Bringing the Cloud to the 4th Industrial Revolution

Oden Technologies
Devon Peticolas

Who Am I?

Devon Peticolas

Sr. Data Engineer

My job is to prepare Oden's data and pipelines for Machine Learning and Data Science.

Who Are We?

Oden Technologies

Founded in 2014

Pre-Series A

Who Are Our Customers?

Manufacturers, people making things.

- Factories
- Multi-Factory Companies
- Focused Teams

$1.7 — 3.2T

Projected impact on global GDP from manufacturing IoT over by 2031
~ McKinsey, 2016

The Product

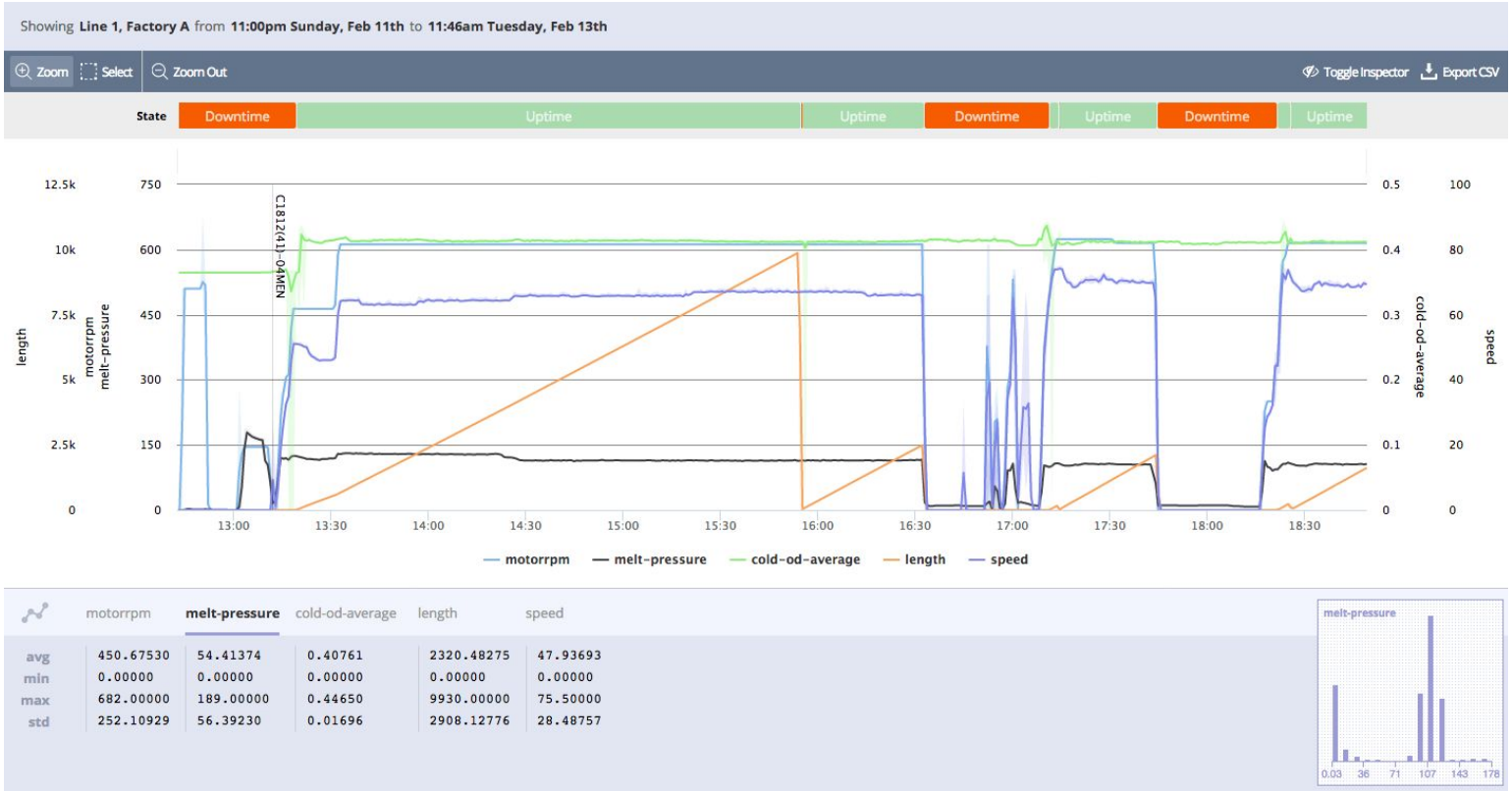
Real Time Overview

Factory

Demo Factory

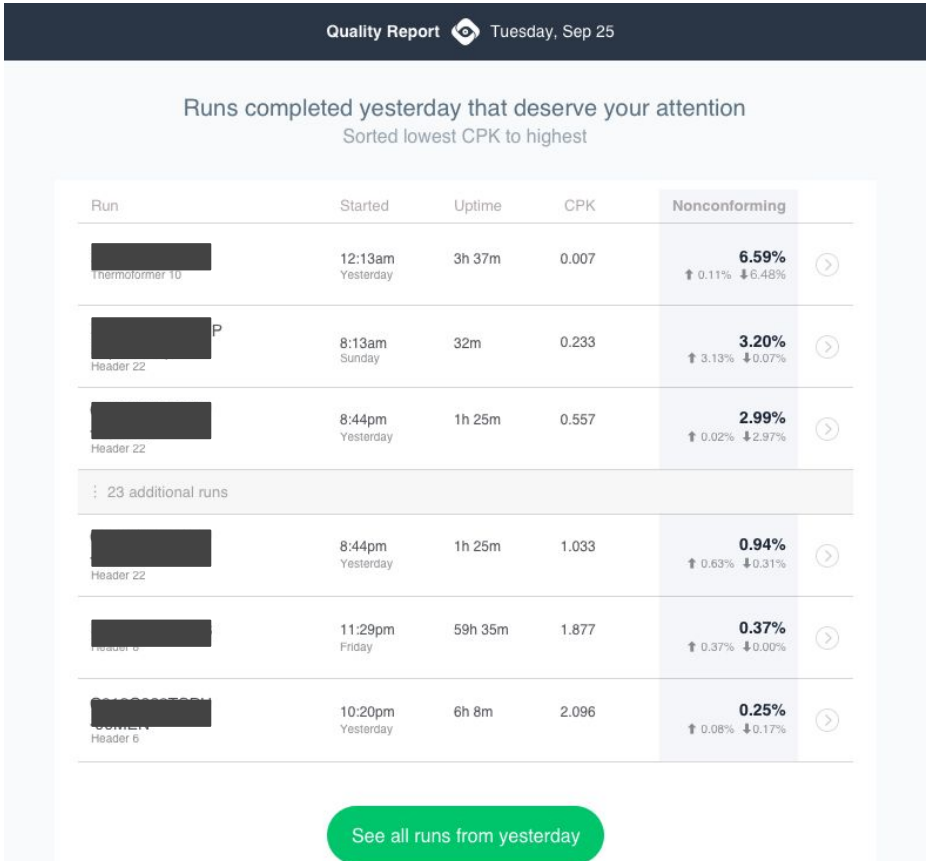
Name	State	Product	Target	Actual
[redacted]	Downtime 18m 39s	[redacted] 25d	130 ±3 machine.speed	122 -8
[redacted]	Running 1h 32m	[redacted] 13d 13h	130 ±3 machine.speed	128 -2
[redacted]	Running 3h 45m	[redacted] 27d 7h	154 ±4 machine.speed	—
[redacted]	Downtime 3h 56m	[redacted] 41d 15h	—	—

Time Series View



Spreadsheet View

Reporting View



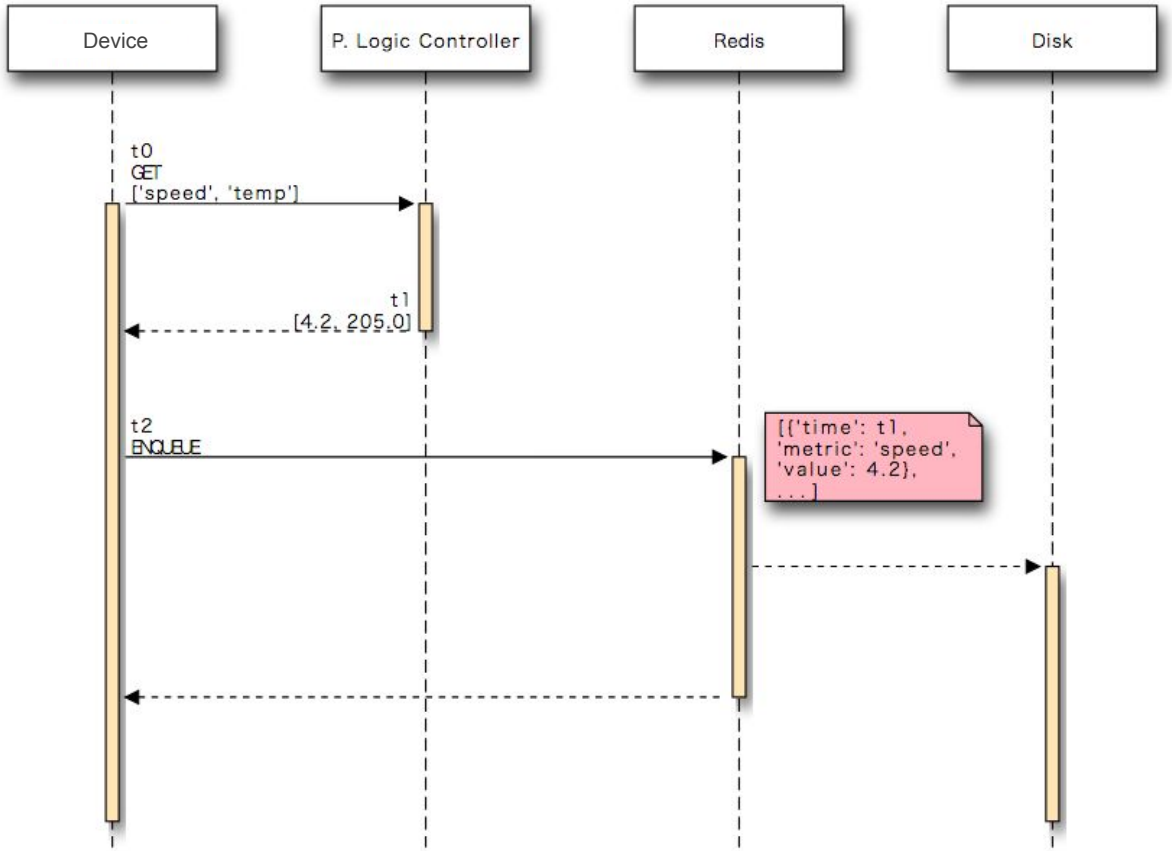
Data Acquisition

Acquisition 1

→ Requests metrics from PLC

→ PLC returns values at t1

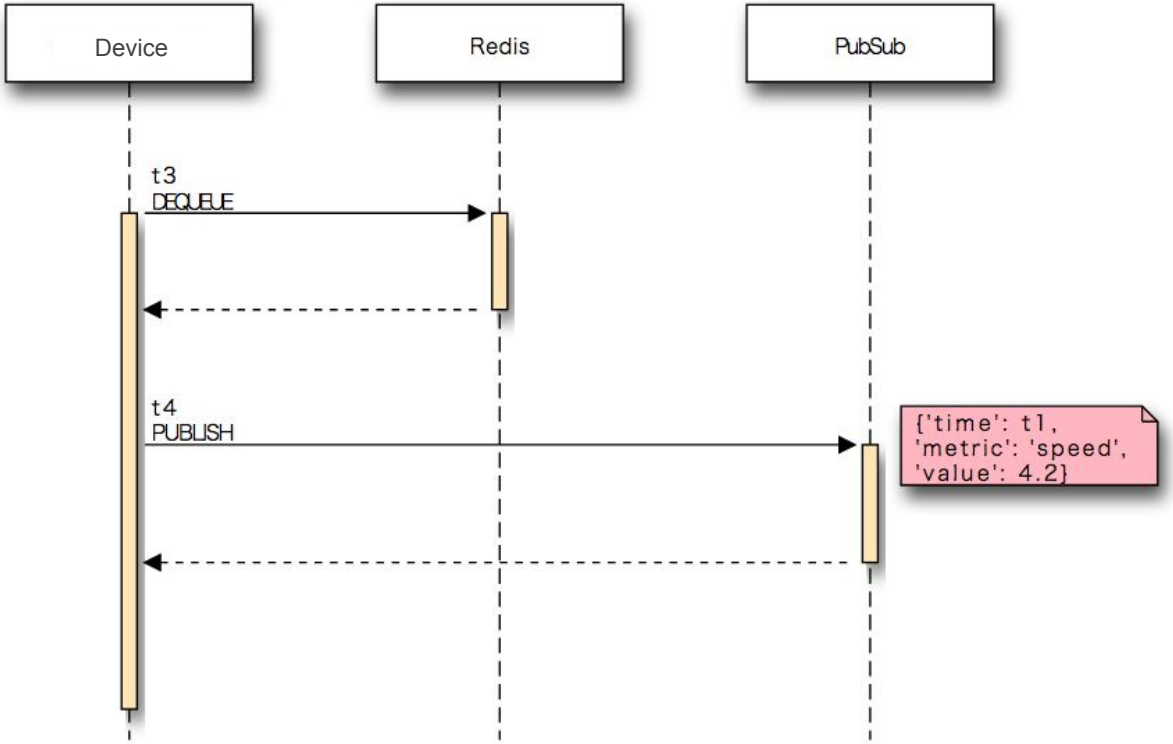
→ Enqueues to disk-backed Redis w/ time t1



Acquisition 2

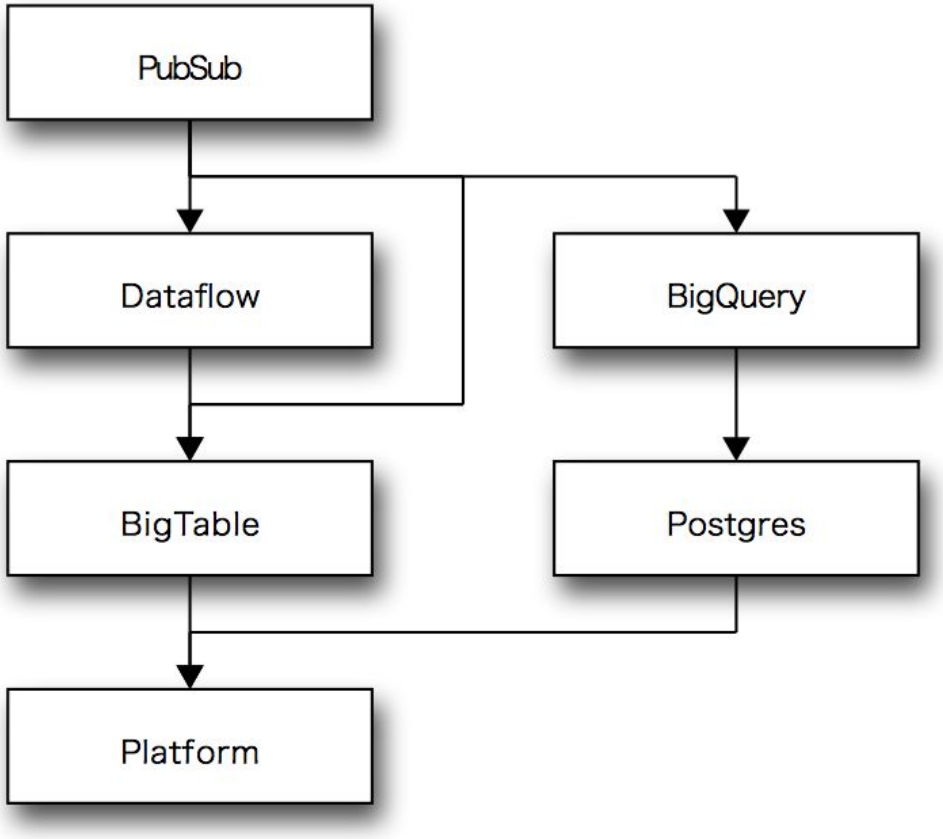
→ Dequeues from Redis

→ Publishes to PubSub



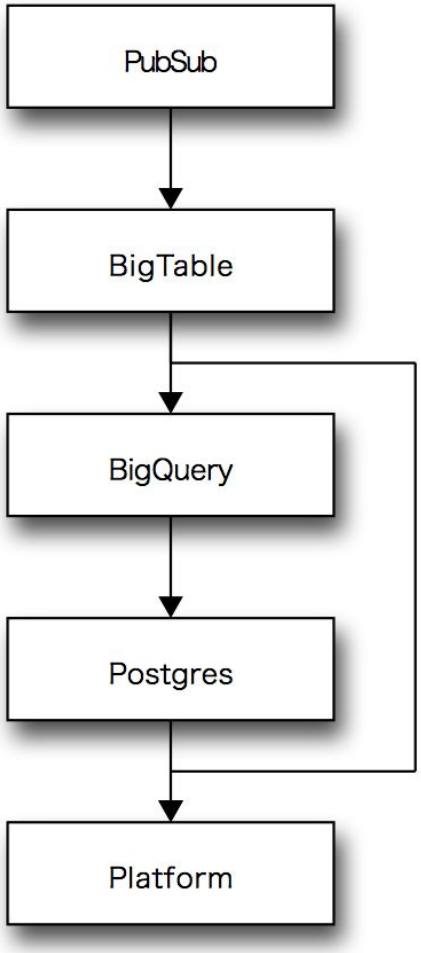
Write And Transform 1

- → PubSub fans out
  - → BT for 1s-res “hot” data
  - → BQ for “slow” data
  - → Df to BT for 60s and 600s “roll-ups”
- → ETLs run rollup BQ to PG
- → PG and BT serve Platform



Write And Transform (future)

- → PubSub fans out
  - → BT for 1s-res “hot” data
  - → Df -> BT for 60s and 600s “roll-ups”
- → BT feeds BQ
- → ETLs run rollup BQ to PG
- → PG and BT serve Platform



Mapping the Factory

Permissions 1

→ User belongs to an Organization

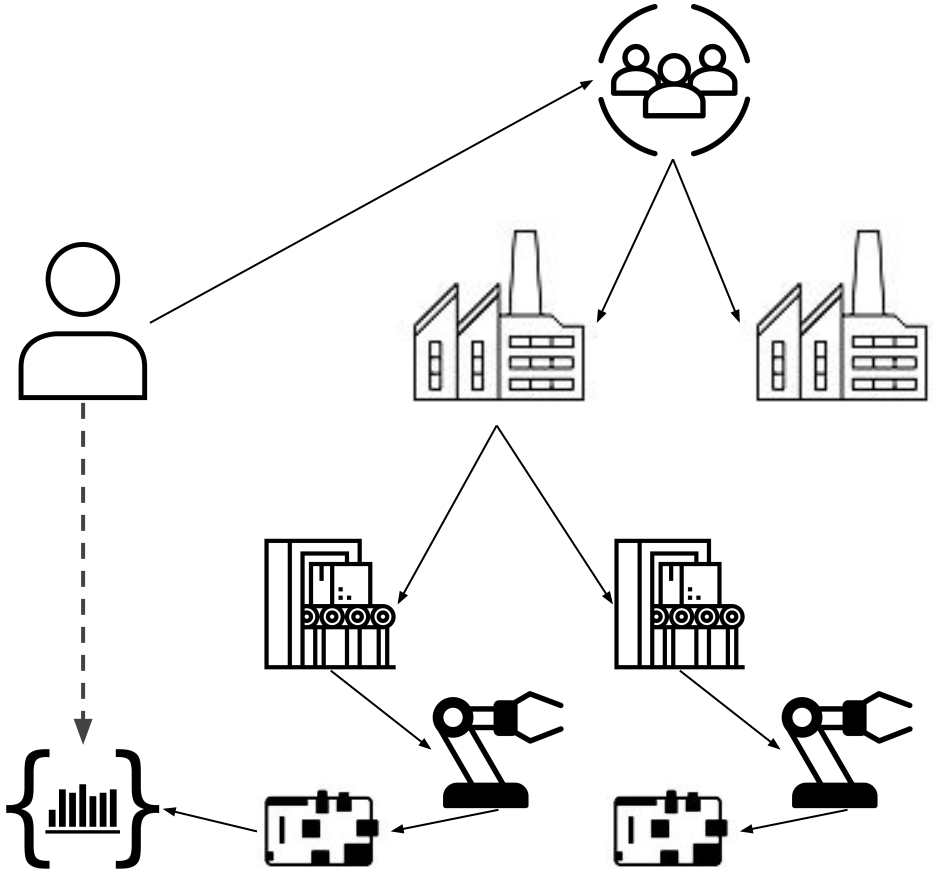
→ Organization owns factories

→ Factories have lines

→ Lines have machines

→ Machines are connected to Oden devices

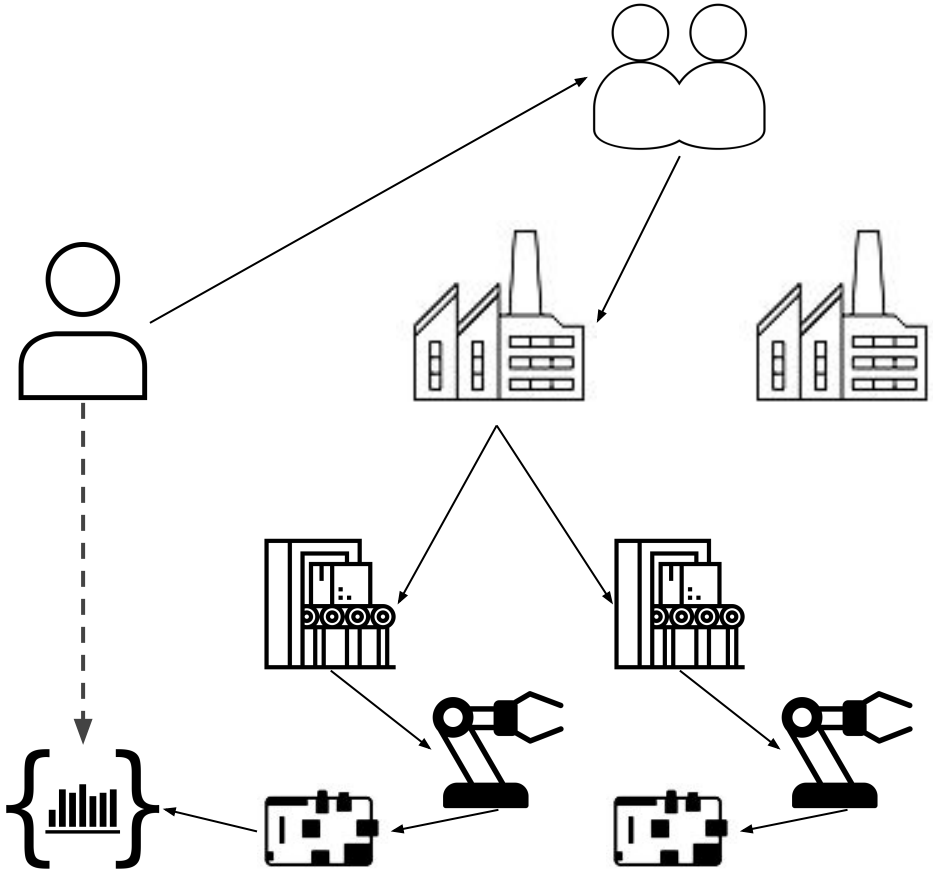
→ Oden devices collect metrics



Permissions 2

→ User belongs to a team

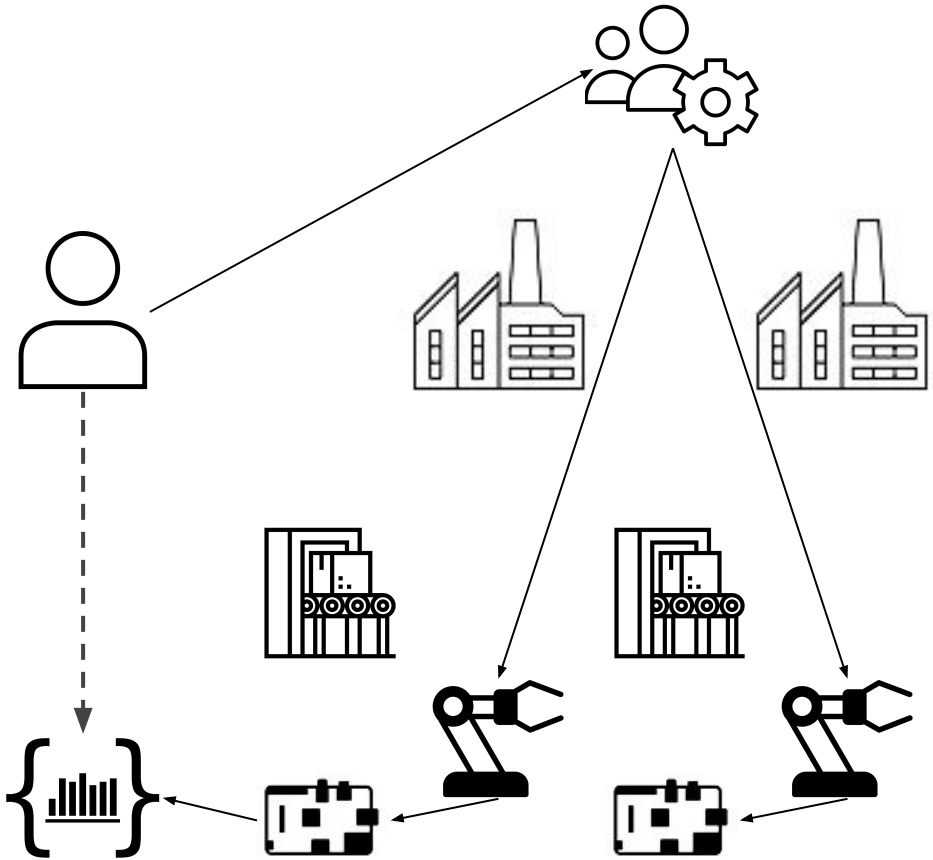
→ Organization can see single a factory



Permissions 3

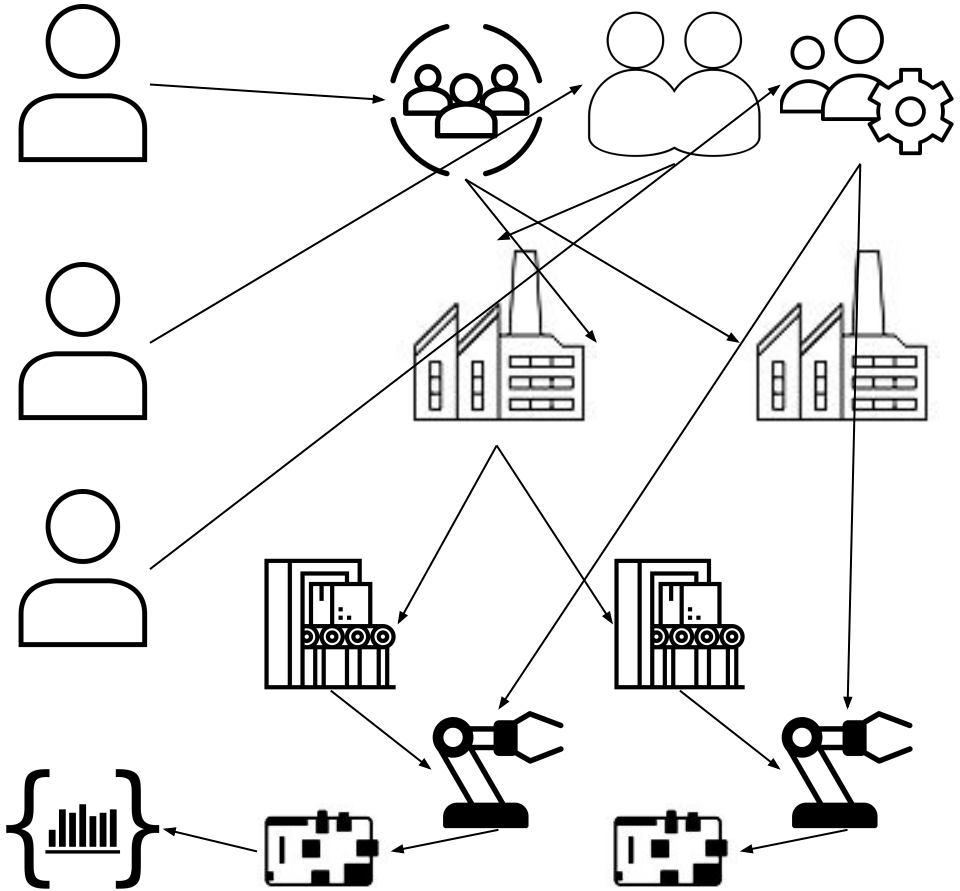
→ User belongs to an Organization

→ Organization maintains machines



Permissions 4

→ All of these relationships live side-by-side simultaneously

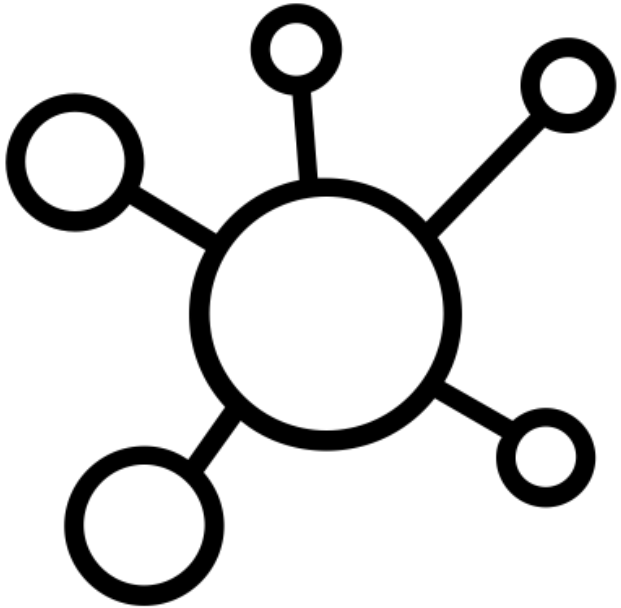


Permissions - Neo4j

→ Answers our arbitrary relationship questions quickly

→ Allows for the currently known relationships between business entities

→ Allows for the currently unknown relationships between business entities



Defining Types of Data

Metric

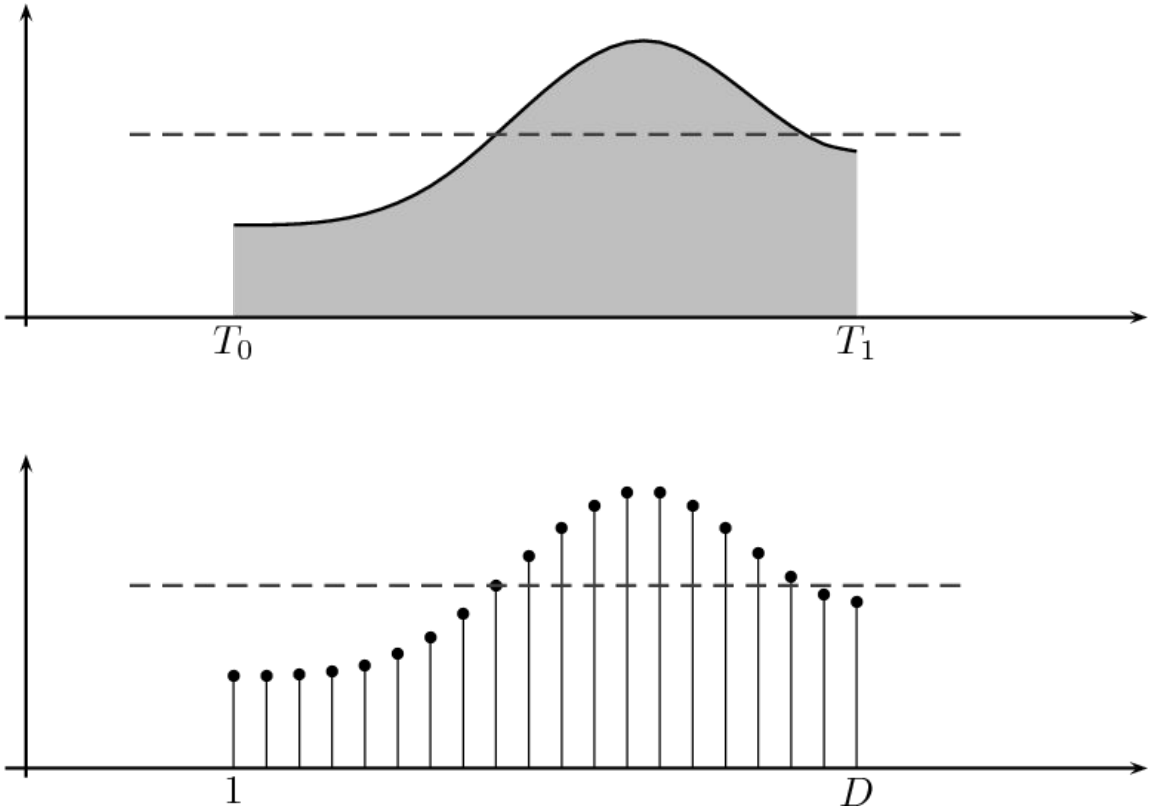
Definition

A metric is an **immutable** floating point value **uniquely** describing the behavior of part of a manufacturing process for some machine, label, and second-resolution timestamp.

Example

“The **melt-temp** on machine **12345** is **4.5** at **2017-12-11T13:30:45**.”

Metric



Interval

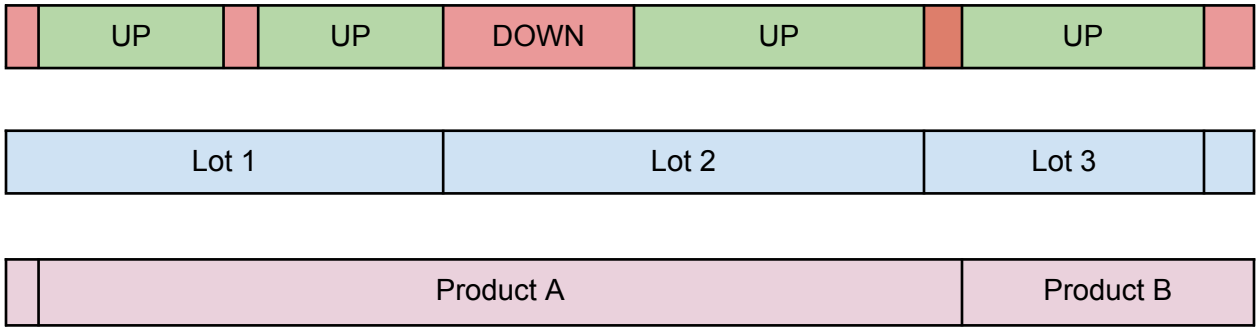
Definition

An interval of some class is a **mutable** and **non-overlapping** time-range that reflects a **business context** for the metrics produced during that time.

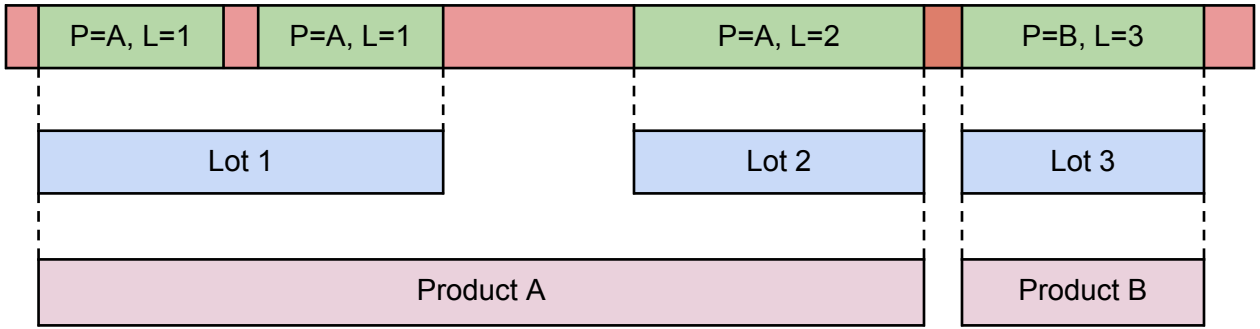
Example

“**Line 13579** is producing **lot 6** of **product A10** during from **2017-12-11T11:00:00** until **2017-12-11T16:20:00**”

Interval 2



Interval 3



Aggregate

Definition

An aggregate is a function of the metrics within an interval that summarizes behavior of the system during that business context.

Example

“When **Line 13579** produced **lot 6** of **product A10** 0.5% of the product produced was not within specification.”

Associative Aggregate

Formal Definition

$(x * y) * z = x * (y * z)$ for all x, y, z in S

Practical Definition

An aggregate function that can be reapplied to include other aggregates.

Example

“Of all lots produced on **Line 13579** today, 0.5% of the product produced was not within specification.”

Associative

Sum

Count

Sums of Squares

Min

Max

First

Last

Non-Associative

Mean*

Mode

Percent*

KS-Statistic

Standard Deviation*

Rate*

C_p C_{pk} C_{pm} C_{pmk}

Associative

Sum

Count

Sums of Squares

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X-\mu)^2}{N}$$

$$\sigma^2 = \frac{1}{n(n-1)} \left(n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_k \right)^2 \right)$$

Non-Associative

Mean*

Mode

Percent*

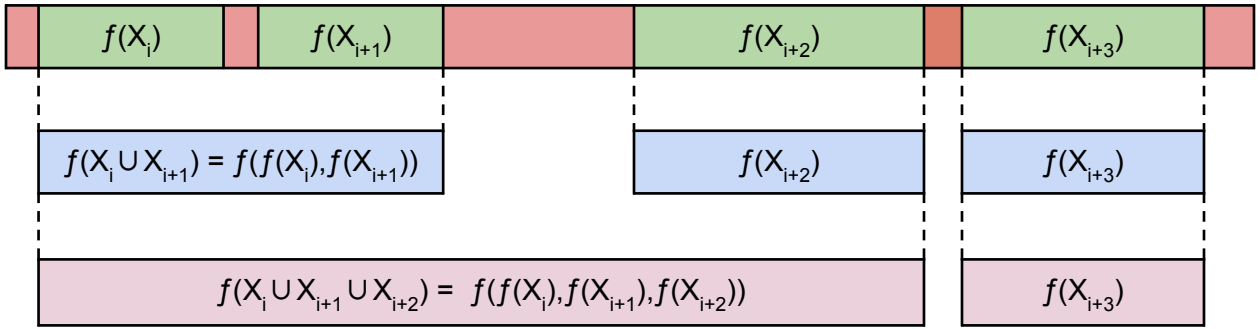
KS-Statistic

Standard Deviation*

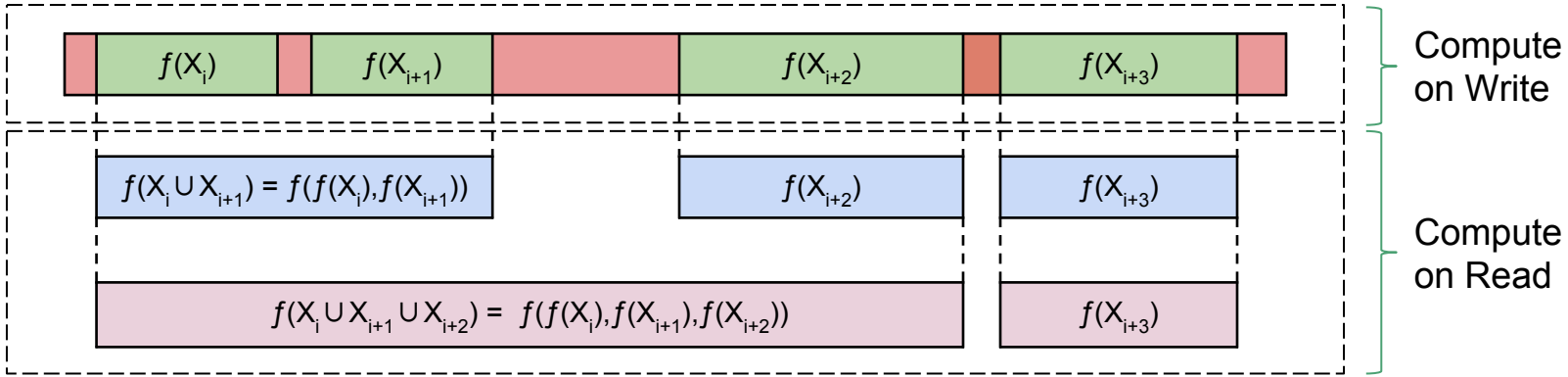
Rate*

C_p C_{pk} C_{pm} C_{pmk}

Associative Aggregate 2



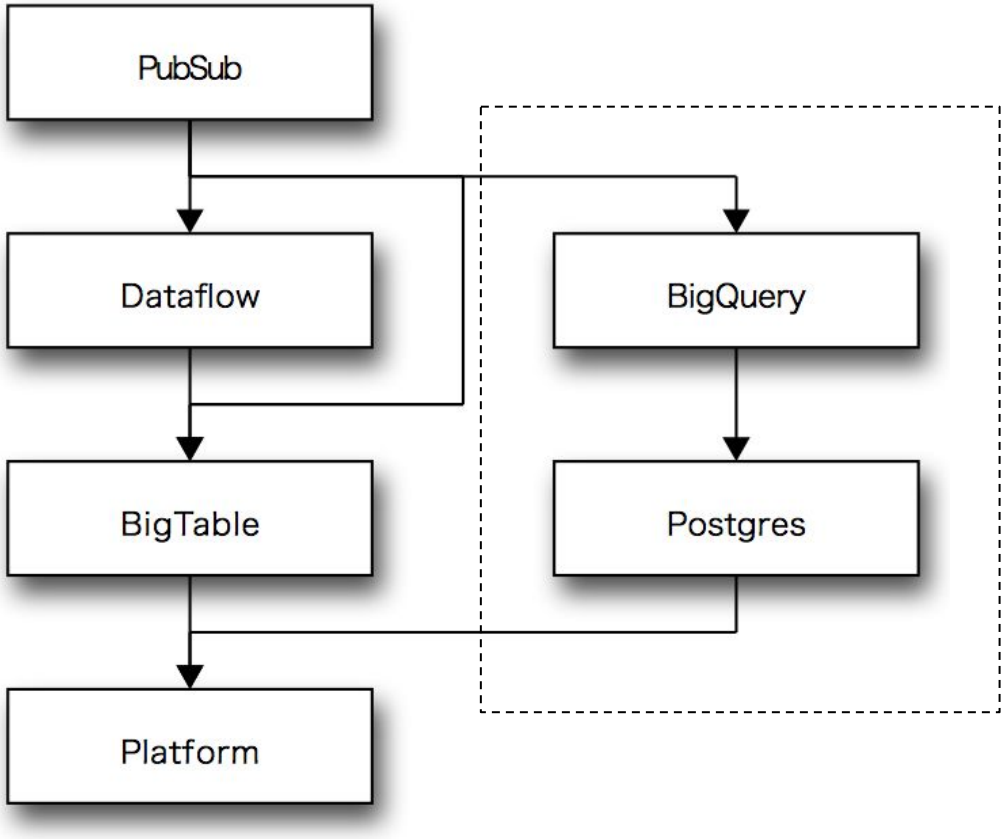
Associative Aggregate 3



Write And Transform (Revisit)

→ Every k minutes visit the last k minutes of “changes”

→ Compute necessary associative aggregates and cache these in Postgres



Spreadsheet View (revisited)

Queries

→ “Which days did I produce the most product in January?”

```
SELECT SUM(rate_produced)
FROM metrics
JOIN days
```

→ “Which products do I produce at the highest quality in terms of being in specification?”

…

Naively From a Traditional Time Series Database

$O(k \times \log(t1 - t0))$

With Precomputed Interval Associative Aggregates

$O(\log(k \times (t1 - t0)))$

How to be “Smart” in Manufacturing

Machine Learning

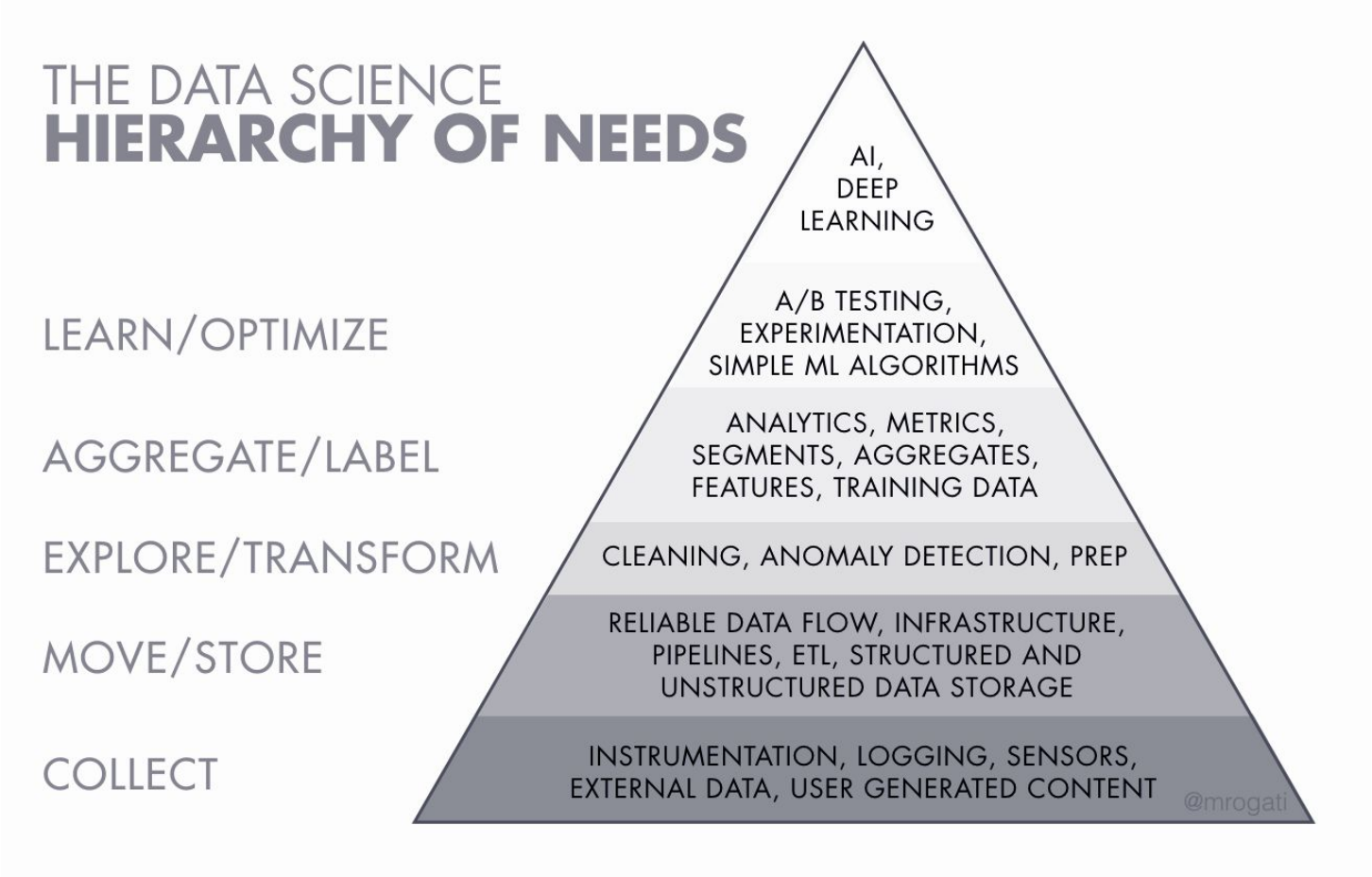
Formal Definition

“A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E." - Tom Mitchell

Informal Definition

“A function with behavior that adapts over time.”

THE DATA SCIENCE HIERARCHY OF NEEDS



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

Where in a factory do we benefit from learning?

Probably Should Not Learn

Metric

Interval

Associative Aggregate

Could Benefit From Learning

Non-Associative Aggregate

Configuration

Attention Gating

Probably Should Not Learn

Metric

Interval

Associative Aggregate

Could Benefit From Learning

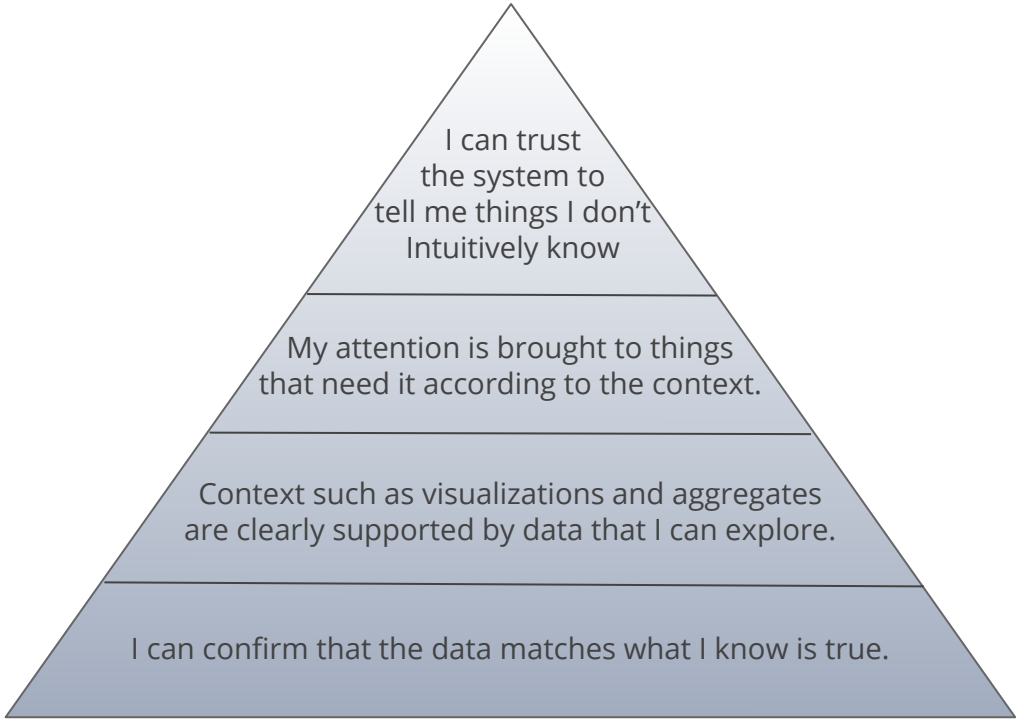
Non-Associative Aggregate

Configuration

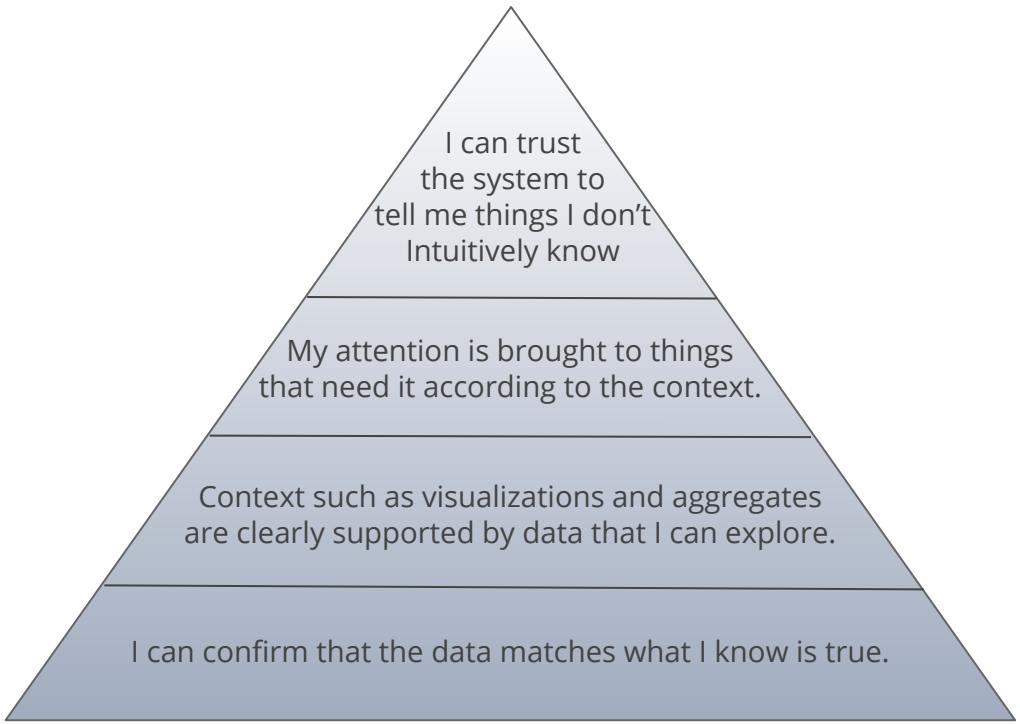
Attention Gating

How should I, as a worker, most optimally spend my time today?

The Manufacturing Analytics **Hierarchy of Trust**



The Manufacturing Analytics **Hierarchy of Trust**



Time is spent on highest value work

Devon Peticolas
devon@oden.io
@dproi