GET YOUR GUIDE

Bringing ML to Production, what is missing?

Mikio Braun, Staff Data Scientist

Applied ML Days, Lausanne 2020

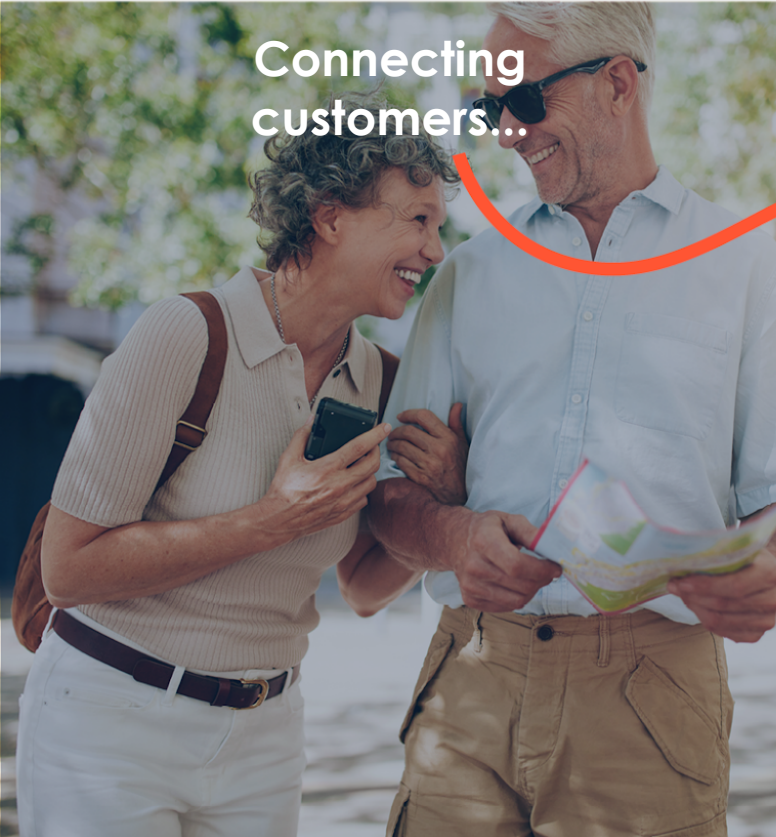With **thousands of incredible experiences around the world**, GetYourGuide is at the heart of travel
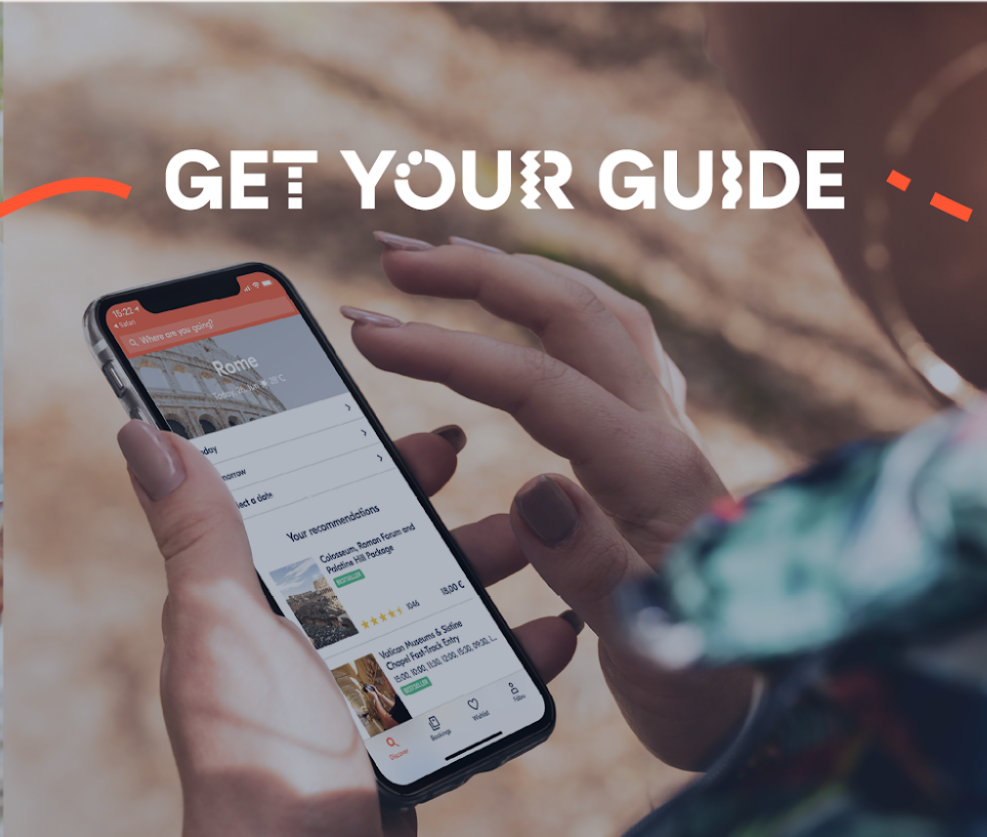

GET YOUR GUIDE

We've built one of the **world's biggest marketplace** for travel activities…

Connecting customers…

GET YOUR GUIDE

…to suppliers around the world

Millions of travelers use GetYourGuide every year

We facilitate the transaction

We offer **more than 40,000 activities** worldwide

GET YOUR GUIDE

Applied ML Days, Lausanne 2020
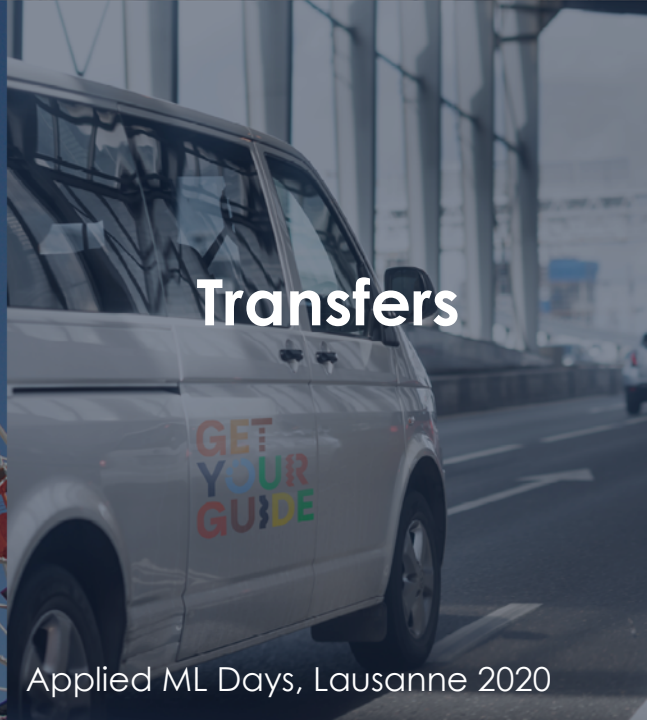
Hop-on hop-off

Attractions

Outdoor activities

Walking tours

Cooking classes

Amusement parks

Transfers

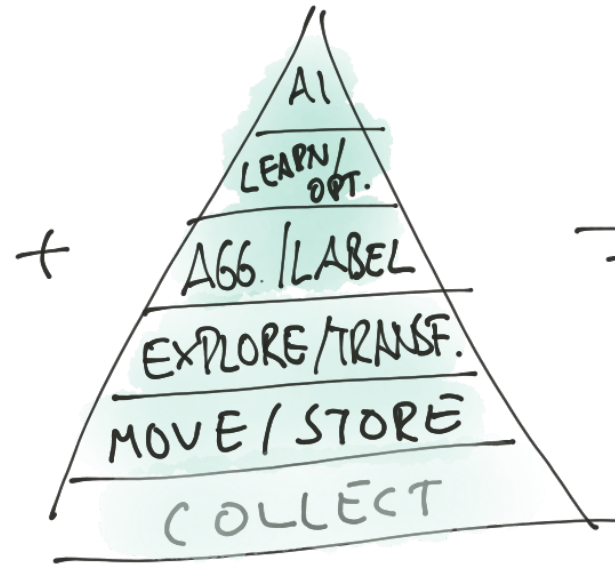Applied ML Days, Lausanne 2020

So How Do You "Do AI?"

Algorithms + Tools + = AI?

Deep Learning
Machine Learning
Statistics
:

Spark
Hadoop
Tensorflow
Jupyter
:

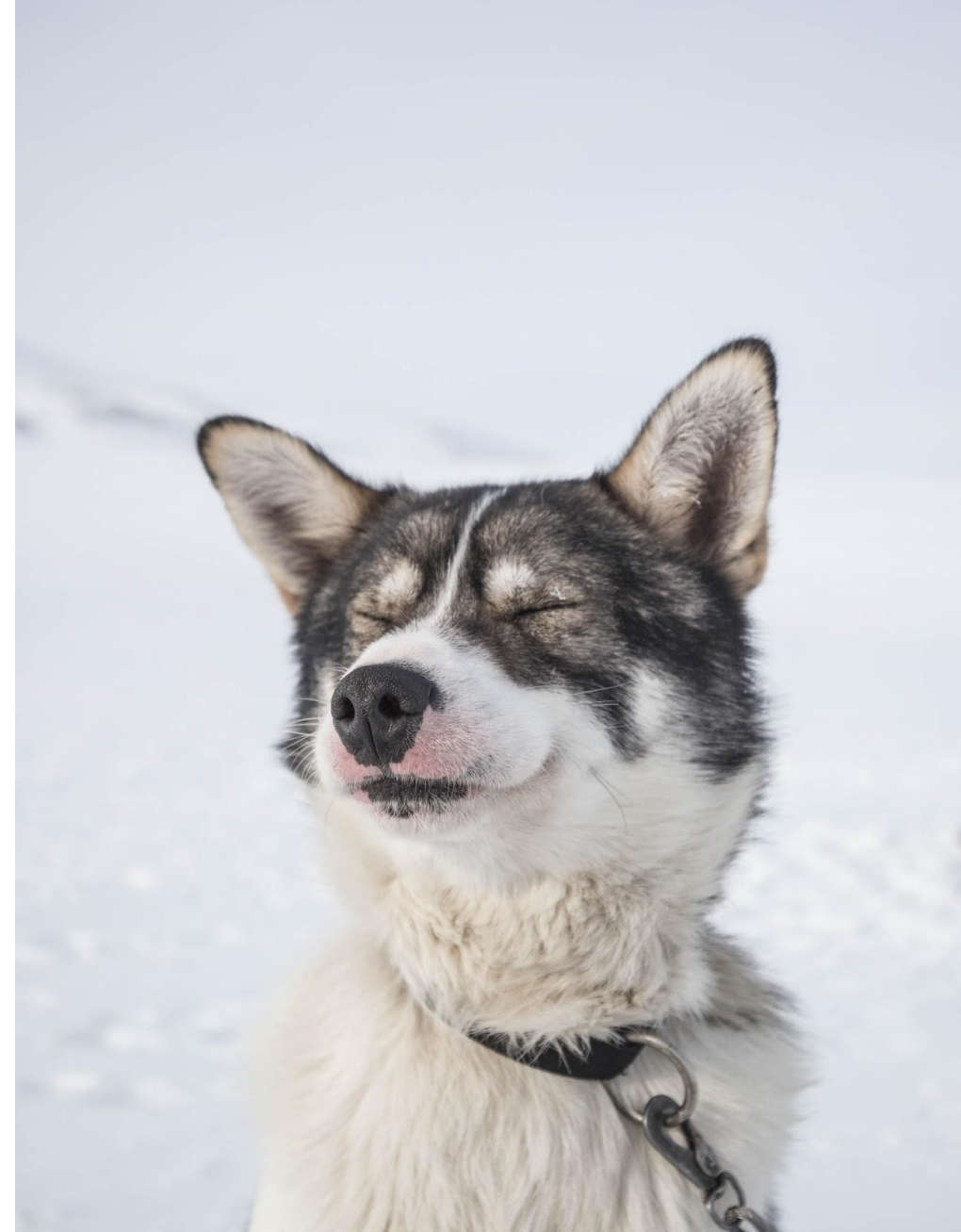Pyramid levels (top to bottom): AI / LEARN/OPT. / AGG./LABEL / EXPLORE/TRANSF. / MOVE/STORE / COLLECT

GET YOUR GUIDE

What is there beyond algorithms and tools?

- Figuring out where to apply ML
- Architecting ML applications
- Teams & Organizations

GET YOUR GUIDE

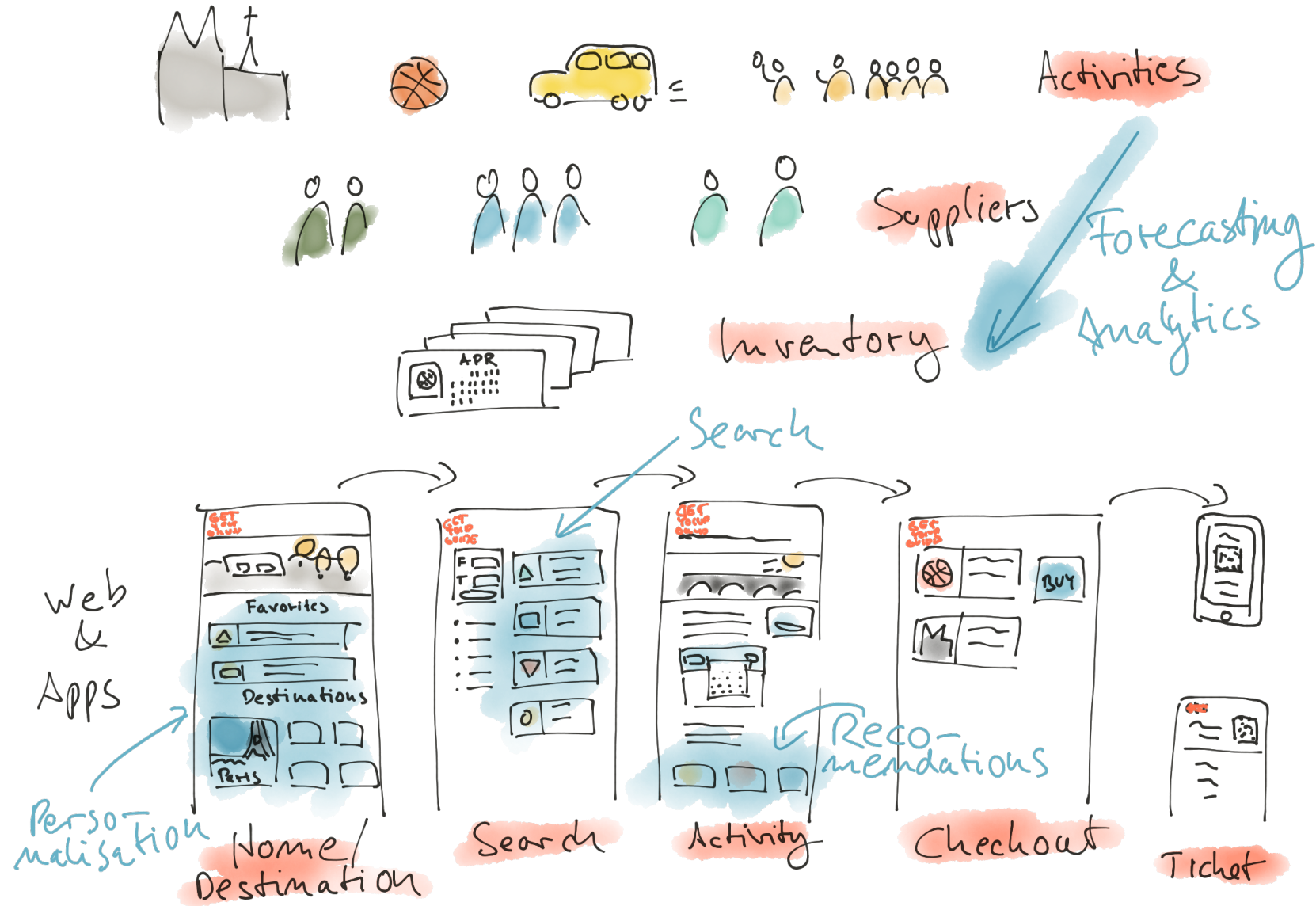# Business Problems to ML Problems

GetYourGuide Overview
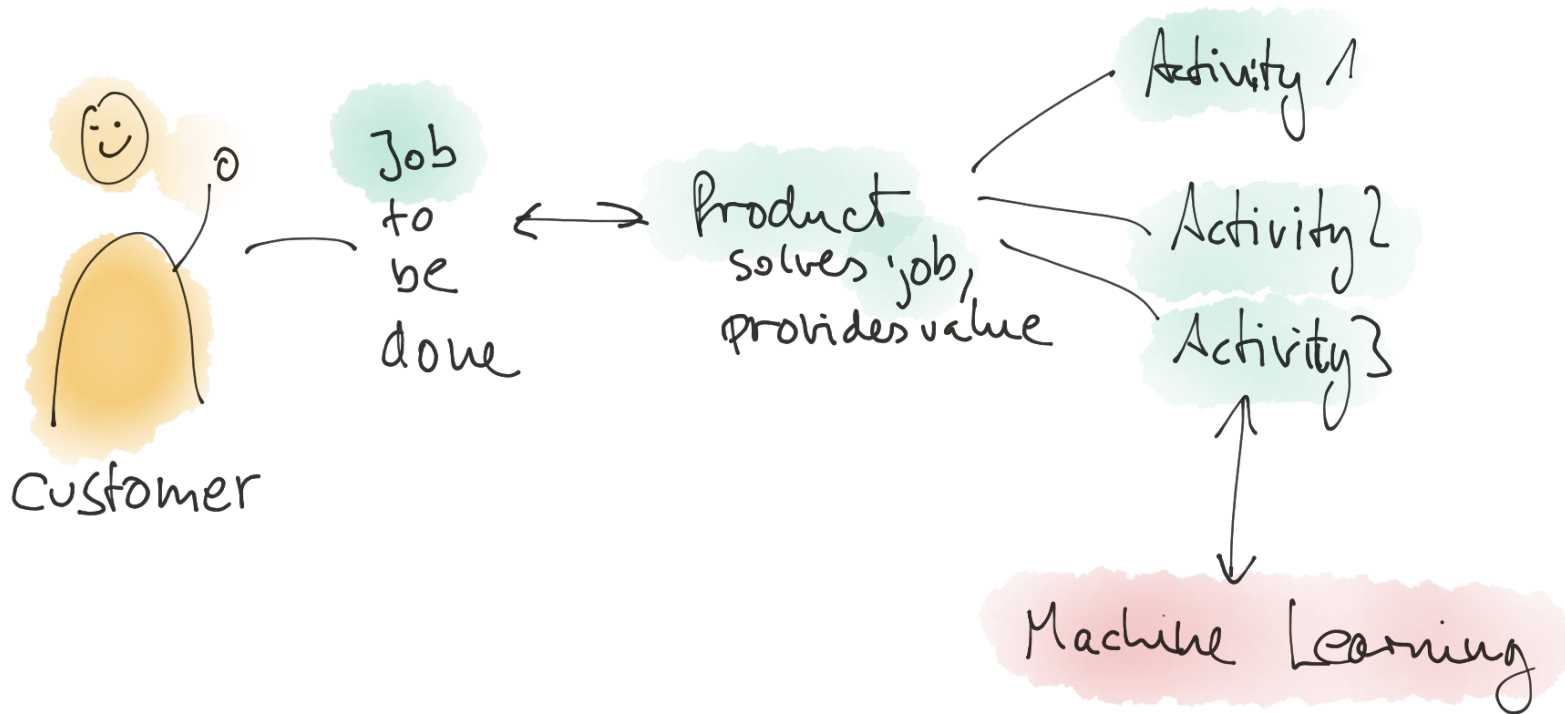

Applied ML Days, Lausanne 2020

GetYourGuide Overview

Activities

Suppliers

Forecasting & Analytics

Inventory

APR

Search

Web & Apps

Favorites

Destinations

Paris

Personalisation

BUY

Recommendations
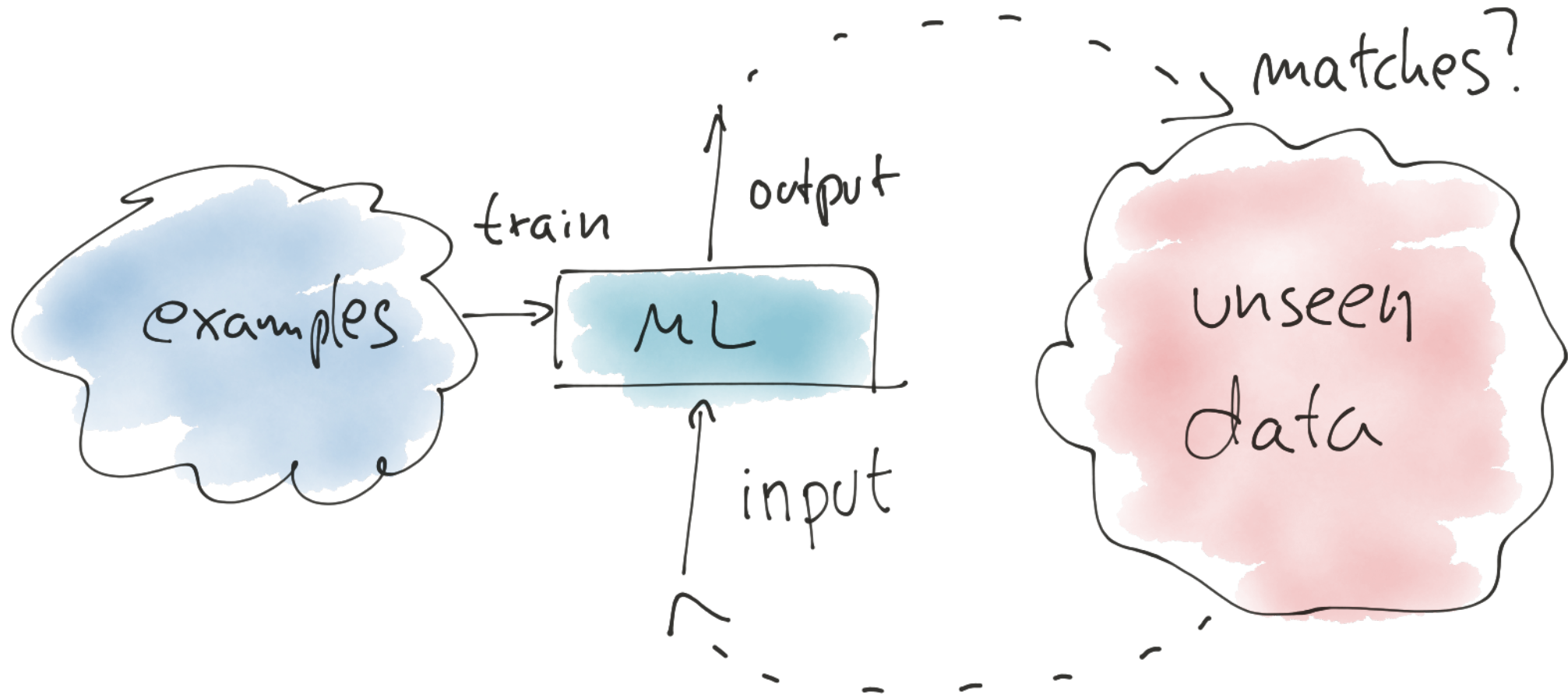
Home/Destination

Search

Activity

Checkout

Ticket

GET YOUR GUIDE

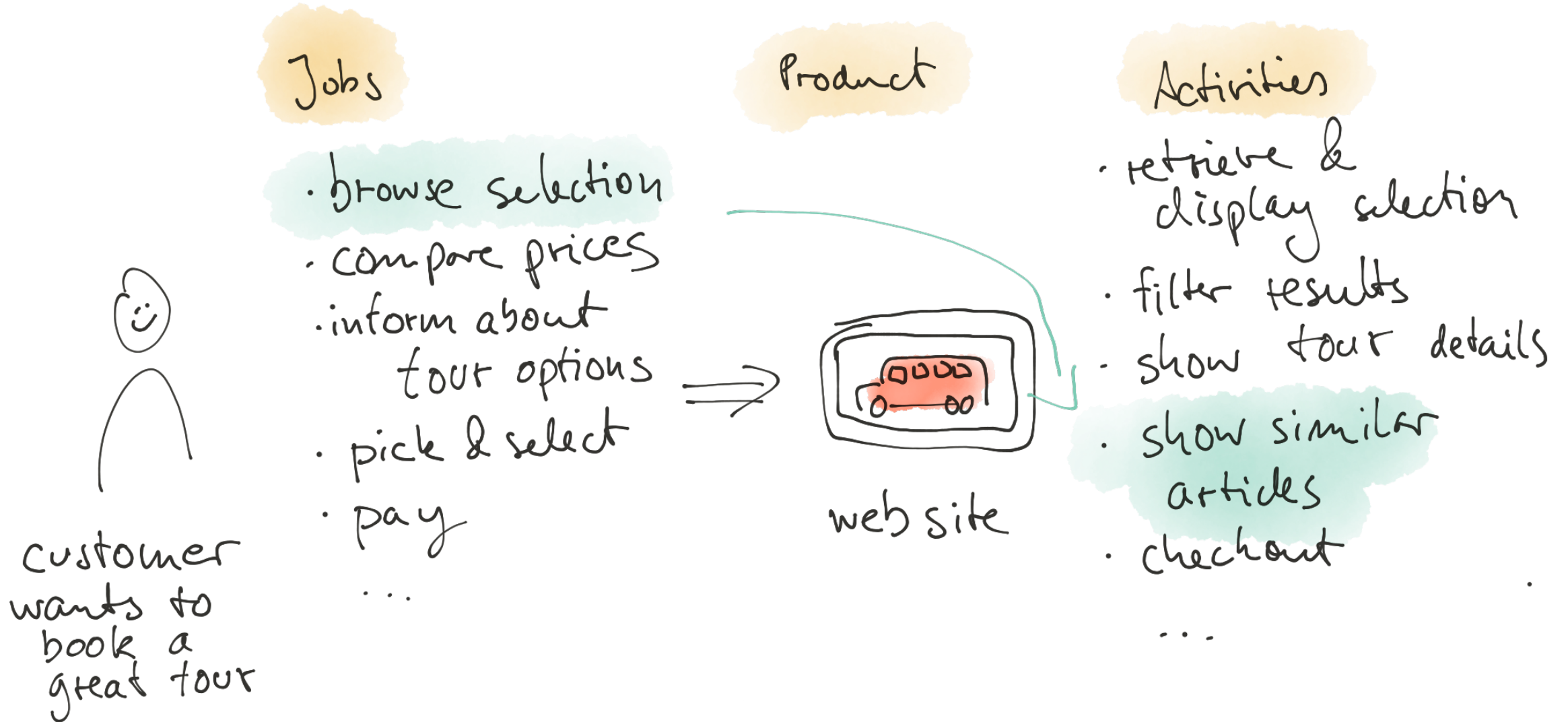# From Business Problems To Machine Learning Problems

- Customer has a Job to be done.
- Business has to give a Solution that adds value.
- Solution consists of Activities, and some of that activities might be solved with ML.
- ML: solve a task by learning from examples.

Machine Learning: Learn From Examples

Example: Recommendations

Jobs

Product

Activities

- browse selection
- compare prices
- inform about tour options
- pick & select
- pay
...

website

- retrieve & display selection
- filter results
- show tour details
- show similar articles
- checkout
...

customer wants to book a great tour

Example: Recommendations

Typical problem for machine learning:

- Hard to specify what exactly means "similar."

- A lot of example data is available.

- Recommendations have to change based on new articles frequently.

Example: Recommendation
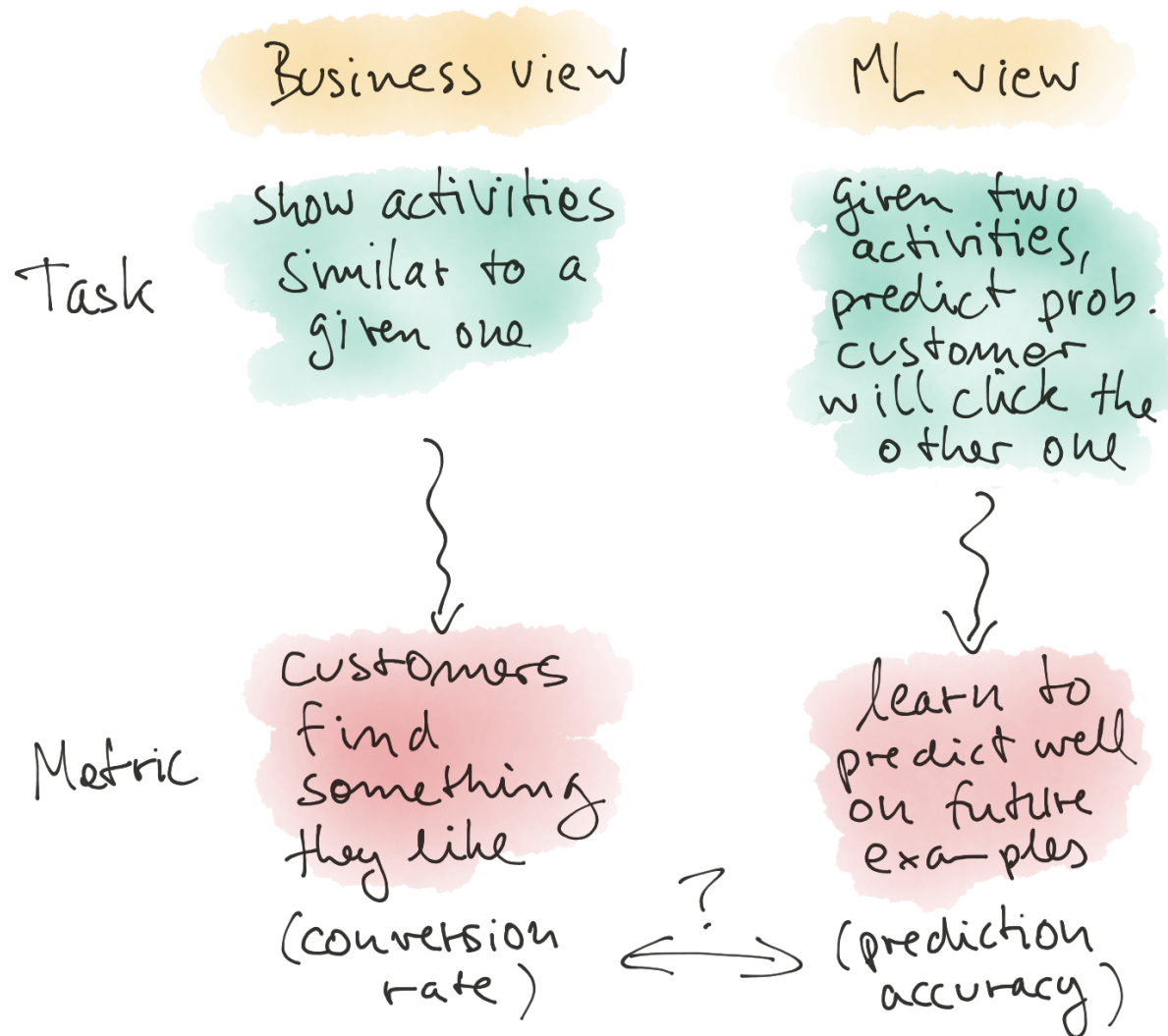
- Business/customer view and machine learning view often have different metrics.
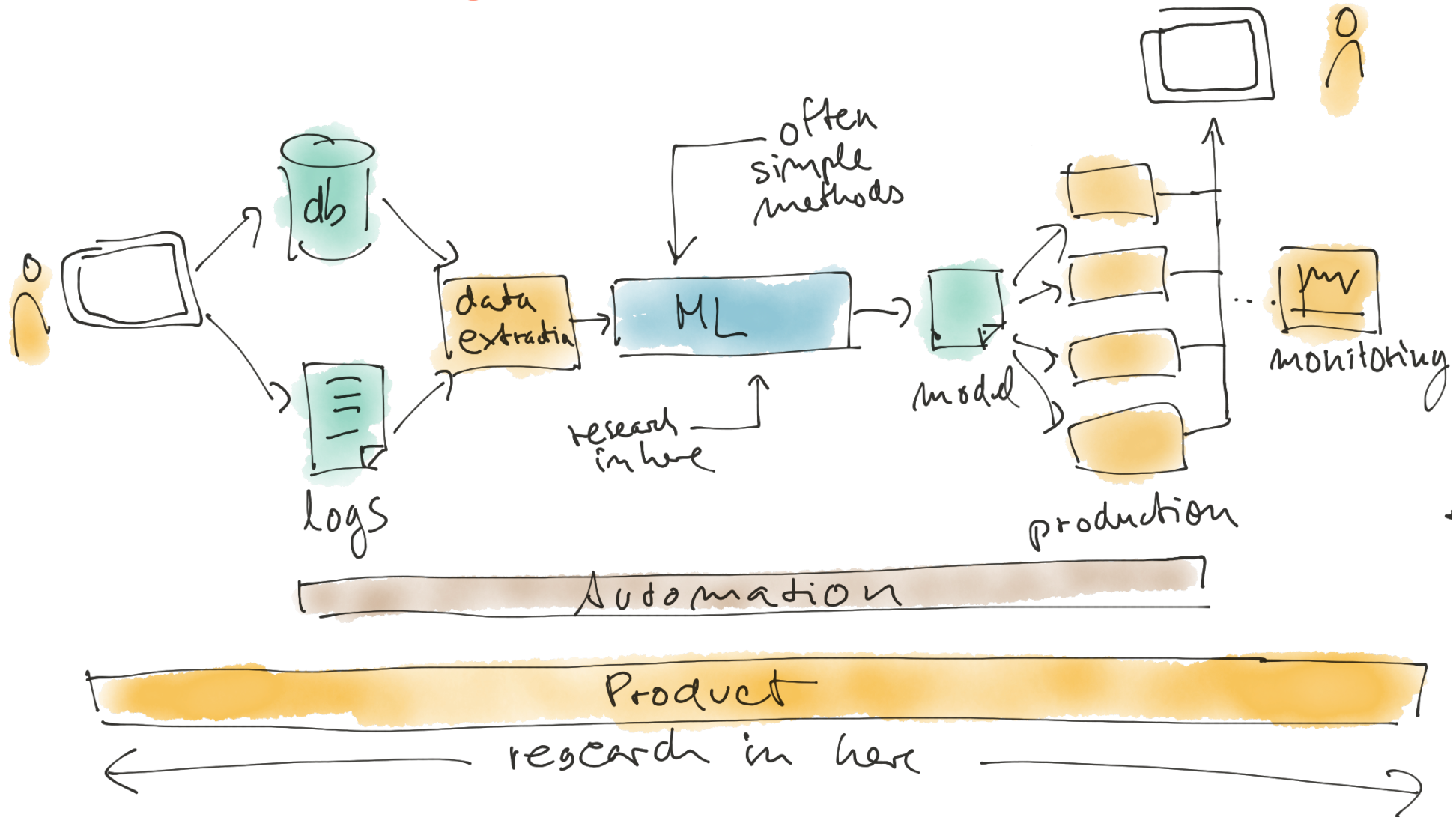- Machine learning does not know about the business view.
- Figuring out the right relationship is not easy.
- Also depends on the available data.

GET YOUR GUIDE

Architectures for AI Systems

Applied ML Days, Lausanne 2020

Machine Learning Systems

Design Patterns for AI Architectures

Core Machine Learning
—how to train, evaluate, etc.

Serving
—access predictions in real-time
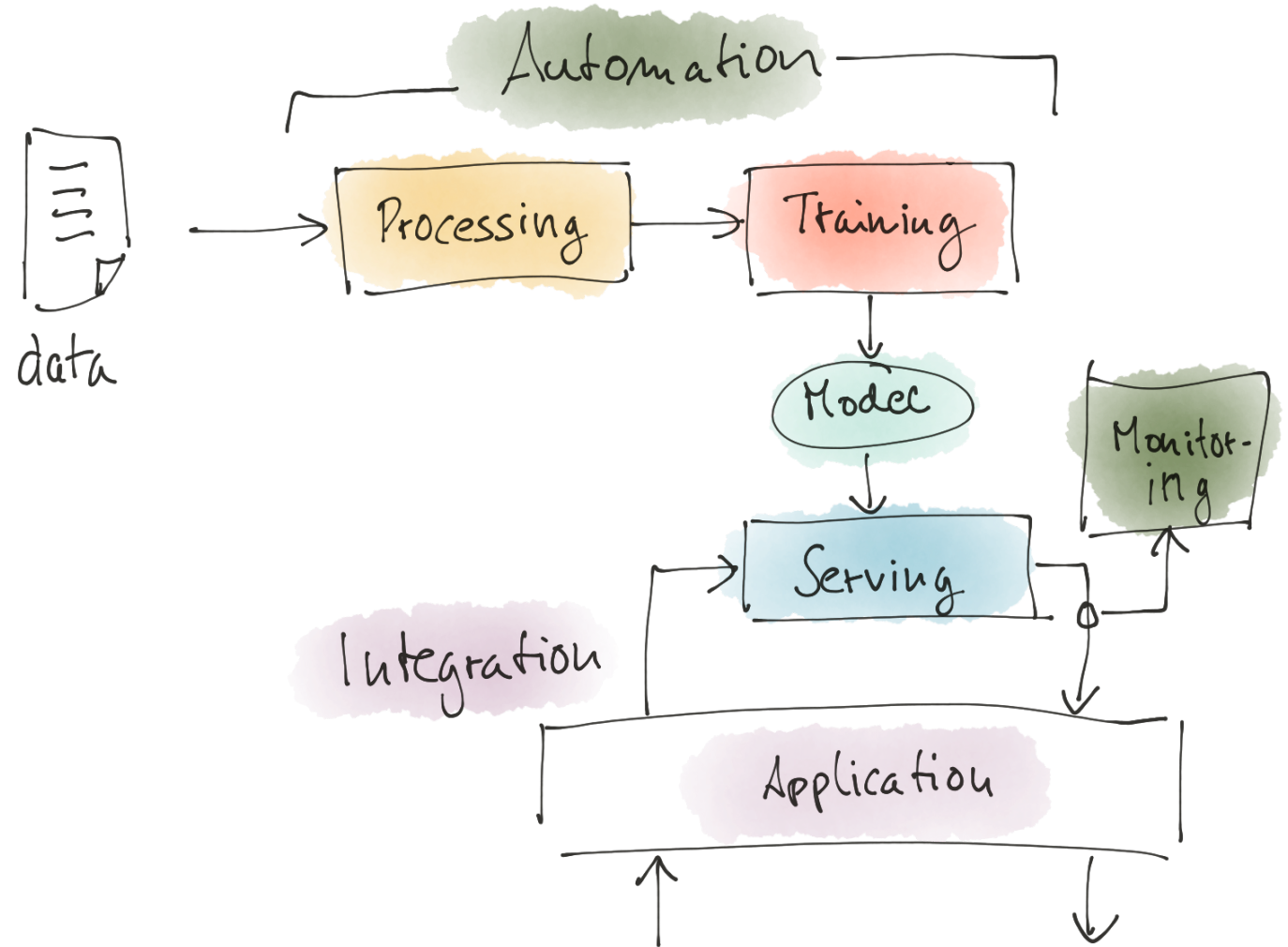
Data Preprocessing and Features
—how to deal with preprocessing

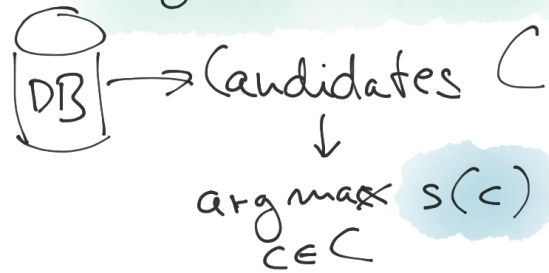**Automation & Monitoring**
—making it more production ready

Machine Learning Integration
—how to fit it into a larger picture

Machine Learning Integration Patterns

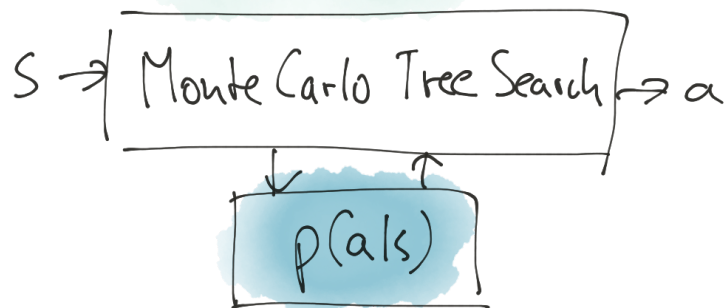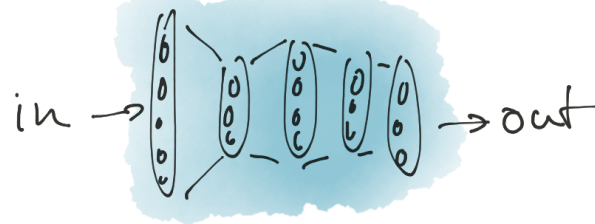- Depending on how ML is used, there are different integration patterns.

- Keep in mind that the ML problem and the application problem are different and have different metrics.

- Beware of interacting ML models if the overall training data is used.
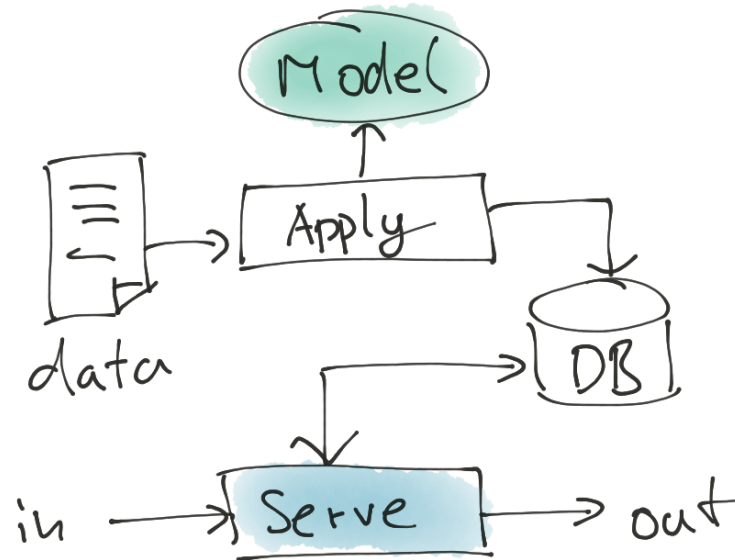
Serving Patterns

- How to provide the predictions of the machine learning model to the application.
- If the domain is small, precomputing predictions might be the easiest way to serve.
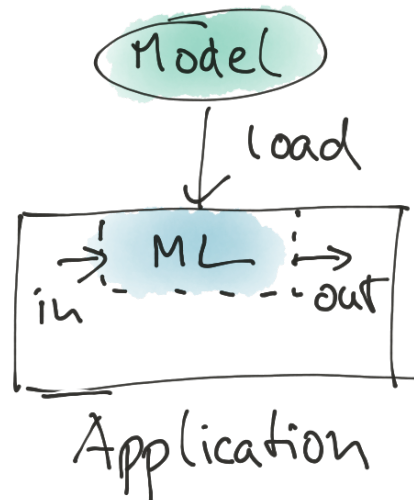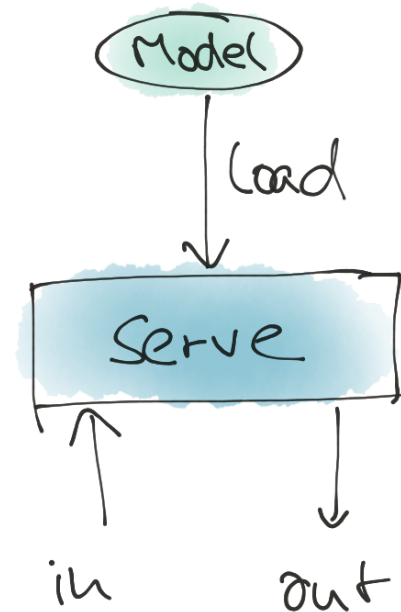
Preprocessing Patterns
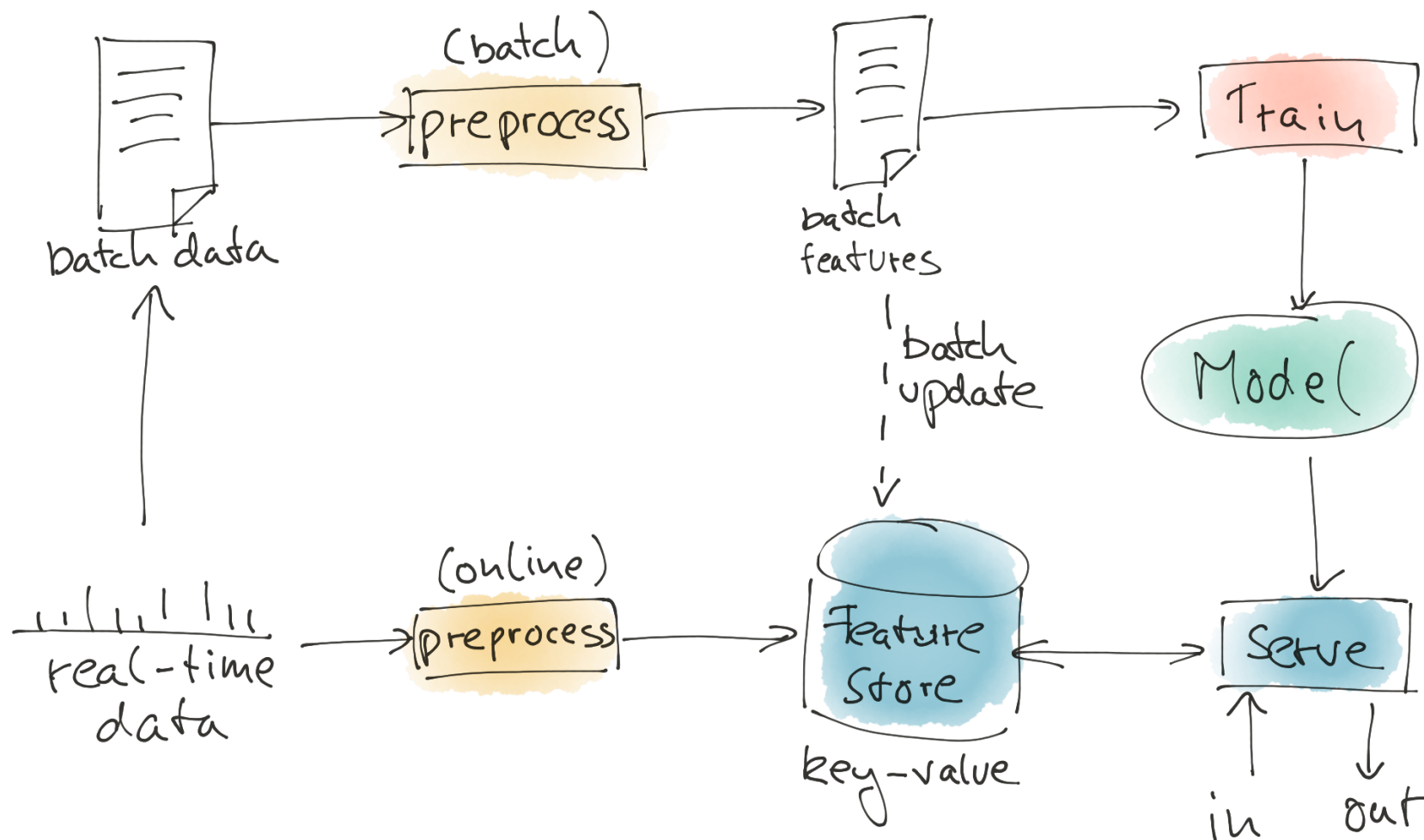

- If predictions are precomputed, no special treatment of feature computation is necessary.

- If it is served online, you can either precompute features in a Feature Store. But you don't have updates for new data.

- Or you can use the Feature Store and update features online as well.

Teams and Organization

GET YOUR GUIDE

How Data Scientists Work
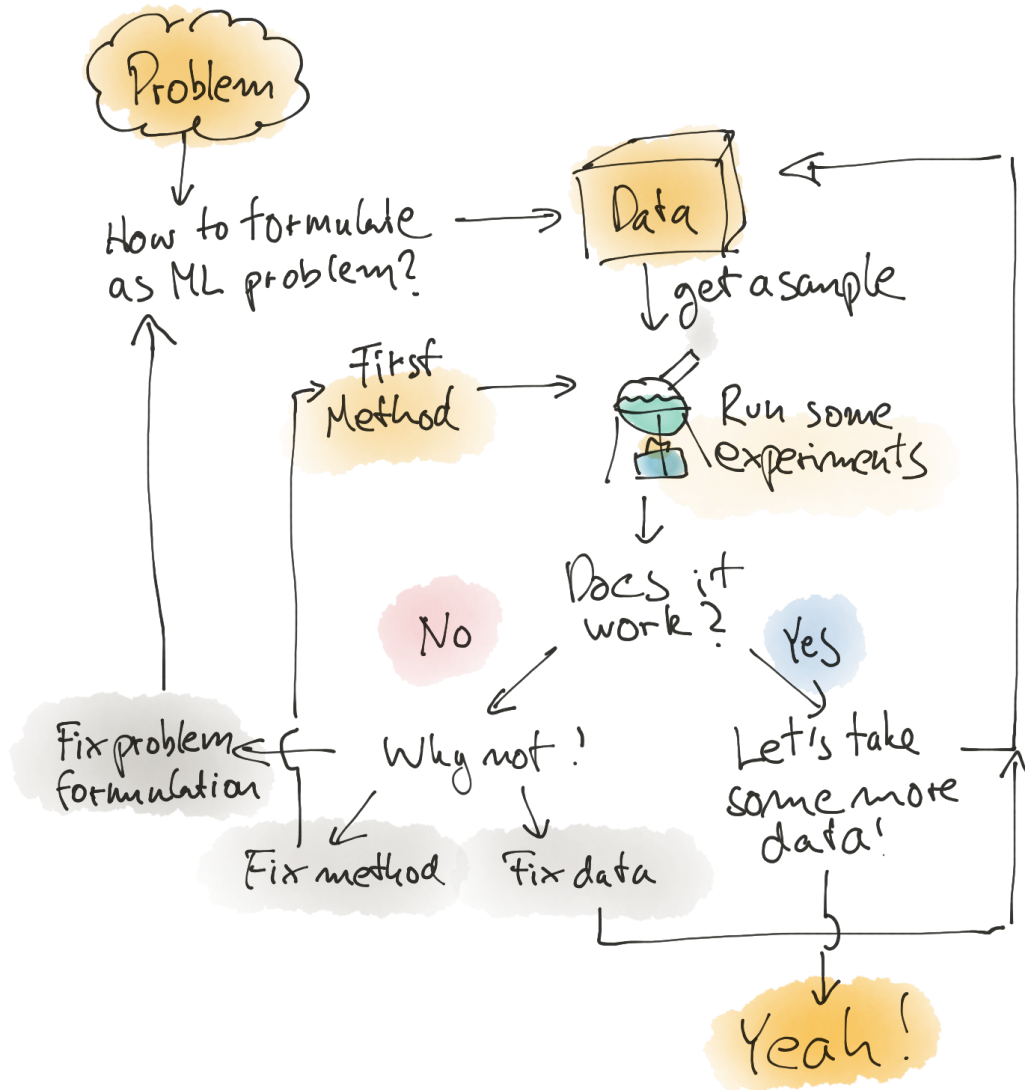
Data Scientists And Developers
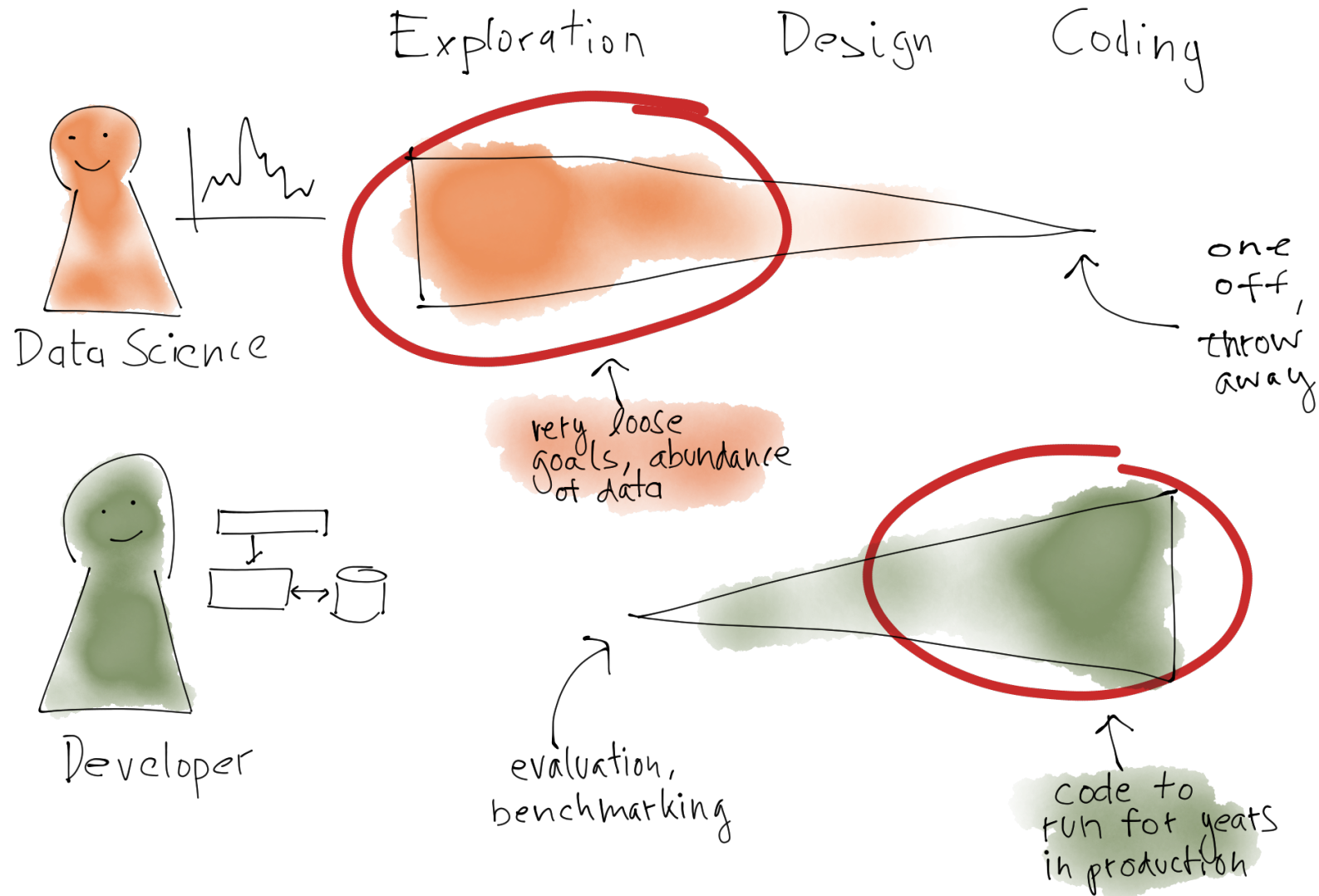
Data Scientists and Developers Collaboration

- What is the most productive way?
- Ideally, interface on code, not just documentation
- Production logs often become data analysis input!

Applied ML Days, Lausanne 2020

Organizing Data Science

Team A

Center of Excellence

Team B

over the fence

Data Scientists   Software Engineers   Ops

Cross functional teams

Data Platforms

THE DATA SCIENCE
HIERARCHY OF NEEDS

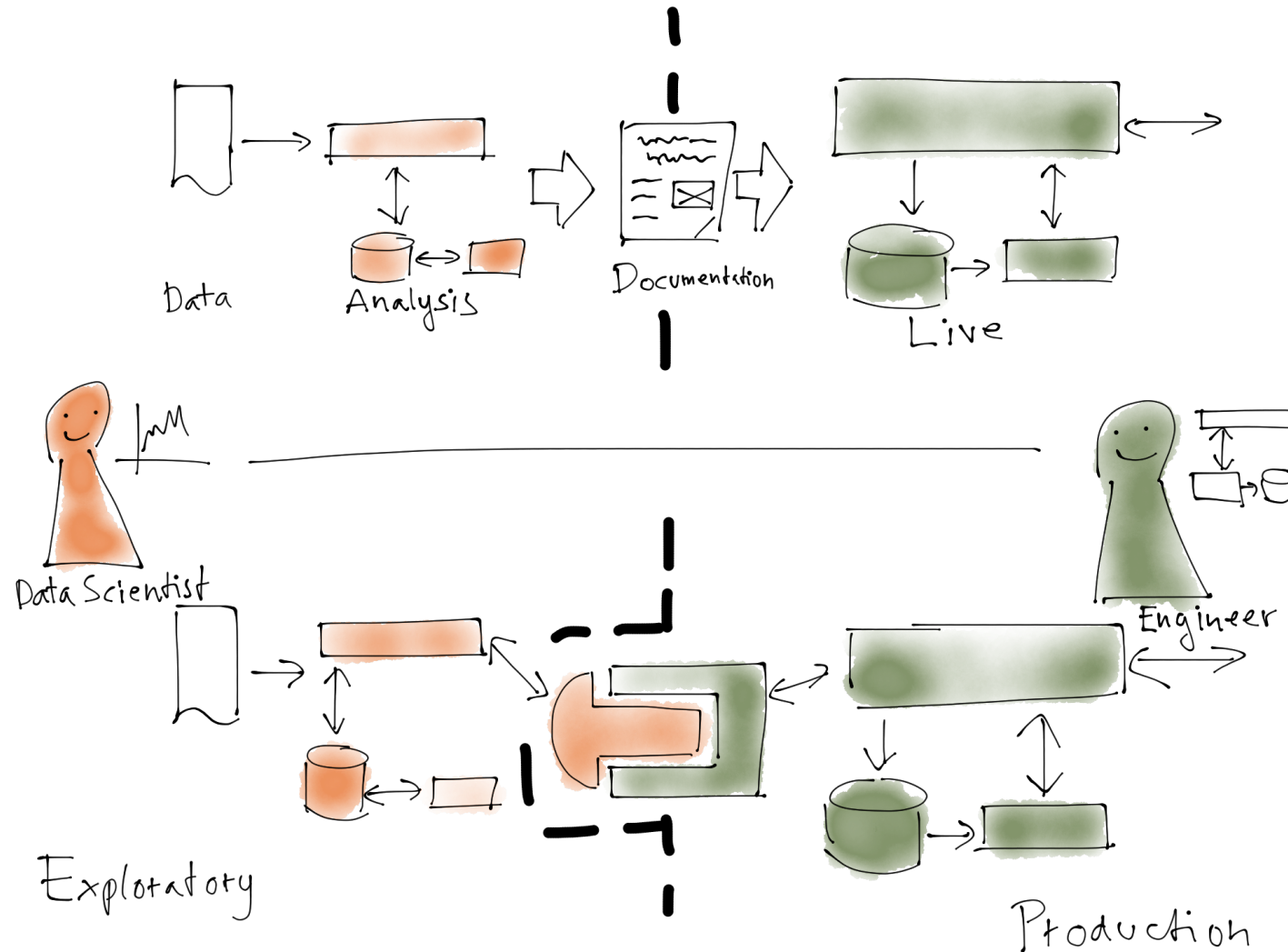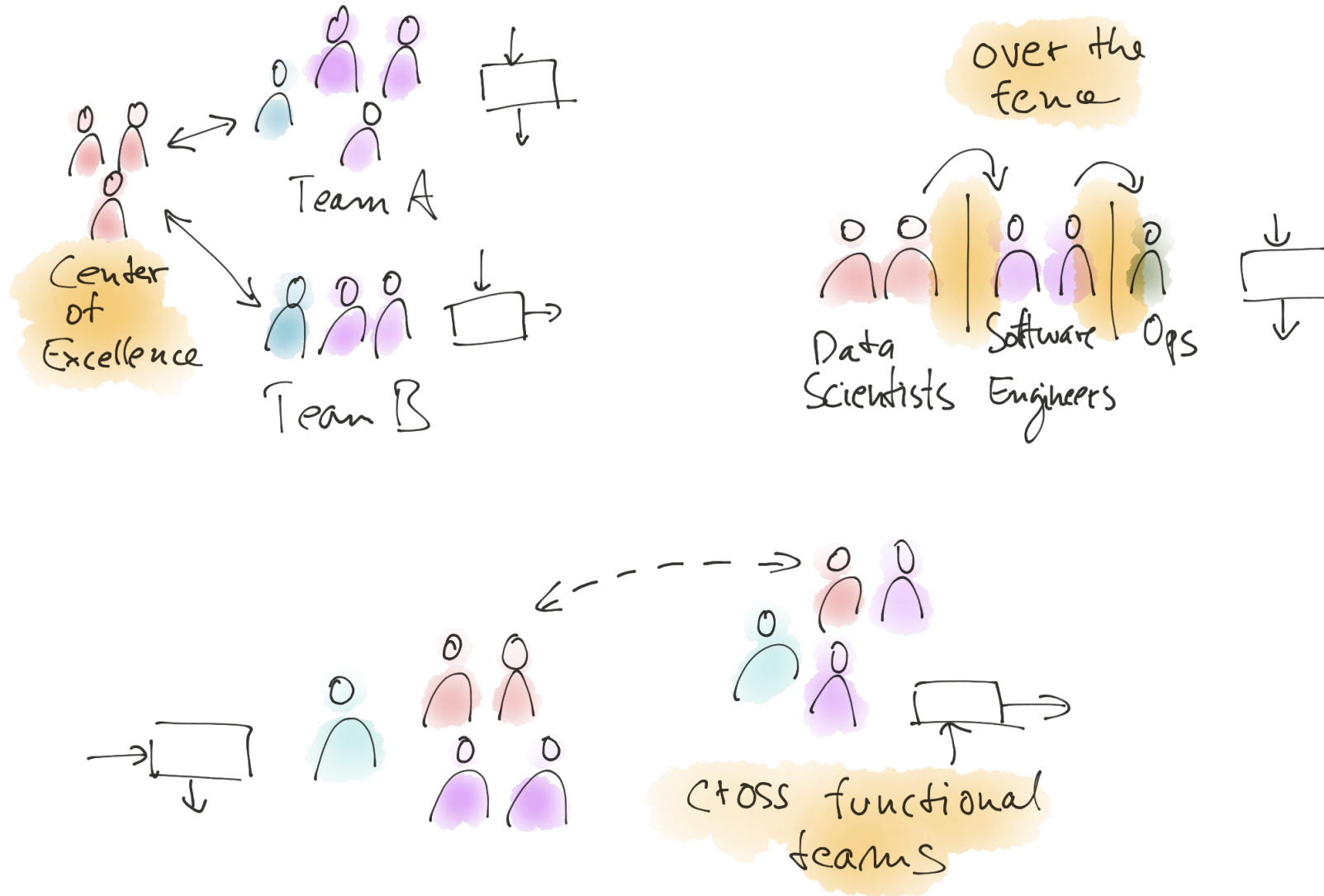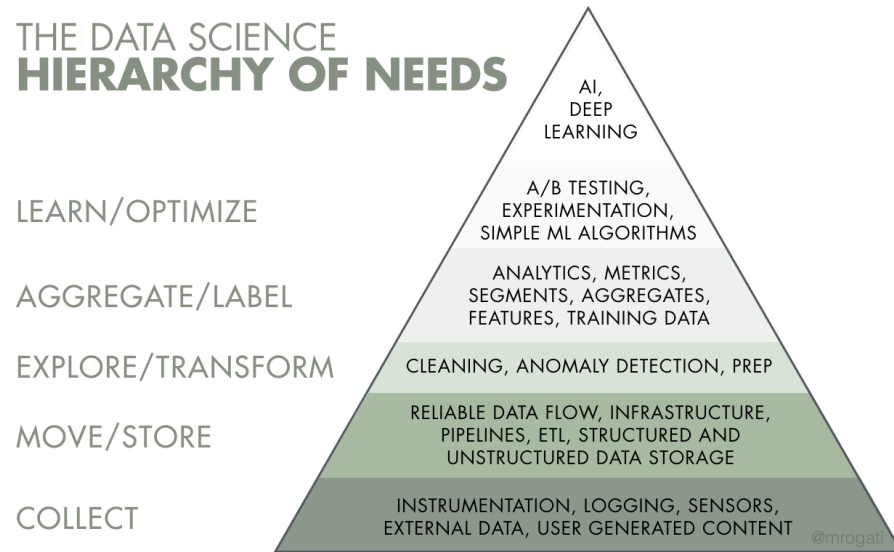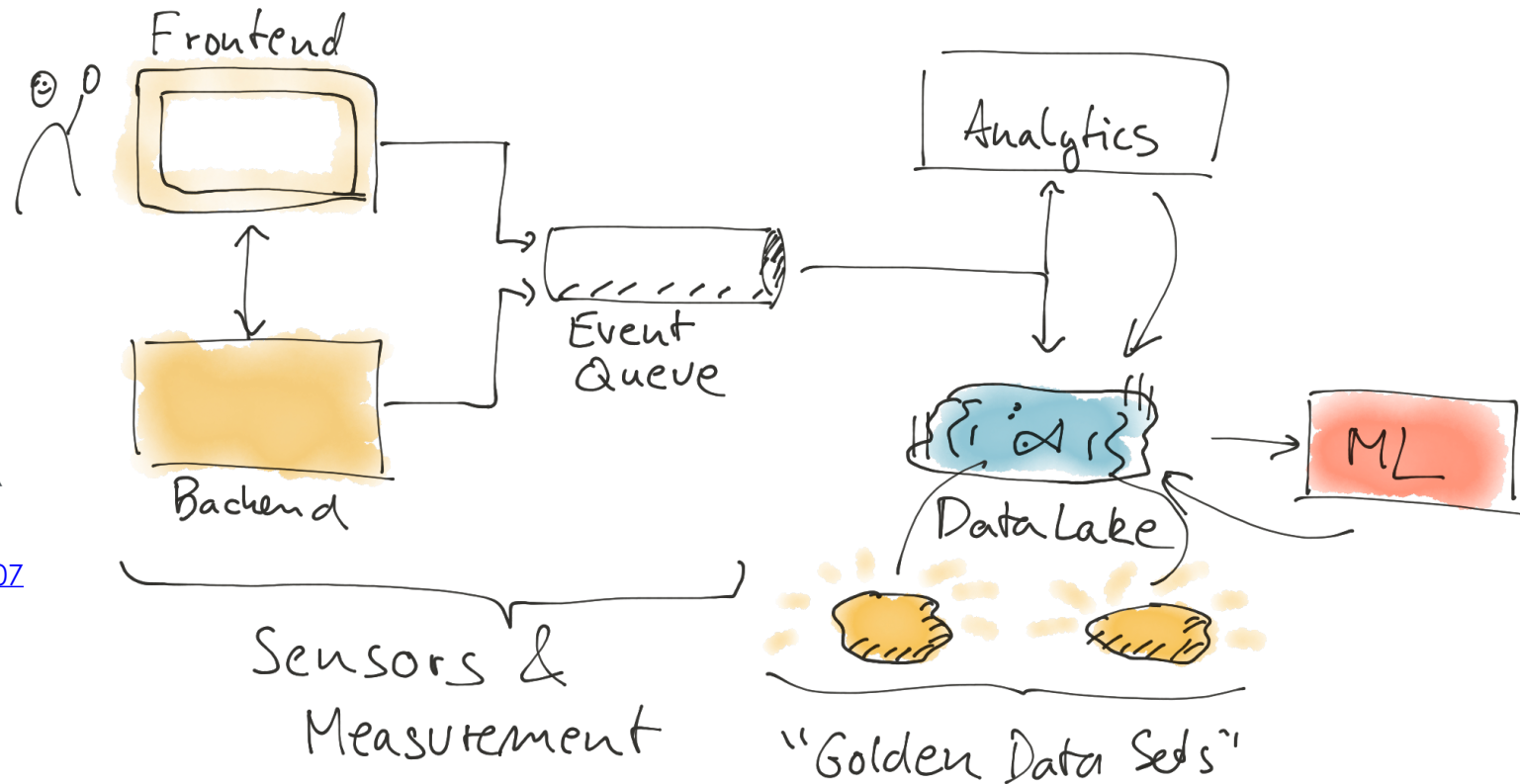LEARN/OPTIMIZE

AGGREGATE/LABEL

EXPLORE/TRANSFORM

MOVE/STORE

COLLECT

AI, DEEP LEARNING

A/B TESTING, EXPERIMENTATION, SIMPLE ML ALGORITHMS

ANALYTICS, METRICS, SEGMENTS, AGGREGATES, FEATURES, TRAINING DATA

CLEANING, ANOMALY DETECTION, PREP

RELIABLE DATA FLOW, INFRASTRUCTURE, PIPELINES, ETL, STRUCTURED AND UNSTRUCTURED DATA STORAGE

INSTRUMENTATION, LOGGING, SENSORS, EXTERNAL DATA, USER GENERATED CONTENT

@mrogati

https://hackernoon.com/the-ai-hierarchy-of-needs-18f111fcc007

Machine Learning: The High Interest Credit Card of Technical Debt

Summary

- Not just about methods and tools.
- How to identify where to use ML?
- Building AI/ML systems.
- How to set up teams and organizations for ML.

Questions?

mikio.braun@getyourguide.com
Twitter @mikiobraun

GET YOUR GUIDE