tray.ai

Five keys to integrating Snowflake and Databricks into your business in the age of AI

Table of contents

The evolution of integration

Data engineering teams are well-acquainted with traditional Extract, Transform, and Load (ETL) processes that have been a cornerstone of data management since the 1990s. Originating with the rise of the data warehouse and the need to integrate and cleanse data for downstream analytics, enterprises adopted ETL tools like Informatica to extract, transform, and load data from operational systems into centralized data warehouses for reporting and analytics.

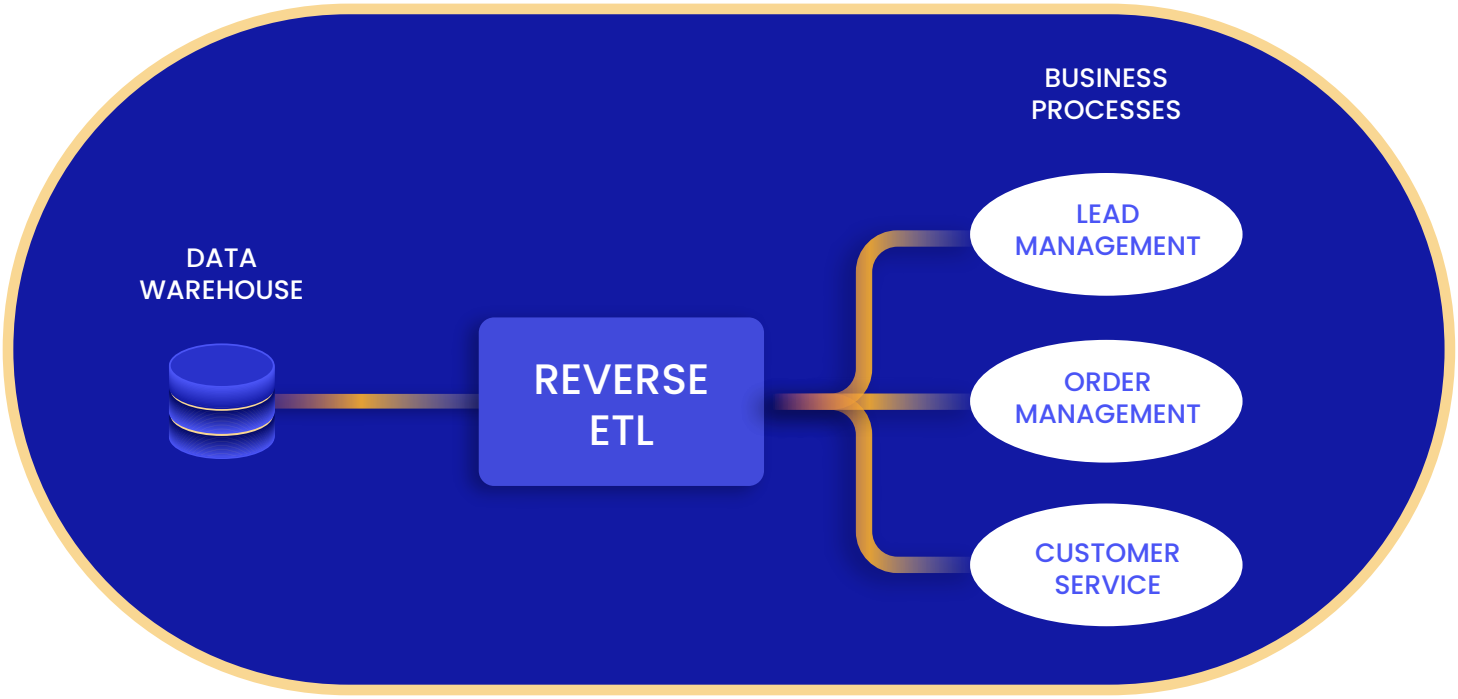
Over time, ETL evolved into ELT (Extract, Load, and Transform), where data is loaded first and transformed later. The sources and targets have changed dramatically, incorporating CRM, ERP, and HCM systems across cloud and on-premises, alongside evolving data warehouse technologies. Yet, the core concept remains the same—data is pulled from various systems such as marketing, sales, finance, and HR, then loaded into a data warehouse designed for fast querying and analysis.

With the need to fuel data lakes and lakehouses, the traditional ETL/ELT integration pattern, whether batch or real-time, isn't going away anytime soon. However, enterprises are looking to get more value out of the data assets they create, moving beyond simply using that data for analytics and instead leveraging it directly within marketing, sales, finance, or service processes. Data warehouses, too, are evolving and becoming multi-model, incorporating vector storage to support Retrieval Augmented Generation (RAG) use cases far beyond their traditional analytics-centric roots.

It's the evolution of the data warehouse from a system of analytics to a system of data intelligence, an asset that doesn't just inform analytics but can also drive the business. These warehouses have essentially evolved into full-fledged cloud Database Management Systems (DBMS).

This critical shift offers the potential for organizations to get more ROI from their data assets, shifting from a single analytics-only use case to driving countless business processes. However, it requires enterprises to think about how to leverage their data warehouse assets in a new way. The benefits are manifold. Enterprises can gain more value from their database and cloud data warehouse investments by no longer using them as single-purpose analytics data stores. Instead, the data can be used to power a variety of business processes with intelligence.

This shift has a profound impact on integration patterns and how integration tooling works. No longer simply a rigid, rarely changing uni-directional data load from systems of record or unstructured sources into the data warehouse. Instead, the data warehouse becomes an active participant in hundreds of business processes and workflows driven by business teams. Now, the data flows in the other direction, potentially millions of times per day, from data warehouses into real-time business processes, whether simply acting as a Customer Data Platform (CDP) for lead and service flows, informing order management processes based on a customer risk lookup, or acting as an asset in a RAG pipeline. The process is commonly termed **Reverse ETL**, requiring a rethink of traditional integration tooling, from what it provides to how it's used.



The operationalization of the data warehouse

With the rise in AI, modern data warehouses like Snowflake, Databricks, Redshift, and BigQuery are increasingly housing a broad array of data, including traditional business data, data enriched by AI and large language models (LLMs), core data for training machine learning models, and unstructured data.

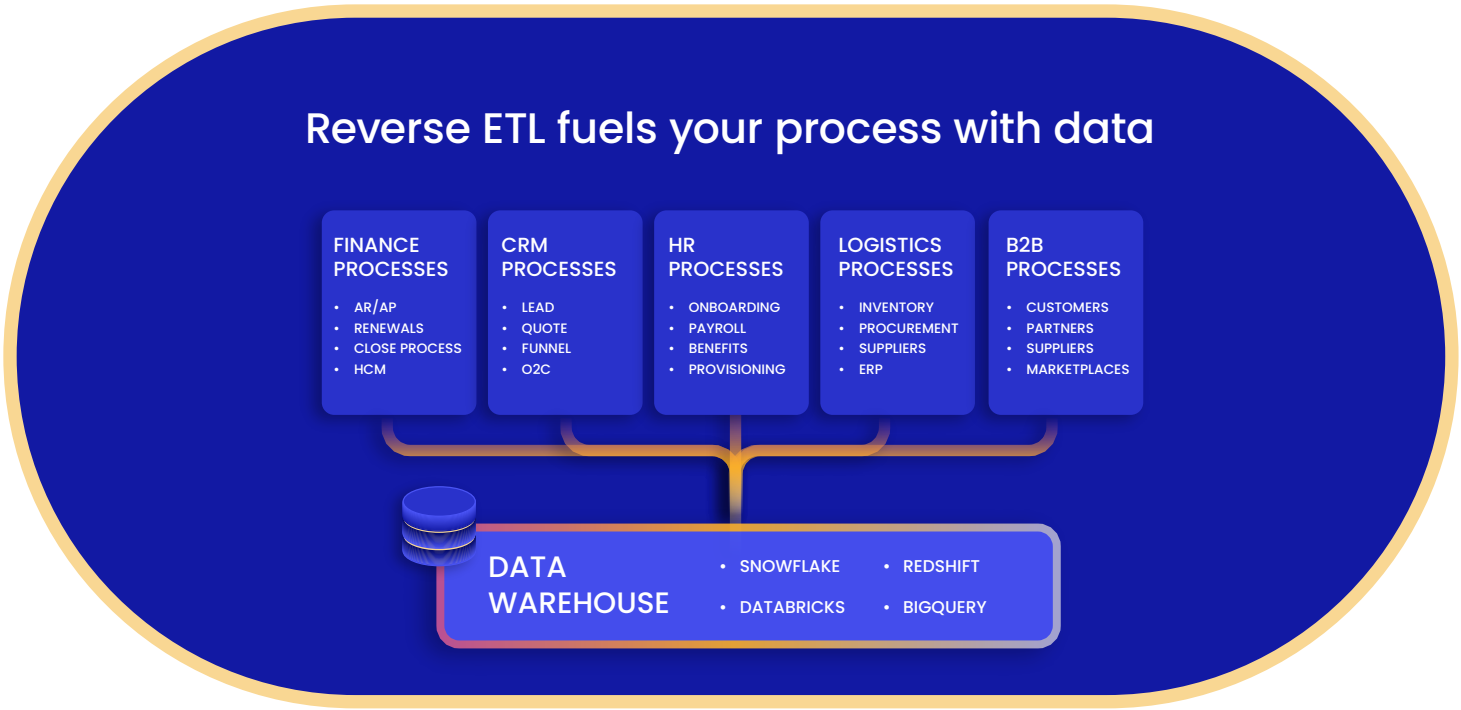
Beyond traditional data, these warehouses often now support a broad spectrum of AI data, catering to the diverse needs of data scientists, engineers, and developers involved in AI and machine learning (ML) projects. Key types of AI data that can be managed and analyzed with Snowflake include structured data, which is essential for traditional machine learning models. These warehouses now extend to handling unstructured data, accommodating various information forms such as text-heavy documents, social media conversations, images, videos, and audio. This is crucial data to drive more advanced AI applications, including natural language processing (NLP), intelligent agents, chat experiences, and computer vision, where such data is needed.

Beyond AI and analytics, these data assets often now serve as CDPs. By acting as the point of consolidated customer data from multiple channels, systems, or data streams to create a unified customer profile, they have the potential to infuse master customer data, customer behavior, and preference information into the business. This unified data can then be used for advanced analytics, personalization, and driving a better customer experience during every interaction.

Reverse ETL: Defined

Reverse ETL is fundamentally about incorporating data from your data warehouse into your everyday business processes, the process automations required by your business teams to run your business. Instead of loading data into warehouses like Snowflake or Databricks for analytics, data now flows out of these warehouses to fuel real-time business processes, driving better decision-making and automating workflows across enterprises.

For example, marketing teams are now using reverse ETL to power their processes for real-time personalization, customer engagement, and lead management processes, enriched by data from Snowflake, BigQuery, or Azure SQL. Where these data warehouses are being used as CDPs serving as a single source of truth for all customer data, they act as the customer master in lead management, order management, customer service flows, or any number of business processes.



Modern data warehouses like Snowflake and Databricks are equipped with rich APIs and elastic scalability, allowing them to handle the high-performance demands of business processes like lead routing and order management in real-time. These platforms, which were initially designed for analytics, have morphed into engines that power automation and standardization across the entire business stack. They are central to driving AI-infused processes and improving decision-making across multiple departments.

Five keys to rethinking your integration strategy

1. Simplify by unifying your Process Automation and Data Integration tooling

No one wants to deploy or maintain multiple platforms for different integration patterns. It crushes efficiency, costs too much, and hurts maintainability. That's why it's now important to take a "Goldilocks" approach to integration when it comes to incorporating Reverse ETL and using your data warehouse data within your process, alongside handling traditional ETL requirements

Most iPaaS (Integration Platform as a Service) are optimized for one use case or another. Some are heavily ETL-focused, with strength in data mapping, schema management, and cleansing. However, traditional ETL/ELT tools often remain complex, expensive to maintain, and focused on batch data integration, requiring data integration specialists to manage data flows. Many of these tools also carry the baggage of a vast array of decades-old arcane functionality that is little used by modern data integration teams. However, more problematic, they are often not designed to visualize business processes or be easy enough for business leads to build business process automation, but rather designed for visualization data merging and transformation.

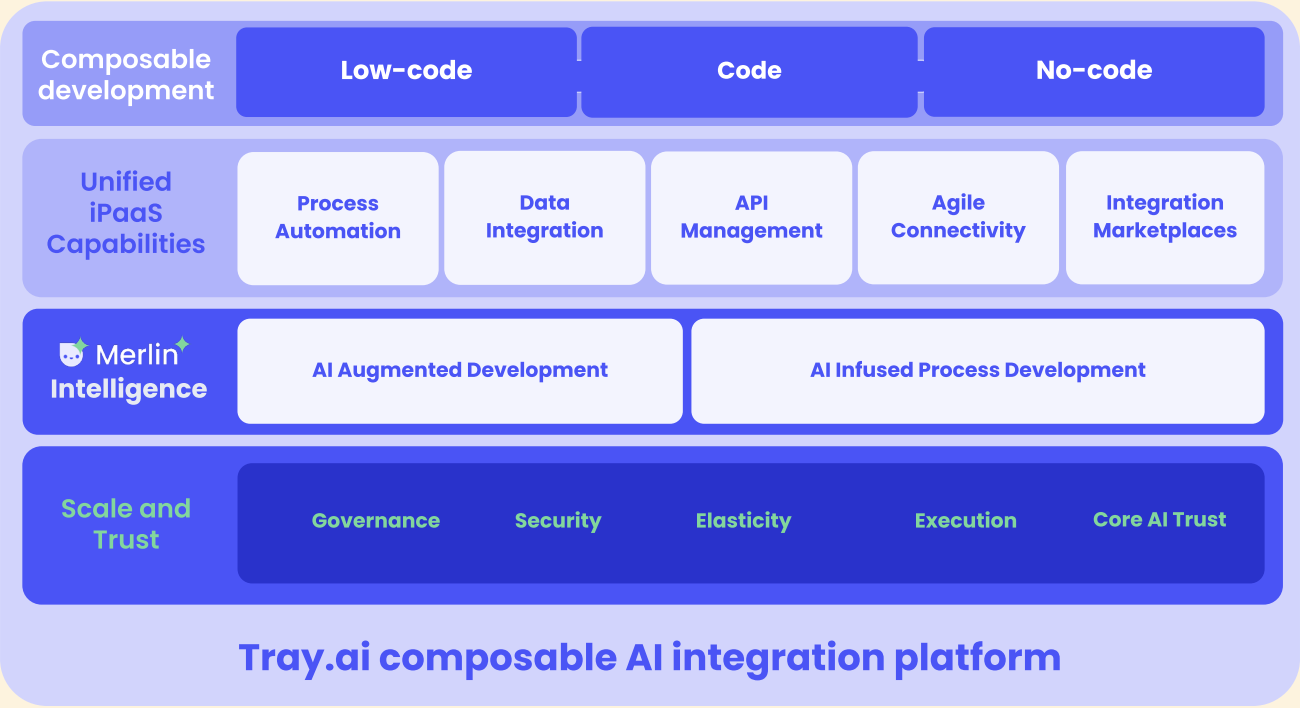
Other iPaaS tools focus heavily on developers

Tip: Use one Integrated Development Environment (IDE) for all integration patterns. Companies can't afford separate developer tooling for app integration, data integration, APIM, and even more tooling to integrate AI. This creates fragmentation, a huge learning curve, and hurts reuse. A recipe for tech debt. Shifting to one IDE – app integration, data integration, APIM, and microservice development, enables teams to achieve composability and maximize reuse and collaboration.

with roots in API management or are a great fit for the lightweight automation of quick-firing business processes but lack the scale to move heavy-weight data workloads for ETL. Without unified thinking, enterprises quickly end up with fragmented tooling—separate tools for separate integration patterns—an integration hairball.

Instead, teams must evaluate their requirements in the context of one unified platform so they can consolidate their architecture and not end up with fragmented tooling. This includes conventional data integration requirements, but also process automation requirements such as event-based integrations and automations and business process visualization, as well as API Management (APIM), and more.

A composable AI integration platform unifies all integration patterns and provides native support for AI capabilities.



AVID Property Group

NUCLEUS RESEARCH CASE STUDY EXCERPT

AVID Property Group takes a unified approach

The company: AVID Property Group is an Australian builder-developer responsible for a diversified selection of residential master-planned communities, land lease communities, completed houses, and apartments across Australia's eastern seaboard. With a record of delivering connected and vibrant communities, AVID brings people together to deliver thriving environments for its customers.

The project: AVID evaluated six different iPaaS vendors, including MuleSoft, Workato, Boomi, Talend, and Tray.ai, to identify the most suitable solution for its user skill sets and functional requirements. Specifically, AVID evaluated each solution's capacity to support each of the following projects: Accounts payable invoice automation, data platform modernization, and removal of legacy technology.

To enable these projects, the organization sought a platform that could scale effectively, capable of integrating various systems, automating processes, and efficiently moving large volumes of data. AVID narrowed its evaluation down to a short list of Boomi, Workato, and Tray.ai. It ultimately selected Tray for its ability to support data integration and workflow automation, whereas Boomi and MuleSoft favored data integration over workflow automation. The organization also required a cloud-native tool. At the time, Tray was the only vendor that could do so with zero physical footprint.

The results: One of the first projects addressed with Tray involved accounts payable invoice automation. This project took two months to implement with an internal team consisting of an enterprise architect, an integration specialist, a data architect, and a business analyst at a 75 percent time commitment and now addresses 2500 to 3000 monthly invoices.

The second project involved supporting AVID's new data platform, integrating various applications to form a single source of truth. Although Tray is not directly responsible for the costs or benefits of this project, the platform was a key factor in its success. This enabled the organization to decommission

its legacy data transfer database and other data sources and implement Google BigQuery, laying the foundation for improved data management and governance, which has proved crucial at scale.

The third project focused on the removal of legacy technology, redesigning AVID's enterprise architecture around new integrations and automations, replacing many hardcoded error-prone integrations between systems.

AVID achieved a 194% ROI, a 7.2 months payback, and avoided $6M in consulting and implementation costs by taking a modern, unified approach.

The full AVID ROI story written by Nucleus Research is available for download on the Tray website.

Download now

Tip: Use a composable integration approach to maximize your data. AVID took a composable integration platform approach. By handling all the patterns in one platform, from process automation to data integration, they were able to maximize delivery efficiency and minimize costs.

2. Empower your users to incorporate data into their processes with low-code and AI-augmented development

Reverse ETL is essentially business process-centric. It's your business users that are using data from your database of choice. It's where your ease of use requirements will differ significantly from conventional ETL tools. Since business teams in marketing or finance are often in charge of automating the lead lifecycle or the invoicing process, they will become a significant user of tooling that enables reverse ETL. They'll be looking to build and take control of change management. This is a sharp contrast with typical ETL, which is often only used by a handful of highly trained specialists

Modern, unified, low-code platforms provide drag-and-drop features that allow users to design complex workflows, business processes, and integrations without extensive coding. This approach significantly reduces the barrier to entry for creating integrations and process automations, enabling business analysts, project managers, and other non-technical users to actively drive or contribute to the development process—essential for infusing your data warehouse data into business projects.

These business users or business technologists will value a more diagrammed business process-oriented experience to build their own integrations and automations faster. Instead, business users are looking to easily compose automations, assembling integrations using business logic that consists of conditionals, loops, branching, nesting, and error-handling, visually tapping into events, apps, databases, or other workflows.

Make it simple for your business users with AI augmentation

Adding AI into this mix further enhances the capabilities of low-code platforms. AI-augmented development assists users in several ways.

GenAI coupled with ML opens up many ways to speed up development, including LLM-powered copilots that can complete 50% or more of the integration automatically so integration developers can focus on the "last mile." GenAI-assisted data mapping can provide a significant upgrade to speed with different fields that can be mapped, while AI-

assisted debugging can quickly enable less technical users to zero in on suspect areas of their integrations.

AI can provide a natural language copilot experience to quickly frame out integrations or provide assistance automating repetitive tasks within the integration process, such as data mapping between different systems. It can provide recommendations and suggestions or explain how a workflow runs, speeding up the development time and reducing the likelihood of errors.

Take a tour to learn how to build your integrations with Tray.

Take interactive tour

3. Make it easy to blend your data with AI services: Ensure your integration and automation tooling is AI-ready

Like nearly every other facet of business, AI is poised to change everything—and reverse ETL is no exception. With the emergence of AI-powered platforms, organizations can now combine the power of Snowflake and Databricks with other AI services to create intelligent business processes.

Business teams are evaluating how to quickly upgrade their business processes with AI – from finance and marketing teams looking to add GenAI, incorporate text extraction, sentiment, text gen, and LLMs. Your Snowflake, BigQuery, Databricks, or Redshift data will play among your corporate data, LLMs, 3rd party vector databases, and other AI services,

like text extraction, classification, or sentiment. Ensuring it's easy to blend your data with this landscape is a way to maximize its value.

AI-ready iPaaS includes a range of AI operations as standard natively, so teams can begin infusing straight away – no dependencies required. Platforms like Tray, an AI-ready integration platform, make this process even easier with an AI Palette, which organizations use to connect Snowflake or Databricks data with AI services such as LLMs, sentiment analysis, and text classification, together with business data. This powerful integration means organizations can not only analyze data but automate decisions and create self-improving processes driven by AI insights.

Example 1: eCommerce order management

An online retailer aims to improve its order fulfillment process to reduce fulfillment errors, drive quicker shipping times, and boost customer satisfaction and loyalty, driving repeat business. They build an order management process automation that integrates the company's ERP data, which includes supply chain and logistics details, with Snowflake or Databricks centralized and historical customer and order information and coupling it with AI to identify the optimal warehouse for order fulfillment, to drive accuracy, speed, and efficiency.

Example 2: High-tech company customer support

A service desk aims to significantly reduce response times and improve customer satisfaction and loyalty. It builds a customer service application that integrates its CRM database, which houses detailed service history, with Snowflake, which incorporates product and order data and calls out to LLMs, as well as a vector data database for real-time RAG-based conversation management. This new app allows the business to offer personalized support based on customer history and predictive AI analytics, resulting in faster issue resolution, a more personalized service experience, and increased customer loyalty.

Tip: Get hands-on and start building. It's better to try before you buy. Not every tool pitched as low-code or AI-ready is easy or fully functional. So get business teams hands-on and see how fast and easy it is for them to build automation while tapping into their data warehouse use case(s). Look for the provider to offer a **free trial**, where you can instantly start building.

4. Plan for different workloads by prioritizing a serverless architecture

Traditional data integration and ETL workloads are predictable, and teams can tightly provision resources based on anticipated demand to meet processing windows. Reverse ETL-driven processes are different. They may vary based on a marketing campaign or spike in orders, with a workflow spiking from thousands per day to millions of requests and then back down. In addition, they are highly concurrent. A typical data integration process will often execute sequentially with a degree of parallelism, while reverse ETL processes may be executing across business teams, fueling processes across marketing, sales, finance, and services. With business teams in the driver's seat building, it can be virtually impossible to provision and manage traditional server or runtime-based integration tools.

A serverless architecture provides a more elastic foundation to support more predictable, high-volume traditional data integration workloads and less predictable, higher-magnitude business processes. With a serverless architecture, there are no worker nodes or runtimes to install and provision. The platform automatically and dynamically requests resources at runtime; there is no upfront provisioning or de-provisioning requirement. It enables business teams to quickly build and deploy business processes and integrations incorporating data from your Snowflake or Databricks deployment without creating operational overhead within the integration tooling.

iPaaS, built using serverless processing, enables teams to completely focus on development and delivery by providing automatic elasticity, enabling more agile project prototyping, and making moving to production simple.

5. Design your reverse ETL processes with trust at the core

Using data from your data warehouses with LLMs, other external AI services, and directly within business processes offers business teams exciting opportunities for innovation and efficiency gains. However, it's critical to ensure the platform you leverage provides strong data controls for maintaining data security, helps ensure compliance with data protection laws, and provides central governance and control over data flow while enabling your users to build in a decentralized way.

Native support for strong data controls is a must. For example, safeguarding personally identifiable information (PII) with sophisticated tokenization techniques ensures anonymity as your data assets interact with AI services. Tokenizing the data streams fed into AI systems helps ensure data privacy throughout the AI processing journey. The data is then detokenized, allowing for the continuation of AI-enhanced processes with an added layer of security.

A complete audit trail is also crucial. This ensures full transparency of the data shared across AI services, which is important for tracking and verifying the proper use and handling of data across every process automations. This level of oversight and accountability is indispensable for businesses aiming to maintain trust and integrity in their AI-driven initiatives.

Tip: Use composability principles to build higher-level orchestrations from foundational-level data controls. Composable development involves creating integrations, process automations, or applications from building blocks or modules. This approach allows business teams to reuse business logic and assemble quickly and flexibly. By using curated building blocks, business teams can use pre-built components and integrations that incorporate compliant data controls while building higher-level orchestrations.

Next steps: Put your data to work with Tray.ai

Your data is an enterprise asset. Maximize its potential in every aspect of your business, and every AI project, with the AI-augmented low-code process automation, data integration, and API management power of Tray.ai. With Tray.ai, it's easy to connect and flow your Snowflake, Databricks, BigQuery, or Redshift data into your marketing stack to supercharge your lead lifecycle processes, create smarter order-to-cash and finance processes, or simply better integrate with your CRM, ERP and HCM data to speed analytic pipelines.

Tray.ai is the fastest way to turn your data assets into an AI powerhouse by instantly connecting them with any large language model and your stack to deliver agents, chatbots, AI-infused processes, and more.

To start trying out integrating apps and business processes to your data warehouse, check out a Tray **trial**, or get a personalized **demo**.

tray.ai

Click below to visit:

Website | Blog