Requirements Gathering and Architecture

Robert Clements

MSDS Program

University of San Francisco

ML Systems Gone Wrong

Cruise says it's not at fault for woman injured in San Francisco crash

By Zak Sos , Allie Rasmus and Andre Torrez | **Updated** October 3, 2023 8:10am PDT | San Francisco | KTVU FOX 2 |

BUSINESS

California allows robo-taxis to expand and emergency responders aren't happy

AUGUST 10, 2023 · 4:44 PM ET

Dara Kerr

NEWS | BAY AREA & STATE

Waymo driverless car set on fire, destroyed by San Francisco crowd

By **Katie Dowd**
Feb 11, 2024

Who do you get requirements from?

- ***All stakeholders***:
  - Customers
  - Project owners/leadership
  - Operations teams
  - Regulators
  - Affected
- Project owners will change their minds often, and will be vague
- Don’t forget safety, responsibility, ethics (hopefully some of this comes from regulators)

What requirements should we collect?

- Goals of the project
  - Measurable metrics for each goal
- Data sources
- Potential risks: this list is huge
  - Risk mitigation strategies
- Outline of architectural components
  - This may be flexible, depending on what’s available to you

What if we were asked to build a Smart Transportation System?

- The SFMTA wants to implement a smart transportation management system to optimize traffic flow, reduce congestion, and lower emissions. Using data from traffic cameras, vehicle sensors, public transit GPS, mobile apps, weather services, and scheduled events, the system would predict traffic patterns and adjust traffic light timings, suggest route alternatives, and manage MUNI schedules. The system should be able to adjust to unexpected events, such as collisions, and ensure equitable service across neighborhoods. It should be live 24/7 and demonstrate measurable improvements in commute times and air quality.

What if we were asked to build a Smart Transportation System?

- Let's focus on just scheduling the lights on ***one street*** according to live traffic from traffic cameras and sensors. In Slack:

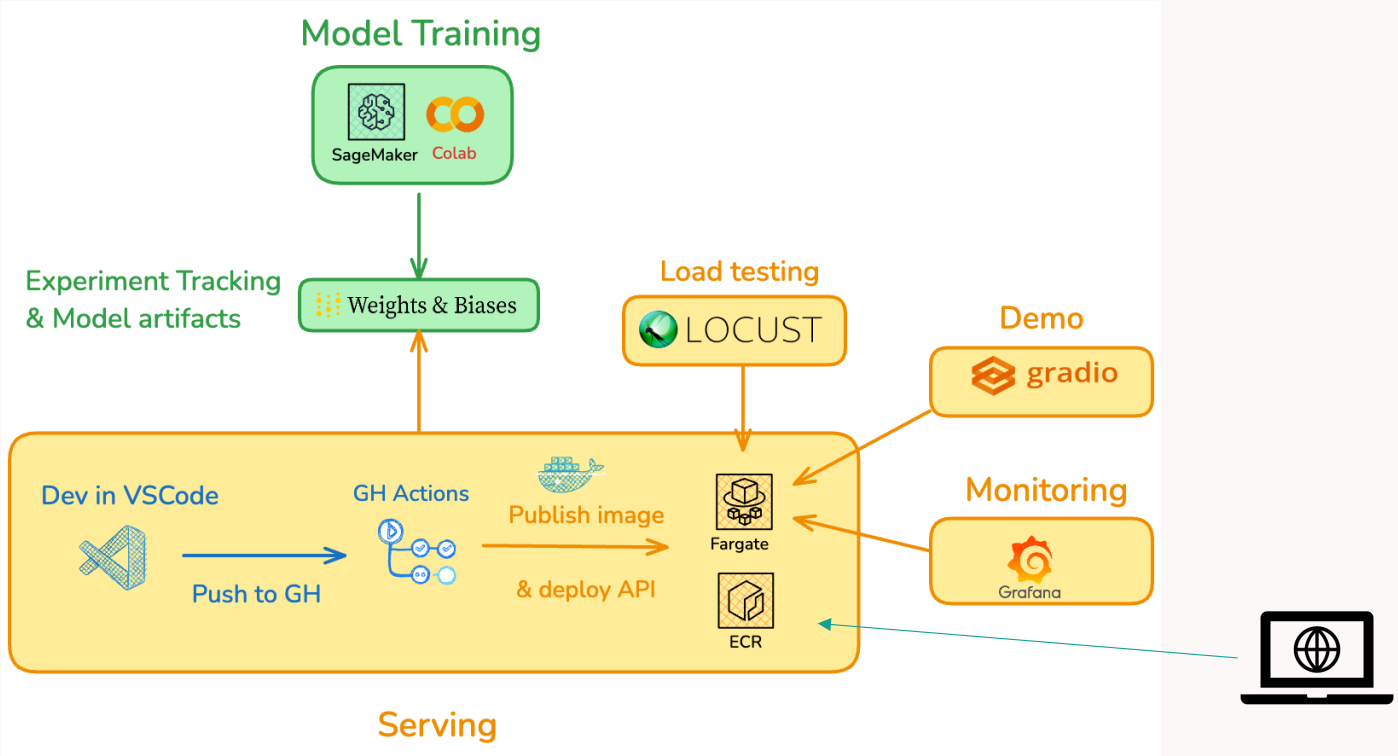
  List ~~two~~ 3 low-level goals? And what metrics can we use to measure our goals?

ML/AI
Devices
Infra
Team

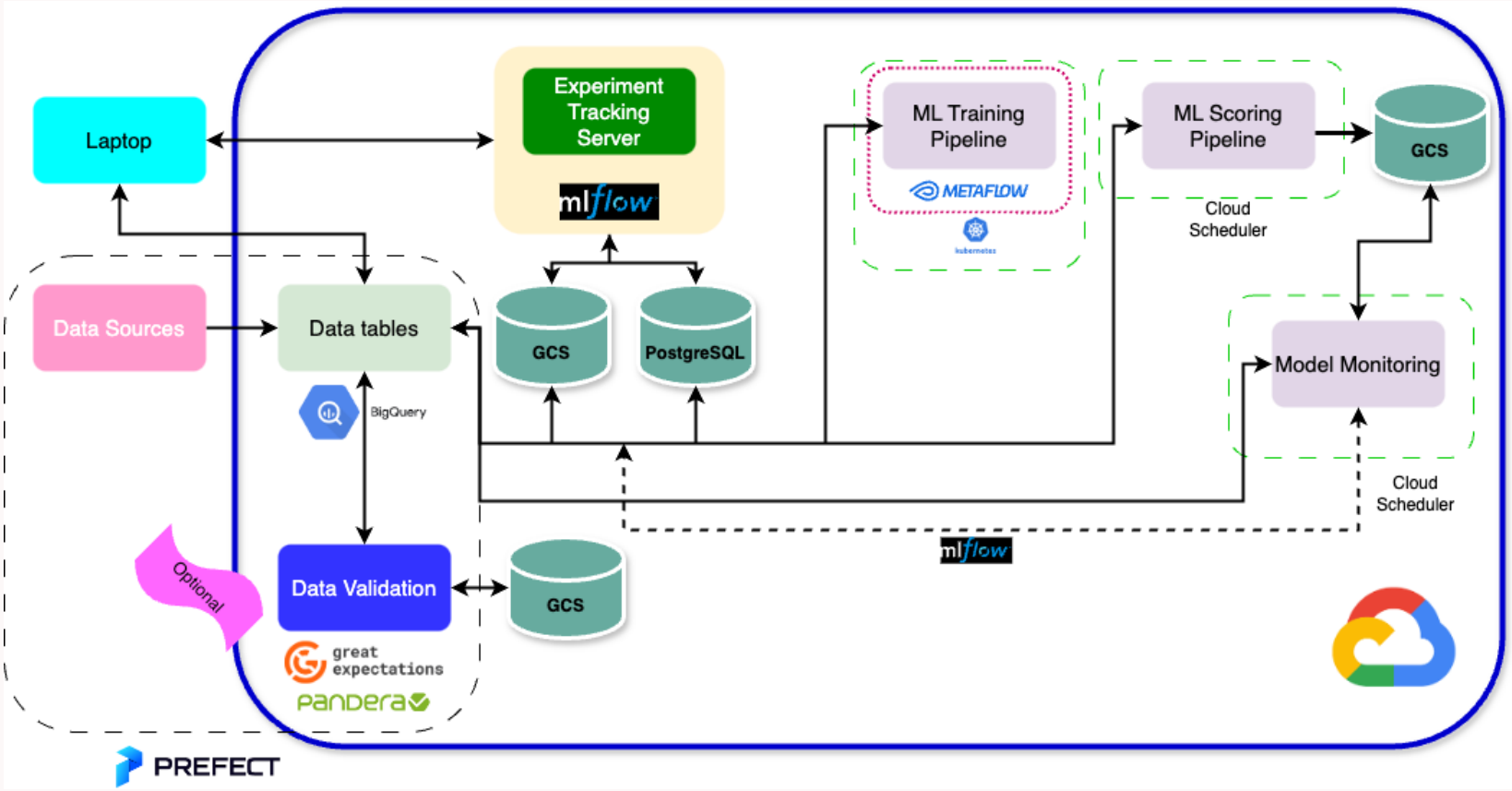
ML System Architecture

Let’s first zoom in on the MLOps architecture

What pieces is the system made of?



What pieces is the system made of?



Environments

- Dev data
  - Exploration
  - Model dev
- Mirror of prod

DEV

- Dev cluster
- Github: use dev branch

Dev. models
EDA
Track exps

SANDBOX ENV

Environments

- Stage data
- Mirror of prod

STAGE

- Stg. cluster (compute)
- Github: merge w/stg branch
- Trigger "build"
- Trigger unit / integration / regression tests

Environments

- Prod data
- Feature store (optional)

PROD

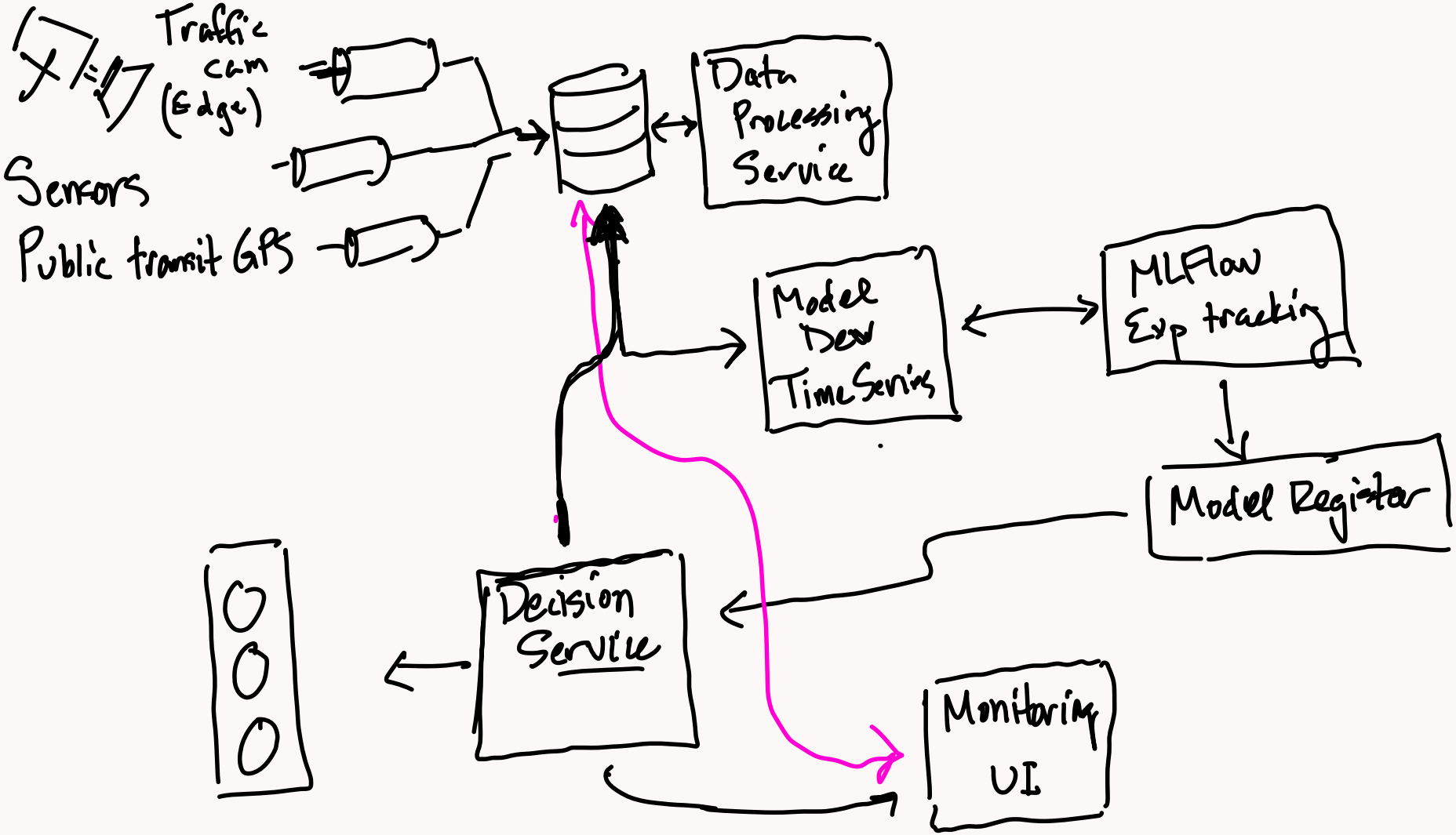
- Prod cluster (compute)
- Github: merge w/main branch

- Deploy model
- Monitor model (inputs, outputs, performance, resources)

Now let’s zoom out and consider the entire system

What if we were asked to build a Smart Transportation System?

- Let's focus on just scheduling the lights on ***one street*** according to ***estimated*** (not live) traffic from traffic cameras and sensors.


Traffic cam (Edge)
Sensors
Public transit GPS
Data Processing Service
Model Dev TimeSeries
MLFlow Exp tracking
Model Registar
Decision Service
Monitoring UI

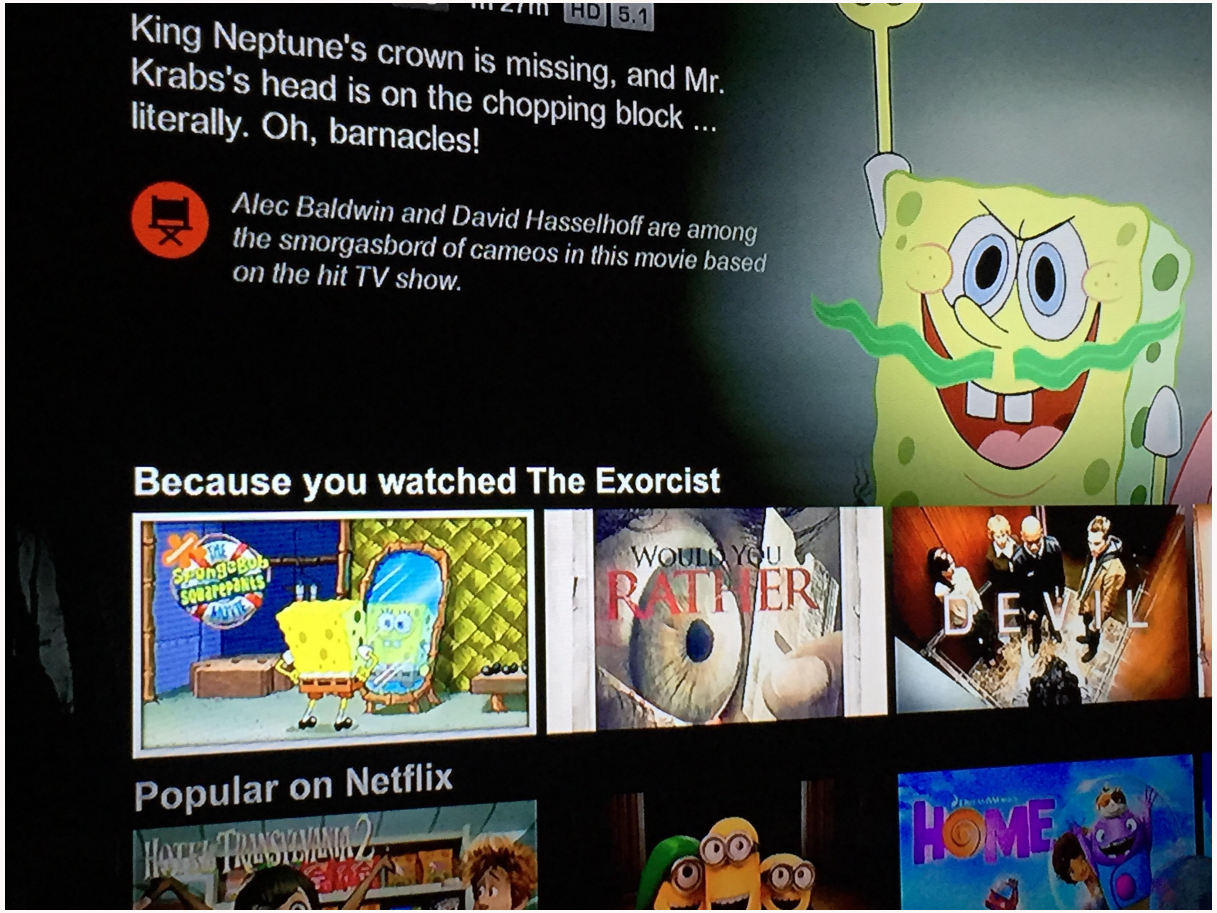
How about these other use cases?

This is next word prediction

More considerations:

- Cost
- Privacy
- Fairness
- Accuracy
- Model size
- Explainability
- Feedback loop

How about these other use cases?



More considerations:

- Cost
- Privacy
- Fairness
- Accuracy
- Model size
- Explainability
- Feedback loop

How about these other use cases?