🧙‍♀️ Mage & 🏎️ DTC Zoomcamp

💫 Making Orchestration Magical with Mage

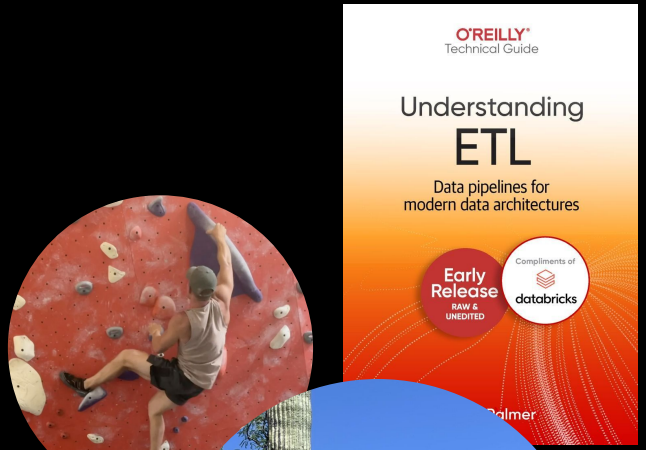
🤦 A bit about me...

↩️ *Previously*

- Analytics engineering — **Storyblocks** 📹
- Product analytics — **AllTrails** 🥾
- Data engineering — **Underline** 💻

🔄 *Now*

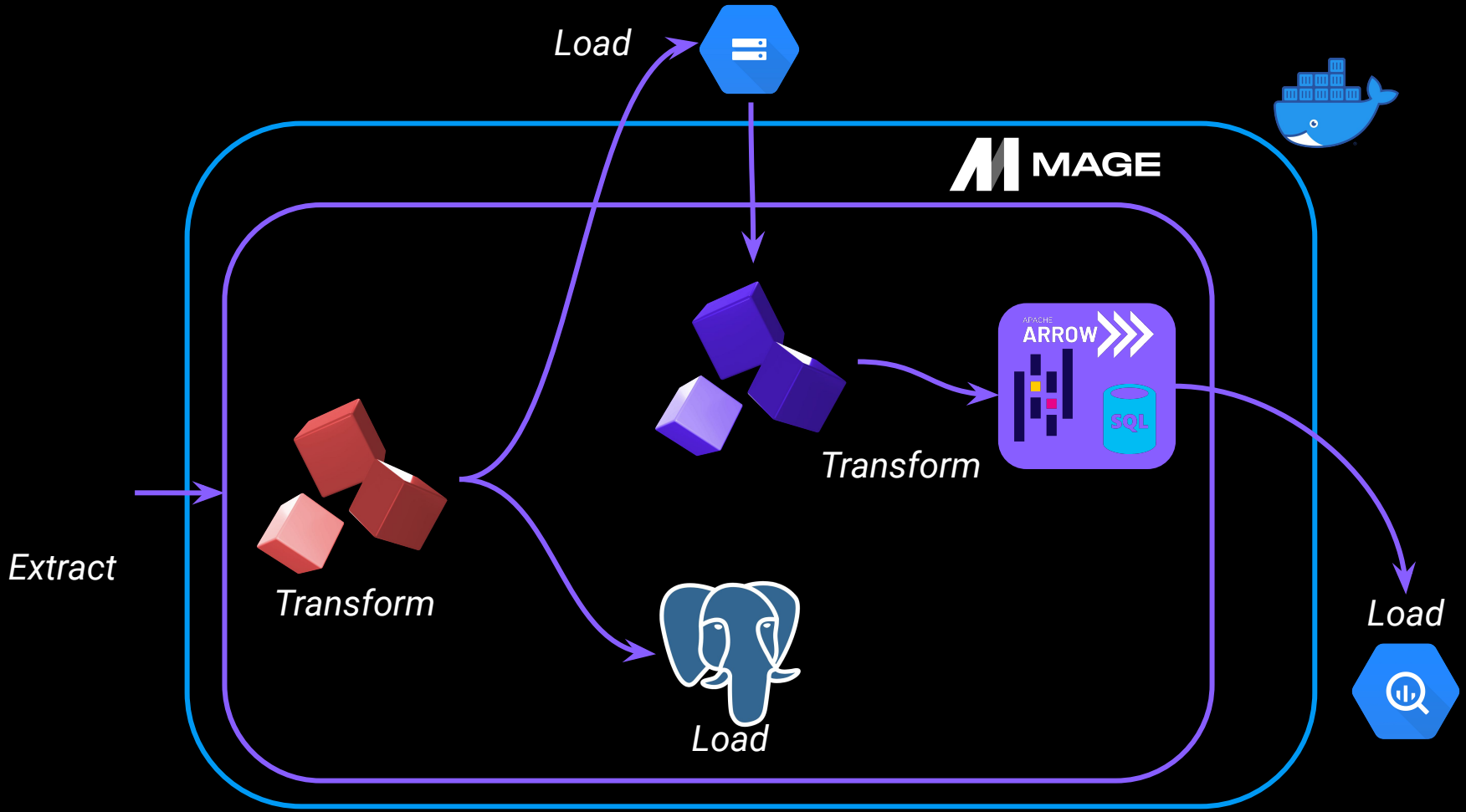
- DevRel (& other stuff) — **Mage** 🧙



🤝 Connect on LinkedIn

🗞️ Check out my Newsletter

What will we build?


Load
MAGE
Extract
Transform
Transform
Load
Load
APACHE
ARROW
SQL

Architecture

Extract

Pull data from a source (API— NYC taxi dataset)

Transform

Data cleaning, transformation, and partitioning.

Load

API to Mage, Mage to Postgres, GCS, BigQuery.

What is orchestration?

What is Orchestration?

A large part of data engineering is **extracting, transforming, and loading** data between sources.

Orchestration is a process of dependency management, facilitated through **automation**.

The data orchestrator manages scheduling, triggering, monitoring, and even resource allocation.

What is Orchestration?

☝️ Every workflow requires sequential steps.

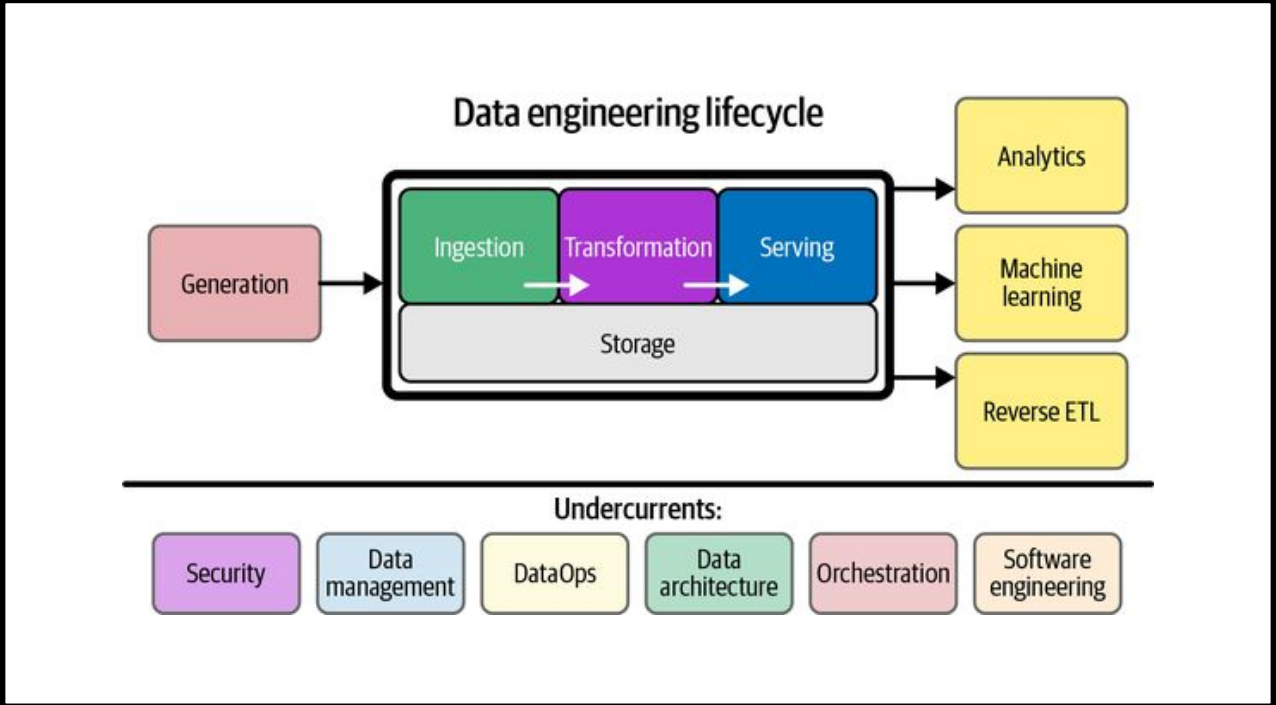
☕ A French press with cold water will only brew disappointment

☔ Poorly sequenced transformations brew a storm far more bitter.

📕 Steps □ tasks

🔄 Workflows □ DAGs (directed acyclic graphs)

What is Orchestration?

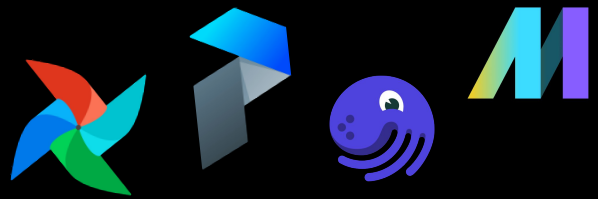


Source

What does a *good* solution look like?

A good orchestrator handles…

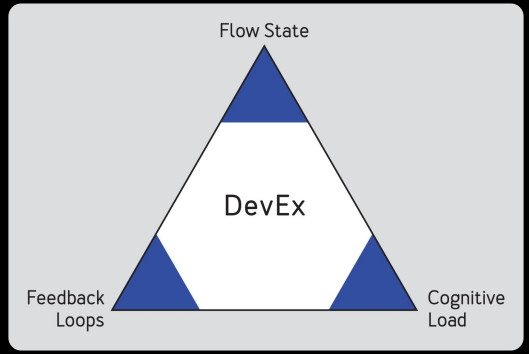
- Workflow management
- Automation
- Error handling
- Recovery
- Monitoring, alerting
- Resource optimization
- Observability
- Debugging
- Compliance/Auditing



A good orchestrator prioritizes….

The *developer experience*

- Flow state 🌊
  - *“I need to switch between 7 tools/services.”*
- Feedback Loops 🔁
  - *“I spent 5 hours locally testing this DAG.”*
- Cognitive Load 🧱
  - How much do you need to know to do your job?



An analogy

An orchestrator is like…

Up next: What is Mage?