

Big Data Pipeline Architecture

ml4devs.com/big-data-pipeline 

Collect

data from internal & external sources

Ingest

data through batch jobs or streams

Store

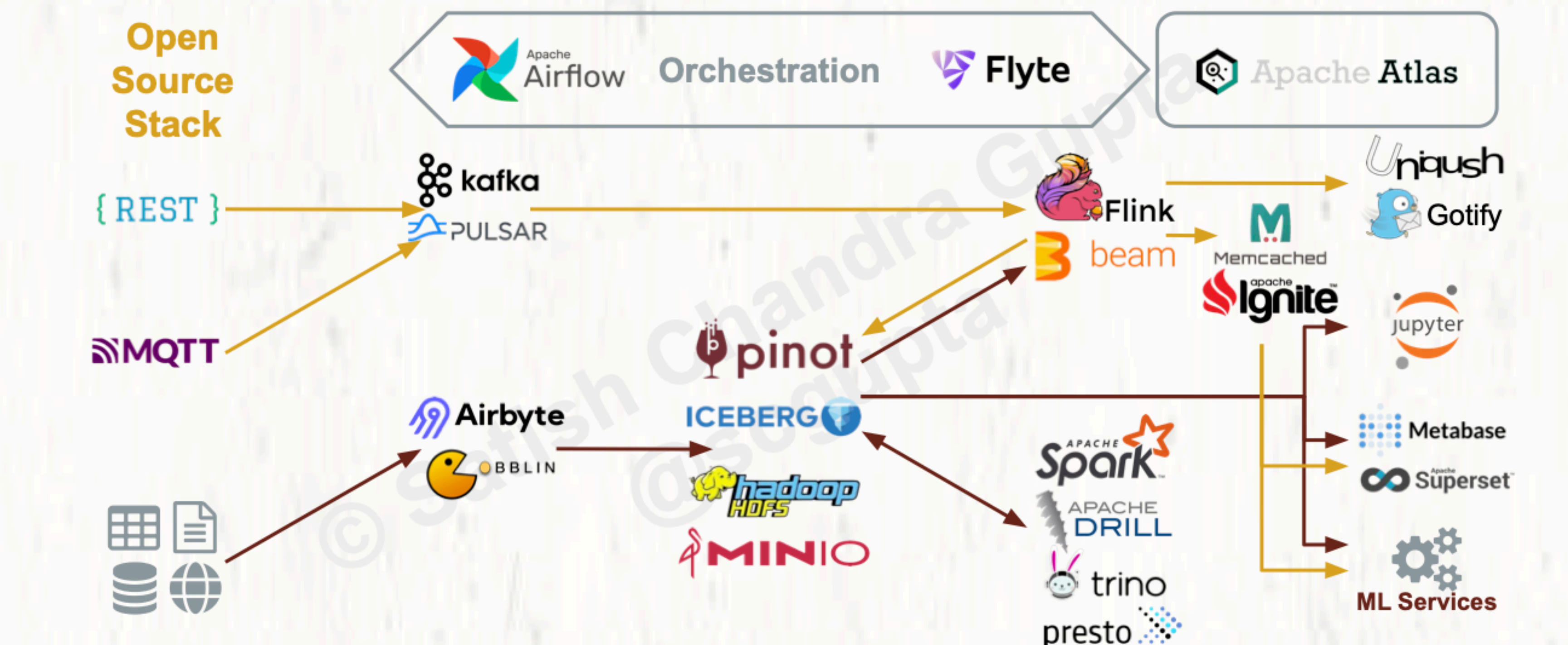
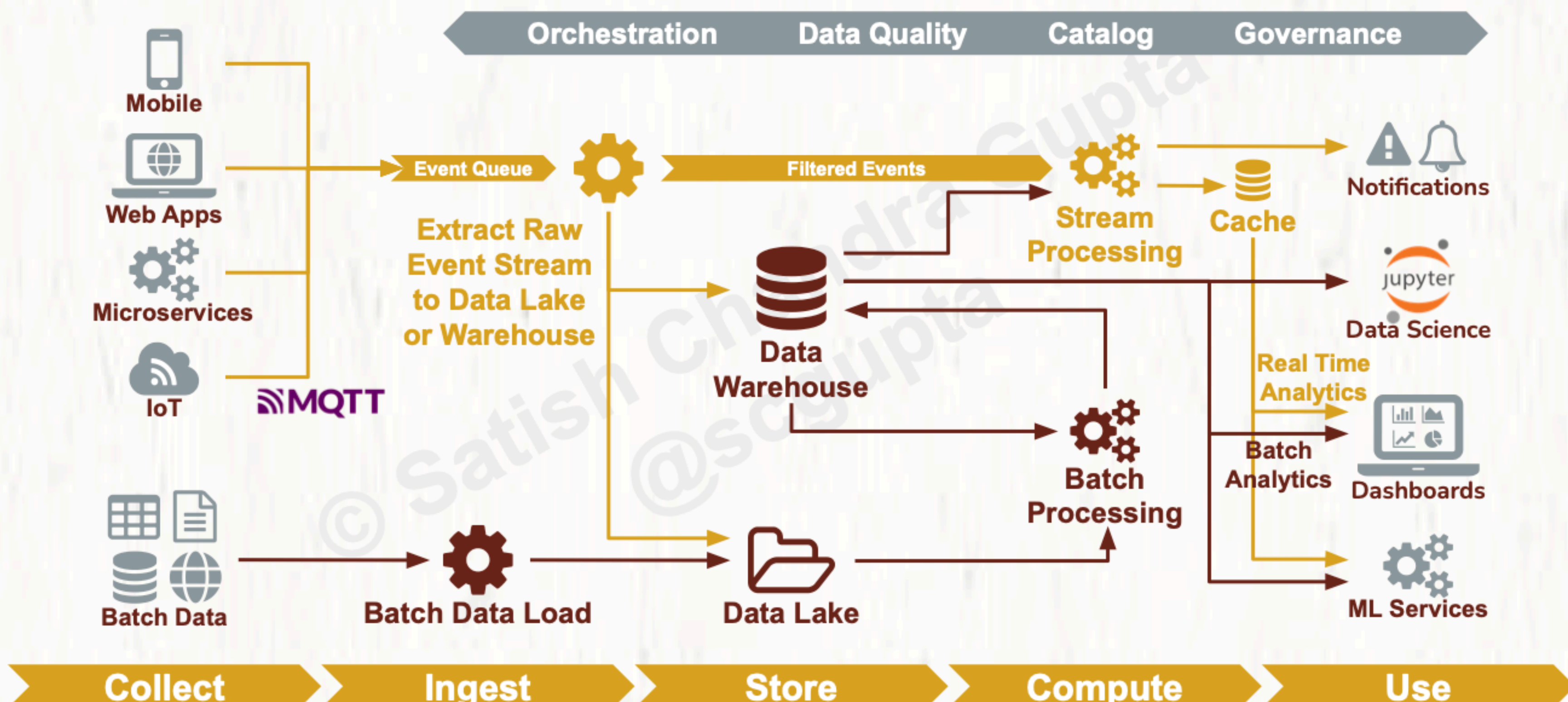
in Data Lake or Data Warehouse

Compute

analytics aggregations, ML features

Use

it in dashboards, data science, ML



Big Data Pipelines on AWS, Azure, and Google Cloud

ml4devs.com/big-data-pipeline

