Informatica | aws

eBook

Accelerate Enterprise-Grade GenAI Adoption with Comprehensive Data Management

How Informatica's data management capabilities enable AI-ready data for GenAI applications built on AWS.

Where data & AI come to LIFE

Contents

Introduction

Businesses are turning to generative AI (GenAI) to increase productivity, reduce operational costs and achieve their business objectives. Across industries, companies continue to take advantage of GenAI to generate content, automate tasks, and enhance the customer experience through fast, accurate responses to questions.

However, as enterprises embrace this emerging technology, they often face challenges in moving from proof of concept (POC) to full-scale production. Key roadblocks include:

- Inconsistent data

- Insufficient data governance

- Poor data quality

- Weak data management practices

- Excessive time building pipelines

To tackle these challenges, enterprises need a modern data management platform that can simplify the creation of data pipelines, facilitate data governance, and enhance application development with critical GenAI capabilities such as vector store integration. This platform should also provide trusted and curated enterprise context, metadata intelligence, and agentic RAG orchestration.

Barriers to Building Enterprise GenAI Applications

Despite the promising potential of GenAI, enterprises encounter substantial hurdles in its deployment:

Data Integrity and Consistency

GenAI's success hinges on overcoming datarelated hurdles. Large language models (LLMs), when not grounded in quality data, can produce incorrect or misleading results, commonly known as hallucinations. Unfortunately, fine-tuning Language Models in your business and enterprise context can be difficult to manage, with high computational costs and increased maintenance overhead. In addition, when data is fragmented and not easily discoverable, organizations only have access to partial information.

Semantic Contextualization

LLMs generally don't understand your specific enterprise terminology and semantics. Without the right level of data management, information is locked within metadata, and enterprises struggle to inform LLMs on how to use that data. Even though LLMs are smart enough to answer questions, they still don't have the context and nuance of the complex nature of the enterprise without access to metadata. Data personas and data communities need to be included in the development process to improve data quality and accuracy.

Data Quality

Not all areas of an enterprise have a data quality strategy in place, and there is often a lack of highquality, curated, trusted data. When deprived of high-quality data and access to metadata, GenAI applications lack reliability, and response accuracy can suffer as a result. Ensuring high-quality data requires detecting and mitigating data bias by dynamically bringing in high-quality data sets.

Complex Development and Deployment

Building GenAI applications can be complex, making it challenging for enterprises to rapidly adopt new and evolving GenAI capabilities. To move from POC to production, developers need the ability to build and deploy without relying on coding and manual tasks. Data is the biggest differentiator in GenAI, but providing comprehensive answers to prompts is difficult when data is fragmented. Developers often spend significant time creating pipelines to bring data from different parts of the enterprise into the vector database in a format that large language models (LLMs) can understand.

Governance and Security

Enterprise-ready GenAI applications require transparent, traceable data. Data must also be contextual and democratized to deliver relevant GenAI applications. Additionally, enterprises need to enforce data access policies and establish cost and usage control to ensure governance and security. This involves making sure the right people in their organizations have access to the right data based on their role even with GenAI applications.

Informatica:
The Data Management Platform for GenAI

Informatica offers the Intelligent Data Management Cloud (IDMC), an AI-powered software as a service (SaaS) solution that runs on Amazon Web Services (AWS). This platform enables a cohesive, incremental and comprehensive approach to data management that minimizes the undifferentiated heavy lifting of building a modern data foundation that is ready for GenAI.

Through IDMC, Informatica is investing in a set of purpose-built tools for GenAI that enable enterprises to onboard a data foundation for designing and deploying enterprise-grade GenAI applications.

Data Pipeline for Vector Databases

The Cloud Application Integration service as part of the IDMC platform offers a no-code, low-code data pipeline for vector stores. This allows developers to bring high-quality data from more than 300 data sources into a format that LLMs easily understand. With these capabilities, developers eliminate the need to invest time and resources in creating data pipelines.

Master Data Management (MDM)

MDM plays a crucial role in supporting GenAI by ensuring that the data used for training and generating responses is accurate, consistent and comprehensive. Informatica MDM enhances the effectiveness and reliability of GenAI by providing clean, consistent and contextualized data, enabling organizations to leverage GenAI capabilities more effectively.

IDMC Platform

- Multi-tenant SaaS platform running on AWS

- Comprehensive data management capabilities for AI-ready data

- Connectors for Amazon Bedrock and Amazon Bedrock Knowledge Bases to build and orchestrate GenAI applications

- Ability to streamline workflows and automate repetitive tasks with agents for Amazon Bedrock

Trusted Context

Data is often fragmented across the organization and available in different data lakes and other enterprise systems. With Informatica MDM, enterprises are able to deliver a single, trusted source of truth, ensuring consistent, high-quality data across the organization to drive informed decision-making, enhance customer experiences, and maintain compliance with governance standards. Bringing this high-quality trusted data as context into GenAI can enhance the accuracy and reliability of responses generated by GenAI applications.

Informatica:
The Data Management Platform for GenAI (continued)

Informatica Metadata Intelligence

Given the complex nature of an enterprise landscape, GenAI applications need to have additional context about the nuance behind the data. For example, data quality, data lineage, business glossary, field / column names for custom data objects etc. Cloud Data Governance & Catalog (CDGC) and MDM services within IDMC enable this by not just acting on the data itself, but also on the intelligence locked within the metadata. This improves the overall accuracy and relevance of LLMs and provides a layer of validation and traceability. For end users, it's the difference between not knowing whether to take a response seriously and making critical decisions based on a response, knowing that it is accurate and originates from data across the enterprise.

Agentic RAG Orchestration

GenAI applications are often built using retrieval-augmented generation (RAG). Most of the use cases that enterprises are trying to solve with GenAI need a combination of different LLM models, data retrieval mechanisms and access to external systems. The IDMC Cloud Application Integration (CAI) service includes no-code, low-code RAG agent framework to build agent based complex applications, while allowing customers to democratize enterprise GenAI adoption to include the data community personas in the development process.

Streamlining GenAI Development with Amazon Bedrock

Amazon Bedrock is a fully managed service that provides a unified API for accessing and utilizing various highperforming foundation models. With Amazon Bedrock, enterprises can take advantage of a broad choice of LLM models (from Anthropic, Cohere, Mistral, etc.) and vectorDB engines (including Amazon OpenSearch, Amazon Aurora, etc.) to build GenAI applications with security, privacy, and responsible AI practices.

Amazon Bedrock delivers the following capabilities:

Data Integrity and Consistency

Amazon Bedrock makes it easier to develop enterprise GenAI applications by giving developers the flexibility to use a unified API while using foundation models (FMs) from different providers. In addition, fully managed agents for Amazon Bedrock extend FM reasoning capabilities to break down tasks, create an orchestration plan, and execute it.

Semantic Contextualization

By taking advantage of Amazon Bedrock's support for RAG, developers can extend the power of FMs with proprietary data. Amazon Bedrock Knowledge Bases enable developers to securely connect FMs to data sources for retrieval augmentation — from within the managed service — making the FM more knowledgeable about a specific domain and organization.

Data Security

Amazon Bedrock offers capabilities to support data security and compliance requirements. For example, Amazon Bedrock supports common compliance standards including Service and Organization Control (SOC), International Organization for Standardization (ISO), and the General Data Protection Regulation (GDPR). Amazon Bedrock is also Health Insurance Portability and Accountability Act (HIPAA) eligible and is certified for CSA Security Trust Assurance and Risk (STAR) Level 2.

A Blueprint for Building Enterprise-Grade GenAI Applications

Using IDMC and AWS, organizations can get started with the blueprint for designing and deploying enterprise-grade GenAI applications into production. This blueprint includes the following steps:

1 Onboard Data Sources

The first step is using the Informatica CDI pipeline to ingest data from internal and external data sources across the enterprise, and then generating vector embedding using an embedding model (e.g. Amazon Titan, Cohere, etc.) and storing it into Amazon Bedrock Knowledge Bases.

2 Prepare Enterprise Metadata

Once you have onboarded data, you can vectorize semantic content including data quality metrics, data access policies and data lineage from Informatica CDGC and store the embeddings into Amazon Bedrock Knowledge Bases.

**3 Implement Agentic RAG Application**

Implement the planning agent, orchestrator, and different executors to fetch data from various systems and the summarizer. For implementation, you can leverage the Informatica's No code/Low code CAI framework and the CAI connector for Amazon Bedrock to enhance and standardize response summaries. Additionally, access control enforcement takes place during data fetching.

4 Develop and Deploy Application Front-End

Implementing a successful architecture requires a front-end framework for building applications, an API management layer to manage rate limiting and advanced authentication, a robust data management foundation to orchestrate retrieval of relevant data and metadata intelligence and finally, Amazon Bedrock, which provides a variety of models to choose from and the security controls essential for the enterprise. The Informatica and AWS blueprint for building GenAI applications also includes Amazon Q, a GenAI–powered assistant. With this fully consumption-based solution, enterprises can now choose the right services to build the GenAI applications that are fit for their business.

Building Generative AI App on AWS with Informatica
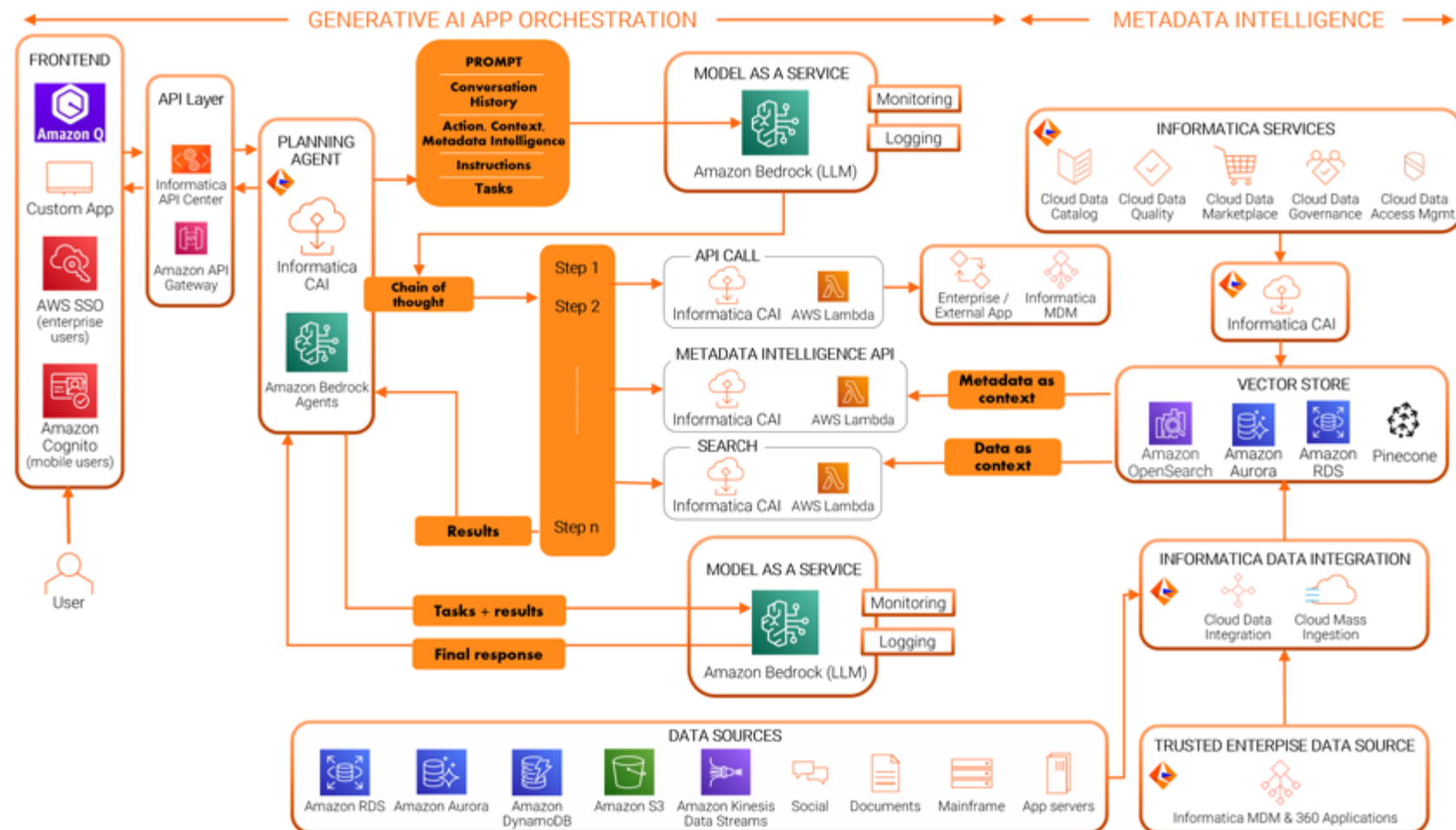
Accelerating Innovation: Streamlining Data Management and Simplifying GenAI Application Development

Taking advantage of the AI-powered, microservices-based IDMC helps enterprises become more data-driven, develop more innovative products and services, and deliver exceptional customer experiences by:

Speeding Up Development

By simplifying data foundations and overall data management and taking advantage of choice of LLM available in Amazon Bedrock Model as a Service, enterprises can streamline RAG orchestration to reduce the time required to design and deploy new GenAI applications. In addition, by increasing scalability and optimizing performance with elastic and serverless computing, enterprises can drive new growth.

Adding More Context to GenAI Data

IDMC allows enterprises to enhance response context by improving response generation through integrating measures such as data completeness, consistency, accuracy, and timeliness. Enterprises can also employ semantic annotations and metadata from glossary terms to enrich the context with domain-specific knowledge while allowing the system to dynamically select the most suitable datasets for contextually relevant responses based on quality measures.

Ensuring Data Accuracy and Transparency

By leveraging high-quality data from Informatica master data management (MDM) and linking to data sources, an enterprise can enhance the overall accuracy of GenAI applications and improve the accuracy of GenAI models by using high-quality, authoritative, trustworthy data.

Using IDMC to develop GenAI applications, enterprises can enable transparency in response generation by detailing the quality indicators and terminology explanations using metadata and indexing from the data catalog.

Reducing Regulatory Risk

An enterprise can lower their risk by ensuring the accuracy and protection of sensitive data and ensuring that data used for responses complies with regulatory requirements and organizational policies to mitigate data breaches. Informatica helps reduce risk with built-in IDMC governance capabilities, as well as integration with Amazon Bedrock Guardrails, which offer an additional layer of safeguards regardless of the underlying FM.

Conclusion

Taking GenAI applications to production is a challenge for most enterprises, as they struggle with data fragmentation, data quality & accuracy issues. To overcome these challenges, enterprises must invest in an integrated data management strategy to create the right data foundation for GenAI development.

One way to support a successful data foundation and accelerate development is by investing in a solution such as Informatica IDMC. With IDMC and AWS, enterprises remove the barriers of GenAI development while rapidly creating the kinds of GenAI applications that can transform their business. Informatica and AWS have built this comprehensive GenAI architectural blueprint to enable successful enterprise adoption of GenAI.

Learn more about the Informatica GenAI blueprint for AWS, how you can leverage it for free and how to build enterprise-ready GenAI applications.

About Us

Informatica (NYSE: INFA), a leader in enterprise AI-powered cloud data management, brings data and AI to life by empowering businesses to realize the transformative power of their most critical assets. We have created a new category of software, the Informatica Intelligent Data Management Cloud™ (IDMC), powered by AI and an end-to-end data management platform that connects, manages and unifies data across virtually any multi-cloud, hybrid system, democratizing data and enabling enterprises to modernize their business strategies. Customers in approximately 100 countries and more than 80 of the Fortune 100 rely on Informatica to drive data-led digital transformation. **Informatica. Where data and AI come to life.™**

Where data & AI come to

LIFE™
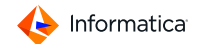
Informatica

Worldwide Headquarters
2100 Seaport Blvd.
Redwood City, CA 94063, USA
Phone: 650.385.5000
Fax: 650.385.5500
Toll-free in the US: 1.800.653.3871

informatica.com
linkedin.com/company/informatica
x.com/Informatica

CONTACT US

IN19-5100-1224