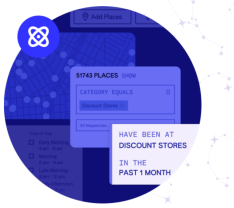
ASTRONOMER

FOURSQUARE

Foursquare Centralizes Orchestration of 9,000+ Data Assets with Astro

Discover how Foursquare reduced data access time 90%, centralized 9,300+ assets, and sped up pipeline development 5× using Astronomer's Astro platform.



9.3K
Data assets orchestrated

5x
Faster pipeline development

90%
Reduction in data discovery and access time

Geospatial Data: Pervasive and Complex

Foursquare delivers value to businesses by harnessing the power of location. Every day, it processes billions of records—including GPS pings, ad impressions, and inputs from first- and third-party sources—which are ingested and modeled to power both internal workflows and customer-facing offerings.

Because of this geospatial focus, Foursquare's tech stack demands specialized processing and storage solutions capable of handling high volumes of location-based data at scale. The platform supports a wide range of use cases, including ETL, model training, real-time dashboards, and disaster recovery—each requiring reliable and flexible orchestration.

Reaching a Crossroads: Unifying Data Under One Control Plane

In 2023, Foursquare's engineering and data teams hit a critical inflection point. A daily firehose of data was being routed through a fragmented landscape of tools, teams, and homegrown systems. CTO Vikram Gundeti recognized the urgent need for a centralized control plane—one that would unify visibility, standardize operations, and bring cohesion to Foursquare's complex data ecosystem.

In a blog post chronicling this journey, Gundeti wrote:

> While we made incremental improvements to these component systems over the years, we lacked central visibility that made it difficult to answer questions like: What are all the datasets we have within our data lake? And which ones are production versus staging? Who has access to each of those data sets? Which data sets does a specific customer-facing data product depend on?
>
> **Vikram Gundeti**
> CTO

This lack of visibility had wide-ranging implications. It hindered productivity, slowed release cycles, and created redundant assets that risked driving up costs.

Strategic Tooling: Data Pipeline Orchestration

To build the centralized data control plane, Foursquare had to make key decisions about tooling—especially around data pipeline orchestration. Historically, teams comprising around 50 engineers relied on a mix of self-hosted OSS Airflow and internal systems built around Luigi to manage a sprawling network of pipelines. Managing these separate systems introduced friction and wasted engineering time, and the systems lacked the support and scalability required to keep pace with evolving needs.

> "We realized that Luigi did not provide the necessary level of support that would be required for long-term use. And with self-hosted Airflow, we did not see a path forward for scalability. We were running into moments where contributors were worried about stomping on each other's work. In the local development environment, when someone would make a change, it would take five to ten minutes to run, and could create issues if workstreams had dependencies that spanned different teams."
>
> **Rob Joseph**
> VP OF INFRASTRUCTURE

Given the time-sensitive nature of its data products, Foursquare also recognized the importance of ensuring maximum uptime and operational resilience.

Why Foursquare Chose Astro to Scale Airflow and Empower Engineers

Foursquare turned to Astronomer's Astro platform to simplify operations, improve uptime, and free up engineering bandwidth. "We selected Astronomer's Astro platform to improve the local development experience and eliminate the burden of having to host Airflow ourselves," said Gundeti.

> "We selected Astronomer's Astro platform to improve the local development experience and eliminate the burden of having to host Airflow ourselves. Astronomer has unlocked a lot of efficiency and productivity."
>
> **Vikram Gundeti**
> CTO

With Astro, Foursquare implemented a robust internal alerting system that delivers real-time insights based on DAG completion and task failures. The team also introduced a "Canary DAG," dedicated solely to validating the uptime of the Airflow scheduler. When issues arise, teams at both Foursquare and Astronomer are notified instantly—ensuring rapid response and uninterrupted service. "Astronomer has unlocked a lot of efficiency and productivity," reports Gundeti.

Results: Visibility, Velocity, and Developer Satisfaction

Foursquare successfully centralized its data operations under a unified control plane, unlocking enhanced visibility and governance across a catalog of more than 9,000 data assets. What once took days to discover or access can now be done in minutes. With operational complexity reduced and productivity tools improved, developer satisfaction has surged—and the team is free to focus on innovation, optimization, and future growth.

FOURSQUARE

ASTRO

North America
Technology
Enterprise

KEY OUTCOMES

9.3K
Data assets orchestrated

5x
Faster pipeline development

90%
Reduction in data discovery and access time

Learn What Astronomer Can Do For You

Foursquare boosted developer productivity and data visibility at scale. See how they did it—and talk to us about achieving similar results with Astronomer.

Book a Demo