Airbyte

Open source powers the modern data stack

Hello!

I am Michel Tricot

Co-Founder & CEO of Airbyte

🐦 @MichelTricot
🐙 michel-tricot
in /in/micheltricot

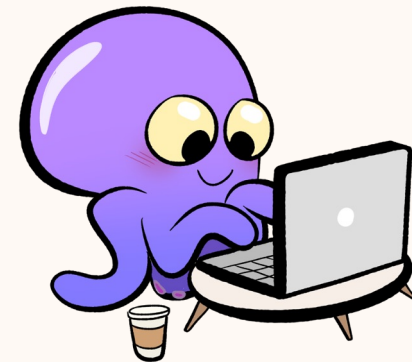## Airbyte

Open-Source Data Integration

15,000 users
6,000 Slack members
6,000 GitHub stars

1

Extract - Load - Transform

A new paradigm for modern data teams

Extract

Source-specific routines to pull selected data from an external system.

Transform

Business logic specific to your organization to serve an analytics or operational use case.

Load

Destination specific routines to push data where it is going to be consumed.

ETL doesn't work in today's world

Inflexible

- Friction when changing an existing pipeline.
- Hard to add new data.
- Most issues force data to be re-extracted.

Lack of Autonomy

- Warehouses made data consumers more autonomous.
- Changes require engineering involvement.

Complex

- Custom DSL.
- Force adoption of a data stack.
- Address 70% of the needs, 30% still built and maintained in-house.

Extract

General-purpose routines to pull selected data from a source.

Load

General-purpose routines to push raw data where it is going to be consumed.

Transform

Business logic specific to your organization to serve an analytics or operational use case with SQL / dbt / ...

ELT fixes the ETL-related issues

Flexibility

- All the data available on the destination.

- Data consumers are free to use what they need for the insights they want.

Autonomy

- Data consumers can leverage SQL queries to transform the data the way they want.
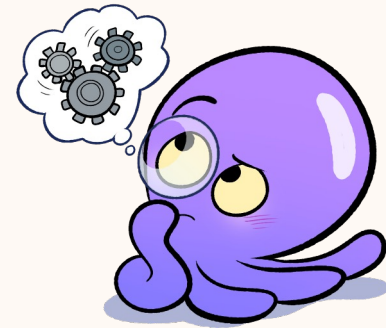
- No need to involve the engineering team.

Future proof

- Issues during transformation don't prevent access to the data.

- Easy to update transformation schemas.
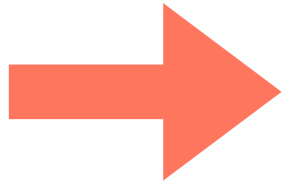
2

The open modern data stack

Composed of the best of breed

The world of data is changing

All companies have custom data needs

- Types
- Skills
- Scale
- Budget
- Security
- Privacy
- Turnaround
- Compliance
- ...

→ **We are moving away from end-to-end solutions**
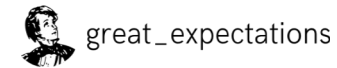
Select the best building blocks

Extract & Load

Airbyte

Fivetran

Data Lakehouse

databricks

Quality

great_expectations

Transformation

dbt

SQL

BI / Visualization

Superset

Looker

Observability

MC MONTE CARLO

Bigeye

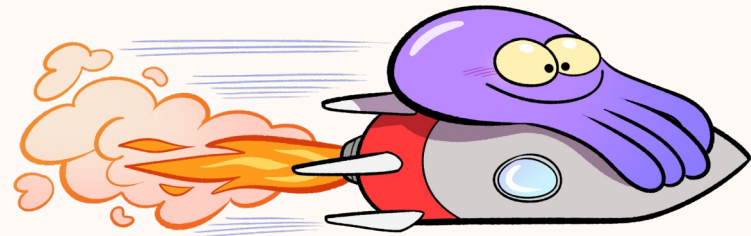**Infrastructure**

docker Azure

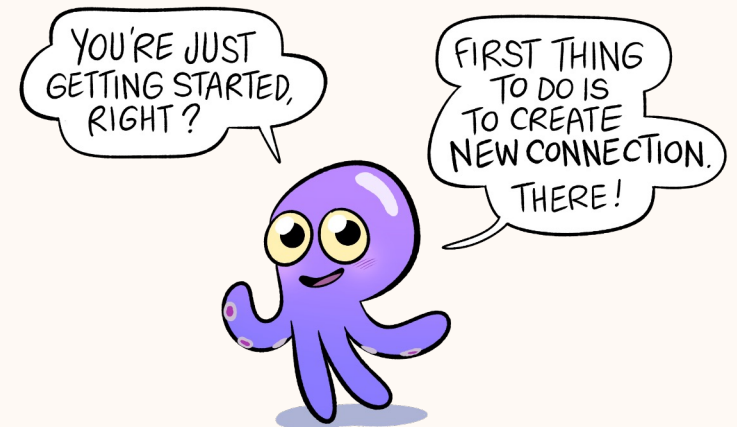amazon web services

3

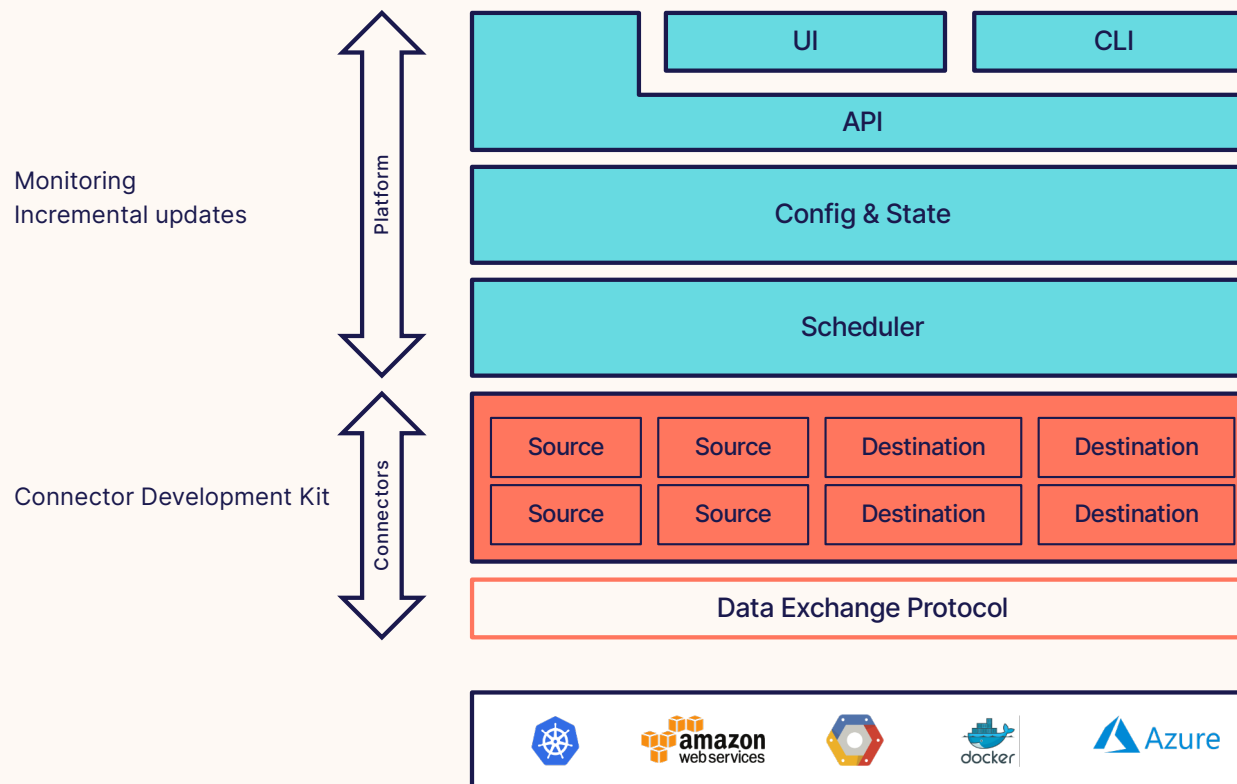# Airbyte-dbt-Superset

To rule them all

Airbyte: Solving Data Integration

- **Open-Source** data integration platform

- **170 high-quality data connectors** (targeting 500 by EOY)

- **Connector Development Kit** to build your own data connector

- Active **Slack community**

- Super fun mascot: **Octavia Squidington III**

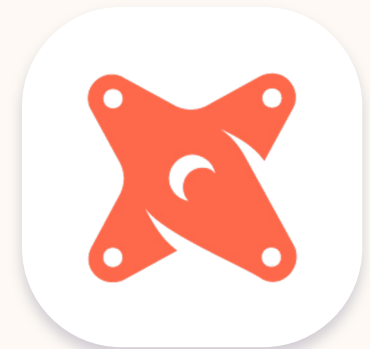# An open data platform for modern teams

dbt: Solving Data Transformation

- **Open-Source** data transformation application
  - **Transform SQL** to tables, views
  - **Cascades** transformations
- Supports most SQL **Databases** & **Warehouses**
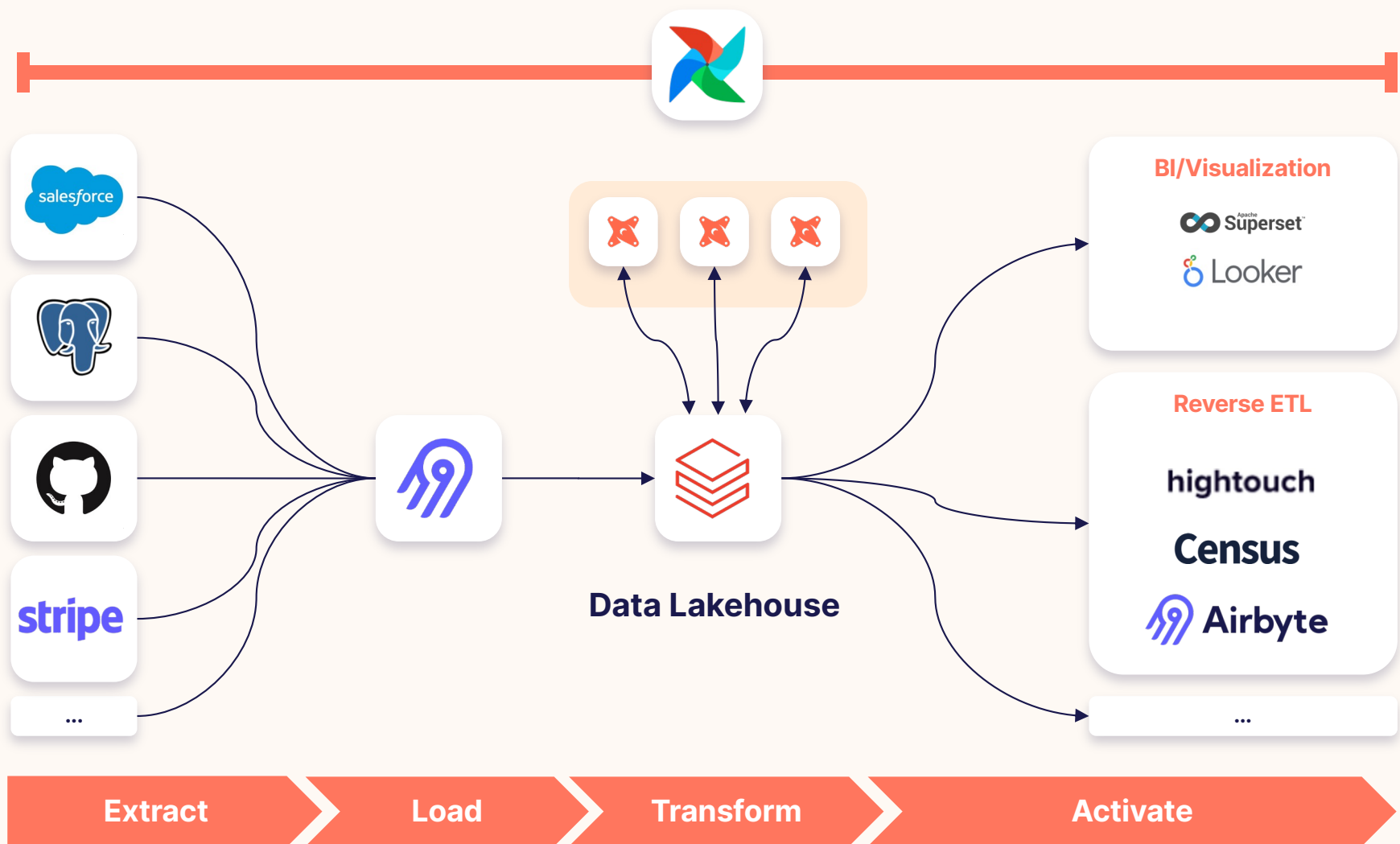- Brings software engineering **best practices** to Analytics
  Engineers
  - Version control
  - Testing
  - Packaging

Superset: Exploration & Visualization

- **Open-Source** data exploration and visualization platform
  - **SQL** to charts
  - **Reusable** queries
- Supports most SQL **Databases** & **Warehouses**

Data Lakehouse

BI/Visualization
Apache Superset
Looker

Reverse ETL
hightouch
Census
Airbyte
...

Extract Load Transform Activate

18

Thanks!

Any questions?

@MichelTricot

slack.airbyte.com (@Michel)

airbytehq/airbyte