

DATA ENGINEERING

# SNOWFLAKE ARCHITECTURE 2026

 Orchestra

# SNOWFLAKE IN 2026

It's been over a year since we published our first guide to Modern Data Stack reference architecture in 2026. Since this time, the data landscape has developed massively with AI lurking in the foreground.

Semantic Layers, Compute Engines, and LLMs are the new battlegrounds for data warehouses. Databricks and Snowflake have both massively expanded their products, and the right way of thinking about the warehouse is now far from obvious.

In this whitepaper, we'll cover a few key topics to watch for those looking to simplify and improve their data architecture. We'll see how the architecture changed since the Modern Data Stack, and offer ideas to make Architecture robust through AI and 2030.

## SNOWFLAKE IN NUMBERS

 \$4.5bn revenue

"Snowflake delivered another strong quarter with product revenue of \$1.23 billion, up 30% year-over-year, and remaining performance obligations totaling \$9.77 billion, up 42% year-over-year," - **Sridhar Ramaswamy**

 9,100 accounts

using Snowflake AI

Intelligence has gone from a nascent offering to an essential capability for almost 2,5001 accounts in 3 months. [Source](#)

 13,325 customers

"We delivered strong new logo momentum, adding 740 net new customers, up 40% year-over-year, while continuing to deepen engagement across our base." - **Brian Robins, CFO**

 125%+ NRR

Introduced 430+ new capabilities in fiscal 2026, reflecting sustained innovation across the AI Data Cloud. Latest innovations include Cortex Code, a cutting-edge AI coding agent

### Snowflake cortex driving growth

1. Over 5,200 accounts (over half the user base) are leveraging cortex
2. Project SnowWork announced in March-26, which is focussed on business users

>60%

Of Snowflake users using dbt to power mission-critical data transformations  
[Source](#)

# ORCHESTRATION LAYER

Why Orchestration is the #1 priority for architects

Data Teams are spending less time building pipelines and more time maintaining systems. AI has reduced the cost of writing code to close to 0, but increased the need for strongly-governed, declarative systems of orchestration

Declarative Systems gaining popularity

1. Many systems following the path dbt set, by enabling declarative orchestration *outside* of Snowflake
- 2.



Orchestrating Snowflake with Orchestra

## Example: AwardCo

Awardco is a comprehensive employee recognition and rewards platform designed to boost company culture, improve employee engagement, and incentivize performance.

“We pull data from various source systems and SAAS Platforms into Snowflake where we transform it into a data warehouse. The data warehouse is used for reporting, analytics, and machine learning models.”

– **Micah Raventos**  
Data Architect, Awardco

“ WITH ORCHESTRATION WE’VE SAVED MONEY ON SNOWFLAKE, IMPROVE DATA FRESHNESS, AND FIX ISSUES FASTER. WE WANT AI TO BE A BIGGER PART OF THE SOLUTION”

• 1 HOUR SET-UP • EASE OF USE



MICAH RAVENTOS  
Data Architect  
Awardco

# AI AND SEMANTIC LAYER

## Understanding AI

Snowflake is allowing users to build their semantic layers natively without Snowflake, creating .yaml files atop data layers that allow Agent Analysts such as Snowflake Cortex to easily answer questions defined using natural language

## Why Data Architects Care

- 1. Speed of resolution:** Agentic Analytics is helping business users serve their own questions, removing load on the Data Team while increasing speed of resolution
- 2. Faster Build Time:** Data Engineering Teams do not lose focus answering questions with Agentic Analytics

Good afternoon, Adam  
What insights can I help with?

Ask Snowflake Intelligence...

Snowflake\_Documentation Sources: Auto

- How do I create a new Snowflake account and set up my first database?
- What are virtual warehouses in Snowflake, and how do I properly size them?
- Can you explain zero-copy cloning and how to clone a database or table?

Snowflake Intelligence

## Example: RxBenefits

RxBenefits is the first and leading pharmacy benefits optimizer (PBO) in the employee benefits industry. Since 1995, we've partnered with benefits consultants to provide pharmacy cost-containment solutions to self-insured employers of all sizes.

"We use Snowflake, dbt cloud, Boomi/Rivery and the AWS Data Stack. Before orchestration we had zero oversight of our pipelines. Now with reliable data stakeholders are happier, and AI analysts are extremely powerful"  
– Justin Frye  
Senior Data Engineer, RxBenefits

" WITH ORCHESTRA  
WE'VE HIT THE  
GROUND RUNNING –  
ALL OUR  
INTEGRATIONS WERE  
SET-UP IN 4 HOURS.  
THE EASE OF USE IS  
UNMATCHED"

4 HOUR SET-UP EASE OF USE



JUSTIN FRYE  
Senior Data Engineer  
RxBenefits, Inc

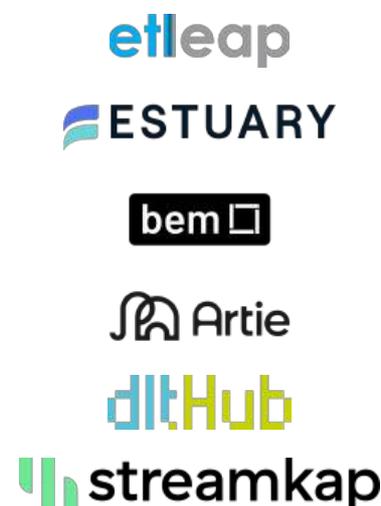
# CONSOLIDATION IN ELT

## Why consolidation is a headache

The industry is consolidation. Fivetran acquired Census and is merging with dbt Labs. Coalesce acquired Castor and Synq. Snowflake acquired other companies like Select Star. This makes it difficult to evaluate vendor risk when looking at tools

## Key Focus Areas for Architects

- 1. Cost:** Companies getting bought may signal higher price increases (e.g. Boomi) or may reduce prices instead
- 2. Vendor management:** minimising the number of vendors can ease the operational load on Data Teams by consolidating vendor management



Next generation ELT providers: etleap (iceberg), Estuary (batch and streaming), Bem (unstructured), Artie (CDC), dlt (build-it-yourself), Streamkap (streaming)



459%

ROI for implementing Fivetran over 3 years<sup>2</sup>



48%+

Time back for data engineers from implementing Fivetran<sup>2</sup>



\$1.5M

Average benefit for an organisation looking to implement Fivetran<sup>2</sup>

Over 80%<sup>1</sup> of companies leverage a data ingestion tool of some kind. The ROI for Data Ingestion tools is clear, and given the renewed focus on speed and simplicity, architects may want to be on the look-out for start-up or less established players looking to deliver next generation solutions at better prices. Ensuring there is easy switchability between tools through orchestrators can also speed up migration work and reduce costs.

ARCHITECTURE

# ARCHITECTURE

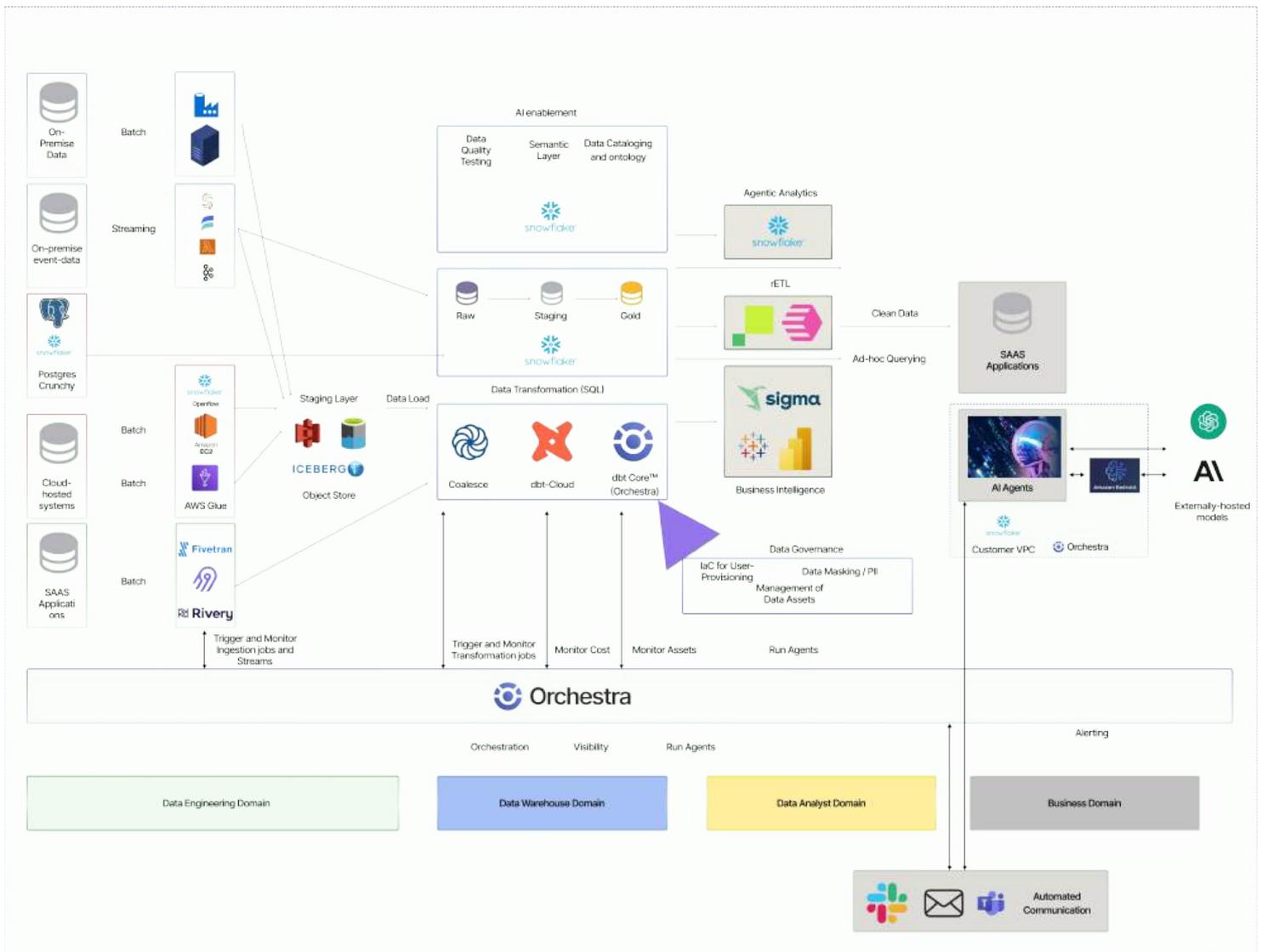
The common architecture patterns

 **Orchestra**

# DATA ARCHITECTURE

You may need to zoom in a bit! See if you can notice the difference vs. the architecture that was implemented last year.

## Snowflake AI and Data Architecture (2026)



# DATA ARCHITECTURE

## Changes to ELT

In 2026 data architects are focussing more on streaming and kappa-style architecture, ensuring systems are simple from the beginning instead of building difficult-to-manage, complex systems spread across tools and infrastructure areas.

Snowflake also consolidated the end-to-end into their platform with their acquisition of Crunchy data, and should soon be able to offer native postgrges that can automatically replicate to Snowflake, similar to Clickhouse's offering.

Finally, Snowflake announced OpenFlow in addition to their native connectors, allowing data teams to seamlessly leverage Apache Nifi for data movement needs – all within the Snowflake Ecosystem.

Systems of transformation should now be **state aware** – meaning that orchestrators store the state of data as it progresses. This reduces the burden on streaming tools like Estuary and data engineers, since data transformations can simply be “set and forget”, instead of spending lots of time coming up with separate schedules across pipeline steps.

## Changes to Governance

In 2026 Snowflake made big changes to Governance, pushing out Snowflake Quality Rules, adding dbt in Snowflake, adding a native cataloging functionality, and a semantic layer too. Governance-related workflows appear to be best-placed living in Snowflake.

This has a direct knock-on effect for AI Architecture. Good AI requires good data. By creating a semantic layer, those who are building for AI should be able to build out that AI in Snowflake

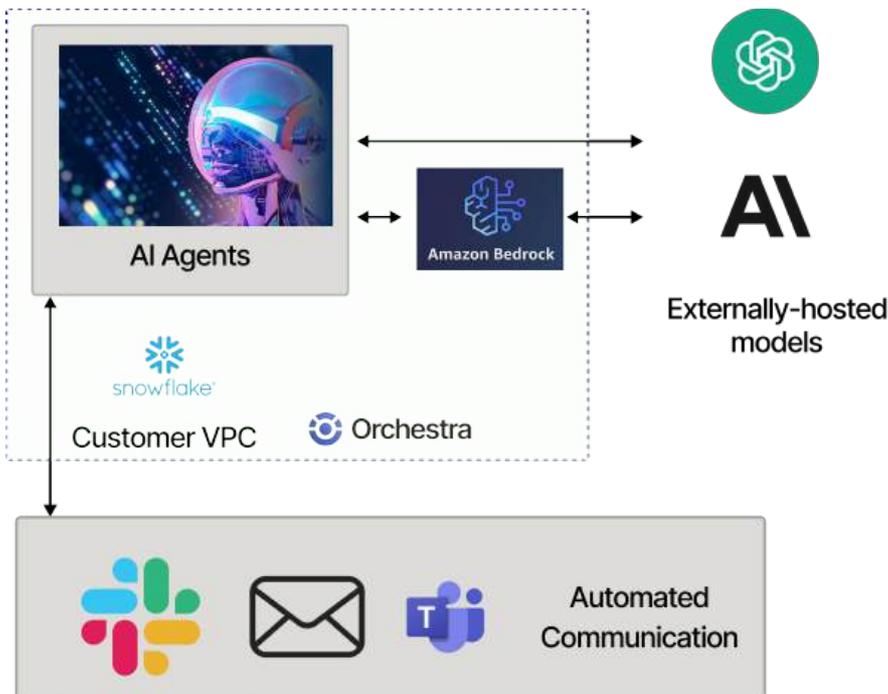
## Data Quality and Semantic Layering

Many Domain Architects and Data Architects could have said in 2023 that for data to be human-searchable, an additional layer like an ontology or semantic layer was necessary. In 2026, Snowflake provides their own (“Semantic Views”). Companies may also choose to use their own, such as cube.dev or TextQL.

# AI ARCHITECTURE

## What is AI Architecture

Data-warehouse centric AI Architecture has emerged in 2026.



## Agentic workflows for self-serve

Snowflake provides the infrastructure for developing applications, hosting LLMs and also enabling “assistant” or “chatbot” type agents. In the context of enabling self-serve analytics, this is an attractive pattern and makes sense – data teams can take advantage of deploying AI infrastructure in the same place as the warehouse without adding another vendor.

## The future: Agentic Orchestration Architecture

Despite the inroads Snowflake has made, many Data Architects prefer a more generalisable pattern. In the more generalisable pattern, Snowflake is but a component of the agentic stack, where the Snowflake APIs and Snowflake Cortex Agents are made available to other agents via API and MCP.

Data Architects are working on “Master” agentic orchestrators; systems that can take in prompts, and non-deterministically choose what to do. This may involve calling LLMs hosted within an enterprise’s VPC (e.g. AWS bedrock), and external provider (Open AI, Anthropic), or fetching data (Snowflake). The agent itself will likely run in an Orchestrator, where tools like monitoring, audit, and memory can be easily added on

# SUMMARY

## The future: Agentic Orchestration Architecture

Despite the growing consolidation in the space, many of the “old names” from previous data architecture additions are still present in the architecture for 2026. This reflects the undeniable appeal of the “Best-in-class” architecture from the Modern Data Stack – where the most important functions like data movement, warehousing, and orchestration warrant their own, separate tool.

Less in-vogue areas from the MDS, such as governance, data quality, semantics, and observability are getting absorbed by companies such as Snowflake.

Data Architects’ biggest challenge to 2030 will be to transition to AI architects. They must design systems of agents that allow anyone in an organisation to reliably and repeatably automate tedious, time-consuming, complex tasks using AI.

While Data Teams are happy to leverage a modern data warehouse rather than build their own, AI is not so simple. AI Agents require secure connectivity to dozens of hundreds of systems; a function only the orchestrator succeeds with today. Data Teams will not build their own LLMs, but will insist on defining the “business logic” – the “brain” of an AI Agent.

Where these agents run is an open question, but enterprises are for the most part, proceeding cautiously with AI and have a strong preference for secure cloud-hosted systems and own-deployed LLMs. This pattern goes against the Snowflake-maximalist principle of running everything in Snowflake.

As we enter the rest of 2026 one thing is true despite consolidation: Data Architecture on Snowflake is still as interesting, nuanced, and complex as it was in 2024. AI is only making the picture more detailed, and it is a great time to be building on Snowflake.

FIND OUT MORE

# CONTACT

For any questions relating to the report, please contact  
[support@getorchestra.io](mailto:support@getorchestra.io)

To find out more about Orchestra, check-out [https://  
getorchestra.io](https://getorchestra.io)



## LEGAL NOTICE

© 2026 Orchestra Technologies Limited. All rights reserved.

This whitepaper and all content contained within it, including without limitation all text, graphics, diagrams, data, analysis, and branding, are the property of Orchestra unless otherwise stated. No part of this publication may be reproduced, distributed, transmitted, displayed, published, stored in a retrieval system, or used in any form or by any means, whether electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of Orchestra.

Requests to reproduce, republish, quote, excerpt, reference, or otherwise use any part of this whitepaper must be submitted to Orchestra in advance for written approval. Where permission is granted, full and clear attribution to Orchestra must be included in every instance of use.

This whitepaper is provided for general informational purposes only. While Orchestra has made reasonable efforts to ensure the accuracy of the information at the time of publication, the content is provided on an “as is” and “as available” basis without warranties of any kind, whether express, implied, statutory, or otherwise, including without limitation warranties of accuracy, completeness, reliability, merchantability, fitness for a particular purpose, non-infringement, or suitability for any specific use. The contents of this whitepaper do not constitute legal, financial, tax, investment, technical, or other professional advice, and should not be relied upon as such. Any reliance placed on this whitepaper is strictly at the reader’s own risk.

To the fullest extent permitted by law, Orchestra shall not be liable for any direct, indirect, incidental, consequential, special, punitive, or exemplary loss or damage arising out of or in connection with the use of, reliance on, or inability to use this whitepaper or any information contained in it.

Orchestra reserves the right to update, revise, or remove the contents of this whitepaper at any time without notice.

For permissions and usage requests, please contact: [legal@getorchestra.io](mailto:legal@getorchestra.io)