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
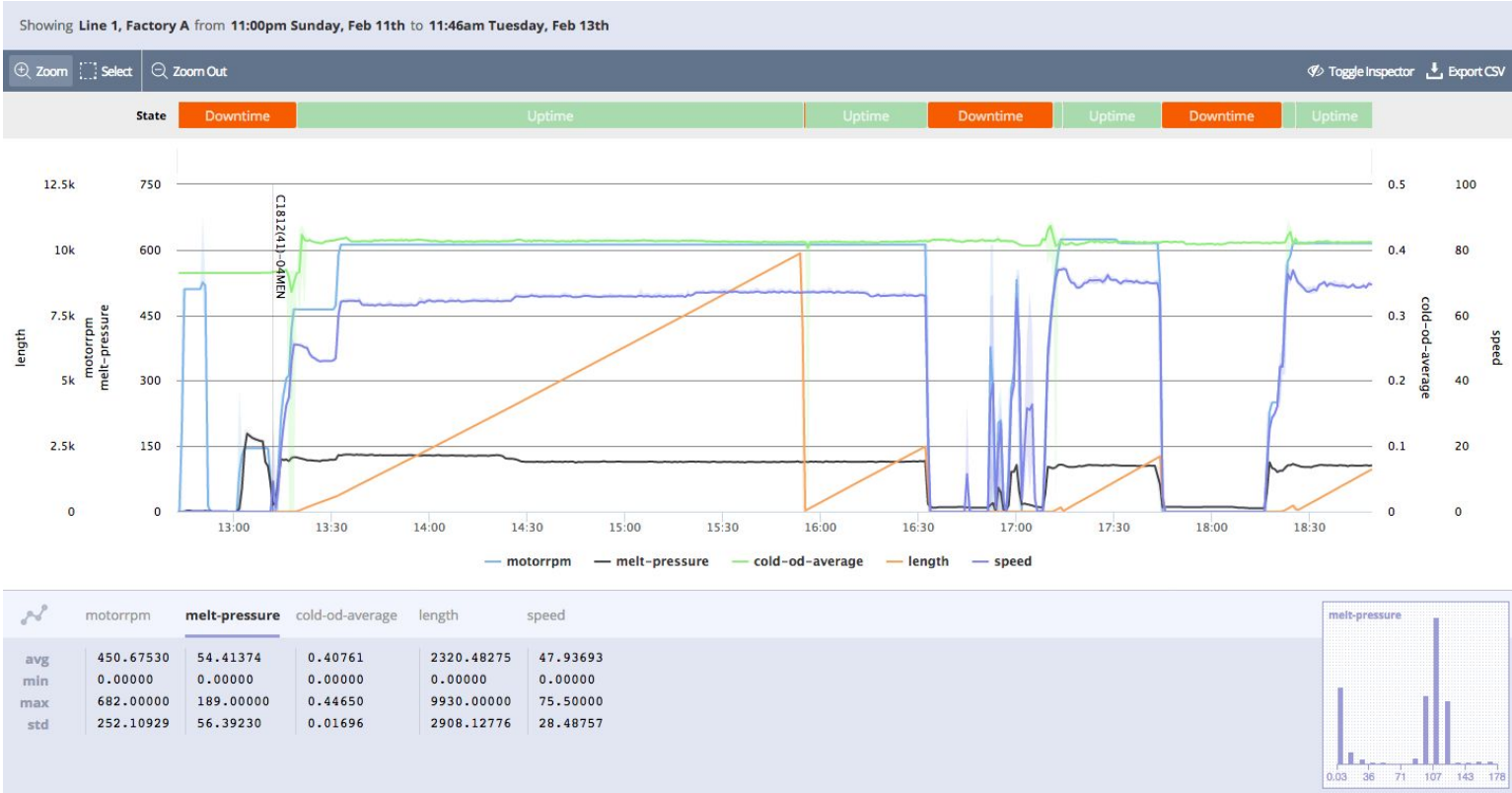~ McKinsey, 2016

The Product

Real Time Overview

Demo Factory

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

Time Series View

Spreadsheet View

Run

📅 Tue Jan 16th - Tue Feb 13th

☀ All week • 12am - 12am
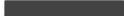
Hide Filters

Product ⌃

All Products

State ⌄
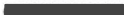
Annotations ⌄

Lines 🟢 ⌄

Search

Product ↓	Line	Annotations	States	From	To	Duration	1-machine-speed (average)	1-total-batch-part-counter-low-acc (deltasum)	➕
▮ / Demo Factory	▮	None	Insert Broken Uptime Downtime	Fri Jan 12th 11:44pm	Tue Jan 16th 3:06am	3d 3h	64.845 ◇ 125 ± 3	159,779	
▮ / Demo Factory	▮	None	Uptime Downtime	Tue Jan 30th 10:19am	Still active	13d 13h	84.215 ◇ 130 ± 3	1,005,754	
▮ / Demo Factory	▮	None	Uptime Downtime Setup/Changeover	Mon Jan 29th 12:59am	Tue Jan 30th 10:19am	33h 19m	62.959 ◇ 130 ± 3	81,001	
▮ / Demo Factory	▮	None	Uptime Downtime Setup/Changeover Punch Pin	Thu Jan 18th 9:53pm	Mon Jan 22nd 11:07am	3d 13h	39.237 ◇ 130 ± 3	122,604	
▮ ⌃ / Demo Factory	▮	None	Setup/Changeover Uptime Downtime No Tubs Available	Mon Jan 22nd 11:07am	Wed Jan 24th 3:35am	40h 28m	118.844 ◇ 120 ± 3	156,948	
▮ ⌃ / Demo Factory	▮	None	Setup/Changeover Uptime	Tue Jan 16th 12:27pm	Thu Jan 18th 11:42pm	2d 11h	118.144 ◇ 110 ± 3	247,145	

Reporting View

Runs completed yesterday that deserve your attention
Sorted lowest CPK to highest

Run	Started	Uptime	CPK	Nonconforming	
▮▮▮▮▮▮ Thermoformer 10	12:13am Yesterday	3h 37m	0.007	**6.59%** ↑0.11% ↓6.48%	›
▮▮▮▮▮▮P Header 22	8:13am Sunday	32m	0.233	**3.20%** ↑3.13% ↓0.07%	›
▮▮▮▮▮ Header 22	8:44pm Yesterday	1h 25m	0.557	**2.99%** ↑0.02% ↓2.97%	›
⋮ 23 additional runs					
▮▮▮▮▮▮ Header 22	8:44pm Yesterday	1h 25m	1.033	**0.94%** ↑0.63% ↓0.31%	›
▮▮▮▮▮ Header 6	11:29pm Friday	59h 35m	1.877	**0.37%** ↑0.37% ↓0.00%	›
▮▮▮▮▮▮▮▮ Header 6	10:20pm Yesterday	6h 8m	2.096	**0.25%** ↑0.08% ↓0.17%	›

See all runs from yesterday

Data Acquisition

Acquisition 1

→ Requests metrics from PLC

→ PLC returns values at t1

→ Enqueues to disk-backed Redis w/ time t1
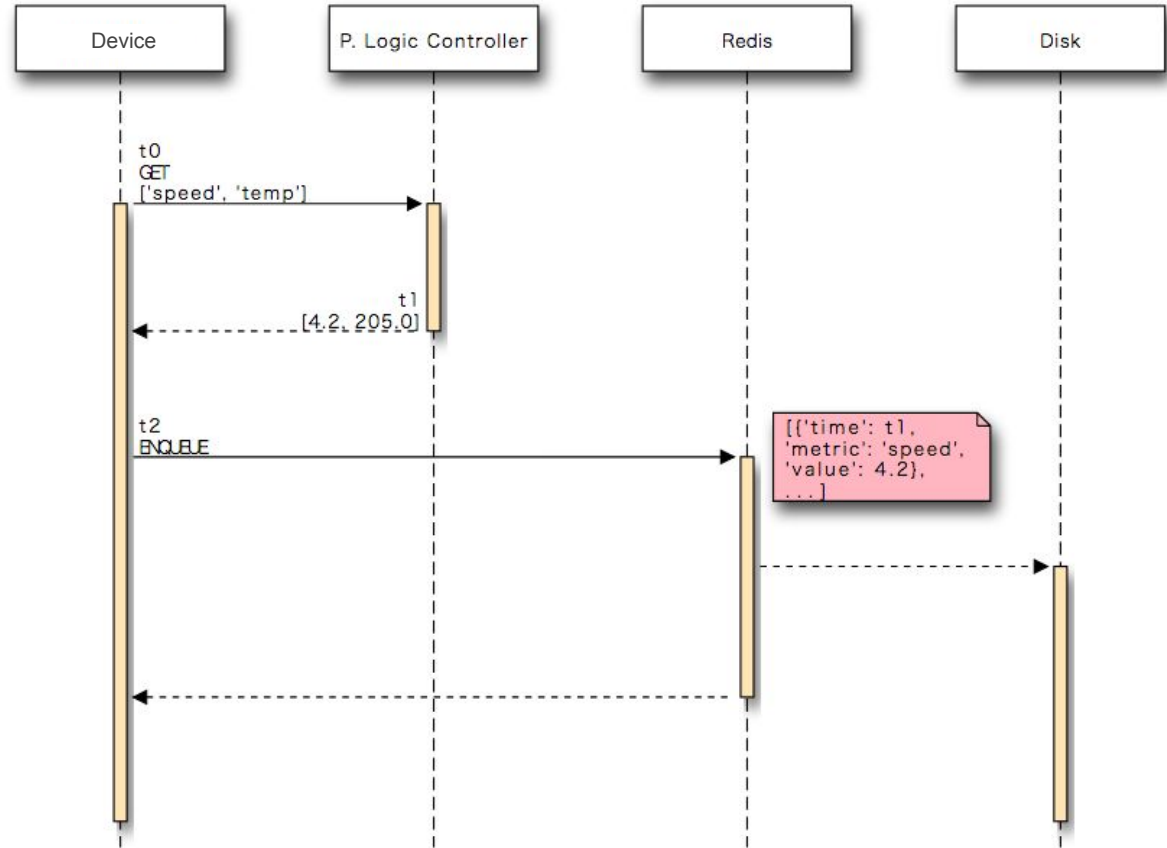
Acquisition 2

→ Dequeues from Redis

→ Publishes to PubSub
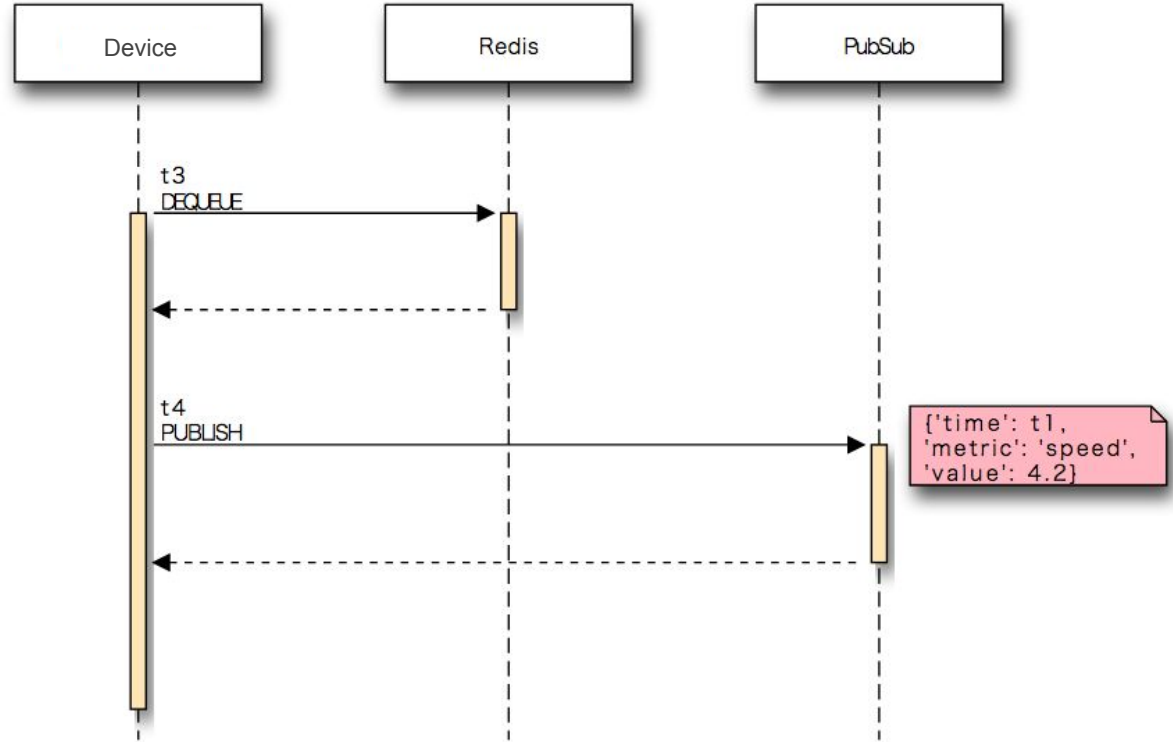
Write And Transform 1

→ 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
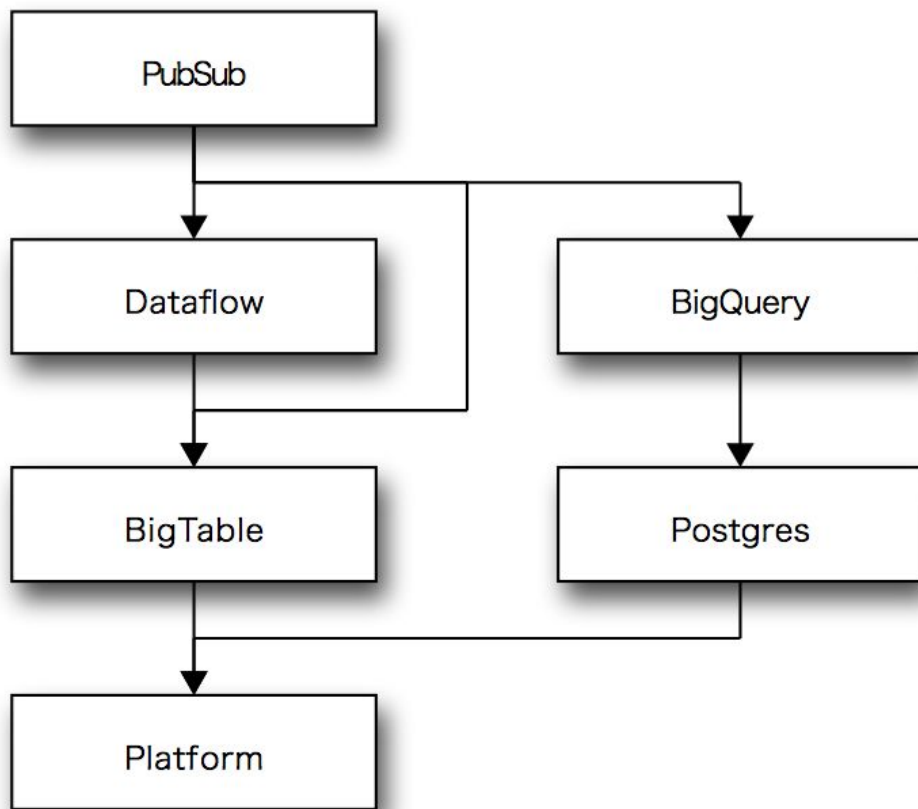
Write And
Transform (future)

→ 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

```
┌──────────┐
│  PubSub  │
└──────────┘
      │
      ▼
┌──────────┐
│ BigTable │
└──────────┘
      │
      ▼
┌──────────┐
│ BigQuery │
└──────────┘
      │
      ▼
┌──────────┐
│ Postgres │
└──────────┘
      │
      ▼
┌──────────┐
│ Platform │
└──────────┘
```
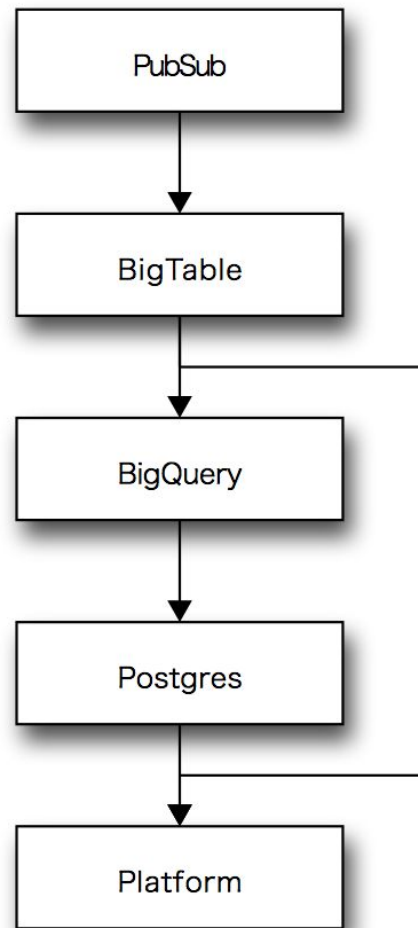
Mapping the Factory

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
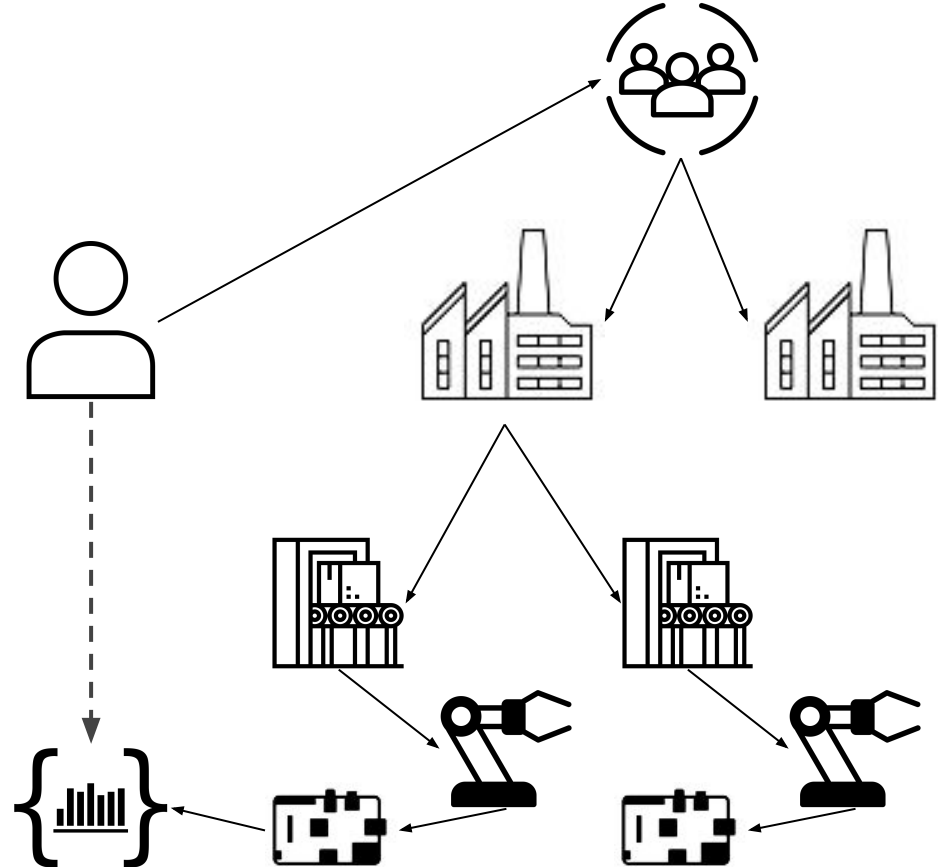
Permissions 2

→ User belongs to a team

→ Organization can see single a factory
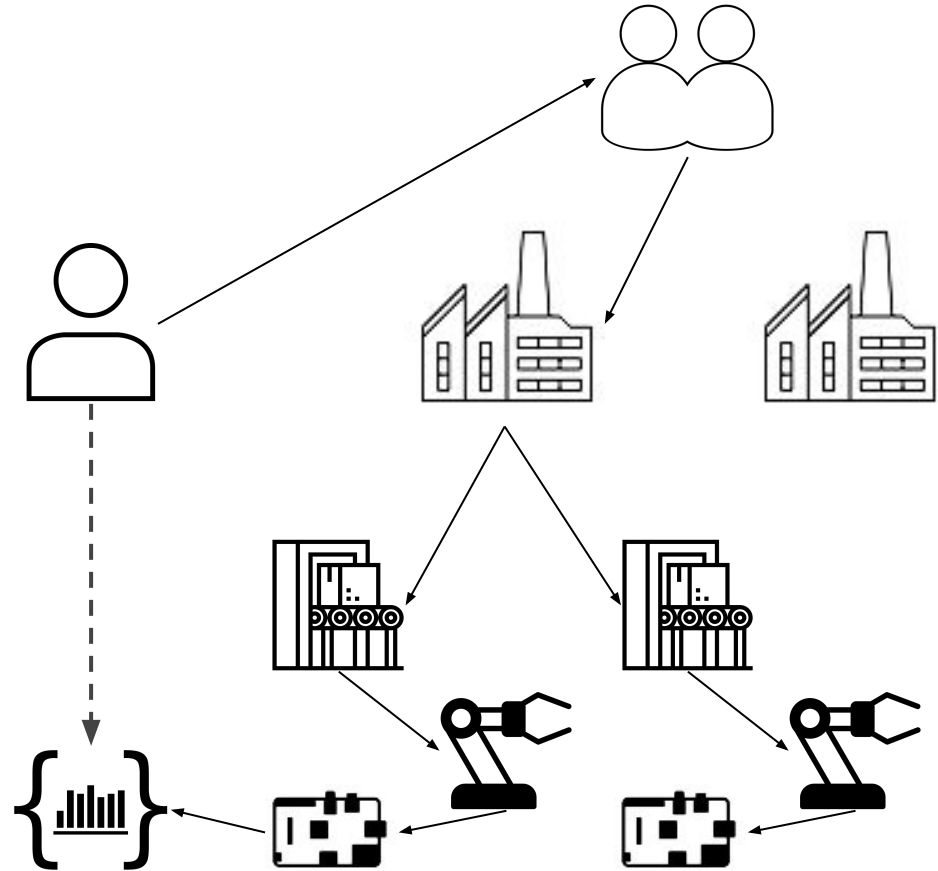
Permissions 3

→ User belongs to an Organization

→ Organization maintains machines
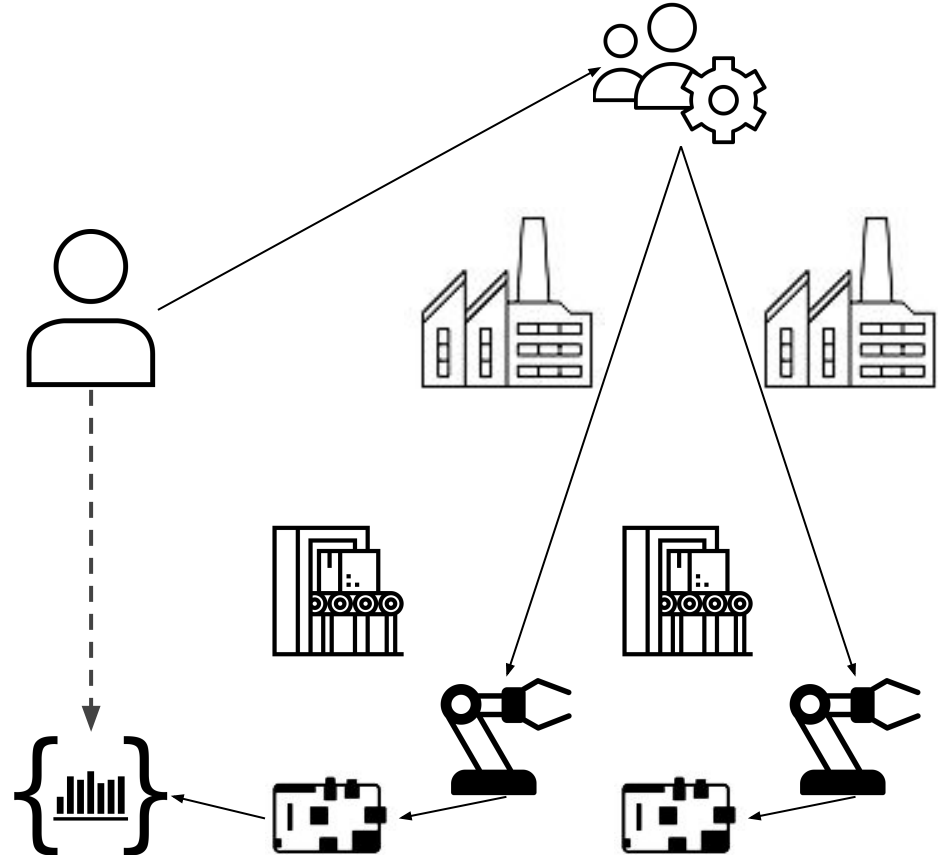
Permissions 4

→ All of these relationships live
side-by-side simultaneously
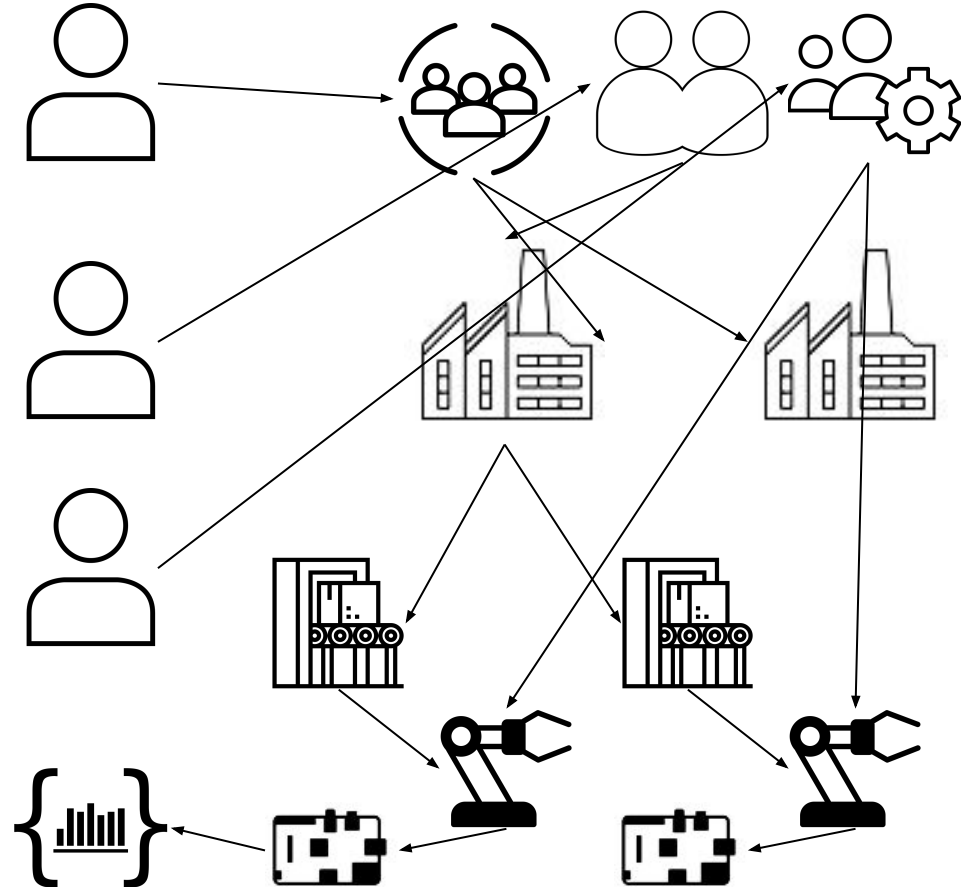
Permissions - Neo4j

→ Answers our arbitrary relationship questions quickly

→ Allows for the currently known relationships between business entities

→ Allows for the currently unknown relationships between business entities
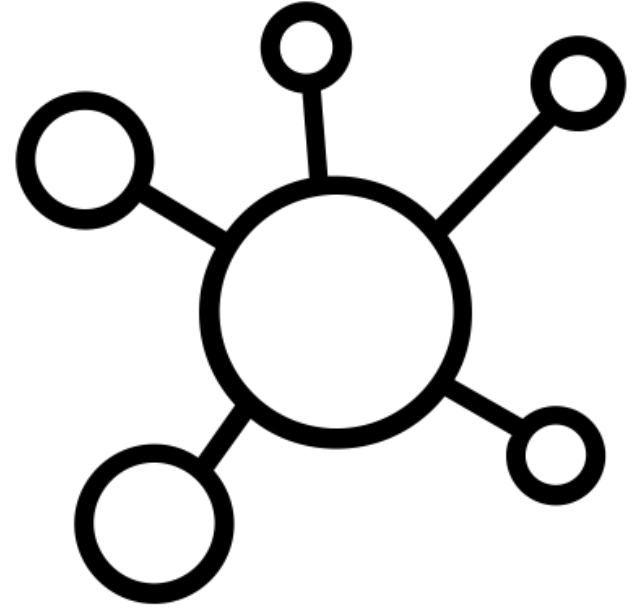
Defining Types of Data

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
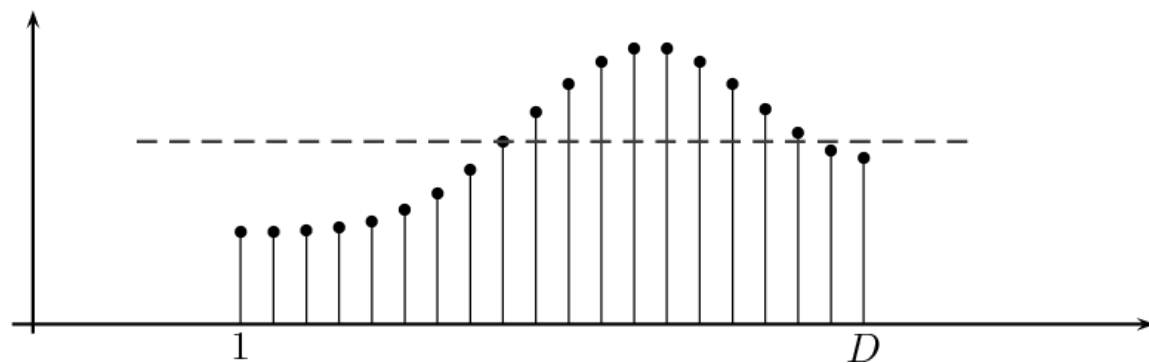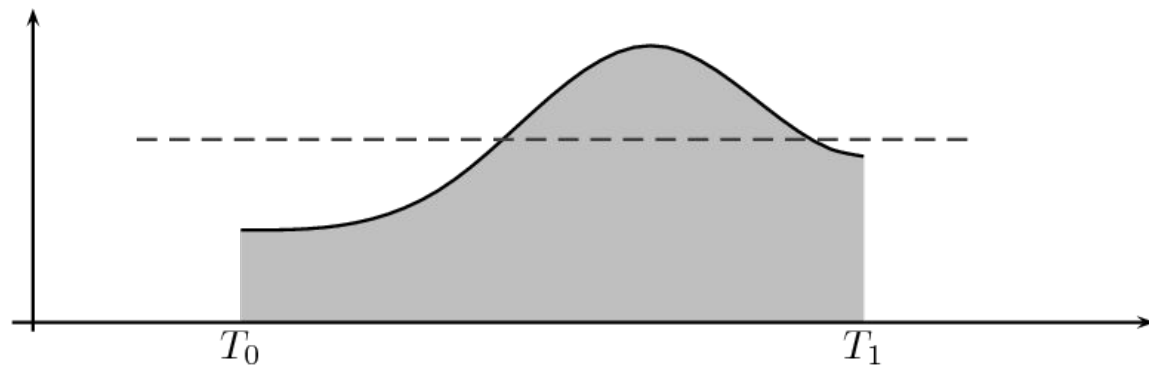
Interval

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
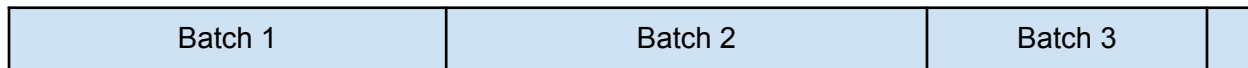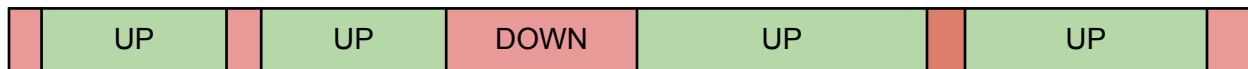
Interval 3
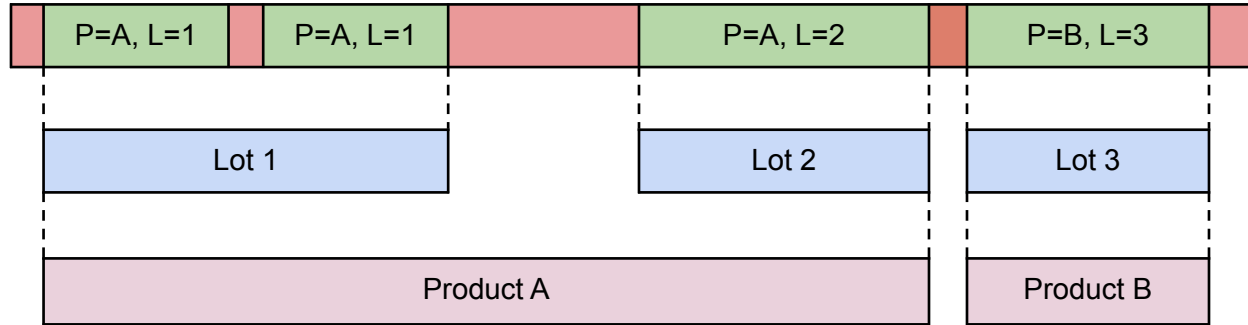
Aggregate

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

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

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

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)

Search
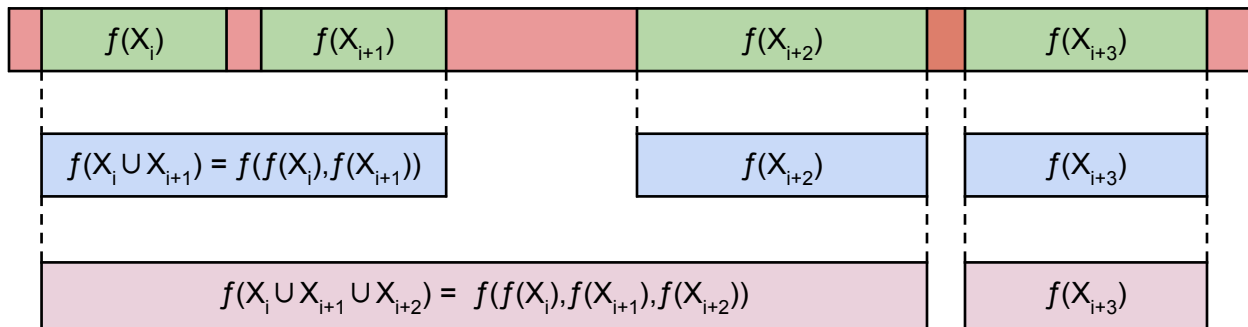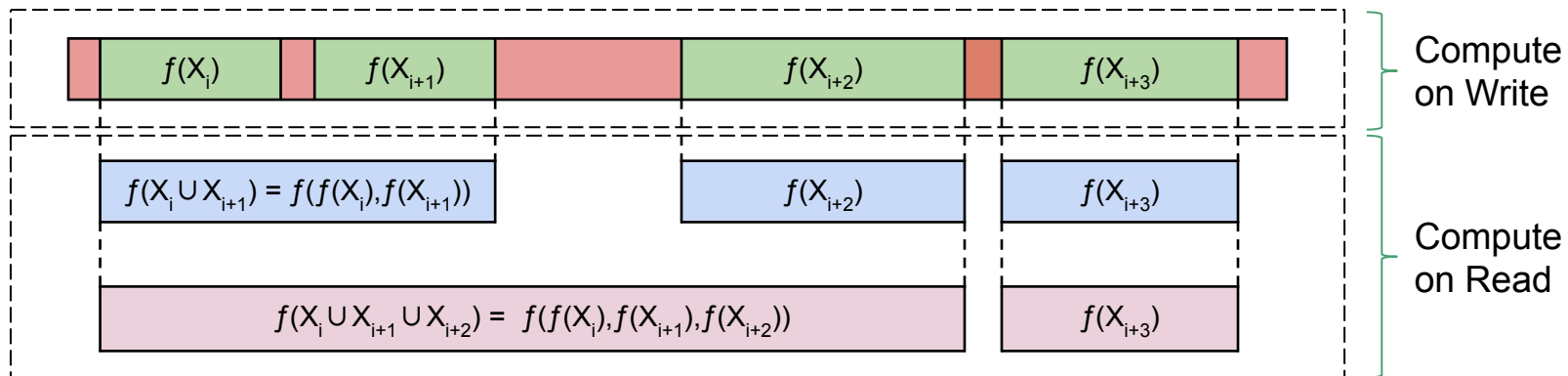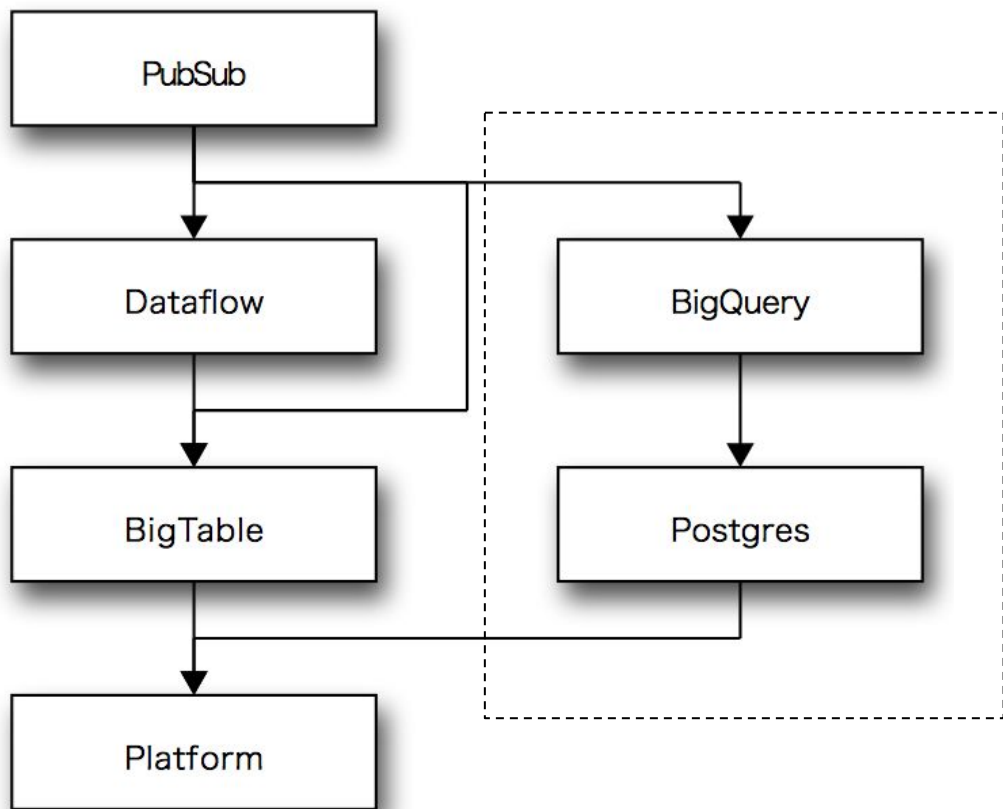
		Product ↓	Line	Annotations	States	From	To	Duration	1-machine-speed (average)	1-total-batch-part-counter-low-acc (deltasum)
Run ▼		▬▬▬	▬▬▬ Demo Factory	None	Insert Broken Uptime Downtime	Fri Jan 12th 11:44pm	Tue Jan 16th 3:06am	3d 3h	64.845 ⟷ 125 ± 3	159,779
📅 Tue Jan 16th - Tue Feb 13th		▬▬▬	▬▬▬ Demo Factory	None	Uptime Downtime	Tue Jan 30th 10:19am	Still active	13d 13h	84.215 ⟷ 130 ± 3	1,005,754
☀ All week • 12am - 12am		▬▬▬	▬▬▬ Demo Factory	None	Uptime Downtime Setup/Changeover	Mon Jan 29th 12:59am	Tue Jan 30th 10:19am	33h 19m	62.959 ⟷ 130 ± 3	81,001
Hide Filters		▬▬▬	▬▬▬ Demo Factory	None	Uptime Downtime Setup/Changeover Punch Pin	Thu Jan 18th 9:53pm	Mon Jan 22nd 11:07am	3d 13h	39.237 ⟷ 130 ± 3	122,604
Product ⌃		▬▬▬	▬▬▬ Demo Factory	None	Setup/Changeover Uptime Downtime No Tubs Available	Mon Jan 22nd 11:07am	Wed Jan 24th 3:35am	40h 28m	118.844 ⟷ 120 ± 3	156,948
All Products		▬▬▬	▬▬▬ Demo Factory	None	Setup/Changeover Uptime	Tue Jan 16th 12:27pm	Thu Jan 18th 11:42pm	2d 11h	118.144 ⟷ 110 ± 3	247,145

State ⌄
Annotations ⌄
Lines ● ⌄

Search

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

AI, DEEP LEARNING

LEARN/OPTIMIZE — A/B TESTING, EXPERIMENTATION, SIMPLE ML ALGORITHMS

AGGREGATE/LABEL — ANALYTICS, METRICS, SEGMENTS, AGGREGATES, FEATURES, TRAINING DATA

EXPLORE/TRANSFORM — CLEANING, ANOMALY DETECTION, PREP

MOVE/STORE — RELIABLE DATA FLOW, INFRASTRUCTURE, PIPELINES, ETL, STRUCTURED AND UNSTRUCTURED DATA STORAGE

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

@mrogati

Monica Rogati
https://hackernoon.com/the-ai-hierarchy-of-needs-18f111fcc007
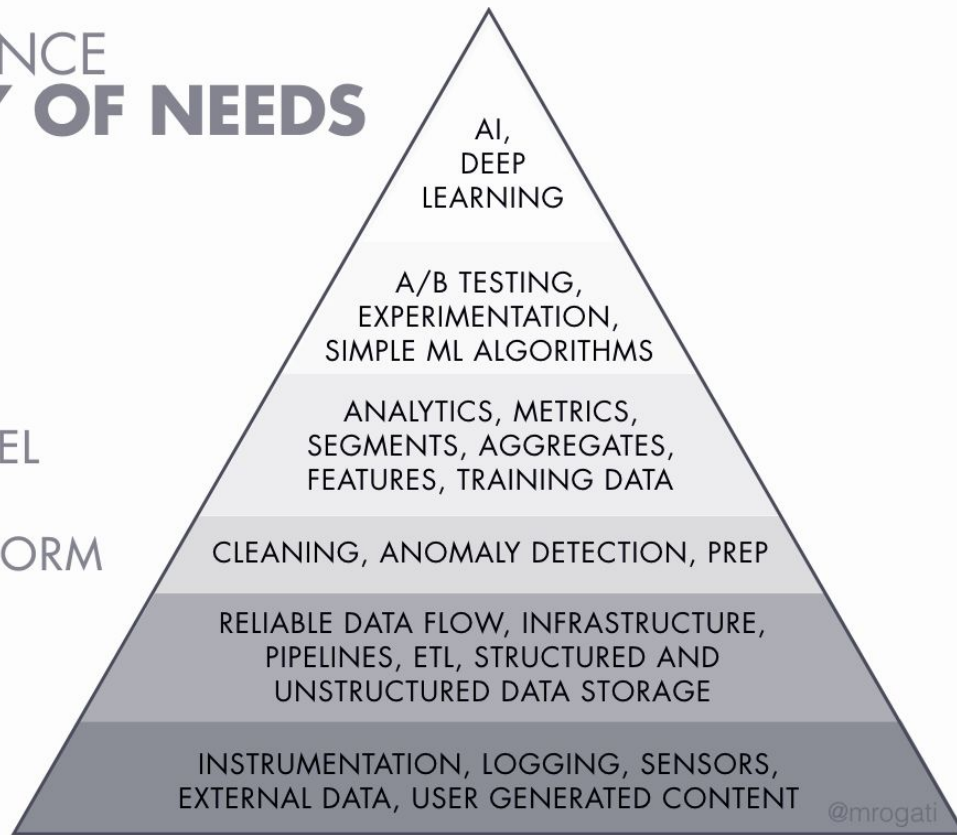
Where in a factory do we benefit from learning?

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

I can trust
the system to
tell me things I don't
Intuitively know

My attention is brought to things
that need it according to the context.

Context such as visualizations and aggregates
are clearly supported by data that I can explore.

I can confirm that the data matches what I know is true.

The Manufacturing Analytics **Hierarchy of Trust**

I can trust
the system to
tell me things I don't
Intuitively know

My attention is brought to things
that need it according to the context.

Context such as visualizations and aggregates
are clearly supported by data that I can explore.

I can confirm that the data matches what I know is true.

Time is spent on highest value work

Devon Peticolas
devon@oden.io
@dproi