

DELHIVERY

CASE STUDY

How Delhivery handles
1 terabyte of logistics
data every day with
Atlan





Atlan made the most complex data concept, which is data governance, look simple and easy to use.

Because of its intuitive UI, **Delhivery users love to return to Atlan for any issues they face with metadata.**

LOKENDAR REDDY KOYA,
SENIOR DATA ANALYST





Our vision is to become the **operating system for commerce in India**, through a combination of world-class infrastructure and cutting-edge engineering and technology.



AKASH DEEP VERMA,
Director of Data Engineering

About Delhivery

Delhivery is India's largest and fastest-growing fully integrated logistics player. The company has fulfilled over 1 billion orders across India, with a nationwide network that services more than 90% of India.

In 2019, Delhivery was awarded the ET Startup of the Year award for connecting remote corners of India.





Delhivery's massive growth was creating **chaos** for its data users. Its data stack was missing a crucial solution to manage the flood of incoming data.



Delhivery has been growing exponentially and we are generating nearly 1 TB of logistics data per day. **It's getting very difficult to answer questions about our data.**



AKASH DEEP VERMA,
Director of Data Engineering



AWS
Cloud



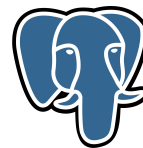
AWS: S3
Data lake



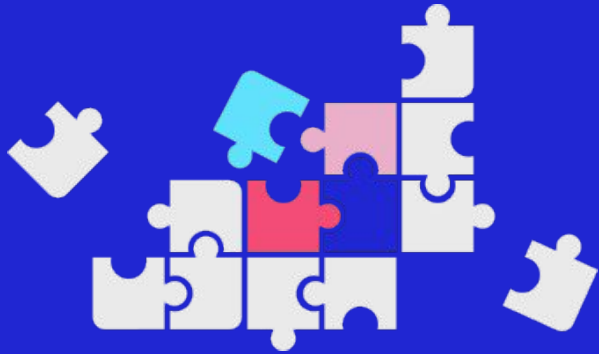
Metadata management &
data governance



Tableau
Business intelligence



PostgreSQL
Database



The Problem

**Delhivery's complex
web of supply
chain data**

Delhivery's microservices architecture created **data silos across verticals** and **bottlenecks for developers**.

LENGTHY ONBOARDING

New hires took too long to understand Delhivery's data.

"The typical onboarding time was **1-2 months** for new employees to learn about Delhivery. **But it kept taking longer**, around 3-4 months, as we kept expanding."

DIFFICULT DATA DISCOVERY

Delhivery's many verticals made data difficult to find.

"Our data users needed a discovery tool... a Google-like interface to **effectively discover data assets and build trust** from there."

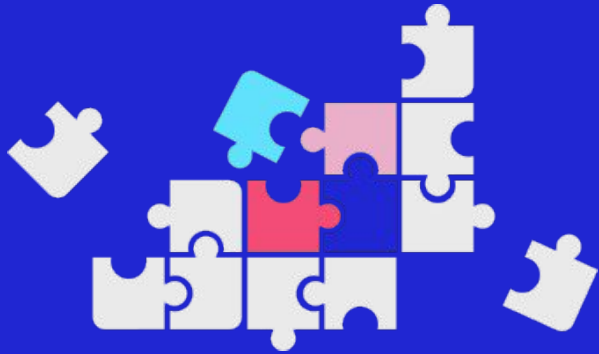
SOLVING FOR DIVERSE USERS

400 people across product and tech relied on data.

"We want to solve for **product managers** so they can get a feel of our data, for **data analysts** to understand how data is structured, and for **developers** so they can understand its lineage."



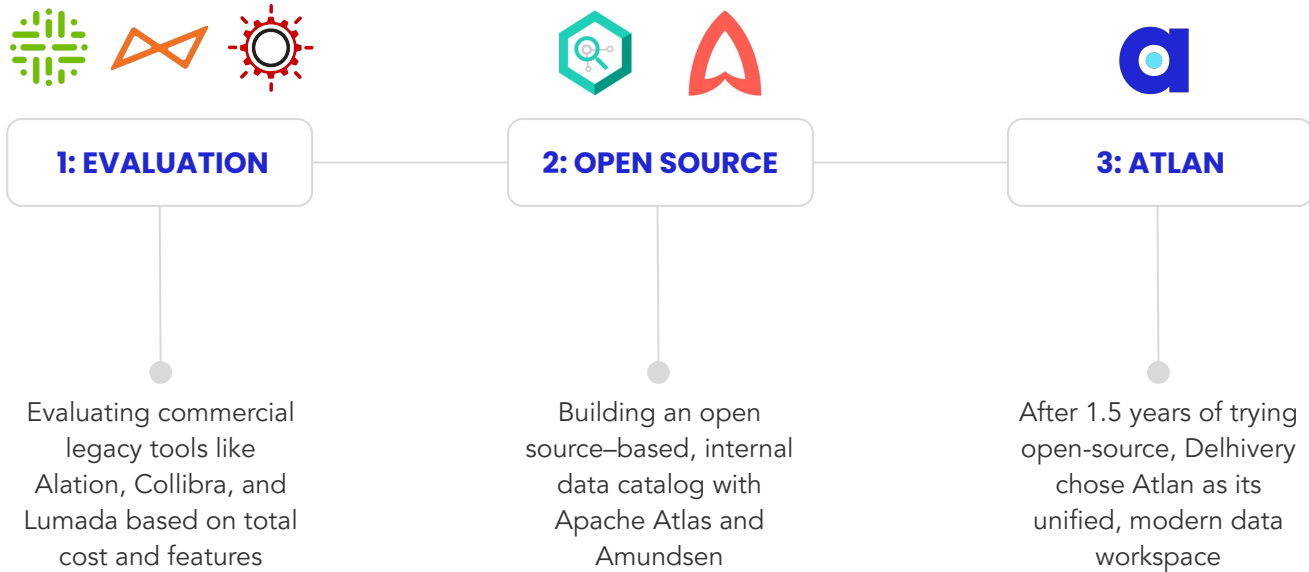
AKASH DEEP VERMA
Director, Data Engineering



The Evaluation

**Delhivery's search
for a modern data
discovery solution**

Delhivery spent **18 months** trying to build an internal catalog on open-source tools. Eventually, it realized that **Atlan would deliver a better solution at a lower cost.**



BUILD OR BUY?

Delhivery began their search a couple years ago by vetting legacy catalogs, but none had the right features at the right price.



ALATION



COLLIBRA



LUMADA



Buying one of these products would have been the simplest fix, but we couldn't find the right solution.

Each one was either **missing non-negotiable features** (such as seeing a data preview or querying data) or the **total cost of ownership was just too high** for us (due to expensive set-up, licensing and professional service fees).



AKASH DEEP VERMA,
Director of Data Engineering

BUILD OR BUY?

Delhivery asked **2 developers** to spend **7 months** building an open-source, internal data catalog on Apache Atlas. They quickly learned that it was too technical for its diverse data users.



APACHE ATLAS



As we started implementing Apache Atlas, we realised we had a problem. It's great for developers, but it has a steep learning curve for non-technical team members.

Its interface and language was highly technical, which made it **difficult to onboard product managers or operations people** who needed to understand our data.



AKASH DEEP VERMA,
Director of Data Engineering

BUILD OR BUY?

To solve the issues with Atlas, Delhivery brought in Amundsen, spent months developing missing features, and finally rolled out its internal data catalog. The problem was, no one wanted to use it.



AMUNDSEN

Delhivery's internal Atlas + Amundsen data catalog

The screenshot displays the Amundsen data catalog interface. At the top, there is a dark blue header with the text "DELHIVERY DATA CATALOG" on the left, a search bar containing "exp", and navigation links for "Announcements" and "Browse" on the right. Below the header, the breadcrumb path is "express_dwh.waybill_weight_cdc_parquet" with a star icon, and sub-breadcrumbs "Datasets • hive_table • glue". Action buttons for "SuperSet", "Loading...", and "Download" are visible. The main content area is split into two columns. The left column contains metadata: "Description" (empty), "Date Range" (From: Jul 04, 2019, To: Nov 12, 2020), "Frequent Users" (two profile icons), "Tags" (cps), and "Owners" (data_governance). A note at the bottom of this column reads "Read-Only information, Auto-Generated." The right column shows a list of fields: "u (username)" (string), "unknown_fields" (string), "iwt" (double), and "predicted_vol (pred_internal_vol)" (double). Each field entry includes a "Tags" section with a "+ New" button.



When we asked people what was missing, they all pointed to the user experience. The catalog didn't integrate into their daily work.

Our solution couldn't win without creating an experience that people wanted to use every day.

AKASH DEEP VERMA,
Director of Data Engineering



Delhivery needed an **intuitive, modern data catalog** that people with varying technical skills could use to understand business context.



Meet **atlan**

The modern data workspace

Why Delhivery scrapped its internal catalog after 1.5 years of work and **chose Atlan as its unified data workspace**

INTUITIVE USER EXPERIENCE

“Atlan not only had all the dream features we were looking for, but the user interface and experience were extremely intuitive. Built around a Google-style search, Atlan is designed to be **user-friendly for technical and non-technical users alike.**”

OPEN BY DEFAULT

“We liked that Atlan has a fundamentally open philosophy... This gave us the ability to customise the product for our end users — bringing together **the best of open-source** with a **polished, off-the-shelf user experience.**”

EXPERT PARTNERS, NOT VENDORS

“We also decided to go with Atlan because we realised that we needed **a partner who was an expert in data discovery.** This would let us rely on their expertise as we rolled out Atlan... We had plenty of our own business problems to tackle.”



AKASH DEEP VERMA
Director, Data Engineering

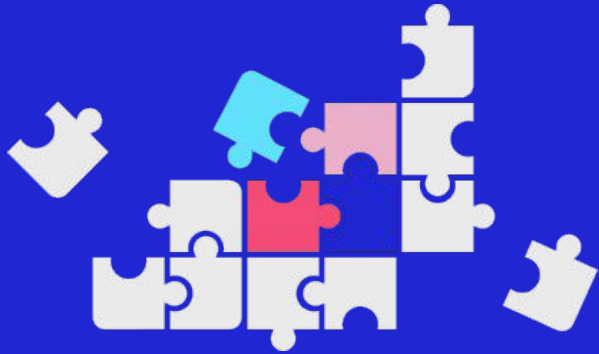


It would take Delhivery **six or seven people and up to two years** to build what Atlan gave us out of the box.

We needed a solution on day zero, not in a year or two.

AKASH DEEP VERMA,
Director of Data Engineering





The Process

**Implementation
through
gamification**

IMPLEMENTING ATLAN

Delhivery used this implementation drive to kickstart a better culture around data documentation and governance.



We need to identify who we define as owners and **make documentation fun** for them. Success will depend on motivation.



RAJIV DINESH,
Head, Data Products

Problem: Identifying assets & owners

Delhivery's multiple verticals meant they had many different assets to catalog, with **each team owning different assets informally**.

They needed to prioritize the most important assets, and incentivize people to take ownership of those data assets.

Solution: Auto-assigning assets based on daily usage

Assets were prioritized based on how frequently they were used. Then, **using query log history and business metadata**, each asset was assigned to the person - a contributor - who queried it most frequently. They could vote to stay as the contributor, add co-contributors, or reassign the asset.



Atlan made cataloging fun at Delhivery by organizing a **4-week long gamification drive** filled with memes, messages, and prizes.



THE RULES

- The team that verifies the most tables wins!
- A table is considered verified if 80% of columns are documented and descriptions have been added.
- Scores are normalized by dividing the total points by the number of people on the team.



THE SCORING SYSTEM

- 1,000 points** per verified table
- 100 points** per table description
- 10 points** per column description



BONUS POINTS

- 10 points** per table classification
- 10 points** per linked glossary



Messages, memes, and prizes: the method to motivation



Nandini Tyagi 7 Dec, 18:07 • Edited

ANNOUNCEMENT - BEST DATA ROCKSTAR TEAMS

Hi team,

Massive congratulations to every single one of you for an **ABSOLUTELY BRILLIANT** contest. Before the contest, we had 60 verified tables and due to this amazing performance by you folks, as of today, we have a **whooping 272 verified tables, this covers all the major priority tables that we use at Delhivery**. This is an **incredible achievement** and will **empower everyone at Delhivery** in their data journey. 🌟❤️

🎉 Now, time to announce the winners of the Atlan-Delhivery contest. 🎉

🏆 **The first prize** goes to

Team Reporting – @Anil Kumar @Masoom Yadav @Pankaj Sondhi @Prakash Choudhary @Sarakshi Jain @Shiv Brat Pandey

This team single handedly verified 100+ tables! Brilliant job team 🌟👏👏

🏆 **The second prize** goes to

Team Orion - @Saurabh Dhillon @Ankit Rai @Jagdeep Yadav

One of the smallest teams with a very high number of priority tables, Orion came in smashing last week and did a great work! Well done! 🌟👏

🏆 **The third prize** goes to

Team LTL - @Rohan Raman @Ali Nawaz Haider Rizvi @Pasupuleti Venkata Shiva @Jatin Malhotra

One of the most consistent teams who is also the **ONLY** team to have verified all the tables they claimed. Fantastic performance by an equally fantastic team 🌟👏

We also want to give a special mention to:

Team Infacility - @Sumit Uppal, b.manojkumar@delhivery.com, tanmay.ss@delhivery.com, chandraprakash.tailor@delhivery.com

Team FMS - @Rakesh Mahuli, sai.krishna@delhivery.com, arshad.ansari@delhivery.com, subhash.vadlamani@delhivery.com, shivam.thakur@delhivery.com, who are also at ~100% completion and made an amazing effort 🌟

Jump to bottom



Atlan announcing the winners of the gamification drive



Nandini Tyagi 7 Dec, 18:28

Also, a special mention to @Koya Lokendar Reddy who has been working very hard last few weeks to prepare for the contest! I am sure I say it on behalf of everyone - great job coordinating the contest Loki 🌟👏

👍 2 🌟 3 🙏 1



Anil Kumar 7 Dec, 18:35

yeah he was reachable 24*7 and able to fix and provide resolution quick ,thanks for effort @Koya Lokendar Reddy you put in team all teams together

❤️ 1 🌟 1



Recognizing the person behind this drive at Delhivery



Koya Lokendar Reddy 8 min

@all

I am very Happy for what our teams have done for last one month. That's something phenomenal 🌟.

Now any newbie in our organization, someone working on any analysis can take full advantage of Atlan portal(<https://atlan.delhivery.com/>) and complete the tasks much more efficiently.

"The ultimate goal is to Democratize entire delhivery data teams. " - this is what I mentioned in my initial thread and I am very much happy that now we are successful at first few steps Democratizing our Data System.

Now everyone in our Organization are equal when it comes to Understanding our Data and its Metadata.

Power is with everyone.

🌟 2



Delhivery's senior data analyst congratulating the team

RESULTS

The gamification drive resulted in a massive culture shift at Delhivery, saving time for data users across the company.



The Atlan team was very active and responsive to the challenges I faced. **The gamification drive got an incredible response!**



LOKENDAR REDDY KOYA,
Senior Data Analyst

250% increase in documentation.

From 60 to 210+ verified tables... and counting!

53% users reported saving time with Atlan.

Of all the users who said they saved time with Atlan, about 70% have been at Delhivery for less than a year.

44% users said Atlan helped them find the best person to talk to about their data.



All data leaders out there must understand the value Atlan adds to an organization in terms of data governance.

It's a tool where you can **collaborate, discuss, organize metadata, and speed up the process of understanding business context.**

LOKENDAR REDDY KOYA,
Senior Data Analyst



Learn how your
data teams can
do their lives' best
work at atlan.com

SEE A DEMO



atlan

 PostgreSQL



 + a b | e a u

aws
