXT GENERATION OF GENAI-POWERED AGENTS AND APPLICATIONS:

The next wave of GenAI will consist of purpose-built agents and applications that enhance or automate many aspects of customer engagement, business operations, and other corporate functions. These agents and apps will proliferate, as LLMs are supported as embedded systems on mobile devices (such as the Apple iPhone 16) and in-memory within modern application stacks. However, developers of such “smart apps” often experience significant data management challenges, such as:

- Connecting to and accessing data from complex, often siloed data sources
- Simplifying and accelerating the iterative development of data products intended to be consumed by these apps, so that this effort does not get in the way of deploying these apps to production

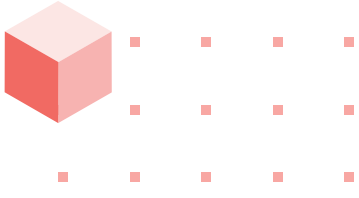
- Accessing real-time data in a scalable fashion, given the app's operating environment
- Developing a unified, rich semantic layer to improve app performance and accuracy
- Providing data security and governance across the different data repositories and technologies
- The Denodo AI SDK solves these challenges, providing developers of AI-powered agents and applications with unparalleled productivity and agility.

The Denodo AI SDK is an open-source software package that includes pre-integrated components and optimized, re-usable code. The SDK can be downloaded and used by Denodo Platform customers to implement RAG based architecture and accelerate the development of AI based agents or applications. The SDK includes streamlined integration with all popular commercial LLMs via configuration parameters. It also includes an example front-end application/chatbot to assist with testing and custom developments.

By offering a set of streamlined APIs purpose-built for Query RAG, the Denodo Platform and the AI SDK empower organizations to easily construct conversational experiences atop any data repositories. This means that it supports analytics-driven questions as well as operational requests that leverage live production data in real time. All the while, the Denodo Platform enforces the necessary data security and governance policies, aligning data access and usage with organizational control, without requiring additional manual intervention.

To benefit developers, the SDK:

- Simplifies integration with multiple data sources, reducing the need for custom connectors and manual data handling
- Streamlines data transformation and merging, automatically preparing data for combined analysis and LLM input
- Reduces orchestration complexity with pre-built APIs that handle embedding, vectorization, and LLM-querying in a single flow
- Enhances security and governance by managing access control and compliance within the SDK, minimizing risks
- Accelerates development timelines by removing repetitive coding tasks and focusing on higher-value features
- Improves data accuracy and consistency, enabling rich, contextually relevant responses from GenAI applications
- Supports scalability and flexibility through open-source components, providing the level of customization to meet specific use cases



How Gen AI Is Enhancing Data Management

GenAI has also been used to improve data management. It introduces highly valuable automation to manual processes that have been prone to human error. Foundational data management tasks that have until now been labor-intensive and error-prone can be handled more efficiently and accurately by AI. According to Gartner's **2025 Planning Guide for Data Management**: through 2026, GenAI will reduce manually intensive data management costs by up to 20% each year while enabling four times as many new use cases.

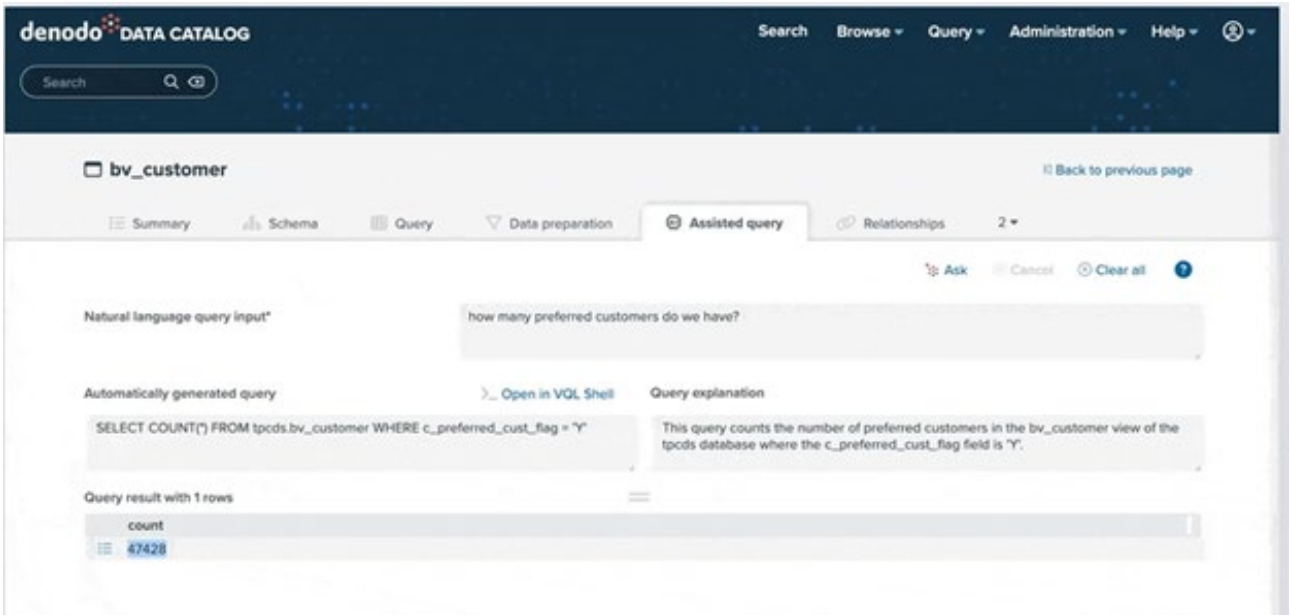
Denodo Assistant delivers on the promise of intelligent data fabric, helping users be more productive. It leverages the rich semantic layer and active metadata maintained by the Denodo Platform, as well as AI and machine learning (ML) techniques, to power all aspects of the Denodo Platform.

> "Denodo's GenAI capability is a transformative innovation that allows anyone to access and interact with data using natural language. It removes the need for knowing technical details, such as database names or the particulars of a query language. By removing friction and barriers, Denodo's GenAI capability can simplify data access, maximize value, and drive data democratization."
>
> — **Mark Thorogood, director, Data Operations and Software Engineering, Perkins Coie**

By leveraging internal algorithms developed by Denodo and by integrating with LLMs offered by industry leaders, Denodo Assistant helps developers and users throughout the data product development cycle by intelligently **answering** questions, offering **suggestions**, and **recommending** solutions.

DENODO ASSISTANT IS COMPRISED OF THE FOLLOWING FEATURES:

Natural language querying. Any user can simply type a question in natural language, which is then converted to SQL code and executed by the platform. This is a big step in **Data Democratization** – enabling business professionals with no knowledge of SQL or BI tools to generate insights from the data. For example, a sales manager may query the data using everyday language, such as, "I need the name, last name, and email of all preferred customers in the United States." A data platform, using GenAI, may translate this natural language query to SQL code, and even generate explanations for the code. The data platform would then execute the code and retrieve the results.



This holds great potential for business professionals, including analysts, marketers, salespeople, healthcare professionals, and more – who can independently access and draw insights from data, without relying on IT and **data professionals**. The Denodo Data Catalog also includes recommended data views and assets to aid data discovery and self-service. This provides business decision makers with the data they need to make informed decisions with much greater agility.

> By 2025, use of natural language as a primary data management API will be the dominant interface leading to a 100x consumption of data across the ecosystem."
>
> — Top Trends in Data and Analytics, 2024, Gartner

Query Optimization. The Denodo Platform has long used AI for query optimizations, including join order, associations, and aggregations. Denodo Assistant includes summary cache (for query acceleration) using ML and active metadata. This helps to optimize query performance with reduced costs.

Data Preparation and Enrichment: The Denodo Platform also features data preparation functionality that includes AI-recommended transformations.

A data user, upon accessing a given data set, is presented with a set of recommended transformations based on prior behaviors using the same or similar data sets. This capability improves with use, as the Denodo Platform learns and updates its metadata about the given data, including when and where it was used, and relationships with other data.

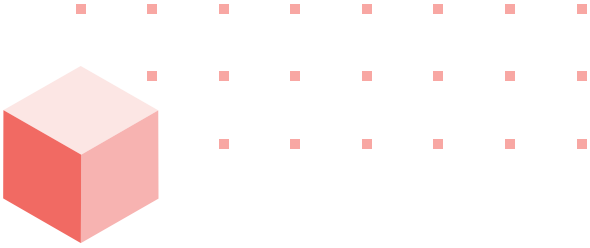
Denodo Assistant also suggests descriptions for views and columns, reducing the burden on data governance professionals. To enhance developer productivity, Denodo Assistant includes features such as Intelligent Autocomplete for SQL and Query Wizard Recommendations.

Preparing for a Gen-AI-Enabled Future

GenAI continues to be a transformative force across diverse industries.

Today, when LLMs are easily accessible, and everyone uses the same foundational models, it is your data that will give you a competitive advantage. From that perspective, nothing has really changed. What has changed is how you need to think about data management to unleash the full potential of GenAI capabilities.

As we stand on the brink of a GenAI-enabled future, it's crucial to ask: Do you have the necessary data infrastructure and data management foundation to not only leverage GenAI effectively, but also to build a sustainable and lasting competitive edge?



denodo

Denodo is a leader in data management. The award-winning Denodo Platform is the leading data integration, management, and delivery platform using a logical approach to enable self-service BI, data science, hybrid/multi-cloud data integration, and enterprise data services. Realizing more than 400% ROI and millions of dollars in benefits, Denodo's customers across large enterprises and mid-market companies in 30+ industries have received payback in less than 6 months.

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