

Accelerate Modern Analytics and AI with a Cloud-Native Cloud Data Integration and Engineering Solution

Modern analytics and artificial intelligence (AI) are transforming enterprises into insight-driven, adaptive organizations. While analytics enables insights and fosters decision-making, AI extends it by enabling machines to uncover hidden patterns beyond human capacity, generating predictive insights and driving proactive decision-making. It moves analytics from reactive and diagnostic to predictive and prescriptive, enabling a deeper understanding and future-oriented action.

Organizations that can adopt and drive analytics and AI can outperform their competitors and grow faster. In addition, 94% of leaders expect analytics to play an important role in enhancing decision-making, and 92% of companies plan to increase their AI investments over the next three years.^{1,2}

Data is the bedrock of both modern analytics and AI. Whereas analytics relies on data to uncover patterns, trends and insights that inform decision-making, AI learns and improves its performance by processing vast amounts of data to make predictions, classifications and drive intelligent actions.

As a result, organizations need a better, faster, more reliable way to integrate and deliver timely data to drive modern analytics and AI.

Leverage Informatica Cloud Data Integration (CDI) to seamlessly design, deploy and manage high-performance, scalable and cloud-native data pipelines with a low-code/no-code graphical user interface (GUI)-based interface across hybrid and multi-cloud environments.

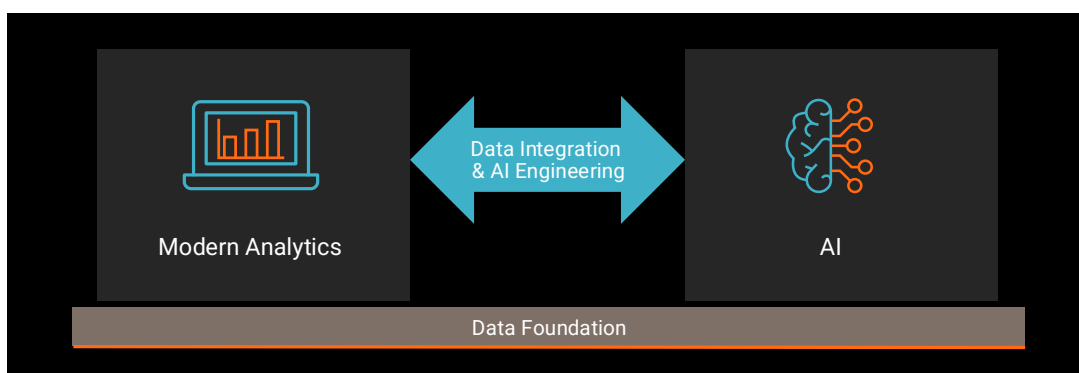


Figure 1: Data as a foundation for data integration and AI engineering to drive modern analytics and AI

¹ <https://www.gartner.com/peer-community/oneminuteinsights/omi-2024-data-analytics-priorities-challenges-insights-field-ycu>.

² <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work>.

Key Benefits and Product Features

High Productivity Development with Intuitive Design and Out-of-the-Box Integrations

Informatica CDI enhances developer productivity by offering an intuitive, drag-and-drop GUI-based design interface that simplifies the development of complex data pipelines for extract, transform and load (ETL) and extract, load and transform (ELT). The CDI low-code/no-code interface enables developers and data engineers to seamlessly build, test and deploy ETL/ELT pipelines with zero hand coding. In addition, it offers rich connectivity with over 300 out-of-the-box connectors, transformation components and pre-built integration templates, accelerating time to value by reducing manual coding and rework.

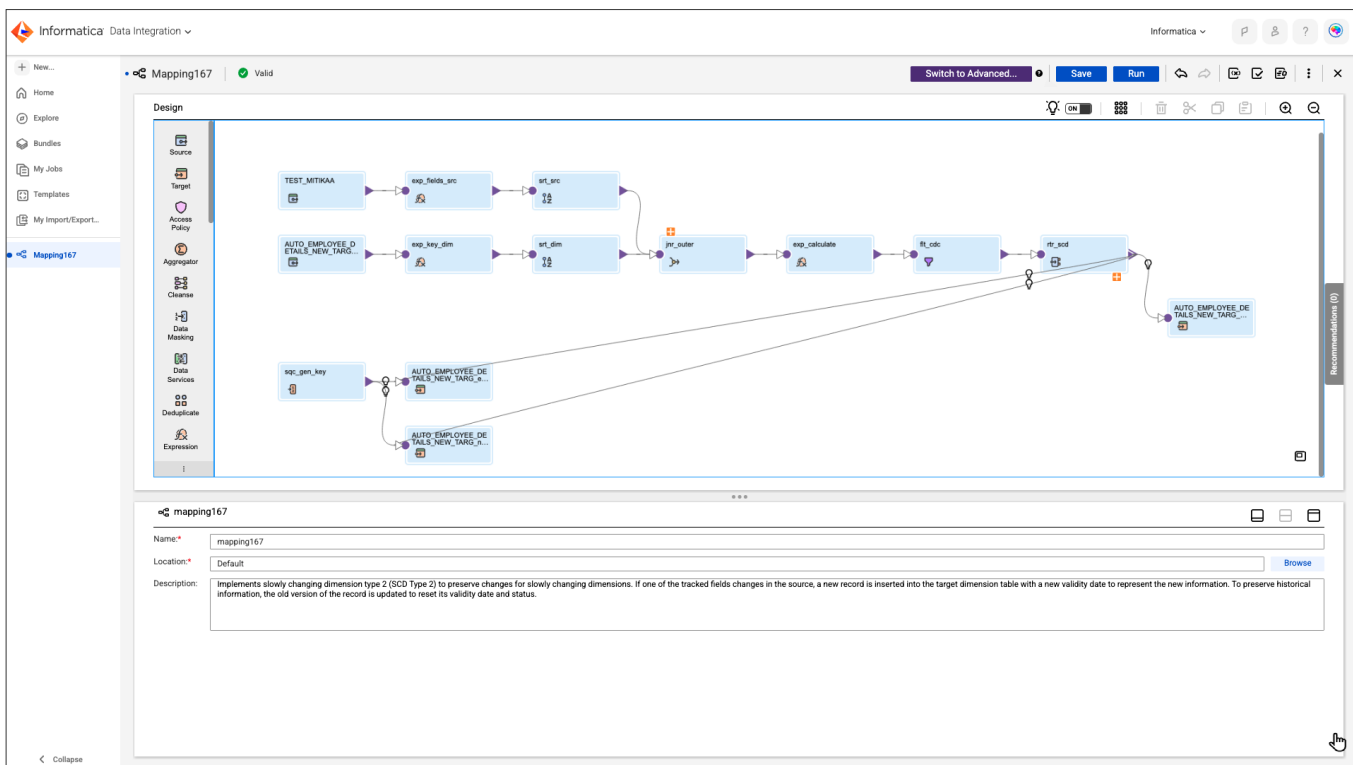


Figure 2: GUI for building ETL/ELT pipelines

Smart Workload Optimization with Flexible Deployment and AI-Powered Execution

Informatica Cloud Data Integration optimizes workloads by offering flexible deployment options – serverless, elastic and pushdown execution, allowing integration tasks to run where they are most efficient and cost effective, whether in the cloud, on-premises or directly in the cloud data warehouse/data lakes – Snowflake, Databricks, Amazon Redshift, Google BigQuery and Microsoft Fabric. The Informatica CLAIRE® AI engine intelligently analyzes metadata and workload patterns to dynamically tune performance, allocate optimal compute resources and offload processing to the most efficient execution environment. This smart workload optimization results in faster data pipeline execution and significantly reduces infrastructure and operational costs.

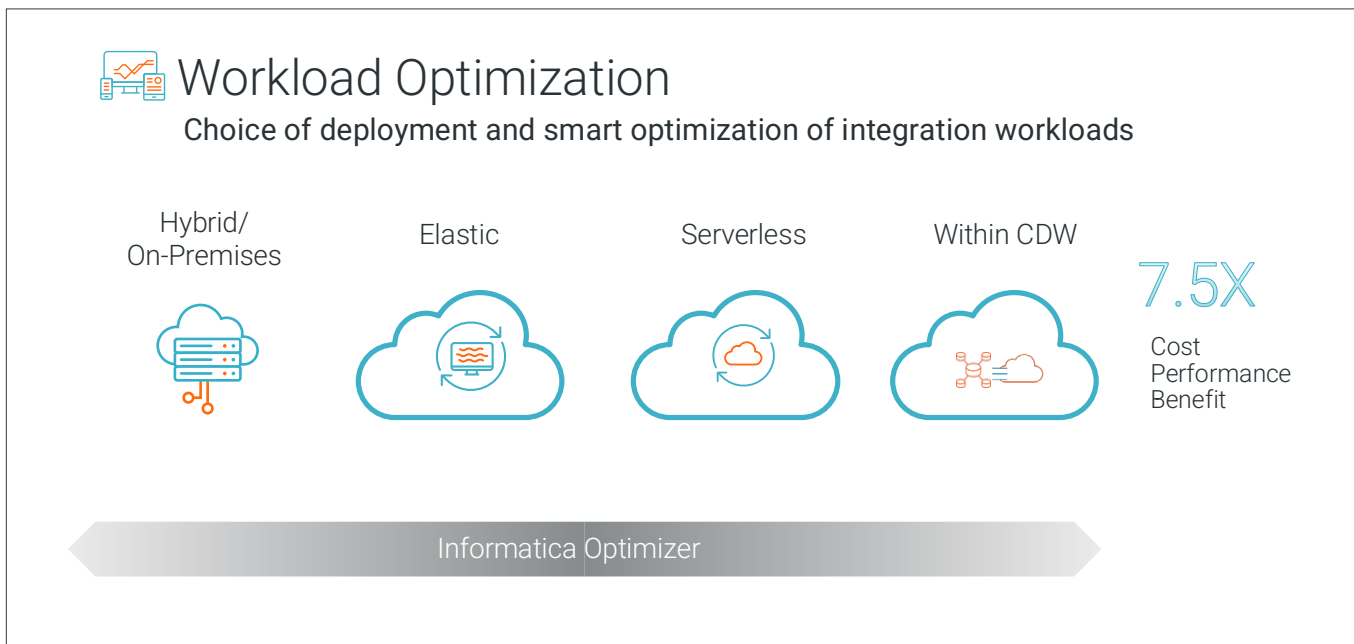


Figure 3: Workload optimization

Future-Proof Data Investments with Reusable Business Logic

Informatica Cloud Data Integration future-proofs data investments by separating business logic from execution through a metadata-driven, platform-agnostic architecture. This enables reusable mappings, templates and workflows that can seamlessly migrate across on-premises, cloud and multi-cloud environments without rewriting them. With CLAIRE-powered automation, existing logic can be intelligently adapted to new technologies, ensuring consistent performance, reduced rework and faster modernization. This “design once, deploy anywhere” approach protects development investments and ensures long-term agility as data architecture evolves.

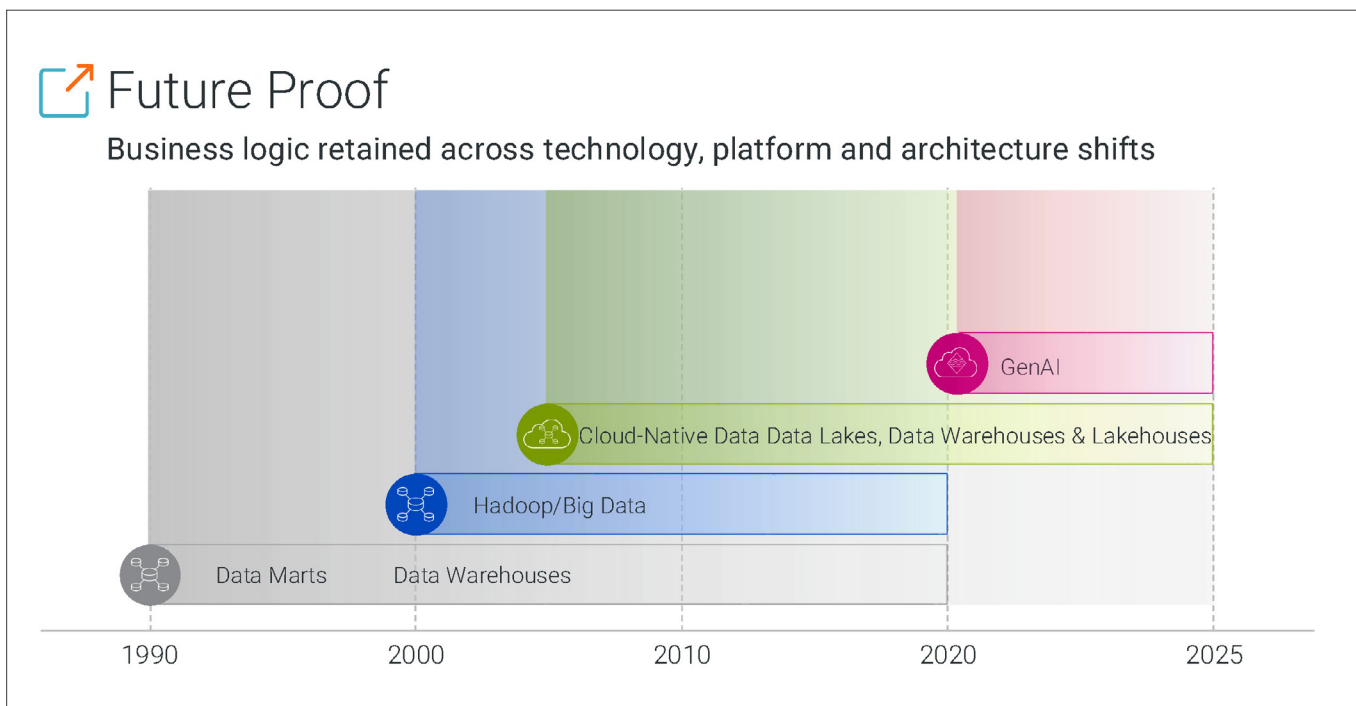


Figure 5: Future-proof your business logic

Deliver Enterprise-Grade Operationalization

Informatica Cloud Data Integration consistently delivers enterprise-grade scalability, availability, reliability, manageability and security to meet the most demanding data integration needs. It enables full data lifecycle management with integrated support of continuous integration and continuous delivery (CI/CD), DataOps, FinOps and intelligent workload management, ensuring agility, cost efficiency and operational excellence. Organizations can seamlessly operationalize multiple pipelines across diverse use cases while maintaining reliable performance at any scale. With a global reach across all major cloud data centers and 99.9% service-level agreement (SLA) guarantees, it empowers modern data enterprises to confidently handle any data integration challenge with resilience, governance and speed.

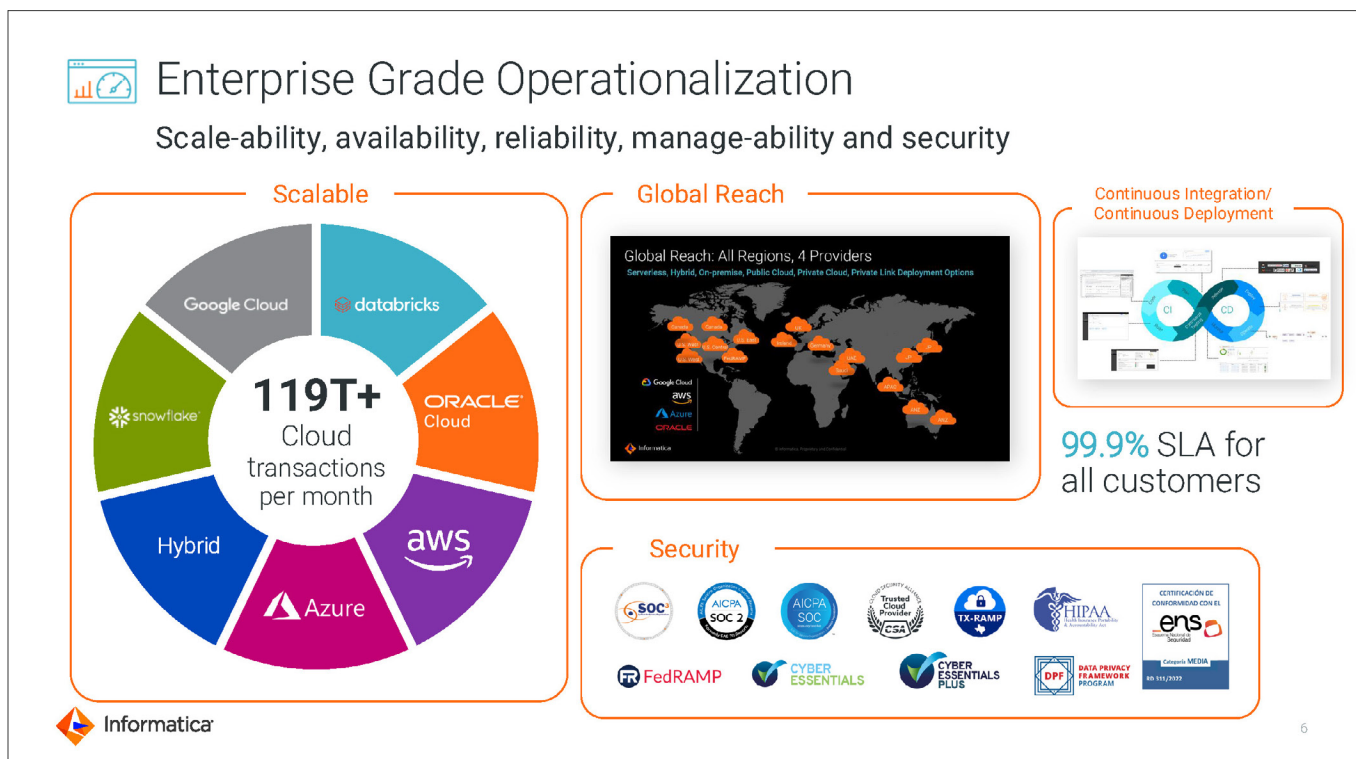


Figure 6: Enterprise-grade operationalization

What's New in Cloud Data Integration

CLAIRE Copilot for Data Integration

CLAIRE Copilot for data integration is the Informatica generative AI (GenAI)-powered assistant that helps users create data ingestion and integration pipelines using natural language prompts. Embedded within Informatica CDI, it enables users to build complex mappings by simply describing their intent in plain English. It can also summarize existing pipelines, recommend best practices and auto-generate documentation.

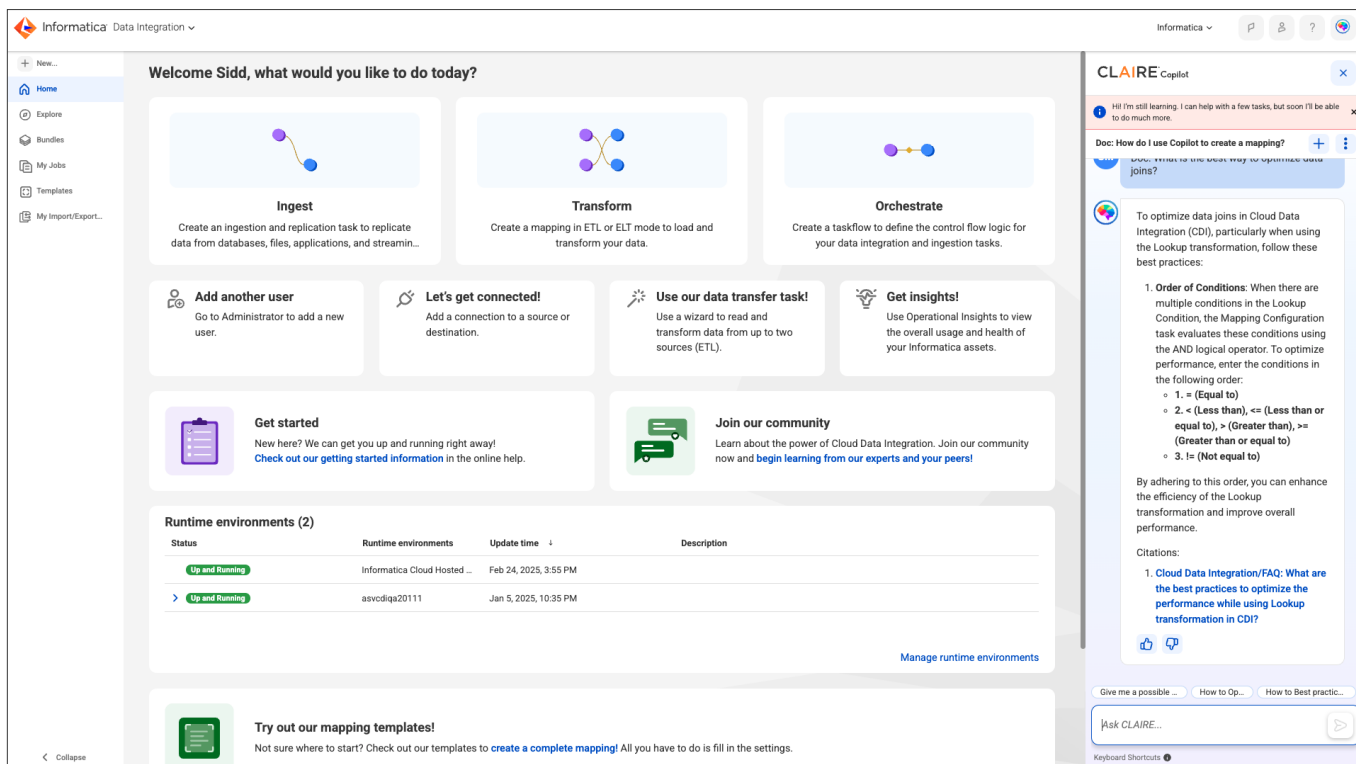


Figure 7: CLAIRE Copilot for data integration

Unified Data Integration

Unified data integration delivers a goal-driven, streamlined experience for data ingestion, replication and transformation across all major cloud platforms. It offers a single pane of glass to manage all integration tasks, with intuitive wizards and inline connection creation that simplify and accelerate end-to-end data pipeline development.

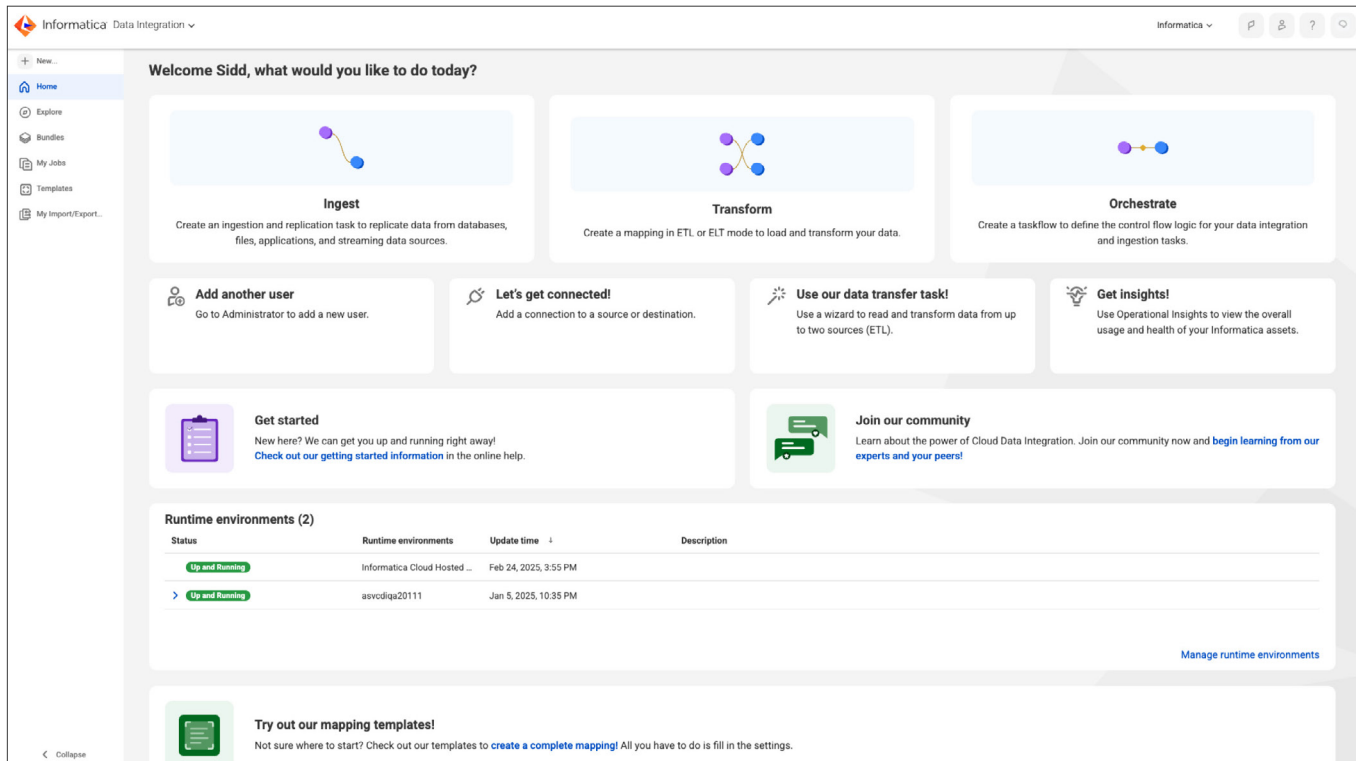


Figure 8: Unified data integration

Support for Open Table Formats

Informatica supports open table formats, such as Apache Iceberg and Delta Lake, by enabling data integration and transformation on them. Informatica Cloud Data Integration allows users to read from and write to open table formats natively, supporting schema evolution, time travel, partition pruning and atomicity, consistency, isolation and durability (ACID) compliance for lakehouses.

Connections

1 Select Connector 2 Connection Form

Configure Connection : Open Table
Follow the setup guide on the right to configure the connection : Use Secret Vault

Connection Name: * OpenTableDemo

Description:

Runtime Environment: * Select

Open Table Format: Apache Iceberg

Catalog Type: REST Catalog

Rest Catalog Type: Polaris Catalog

Catalog Endpoint: *

Catalog Authentication Type: None

Credential Vending: Enabled

Storage Type: Amazon S3

Connect to Open Table

Let's configure the Open Table connection properties to connect to AWS Glue Catalog or Hive Metastore.

Before you begin

Before you get started, you will need to add the Amazon Athena or Hive JDBC driver to the Secure Agent machine and configure the authentication-specific prerequisites.

Permanent IAM Credentials authentication requires the access key and secret key values of the IAM user. Keep the access key and secret key handy before creating the connection. For more information about creating an access key and secret key, see the AWS documentation.

To configure Service Principal authentication, you need the Azure account name, client secret, client ID, and tenant ID for your application registered in the Azure Active Directory. Keep the Azure account name, client secret, client ID, and tenant ID handy before creating the connection. For more information about Azure account name, client secret, client ID, and tenant ID for your application, see the Microsoft Azure documentation.

Check out [Prerequisites](#) to learn more about how to configure policies and role to access Apache Iceberg or Delta Lake tables.

Open Table formats with associated catalog and storage types

You can choose the Open Table format that you want to use and its associated catalog type and storage type to interact with data.

The following table summarizes the Open Table formats that you can use, their catalog types, storage types, and the authentication options available for each storage type:

Open Table format	Catalog type	Storage type	Authentication type
Apache Iceberg	AWS Glue Catalog	Amazon S3	Permanent IAM Credentials authentication
	Hive Metastore	Amazon S3	Permanent IAM Credentials authentication

Figure 9: Open table formats

Unstructured Data Processing

Informatica processes unstructured data like PDFs, images and emails with AI-powered intelligent parsing, classification and transformation, turning raw data into actionable insights. Powered by CLAIRE Intelligent Structure Discovery (ISD), it automatically identifies patterns, extracts elements, infers relationships and detects field types – turning unstructured content into structured, usable data.

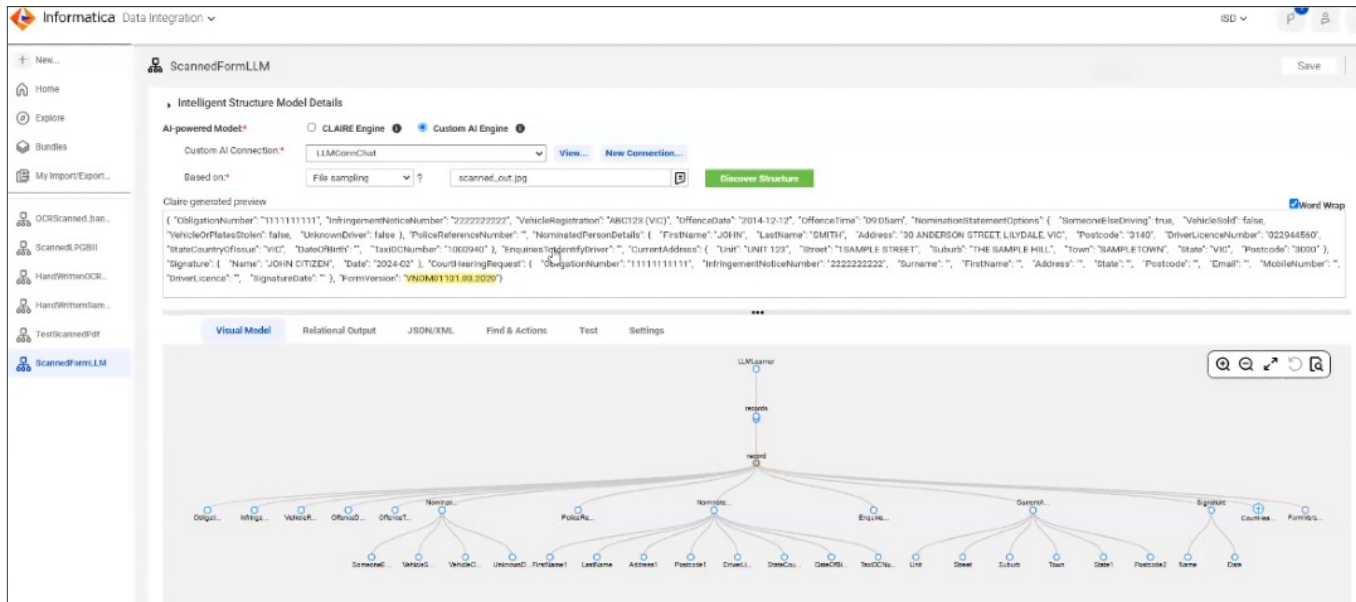


Figure 10: Unstructured data processing

Simplified Multi-Pattern ELT for AI

Informatica processes high volumes of data by leveraging the native compute engines of modern cloud platforms such as Databricks Delta Lake, Google BigQuery, Snowflake, Amazon Redshift, Microsoft Fabric and Oracle. This ecosystem pushdown optimization ensures scalable, high-performance transformations within the cloud data warehouse, reducing ingress and egress cost. In addition, Informatica enables seamless integration with ecosystem large language model (LLM) functions. This allows users to invoke services like Snowflake Cortex or Databricks Foundations Models directly within data integration pipelines to enrich, summarize or analyze data using GenAI.



Figure 11: Multi-pattern ELT for AI

Accelerate Modern Analytics and AI with a Cloud-Native
Cloud Data Integration and Engineering Solution

Where data & AI come to



Worldwide Headquarters

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

Informatica (NYSE: INFA), a leader in AI-powered enterprise cloud data management, helps businesses unlock the full value of their data and AI. As data grows in complexity and volume, Informatica's Intelligent Data Management Cloud™ delivers a complete, end-to-end platform with a suite of industry-leading, integrated solutions to connect, manage and unify data across any cloud, hybrid or multi-cloud environment. Powered by CLAIRE® AI, Informatica's platform integrates natively with all major cloud providers, data warehouses and analytics tools – giving organizations the freedom of choice, avoiding vendor lock-in and delivering better ROI by enabling access to governed data, simplifying operations and scaling with confidence.

Trusted by about 5,000 customers in nearly 100 countries – including over 80 of the Fortune 100 – Informatica is the backbone of platform-agnostic, cloud data-driven transformation.

Informatica. Where data and AI come to life.™

IN17-5231-0825

© Copyright Informatica LLC 2025. Informatica and the Informatica logo are trademarks or registered trademarks of Informatica LLC in the United States and other countries. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners. The information in this documentation is subject to change without notice and provided "AS IS" without warranty of any kind, express or implied.