



eBook

# Simplify Data Integration for Cloud Data Warehouses with ELT

Unlock efficiency and scalability with a seamless data integration solution

Where data  
& AI come to **LIFE**



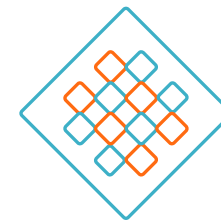
# Contents

|                                                                            |    |
|----------------------------------------------------------------------------|----|
| Introduction                                                               | 3  |
| Chapter 1: Evolution of ELT                                                | 4  |
| Chapter 2: Informatica Ecosystem SQL ELT: A Low-Code, No-Code ELT Approach | 5  |
| Chapter 3: The Top Five Applications of Ecosystem SQL ELT                  | 7  |
| Chapter 4: Why Choose Informatica?                                         | 8  |
| Empowering Businesses with Data                                            | 14 |
| About Informatica                                                          | 15 |

# Introduction

More than 400 million terabytes of data are created each day. Considering this, the challenge of managing vast amounts of data has never been more pressing, especially with the rise of AI and generative AI (GenAI). The evolution of **data management** systems is critical to ensure that data-hungry applications like large language models (LLMs) can access the right data at the necessary scale. With the global AI data management market projected to grow from \$25.1 billion in 2023 to \$70.2 billion by 2028,<sup>1</sup> businesses face the critical task of handling and processing high volumes of data from numerous sources to fuel decision-making.

In response to these challenges, data warehouses have evolved from costly and complex on-premises systems to more flexible, cost-effective and cloud-based solutions. Today, AWS, GCP, Azure, Databricks and Snowflake offer storage and computing separately, allowing elastic scaling and efficient resource utilization. The advent of the cloud also led to **cloud-native data integration solutions** by Informatica. Its modern cloud-native architecture made complex data readable, accessible and actionable for both technical and business users in the most efficient, easy-to-use and intelligent way.



According to the latest estimates, 402.74 million terabytes of data are created each day<sup>2</sup>

<sup>1</sup> <https://www.marketsandmarkets.com/Market-Reports/ai-data-management-market-69639242.html>

<sup>2</sup> <https://explodingtopics.com/blog/data-generated-per-day>

## Chapter 1

# Evolution of ELT

### ELT and ETL: Understanding the Differences

“Extract, transform, load” (ETL) and “extract, load, transform” (ELT) are two distinct **data integration methodologies** that serve different purposes based on organizational needs.

ETL transforms data before loading it into a target system, which is ideal for structured data and environments where **data quality** must be ensured before storage. Conversely, ELT loads raw data into a target system before applying transformations, leveraging the processing power of modern data warehouses. ETL is great for creating environment-agnostic data pipelines, especially when SQL alone isn't enough for data transformation. However, when the source and target are within the same system and SQL can handle the task efficiently, ELT can be a more suitable option. It's advisable to keep the pipelines flexible and avoid locking them into a single environment. A low-code, no-code tool that enables you to design system-agnostic pipelines while also leveraging the strengths of the ELT pattern through endpoint commands offers the best of both worlds. ETL is great for structured data with strict schema adherence, while ELT excels in agile environments with diverse data types and a need for rapid data availability.

Ultimately, neither ETL nor ELT is inherently better; the choice between them depends on the specific use case, data characteristics and organizational requirements.

### Move Towards SQL-Based Transformations for ELT

The evolution toward SQL-based transformations represents a natural progression in data integration methodologies. The landscape evolved further when Informatica introduced ecosystem-based SQL **ELT**, which fully embraced native ecosystem functions, restricting transformations to those supported by the ecosystem and providing guaranteed pushdown optimization. This evolution reflected the growing recognition that modern data warehouses possess processing capabilities that should be harnessed rather than bypassed.

### SQL ELT in the Cloud Data Warehousing Ecosystem

The cloud computing paradigm has significantly influenced the adoption and implementation of SQL ELT approaches. Cloud data warehouses like Snowflake, Google BigQuery and Amazon Redshift provide the computational resources, scalability and SQL processing capabilities that make ELT approaches increasingly feasible. The elasticity of cloud resources aligns perfectly with the variable computational demands of ELT processes, where transformation workloads may fluctuate based on data volumes and complexity. The SQL ELT ecosystem represents a cohesive approach to data processing that offers numerous benefits to enterprises.

## Chapter 2

# Informatica Ecosystem SQL ELT: A Low-Code, No-Code ELT Approach

**As data landscapes become increasingly complex, the need for agile, scalable and cost-effective data integration solutions has never been greater. Informatica's ecosystem SQL ELT approach offers a modern alternative, leveraging cloud-native technologies and a low-code, no-code interface to streamline data workflows.**

Informatica's Ecosystem SQL ELT solution represents a cohesive strategy to data processing that offers numerous architectural advantages. At its core, Informatica's SQL ELT simplifies this data processing architecture by automatically converting data pipeline flows into SQL ELT operations that execute within the target environment. This architectural streamlining eliminates the need for separate transformation engines or intermediate staging areas, reducing both complexity and points of potential failure. The simplified architecture creates a more direct path from data sources to analytical insights, with fewer components to manage and maintain. This consolidation of the data processing stack allows organizations to focus more resources on extracting value from their data rather than managing the infrastructure that processes it.

### Cloud-Native Scalability

Modern leading data warehouses offer native functions and optimization techniques specifically designed for efficient data processing at scale. SQL ELT leverages these capabilities to deliver fast and scalable data transformations. The approach further harnesses the elastic capabilities of cloud platforms to dynamically scale resources based on workload demands.

### Fast Data Processing

Ecosystem ELT reduces data processing latency by performing transformations directly within the cloud data warehouse or lake. By leveraging powerful cloud-native engines such as Snowflake, Databricks or Google BigQuery, it accelerates data preparation for analysis, empowering teams to derive actionable insights quickly.

### Cost Efficiency

By transforming data in place, ELT reduces the need for intermediary storage or compute resources. This approach takes full advantage of the cost-efficient storage and processing options provided by cloud platforms, ensuring organizations only pay for the resources they use. Importantly, this method also minimizes egress charges. Modern cloud data warehouses often levy charges for the external movement of data. By limiting data movement, businesses can achieve substantial cost saving.

## Chapter 2

# Informatica Ecosystem SQL ELT: A Low-Code, No-Code ELT Approach (continued)

### Support for Structured and Unstructured Data

Today's organizations deal with a variety of data formats, from structured database tables to unstructured text, images and logs. This can make it difficult for businesses that want to achieve a holistic overview of their data. Informatica's ELT approach integrates seamlessly with both structured and unstructured data, enabling businesses to unify diverse data types and make them available for downstream processes like AI and GenAI training, which require comprehensive datasets for optimal performance.

### Low-Code and Pro-Code Support

Informatica's ELT solution is designed to cater to both technical and non-technical users. The low-code, no-code interface allows data engineers, analysts and other users to design and deploy data pipelines with minimal effort, while advanced pro-code capabilities ensure developers can fine-tune or customize workflows to meet unique requirements. This dual approach fosters collaboration across teams and accelerates project timelines.

In addition to these capabilities, Informatica's solution also leverages its ecosystem-specific functions, empowering users to utilize pre-built integrations and features provided by ecosystem vendors directly within the data integration pipeline. This not only enhances functionality but also streamlines processes by eliminating the need for additional third-party tools or custom coding, thereby adding significant value to the overall data management strategy.

By combining these capabilities, Informatica's Ecosystem SQL ELT not only simplifies the data integration process but also empowers organizations to unlock the full potential of their cloud data warehouses and lakes, ensuring they stay agile and competitive in today's data-driven landscape.



## Chapter 3

# The Top Five Applications of Ecosystem SQL ELT

The versatility of Ecosystem SQL ELT enables organizations to unlock the full potential of their data for a variety of modern use cases, particularly those involving AI and GenAI. Here we present a deeper look at the top five scenarios where this approach excels:



### Data science and analytics

Ecosystem SQL ELT enables efficient management and analysis of large datasets within the ecosystem and external data lakes. This facilitates the extraction of valuable insights for business decision-making. By processing data within the ecosystem organizations can achieve fast time-to-insight for analytics projects, optimizing data engineering resources.



### Data integration within the same ecosystem

Ecosystem SQL ELT simplifies data integration by extracting data from various sources within the same ecosystem and loading it into a central data warehouse for unified analysis. It supports the creation of a unified data repository for business intelligence, reporting and data analysis, ensuring efficient data storage and fast query performance.



### Real-time analytics

ELT is ideal for big data environments where large volumes of unstructured data are processed. It allows for real-time analytics by making data available immediately for transformation as needed. SQL ELT supports scalable cloud storage, enabling organizations to handle vast amounts of data efficiently and flexibly.



### AI and ML model training

Ecosystem SQL ELT simplifies the process of preparing data for AI and ML models by enabling transformations directly within cloud platforms. Large language model (LLM) functions available within the ecosystem can be leveraged to drive ML model training. It can incorporate native functions directly within the ecosystem, enabling AI use cases without moving data out of the ecosystem.



### Business applications integration

Ecosystem SQL ELT facilitates the extraction of data from various business applications (CRM, ERP, HR systems, etc.) and loads it into a centralized cloud platform for analysis. It efficiently loads data from different source systems into data warehouses for reporting and analysis within the same data ecosystem.

## Chapter 4

# Why Choose Informatica?

**Informatica stands out as a leader in data integration and management, offering a comprehensive platform that addresses the needs of modern, data-driven organizations. Several factors make Informatica the ideal choice for businesses looking to simplify and optimize their ELT processes.**

### Flexibility

Maintain a consistent pipeline definition process whether you're employing ETL with a local engine, utilizing ELT with endpoint commands or executing your workloads with Spark. This flexibility ensures that no matter where your data resides — be it a data warehouse, a data lake, a lake house or even in open table formats — you can leverage the same mapping paradigm to define your data pipelines seamlessly. This capability streamlines data management and integration, reducing the complexity often associated with adapting to varying data environments and processing frameworks. By eliminating the need to redefine processes for different architectures or execution engines, Informatica enhances efficiency and scalability, allowing organizations to focus more on deriving insights and driving business value from their data.



## Chapter 4

# Why Choose Informatica? (continued)

### Ease of Use

Informatica's intuitive low-code, no-code interface empowers users across all skill levels. Data engineers, analysts and even non-technical business users can quickly design, deploy and manage ELT pipelines without the need for extensive coding knowledge (see Figures 1a, 1b and 2). This democratizes data integration, enabling faster adoption and collaboration across teams.

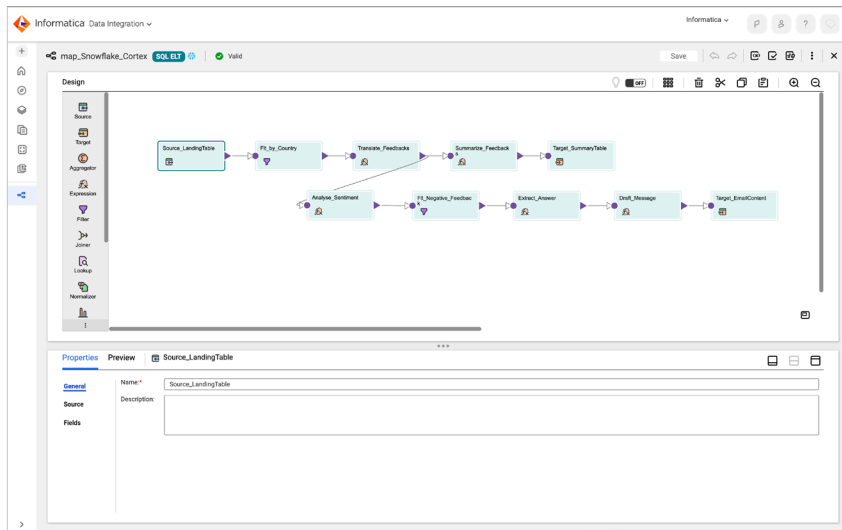


Figure 1a: Generate SQL query

The screenshot shows a 'SQL Query' preview window. The window has a title bar with 'SQL Query' and standard window controls. The main area displays the following SQL query:

```
1  /** SQL Query 1 **/  
2  INSERT INTO "SALES"."CORTEX"."PRODUCTS_REVIEW"  
3    ("PRODUCT_NAME", "REVIEWER_NAME", "REVIEW_TITLE", "REVIEW_TEXT", "COUNTRY")  
4    SELECT PRO2.PRODUCT_NAME,  
5           PRO2.REVIEWER_NAME,  
6           PRO2.REVIEW_TITLE,  
7           PRO2.REVIEW_TEXT,  
8           PRO2.COUNTRY  
9  FROM (  
10     SELECT PRO0."PRODUCT_NAME",  
11            PRO0."REVIEWER_NAME",  
12            PRO0."REVIEW_TITLE",  
13            PRO0."REVIEW_TEXT",  
14            PRO0."COUNTRY",  
15            SNOWFLAKE.CORTEX.TRANSLATE (PRO0."REVIEW_TEXT", 'ja',  
16            'en') ::VARCHAR (65000),
```

The window also includes a 'Close' button in the bottom right corner and a question mark icon in the bottom left corner.

Figure 1b: Preview SQL query during design time

## Chapter 4

# Why Choose Informatica? (continued)

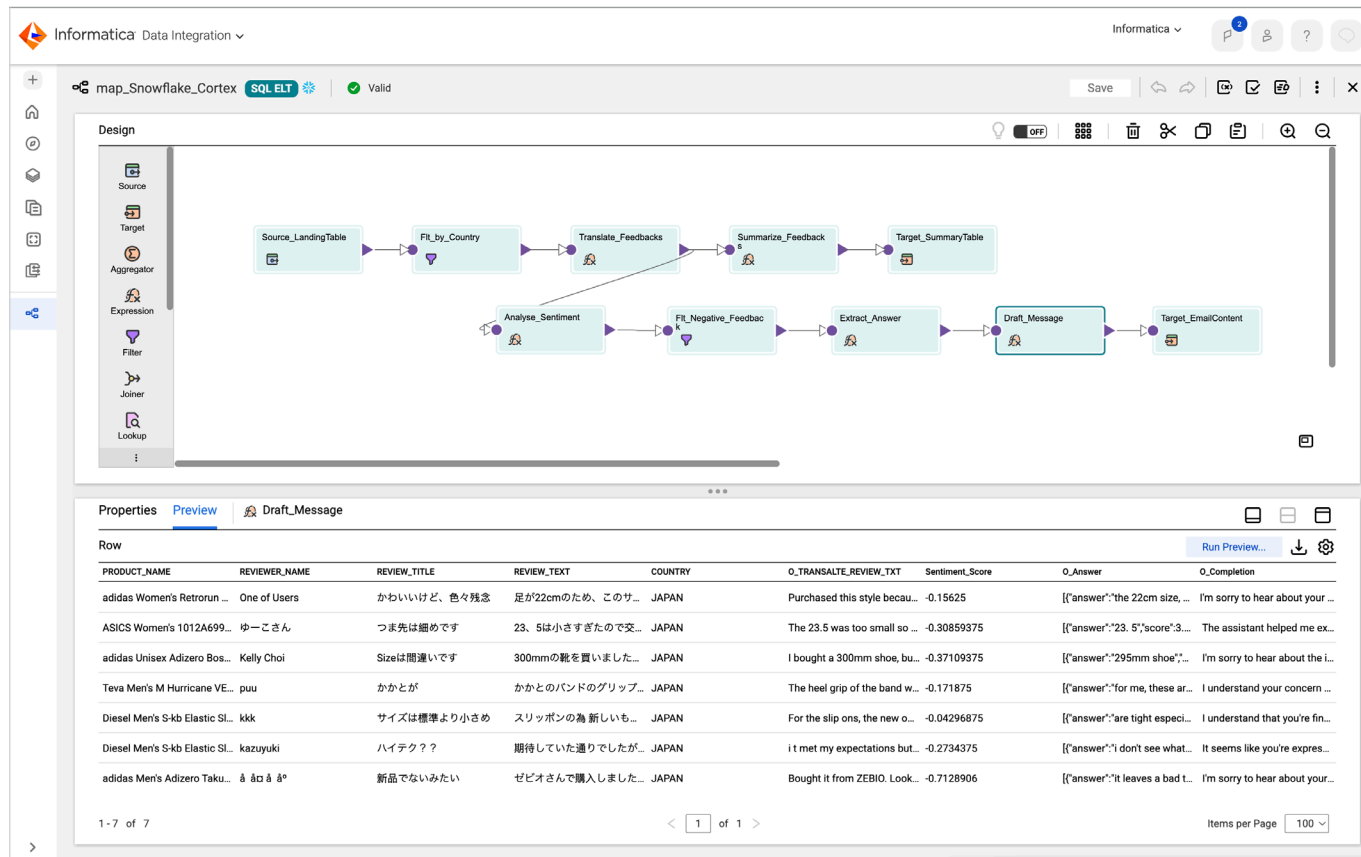


Figure 2: Leverage midstream data preview to troubleshoot and debug data transformation

## Chapter 4

# Why Choose Informatica? (continued)

### Integration with Leading Technologies

Informatica seamlessly integrates with a wide range of cloud platforms, data warehouses and analytical tools, including Snowflake, AWS, Microsoft Fabric, Google Cloud and Databricks (see Figures 3a, 3b, 3c and 4). This interoperability allows businesses to harness the strengths of multiple technologies while maintaining a single, unified data integration workflow. Organizations benefit from a best-of-breed approach without the inefficiencies of switching between tools or managing siloed systems.

#### Snowflake

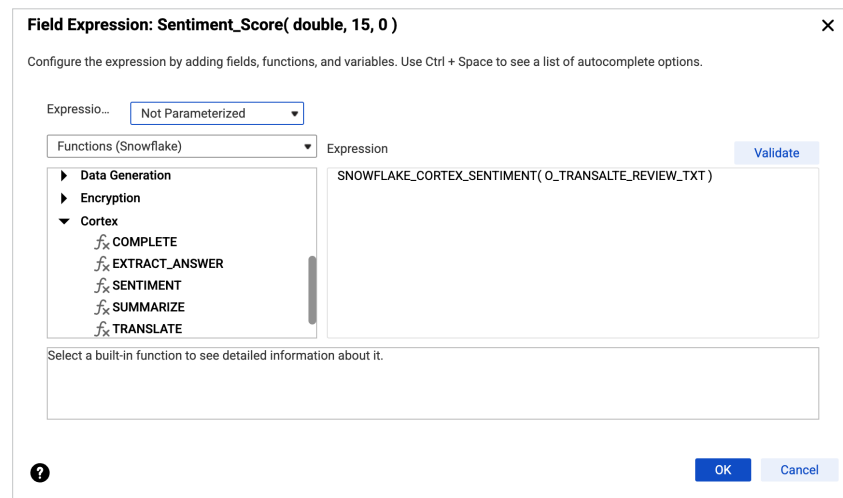


Figure 3a: Ecosystem-native functions

#### Databricks

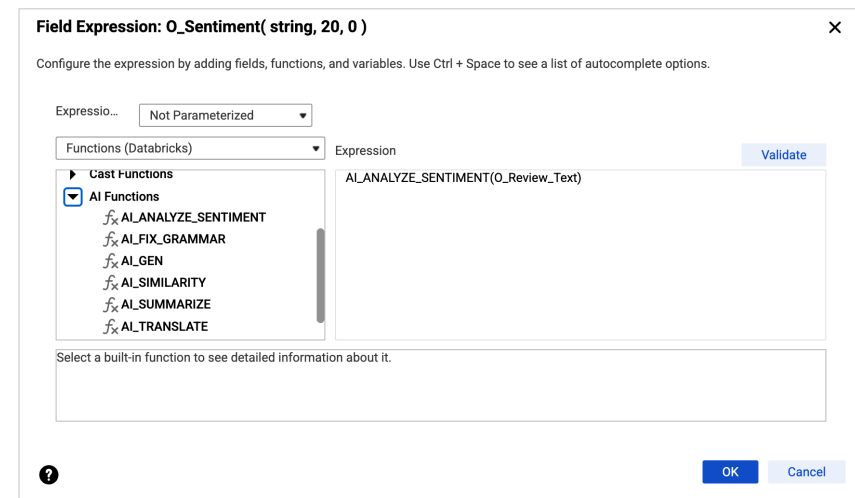


Figure 3b: Ecosystem-native functions

## Chapter 4

# Why Choose Informatica? (continued)

### Amazon Redshift

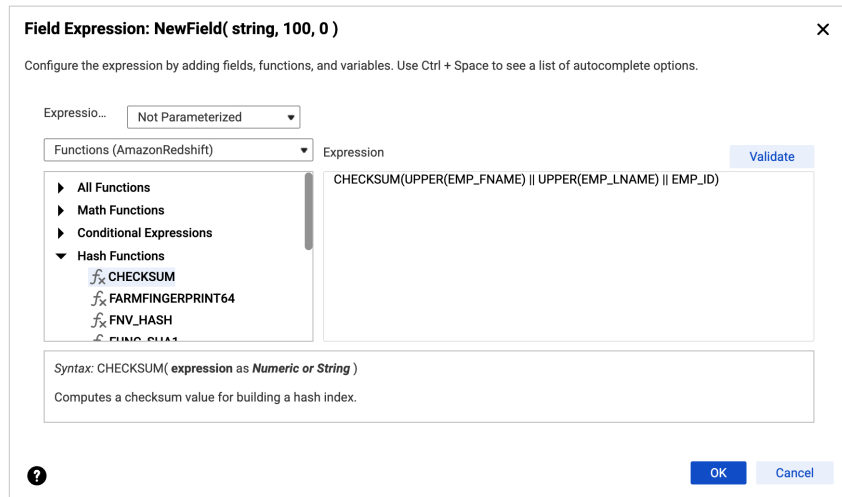


Figure 3c: Ecosystem-native functions

### Scalability

Whether you're managing a handful of data sources or integrating data from hundreds of systems, Informatica is built to scale effortlessly. Its robust architecture supports both cloud data warehouses and cloud data lakes, ensuring organizations can handle growing data volumes and increasing complexity as their needs evolve. Informatica's cloud-native design also means it can dynamically scale resources, providing high performance without overburdening budgets.



## Chapter 4

# Why Choose Informatica? (continued)

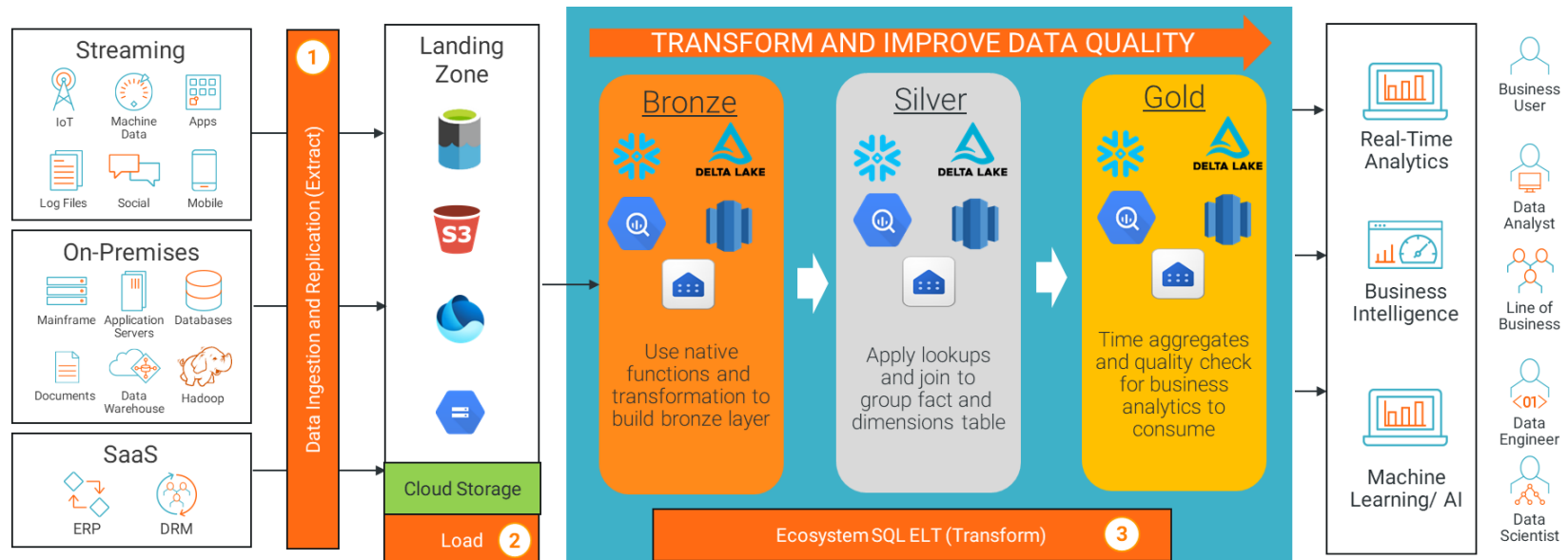


Figure 4: Ecosystem SQL ELT architecture

## AI and Automation

Informatica incorporates advanced AI capabilities powered by its **CLAIRE® engine** to enhance and streamline data workflows. Automation features reduce manual effort, such as identifying data anomalies, suggesting optimizations and enabling intelligent mapping. These AI-driven insights not only accelerate pipeline development but also improve data quality and operational efficiency.

By choosing Informatica, businesses gain a future-ready data integration solution that combines ease of use, scalability and cutting-edge technology. Whether you're driving AI initiatives, modernizing your data architecture or simply improving operational efficiency, Informatica offers the tools and flexibility to meet your goals.

# Empowering Businesses with Data

**In today's data-driven world, the ability to efficiently manage and leverage data is a key differentiator for businesses striving to stay ahead. Informatica simplifies this journey by offering a powerful yet user-friendly platform that removes the complexities of data integration and transformation.**

Whether dealing with structured or unstructured data, from a single source or multiple disparate systems, Informatica enables businesses to seamlessly integrate, transform and analyze data in cloud data warehouses and lakes. By adopting Informatica's low-code, no-code ELT approach organizations reduce their reliance on technical expertise, democratizing data processes and empowering teams across all levels.

Beyond streamlining ELT workflows, Informatica enhances innovation by integrating with leading technologies and enabling the use of generative AI (GenAI). These capabilities help businesses extract deeper insights, make faster decisions and unlock new opportunities, all while maintaining scalability and cost-efficiency.

With Informatica, businesses can focus on what truly matters: turning data into actionable insights, driving growth and staying competitive in an increasingly dynamic market.

**Want to learn more?**



Refer to our  
Ecosystem SQL  
ELT user guide.

**READ GUIDE**

---

**Get started with SQL ELT. Sign up today!**

---



# 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.™**

IN19-5177-1225

© 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.

[informatica.com](https://www.informatica.com)

## Where data & AI come to



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](https://www.informatica.com)  
[linkedin.com/company/informatica](https://www.linkedin.com/company/informatica)  
[x.com/Informatica](https://www.x.com/Informatica)

**CONTACT US**