Introduction

Introducing Eric

Eric Kriner

Customer Engineer

eric.kriner@soda.io

About Soda

Our Mission

Soda's mission is to **empower** and **incentivize** everyone in your organization to **share** and **use** reliable, high-quality data.

Founded
in 2018

40+ Employees
in Europe and
North America

100+ active
customers around
the world

Practitioners and
experts in data
and engineering
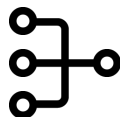
Today's Agenda

SODA

01

Soda's Approach

Enabling data teams to deliver and use trustworthy data

Test data in development

Avoid merging bad-quality data into production.

Test & monitor data in production

Prevent downstream issues by shifting left your data quality tests

Self-serve data testing

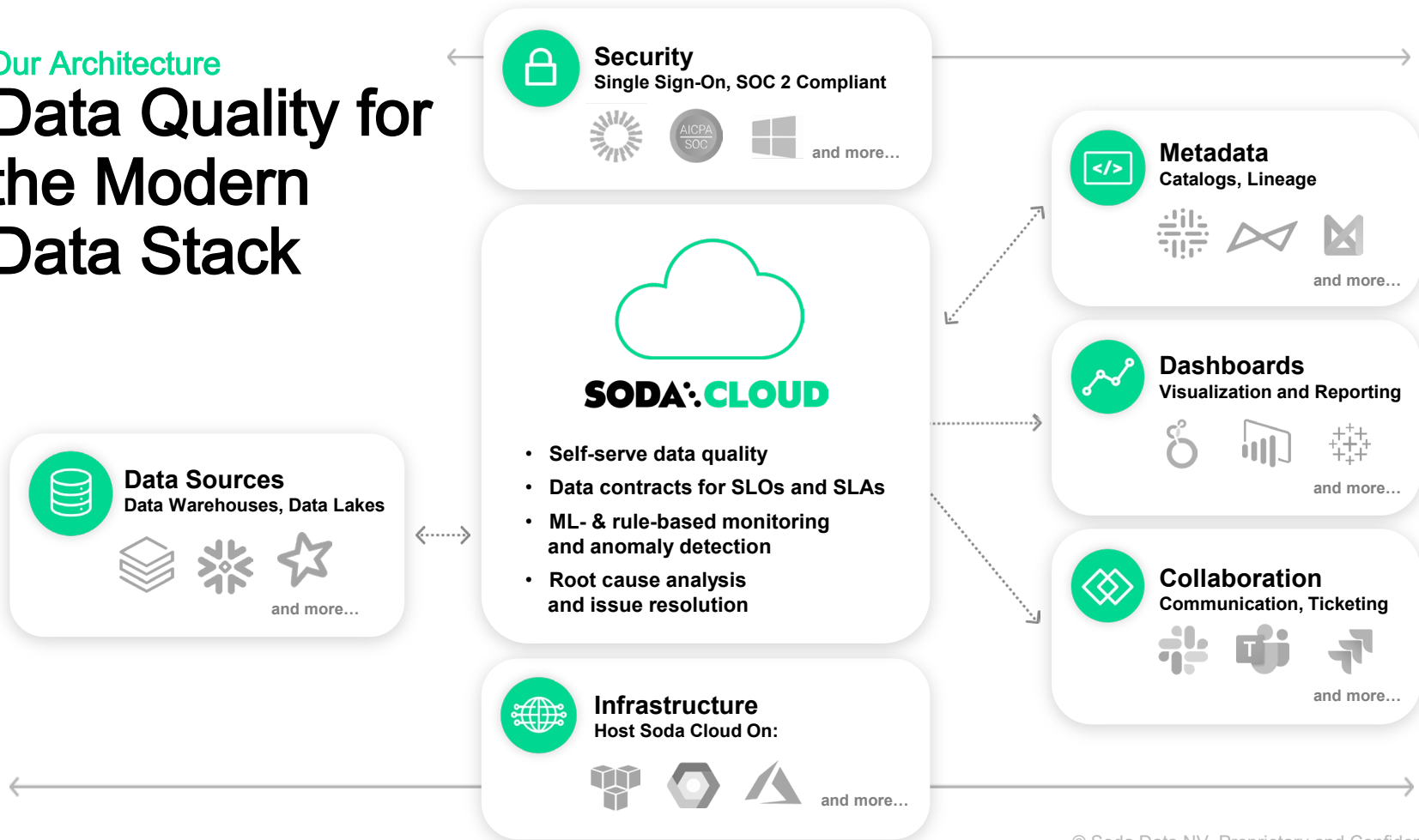Agree and define data quality expectations.

API-first

Leverage our integrations and extensive Python and REST APIs.

Our Architecture

# Data Quality for the Modern Data Stack

Security
Single Sign-On, SOC 2 Compliant
and more…

SODA CLOUD

- Self-serve data quality
- Data contracts for SLOs and SLAs
- ML- & rule-based monitoring and anomaly detection
- Root cause analysis and issue resolution

Metadata
Catalogs, Lineage
and more…

Dashboards
Visualization and Reporting
and more…

Data Sources
Data Warehouses, Data Lakes
and more…

Collaboration
Communication, Ticketing
and more…

Infrastructure
Host Soda Cloud On:
and more…

Built to integrate with your ecosystem

Data Sources

amazon ATHENA	amazon REDSHIFT	APACHE Spark
Azure Synapse Analytics	ClickHouse	databricks
denodo	dremio	DuckDB
Google BigQuery	IBM DB2	Microsoft SQL Server
MySQL	ORACLE	PostgreSQL
trino	snowflake	

Orchestration

- Airflow
- dagster
- dbt
- PREFECT

Messaging

- Microsoft Teams
- slack

Ticketing

- Jira Software
- Opsgenie
- PagerDuty
- servicenow

Dashboards

- Looker
- Power BI
- tableau

Metadata

- Alation
- atlan
- Collibra
- metaphor

Infrastructure

- amazon web services
- Google Cloud Platform
- Microsoft Azure

Security

- Azure Active Directory
- okta

SODA:

SODA

02

Soda in Action:
Product Showcase

Data Manager

Notify

Get notified about failing records related to me

Resolve

Fix issues directly in the operational system

1

Connect

Start capturing critical operational data

2

Check

Define checks to verify the quality of the data

3

Assign

Assigning of failing records to the relevant people

4

5

Report

Track progress of your data quality program

Data Steward

Data Consumer

Agree

Align on expectations using data sharing agreements

Analyze

Take ownership of alerts

Investigate next step

Report incident

Report

Update consumers

Update executives

1 **2** **3** **4** **5**

Prioritize

Root cause analysis

Notify stakeholders

Prioritize incident

Resolve

Distribute tasks

Fix data

Fix code

Prevent

Circuit breakers

Business review flows

Data Producer

Test data quality in a data pipeline

Airflow Pipeline

Ingest raw datasets

Build dbt model to transform

fact_product_category.sql

Build dbt model to transform for report

report_category_sales.sql

report_subcategory_sales.sql

Export to reporting or visualization tool

AdventureWorks
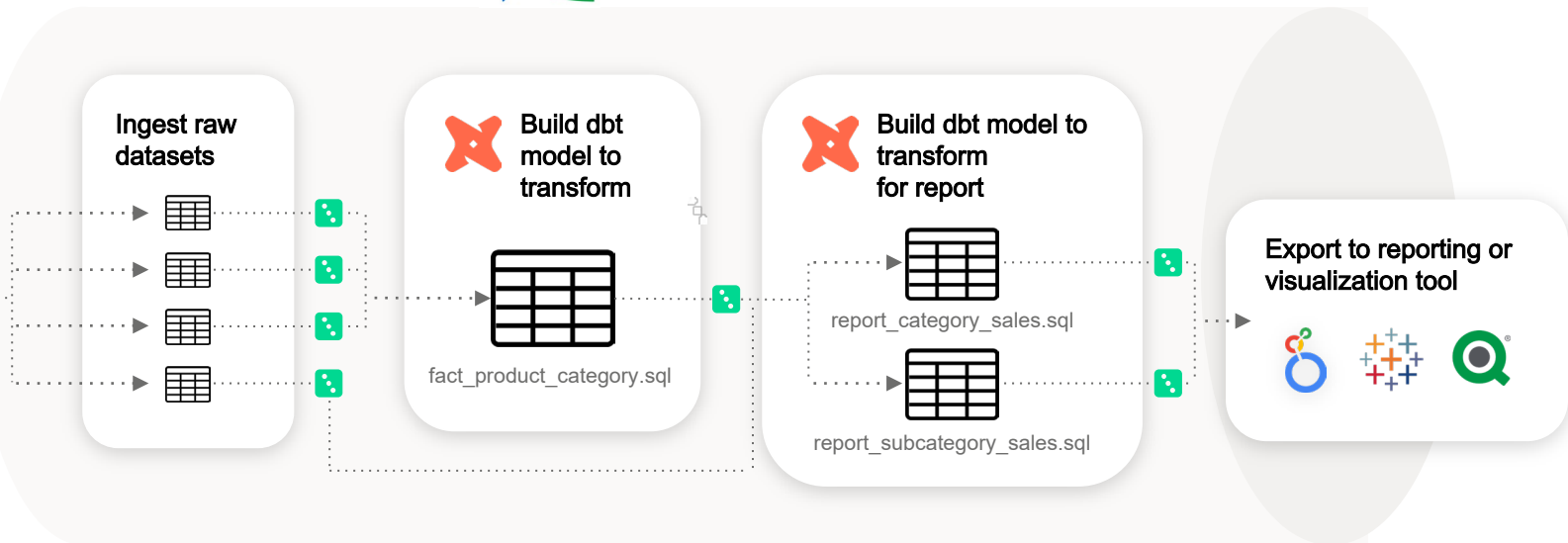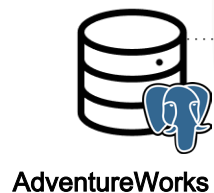
Soda data quality checks

SODA

Test data quality in development

soda-actions [bot] commented 2 weeks ago

🔴 Soda scan completed with the following results:

State	Number of checks
Not evaluated	0
Pass	3
Warn	0
Fail	1

View the full scan results

New branch + new code

New pull request

Automated build + test

Soda data quality tests

✕

`</>`

Fix + commit

Automated build + test

Soda data quality tests

✓

Review

Publish/ Merge to production

Continuous Integration

Continuous Delivery

Use-case:
Trustworthy Data Products
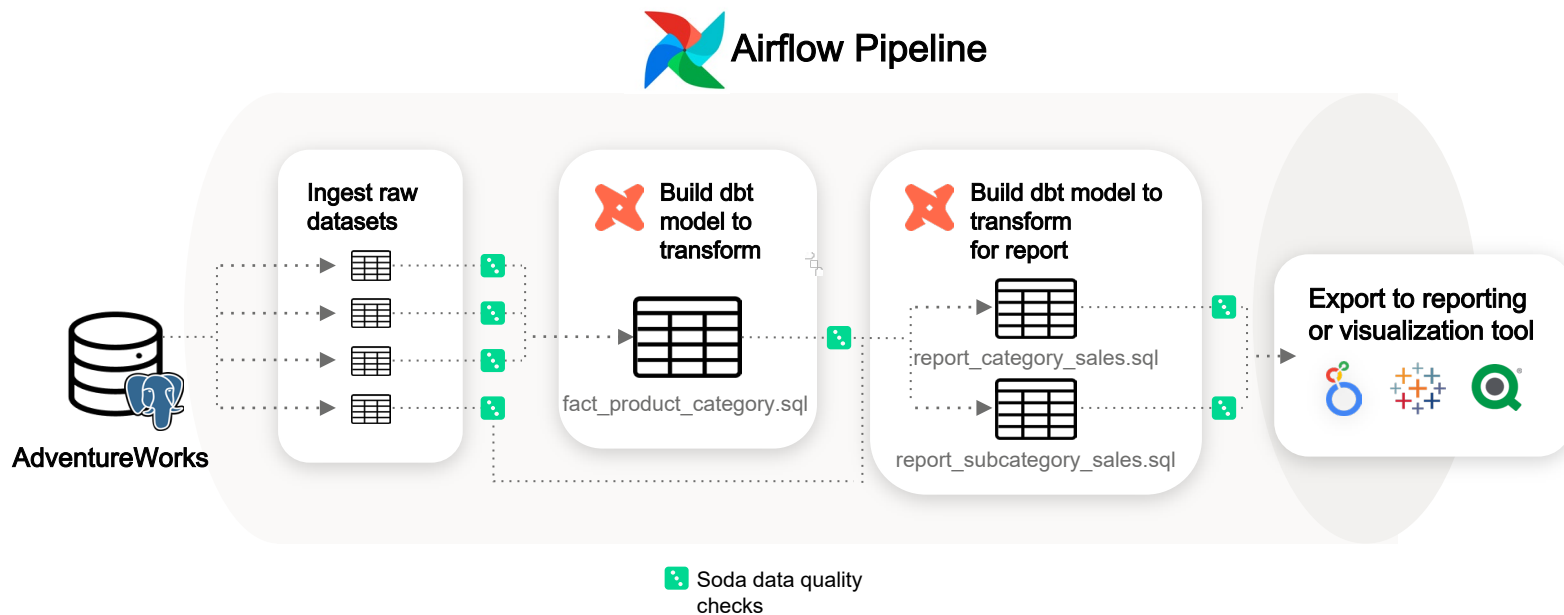
Use-case:
Successful Data Migrations

Ingest raw
data

Push to
staging

Push to
production

Validate
schema

Compare
to source

On-Premise

Data Migration

Cloud
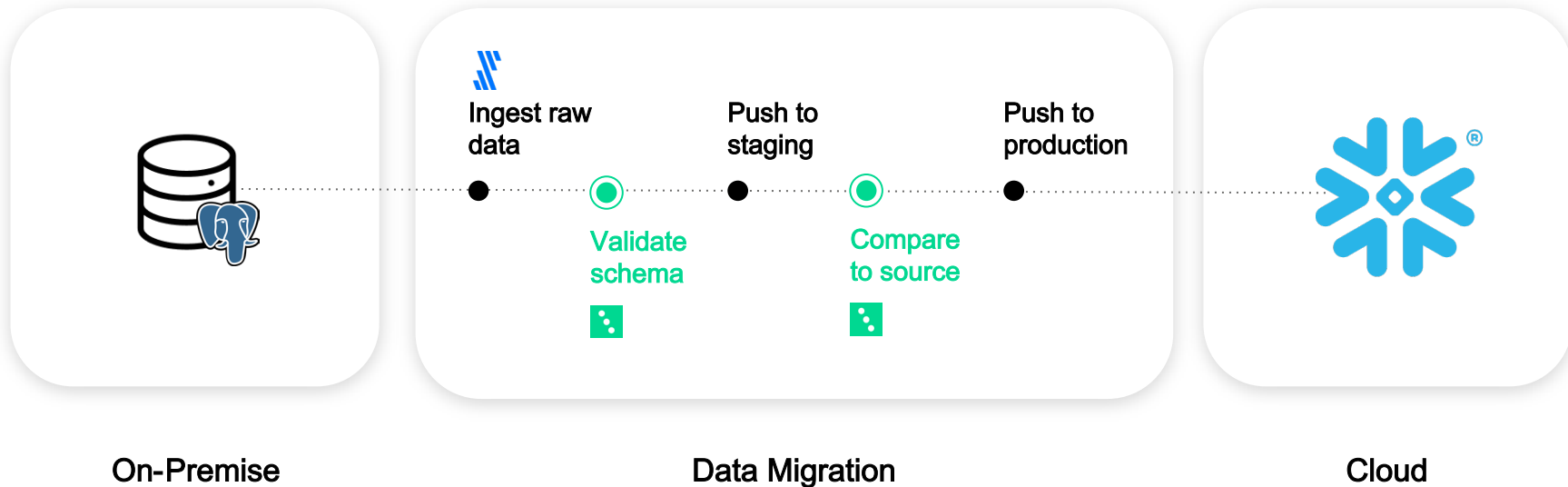
SODA

05

Getting Started with Soda

Help

Join the Community
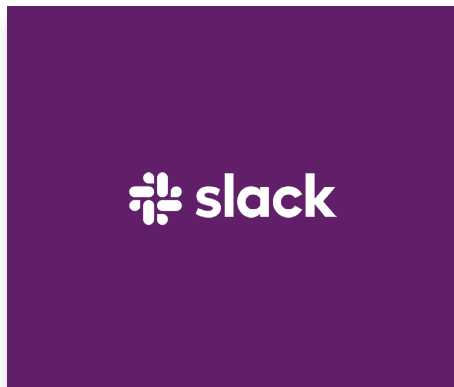
slack

community.soda.io/slack

Free 45-Day Trial

Your Data in Soda
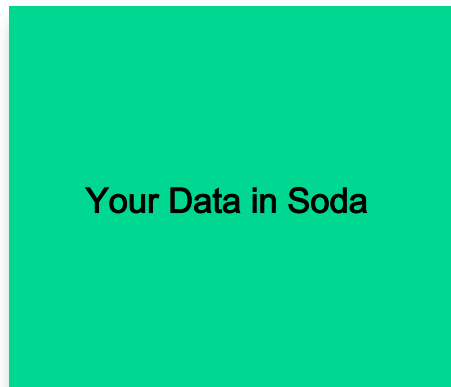
soda.io/

Learn more in our Help Center

SODA. | Help Center

Home
Get started
Connect a data source
Soda Agent
Soda Cloud
SodaCL
Soda Core
Integrate Soda
Reference
API Documentation
Release notes

Introducing self-serve data quality management. Serve yourself!

Hi, how can
we help you?

Search Soda documentation
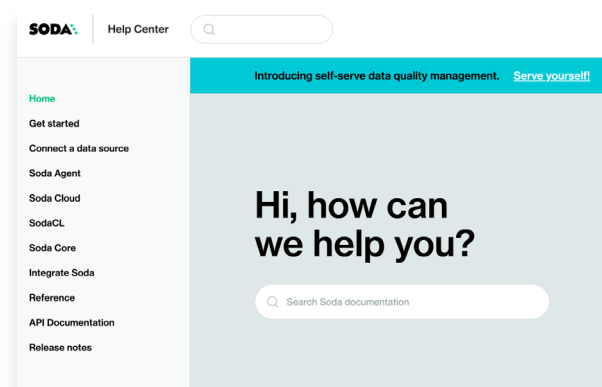
docs.soda.io

SODA

Q&A

Thank you.

SODA