



24 Sep, 2024

LLMs Orchestration with Informatica Boosting AI Efficiency

- Varun Sharma, Solution Architect, IPS
- Pavan Kumar Soma Prabhu, Principal Services Solutions Architect, IPS
- Andrew Phan, Consultant, IPS

Where data & AI come to 

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.

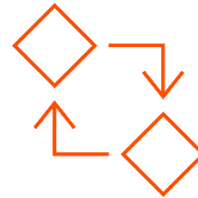
Feature Rich Success Portal



**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

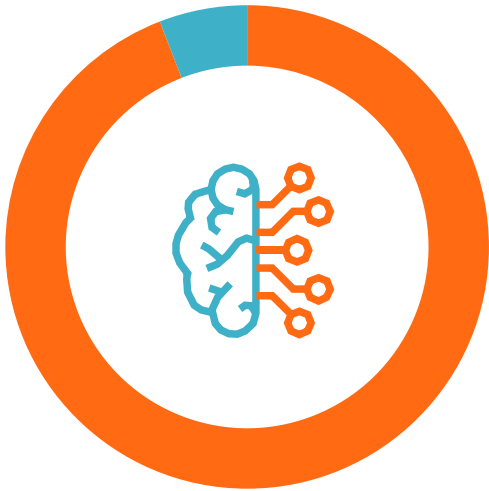
The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

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

- Large Language Models
- What is LLM-Ops ?
- LLM-Ops Architecture
- Implementation using Informatica IDMC
- Demo

Gartner CTO's Guide to the Generative AI: AI Is Top of Mind!



By 2026, more than **30%** of the increase in demand for application programming interfaces (APIs) will come from AI and tools using **LLMs**.

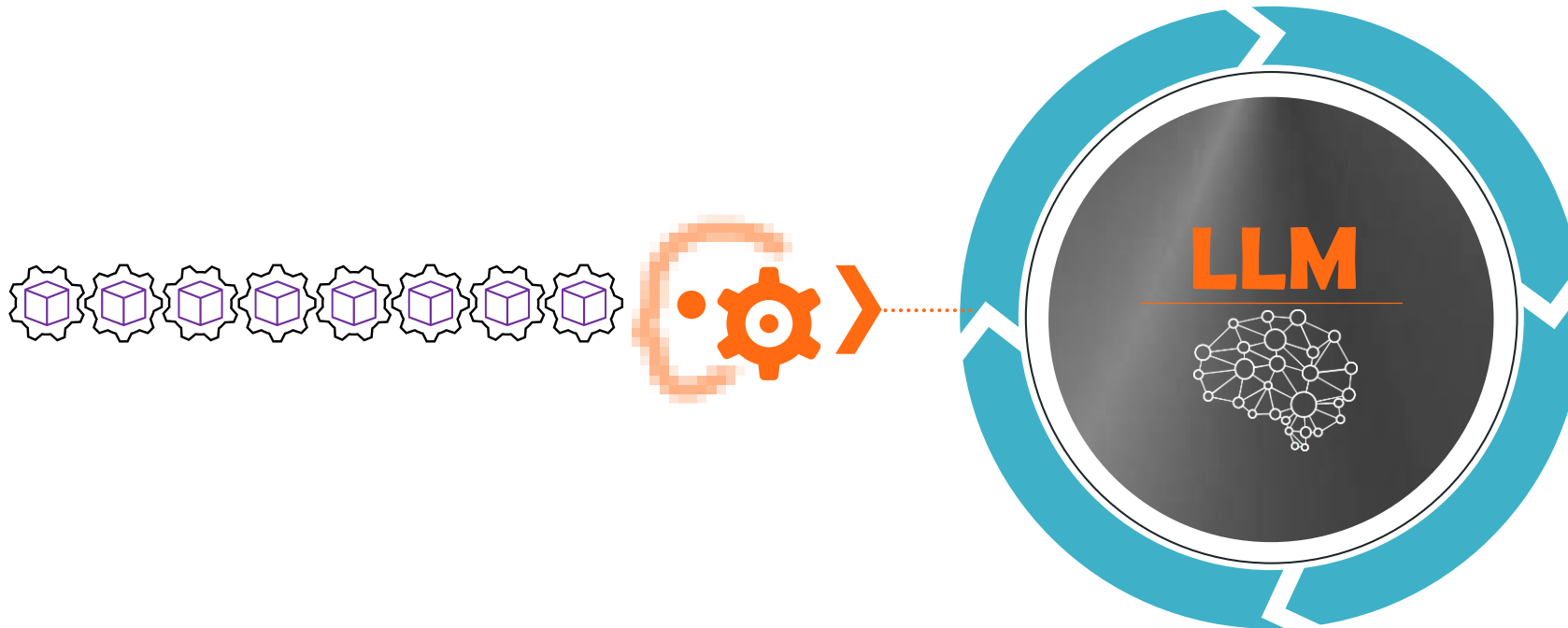
Over **80%** of independent software vendors will have embedded GenAI capabilities in their enterprise applications, a significant increase from less than 5% today

Gartner: A CTO's Guide to the Generative AI Technology Landscape

Large Language Models

Large language models

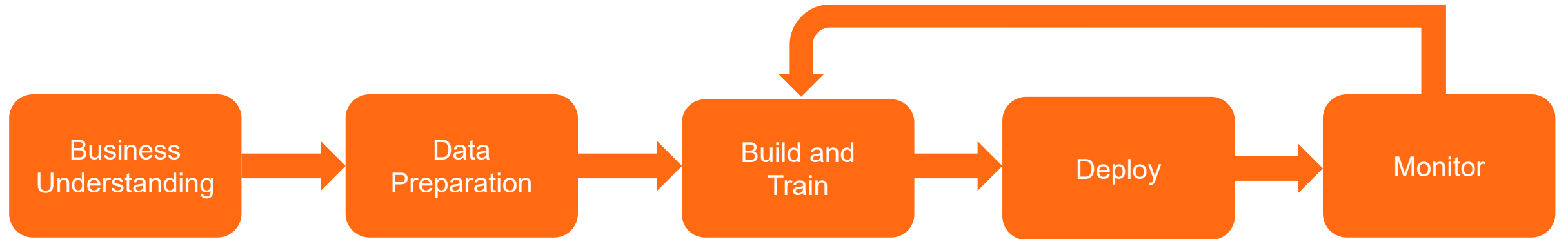
An LLM is a deep learning algorithm that can perform various natural language processing (NLP) tasks, such as text generation, translation, and summarization. These models are called “large” because they are trained on massive datasets and consist of billions of parameters



How Informatica Helps in GenAI Use Cases

Use Cases	Prompt Engineering	Retrieval Augmented Generation	AI Agent / Multi-Agent Systems	Fine Tuning
Definition	<ul style="list-style-type: none"> The art and science of crafting clear instructions (prompts) for AI language models. Enables generation of accurate and helpful responses from AI models. 	<ul style="list-style-type: none"> Incorporating authoritative or proprietary data sources into GenAI Models Improves accuracy, trustworthiness, and context of LLM responses. 	<ul style="list-style-type: none"> An AI program capable of autonomous decision-making and acting on behalf of users, systems, or other programs. 	<ul style="list-style-type: none"> Tailoring GenAI models through focused training on specific datasets to enhance their performance in tackling particular tasks.
Informatica Enables	<p>Simplified, low-code, orchestrated and governed Prompt Engineering with any LLM models</p>	<p>Automated, low-code pipelines with trusted knowledge using any LLM models</p>	<p>Scalable, low-code, governed AI Agent(s) trained on authoritative data using any LLM models</p>	<p>High-quality data to tune GenAI models through cleansing, API management & governance using any LLM models</p>
Informatica IDMC Services for GenAI	<ul style="list-style-type: none"> API & APP integration API management Out-of-the-box connectors Prebuilt Prompt Recipes Governance & Privacy for AI 	<ul style="list-style-type: none"> MDM Data Quality API & APP integration API management Data Engineering & Integration Prebuilt Recipes for RAG pipelines Governance & Privacy for AI 	<ul style="list-style-type: none"> API & APP integration API management Data Engineering & Integration Prebuilt Recipes for AI Agent Governance & Privacy for AI 	<ul style="list-style-type: none"> Data Quality Data Engineering & Integration <ul style="list-style-type: none"> Prebuilt integration pipeline templates Headless data management, INFACore API & App Integration Governance & Privacy for AI

LLM Ops - Architecture



Objective and Value Benefits



Objective

- Explore LLMs
- Starter Kit
- Scalable Operations
- Enhance Collaboration








Value Benefits

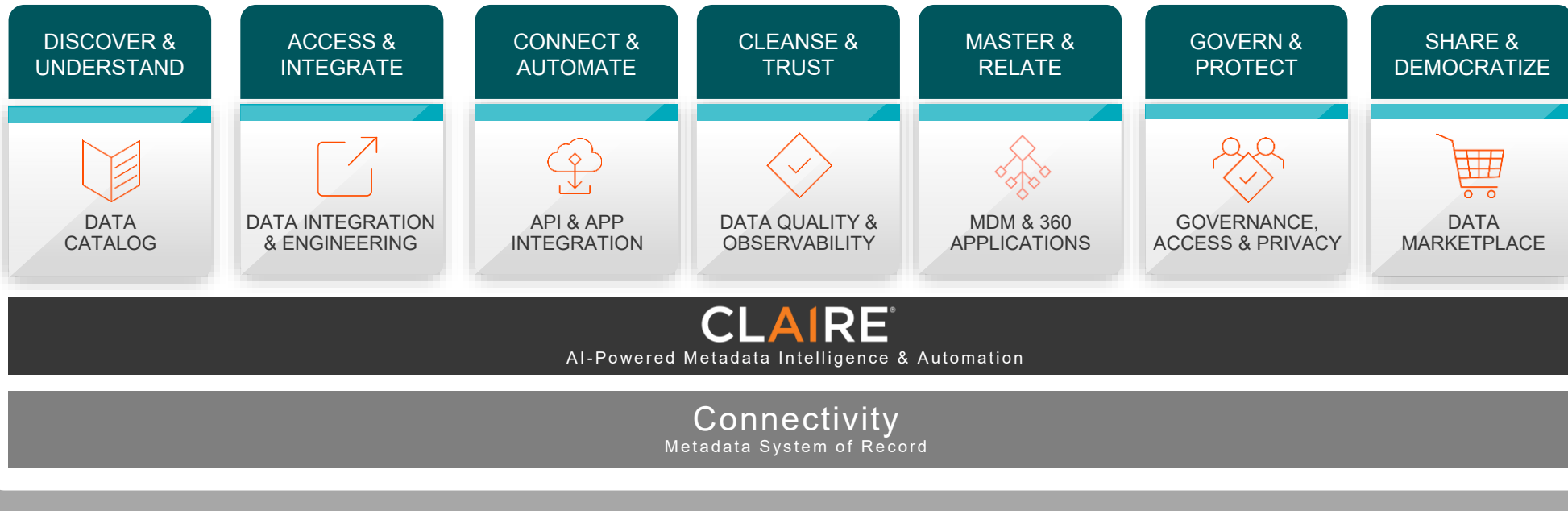
- Faster Time to Market
- Cost Efficiency
- Improved Performance & Accuracy
- Maximize potential using automation

IDMC Architecture

DATA CONSUMERS

-  ETL Developer
-  Data Engineer
-  Citizen Integrator
-  Data Scientist
-  Data Analyst
- Business Users

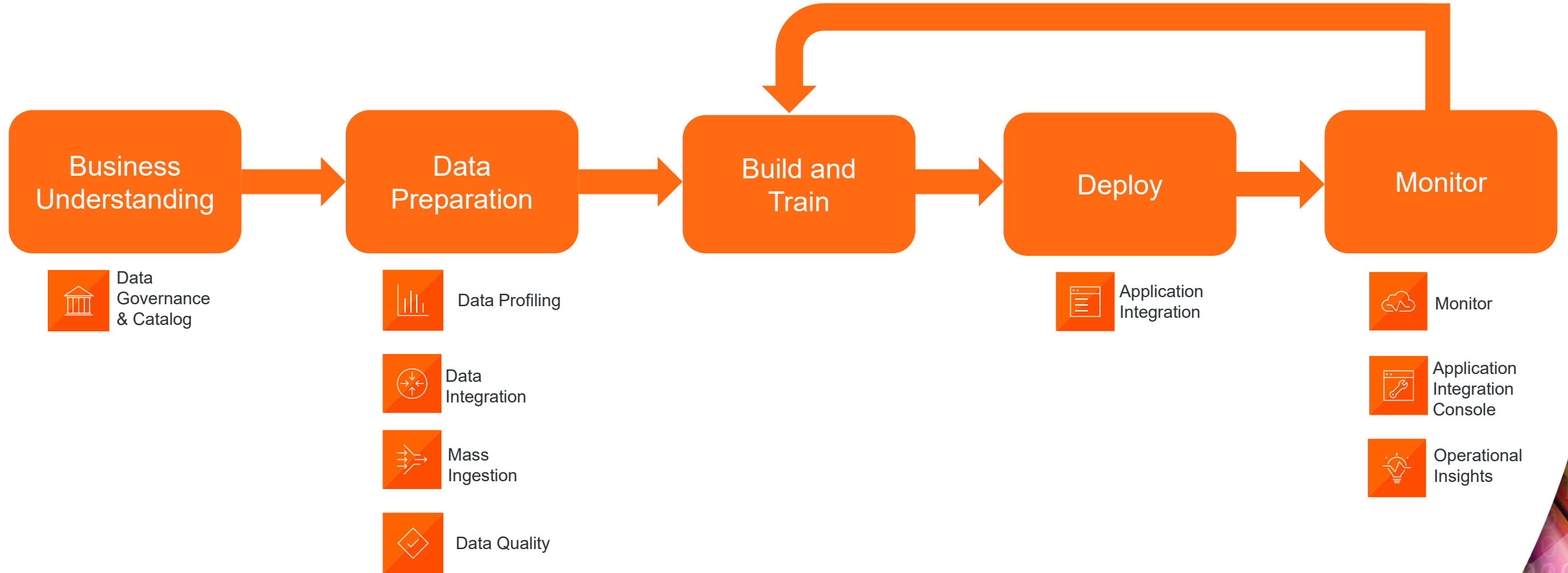
Intelligent Data Management Cloud



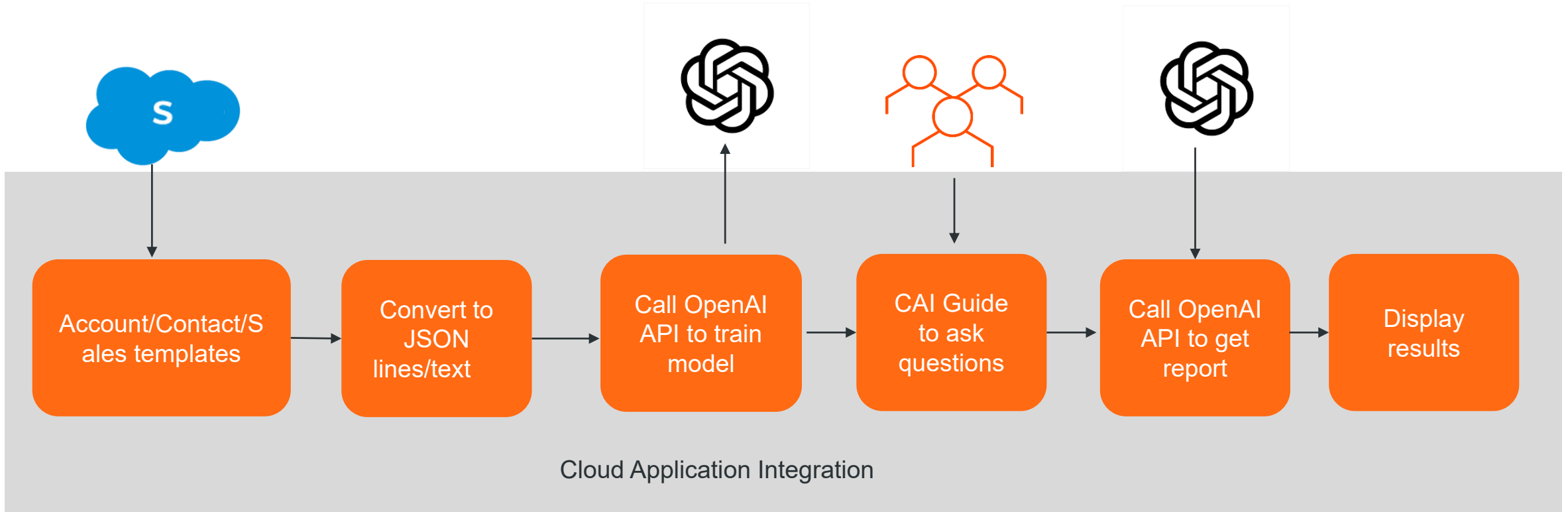
DATA SOURCES

- 
- 
- 
- 
- SaaS Apps Sources
- +
- 
- Mainframe
- 
- Applications
- 
- Databases
- On-premises Sources
- +
- 
- IoT
- 
- Machine Data
- 
- Logs
- Real-time / Streaming Sources

LLM Ops with IDMC



Using CAI to orchestrate LLM calls(CAI as AI Agent)



Demo