



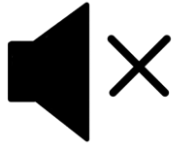
07 Aug, 2025

# Simplify Data Integration with a CLAIRE Co-pilot

- Srinivasa Gopal, Sr Principal Customer Success Architect, CSA
- Krishnendu Chaklader, Principal Solutions Architect, CSA

Where data & AI come to **LIFE**

# Housekeeping Tips



- Today's Webinar is scheduled for **1 hour**
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the **Q&A option** and we will respond at the end of the presentation
- The webinar is **being recorded** and will be available on our [Success Portal](#) - where you can download the **slide deck** for the presentation. The link to the recording will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

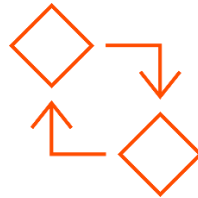
# Feature Rich Success Portal



Bootstrap trial and  
POC Customers



Enriched Customer  
Onboarding  
experience



Product Learning  
Paths and Weekly  
Expert Sessions

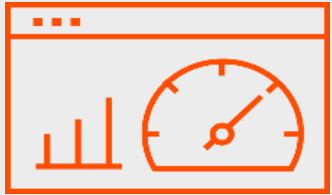


Informatica  
Concierge



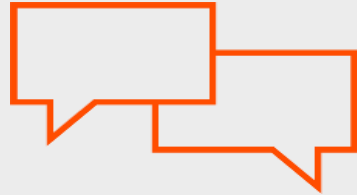
Tailored training and  
content  
recommendations

# More Information



## Success Portal

<https://success.informatica.com>



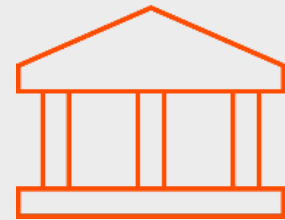
## Communities & Support

<https://network.informatica.com>



## Documentation

<https://docs.informatica.com>



## University

<https://www.informatica.com/in/services-and-training/informatica-university.html>

# Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

# Agenda

1 Market Drivers

2 GEN AI for Data Integration

3 Introduction to CLAIRE CoPilot

4 UseCases and Best Practices

5 Demo

6 Q&A

# Market Drivers

Unprecedented growth in data diversity & volume, emergence of fusion data teams, and need to accelerate data operationalization



Gen AI – market disruption

---

AI-enhanced workflows incorporated in DI tools will **reduce manual intervention by 60%** by 2027



Slow GenAI value realization

---

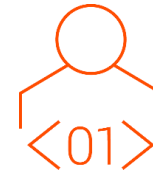
**Only 1% of C-Suite<sup>[1]</sup>** respondents describe their Gen AI rollouts as Mature



Chronic shortage of data engineers

---

**35% annual growth<sup>[3]</sup>** in open data engineering positions



Data engineering democratization

---

**500 million** business data users. Need for self service data pipeline generation is needed



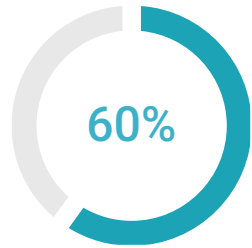
Assisted dev and vibe coding

---

**30-40% reduction<sup>[6]</sup>** in development time through code gen and optimization

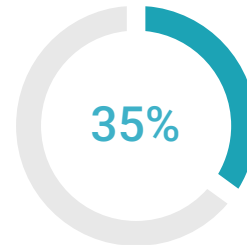
# Businesses Struggling with Value Realization, Complex Architecture, Lack of Resources, Unrealized Costs

## Complexity



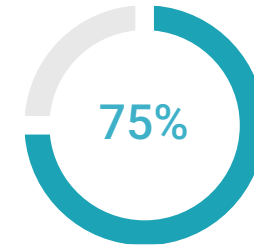
of organizations lack a unified data strategy, leading to fragmented data architectures and hindering effective data management

## Resourcing challenges



year-over-year growth in the number of open data engineering positions

## Cost Overruns

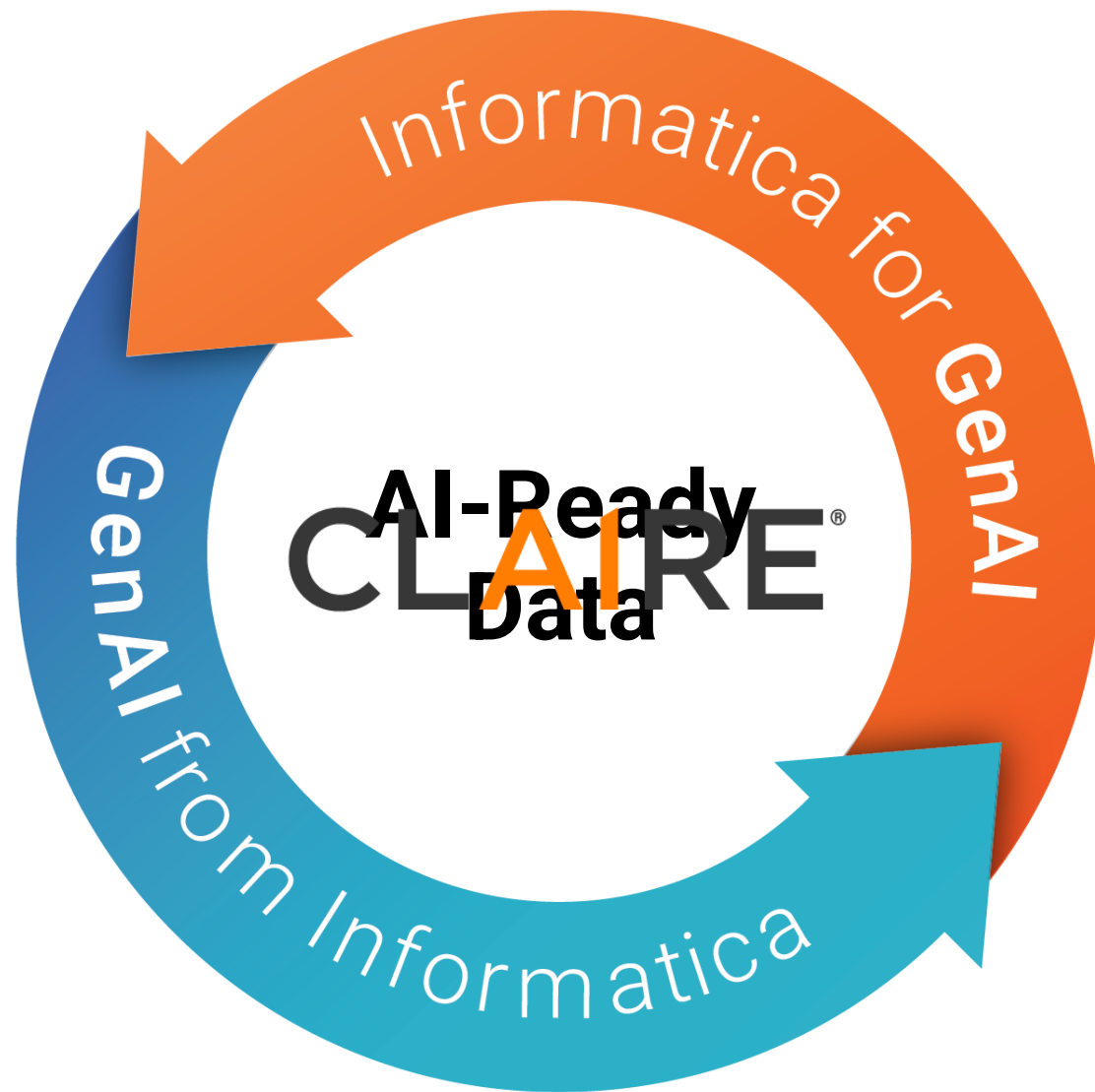


of enterprises piloting Gen AI projects are struggling to demonstrate clear ROI beyond initial experimentation

- 75% of point products don't integrate and interoperate
- Data Engineers and IT service groups spend hundreds of hours creating repeatable code
- ~50% of organizations challenged by **data quality**

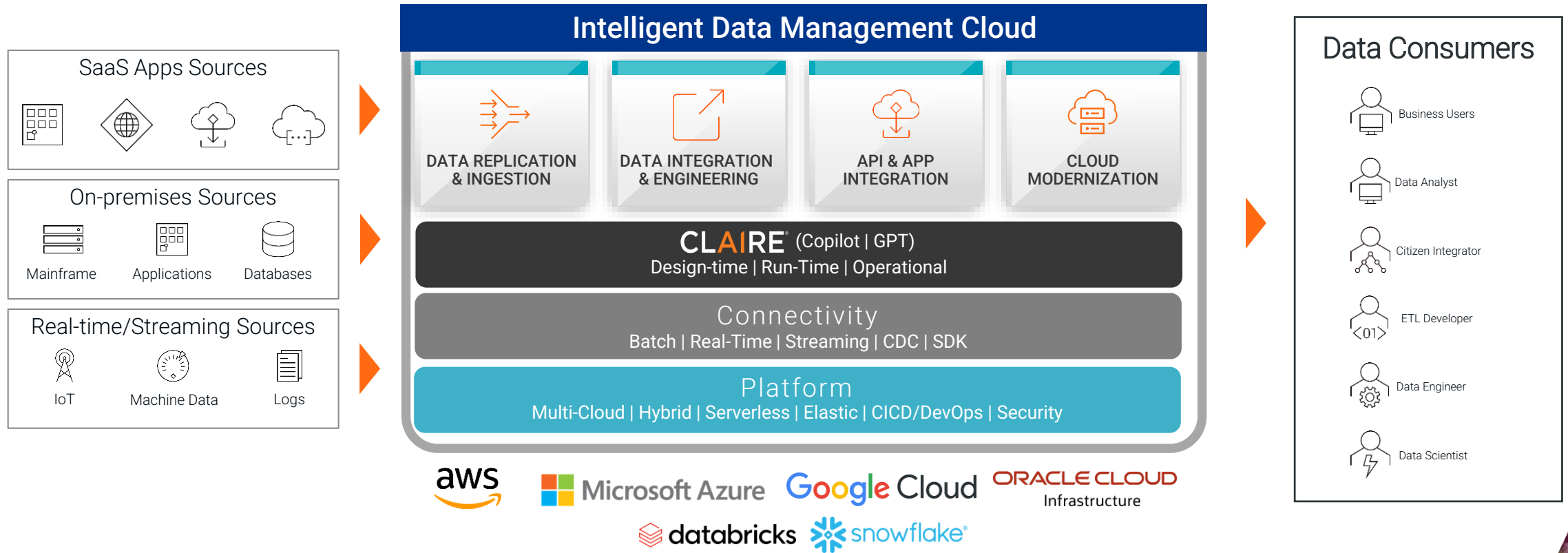
- Difficulty finding specialized skills fast
- Skilled engineers spend up to 70% of their time in data prep and pipeline maintenance, leaving only 30% for analysis and innovation
- Lack of self-service access for non-tech users delaying rapid innovation
- Lack of automation impacting the ability to scale

- Difficulty predicting compute costs
- Lack of visibility and control of users and usage
- 70% of enterprises piloting Gen AI projects are struggling to demonstrate clear ROI beyond initial experimentation



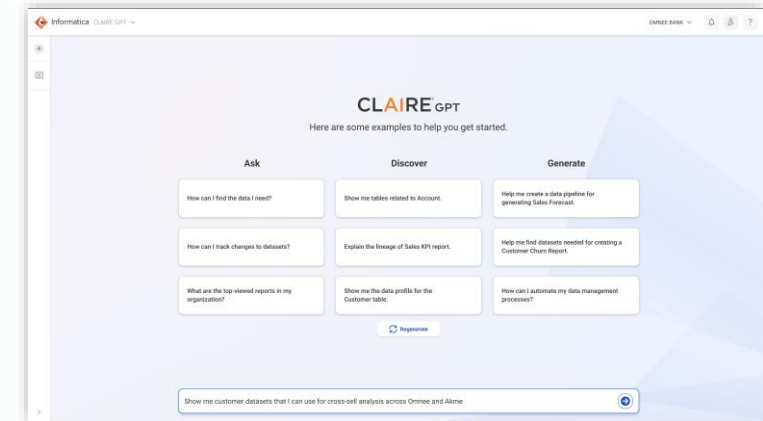
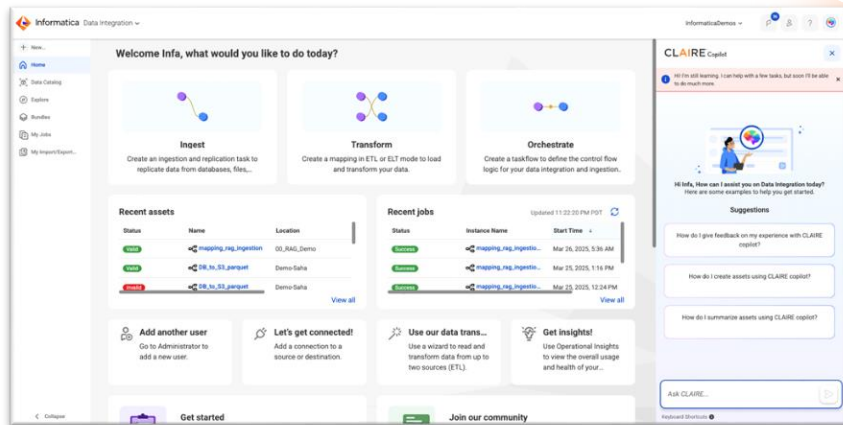
# Data Integration and Engineering

## Supporting Any Integration and Any Data Management Pattern



# CLAIRE<sup>®</sup> Copilot: GenAI for Data Integration

Expands data Integration and engineering – use cases, users and data



- NLP driven interface specialized in performing data integration centric tasks
- Primary focus is to empower integration users to leverage the rich metadata driven gen AI capabilities to simplify integration tasks

- NLP driven interface with capability across all IDMC services, including Data Integration
- Primary focus is to provide a IDMC wide, metadata based, gen AI powered data management capabilities

# Ask Copilot

## What it can do today

### Feature

- Ask for help on any products or services offered by Informatica
- Find out responses to your questions by leveraging info available in the INFA documentation portal, KB articles, and How-to-articles

### Benefits

- Increased literacy of IDMC products
- Usage, gaps, issues and best practices

### Who should use this?

- New Data Engineers, Data Analysts, Architects



**Informatica Data Integration**

Welcome Infa, what would you like to do today?

**Do you use a cloud data warehouse as your primary destination?**  
If yes, we can set you up and optimize your processing so that it is faster and cheaper!

[Yes, let's go](#) [Remind me later](#)

**Ingest**  
Create an ingestion and replication task to replicate data from databases, files,...

**Transform**  
Create a mapping in ETL or ELT mode to load and transform your data.

**Orchestrate**  
Create a taskflow to define the control flow logic for your data integration and ingestion.

**Recent assets**

Status	Name	Location
Valid	mapping_rag_ingestion	00_RAG_Demo
Valid	DB_to_S3_parquet	Demo-Saha
Invalid	DB_to_S3_parquet	Demo-Saha

[View all](#)

**Recent jobs** Updated 11:45:51 PM PDT

Status	Instance Name	Start Time
Success	mapping_rag_ingestio...	Mar 26, 2025, 5:36 AM
Success	mapping_rag_ingestio...	Mar 25, 2025, 1:16 PM
Success	mapping_rag_ingestio...	Mar 25, 2025, 12:24 PM

[View all](#)

**CLAIRES Copilot**

Hi! I'm still learning. I can help with a few tasks, but soon I'll be able to do much more.

**How to create a SQL ETL mapping**

How can I create a SQL ETL mapping

To create a SQL ETL mapping, follow these steps:

1. In **Data Integration**, click **New Mappings** and then **Mapping**.
2. Click **Create**.
3. In the **New Mapping** dialog box, select **Mapping - SQL ETL** and then click **Continue**.
4. If you have configured a primary cloud data warehouse, the **Mapping Designer** will open in SQL ETL mode with the target connection set to your primary cloud data warehouse.
5. If you haven't configured a primary cloud data warehouse:
  - Choose your cloud platform and click **Select**.
  - **Data Integration** will prompt you to choose the target connection or create a new one.
  - Select a target connection. If your organization has no target connections for the selected platform, or if you want to use a connection that doesn't appear in the drop-down list, click **Add Connection** to create a connection.
  - Click **Next**. The **Mapping Designer** will open in SQL ETL mode with the target connection configured with the connection you selected.
6. Configure the target object, configure the source, and add any required transformations to the mapping.
7. Click **Save** to save the mapping.

Note that the **Mapping Designer** for mappings in SQL ETL mode is different from other types of mappings, and you cannot copy mappings in SQL ETL mode to other modes, or vice versa.

Citations:

1. Cloud Data Integration/Mappings/Mappings in SQL ETL mode
2. Cloud Data Integration/Mapping Configuration in SQL ETL Mode/Creating a mapping in SQL ETL mode

[Start Over](#) [Summarize what you have](#)

Ask CLAIRES...

Keyboard Shortcuts

# Automated and Augmented Generation

## What it can do today

### Feature

- Generate mappings and Mass Ingestion tasks (CDIR) using Natural Language
- Interactive interface to enable appropriate functionally complete ingestion assets and mappings

### Benefits

- Save time and effort and standardize data asset creation generation in Data Integration

### Who should use this?

- Data Engineers, Data Analysts

The image displays two screenshots of the Informatica Data Integration interface. The top screenshot shows the 'Auto Ingest Salesforce To ADLS Gen2' workflow configuration. It features a 'How do you want to transform your data?' section with a visual representation of data flow from a 'Source' (Salesforce) through '2 Tables Selected' to a 'Destination' (ADLS Gen2). A 'CLAIRE Copilot' chat window on the right provides assistance, listing available Salesforce connections and frequently used ADLS Gen2 connections.

The bottom screenshot shows the 'm\_datalake\_enrichment' mapping design. The 'Design' view displays a sequence of tasks: 'ADLSGEN2\_ORDER', 'SORT\_SOURCE', 'FIL\_QUARTER', 'AGG\_SALES\_SUMMARY', and 'EXP\_FORMAT', leading to two target destinations. The 'Properties' panel for the 'Source' task is visible, showing the name 'm\_datalake\_enrichment', location 'Default', and a detailed description: 'This mapping identifies the total spends for each account in the quarter. The data is enriched with the account and contact details. It will also calculate the average time between when the lead was generated and the time when the deal was closed.'

The 'CLAIRE Copilot' chat window on the right provides a summary of the workflow and a list of steps: '1. Read the Orders data', '2. Filter to only show data from latest quarter', '3. Aggregate based on account number and the quarter', '4. Lookup the Account Manager and the team details from Opportunity table to enrich the data', and '5. Write the data to ADLS Gen2 Sales Summary folder'. It also offers options to 'Create mapping', 'Save and run', and provides a confirmation message: 'Sure! The mapping has been saved in the "Default" folder. The Mapping Task "m\_datalake\_enrichment\_mct" has been created and a job has been initiated using "MSFT\_RTE\_Gartner" runtime environment.'

# Mapping Descriptions and Summarizations

## What it can do today

### Feature

- Generating mapping descriptions based on the business, technical and syntactical context

### Benefits

- Enforce best practice across organization
- Identify what a mapping does without breaking a sweat
- Identify what you have and make it easy for your users

### Who should use it?

- Data Engineer, Data Architects, Testers, Data Analysts



The screenshot displays the Informatica Data Integration (D.I.) interface. The main window shows a mapping design for 'm\_oracle\_snowflake'. The design consists of a sequence of steps: 'oracle\_customers' (Source), 'fil\_usa' (Filter), 'Expression' (Expression), 'Sorter' (Sorter), and 'Router' (Router). The 'Router' step is configured with two output groups: 'DEFAULT1' and 'PORTLAND'. The 'PORTLAND' group is filtered by 'CITY=PORTLAND'. Below the design, the 'm\_oracle\_snowflake' configuration panel is visible, showing the Name, Location (Default), and Description fields.

On the right side, the 'CLAIRe Copilot' interface is open, displaying a summary of the mapping. The summary is titled 'Summarize' and includes a detailed description of the mapping process. The summary text is as follows:

Here's a summary of the mapping as of January 15, 2025:

Detailed Summary:

- Sourcing data "CUSTOMERS" from an Oracle connection.
- Filtering records where "COUNTRY='USA'".
- Applying an expression "RTRIM(LTRIM(CONTACT\_NAME))" on "CONTACT\_NAME" to trim whitespace.
- Sorting data by "CUSTOMER\_ID" in ascending order before insertion.
- Using a router with an output group for "DEFAULT1" and "PORTLAND" (filtering by "CITY=PORTLAND").
- Inserting data into two new Snowflake targets: "CUSTOMER\_ALL" and "CUSTOMER\_PORTLAND" with field mappings by name.

At the bottom of the CLAIRe Copilot interface, there is a text input field labeled 'Ask CLAIRe...' and a 'Keyboard Shortcuts' button.

# Use cases & Best Practices

## Examples from our customers

Create the first draft of the mapping and then use it to modify.  
*Faster 0-1 mapping generation*

1

2

Replicate common patterns using natural language – *create copies of the same prompt and run recursively*

Empowering non-technical users to self service ingestion –  
*Data Analyst creates the data ingestion and replication task using NL*

3

4

Generate short and detailed summaries of mappings saving time and energy to identify the details of the mapping.

# Summary



## Autonomous DI

Automate 75% of the DM tasks, enabling all user personas to self-serve. Reduce 0-1 mapping generation time from hours to a few minutes.



## Design, Run & Operations Time

Automate all the repetitive & manual steps across the three phases of Data Pipeline Management



## Automate Design Time DM

Given Source(s), Target(s) and transformation req, autonomously generate the data pipeline for bulk load or batch load

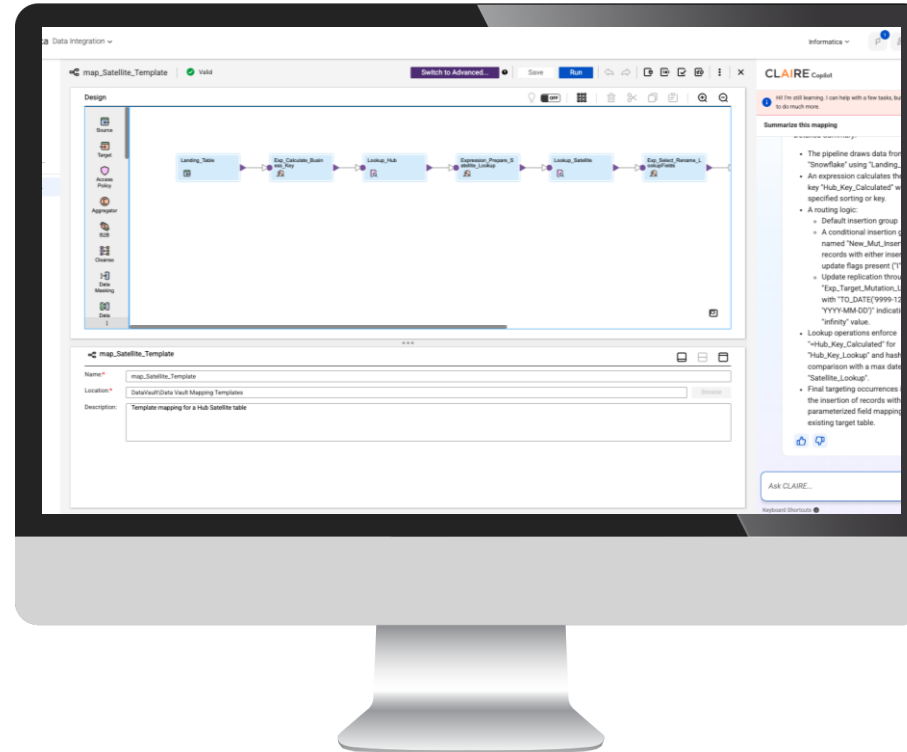


## Enhanced Productivity & Reduced TCO

Leverage Generative AI Natural Language and in-context embedded AI to help accelerate Data Management

# DEMO

DI Copilot



# Want To Learn More?

## Helpful Links

### Videos & Docs

- [Data Integration Co Pilot](#)
- Cloud Application Integration:
  - [Sample Prompts](#)
  - [Example Prompts](#)
- [CAI Co-Pilot](#)
- [CDI / CDI\(R\) Co-Pilot](#)

### Content

- [EU AI Act Blog](#)
- Cloud Data Integration:
  - [Sample Prompts](#)
  - [Improving Prompts](#)
- Cloud Data Ingestion and Replication:
  - [Sample Prompts](#)
  - [Examples DB and Replication Tasks](#)
  - [Example App Ingestion and Replication Tasks](#)

# Where data & AI come to

