SINTEF

# AI in the process industry

## Why, how and examples of use

Sølve Eidnes, SINTEF Digital

November 22nd 2022

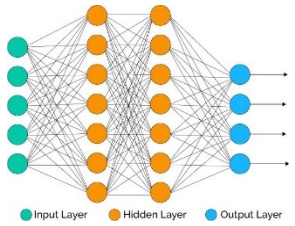
Technology for a better society

# Enabling technologies in a Digital System

- Industrial robotics
- Collaboration
- Drones

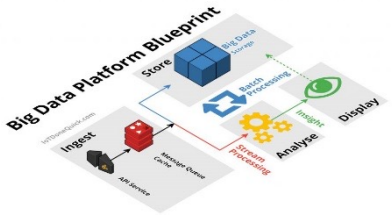
## Autonomy

## AI




- Industrial AI
- Hybrid Analytics
- Explainable AI

## Big Data




- Platforms
- Secure Interopability
- Heterogenous data

## IoT

- Sensors
- Connectivity
- 5G

SINTEF

# AI in the industry

## The examples you hear of…

## … and the real life challenges

### Internet data and Images

- Daily interactions
- Big volume of data
- Structured data
- No noise

### Games

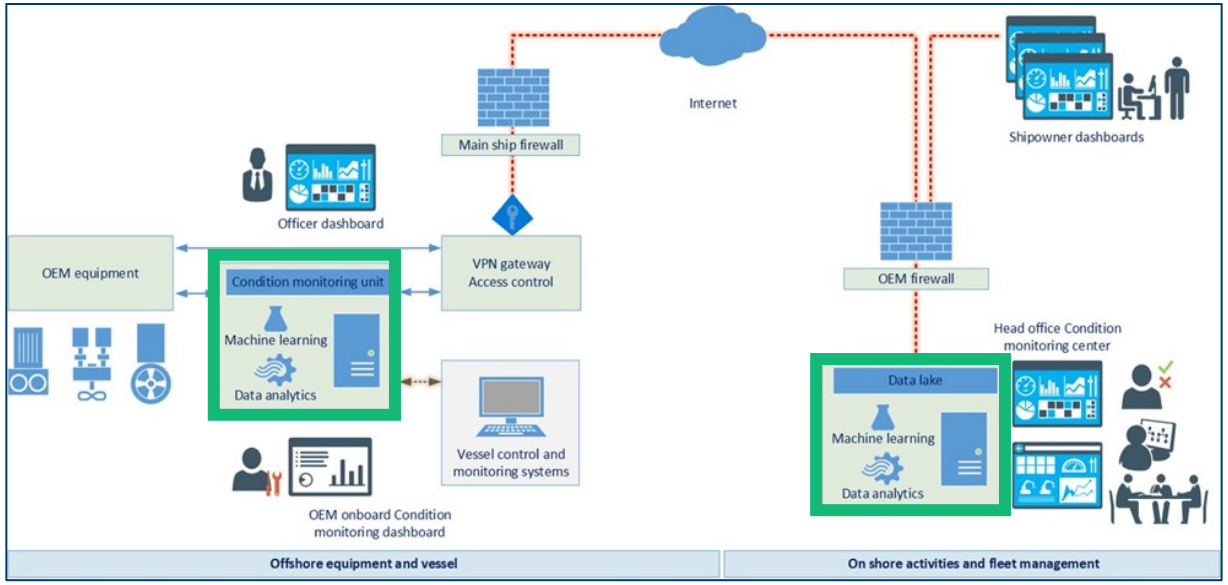
- Clear rules
- Limited environment
- Defined states
- Breakthroughs well covered in media


### Industrial data

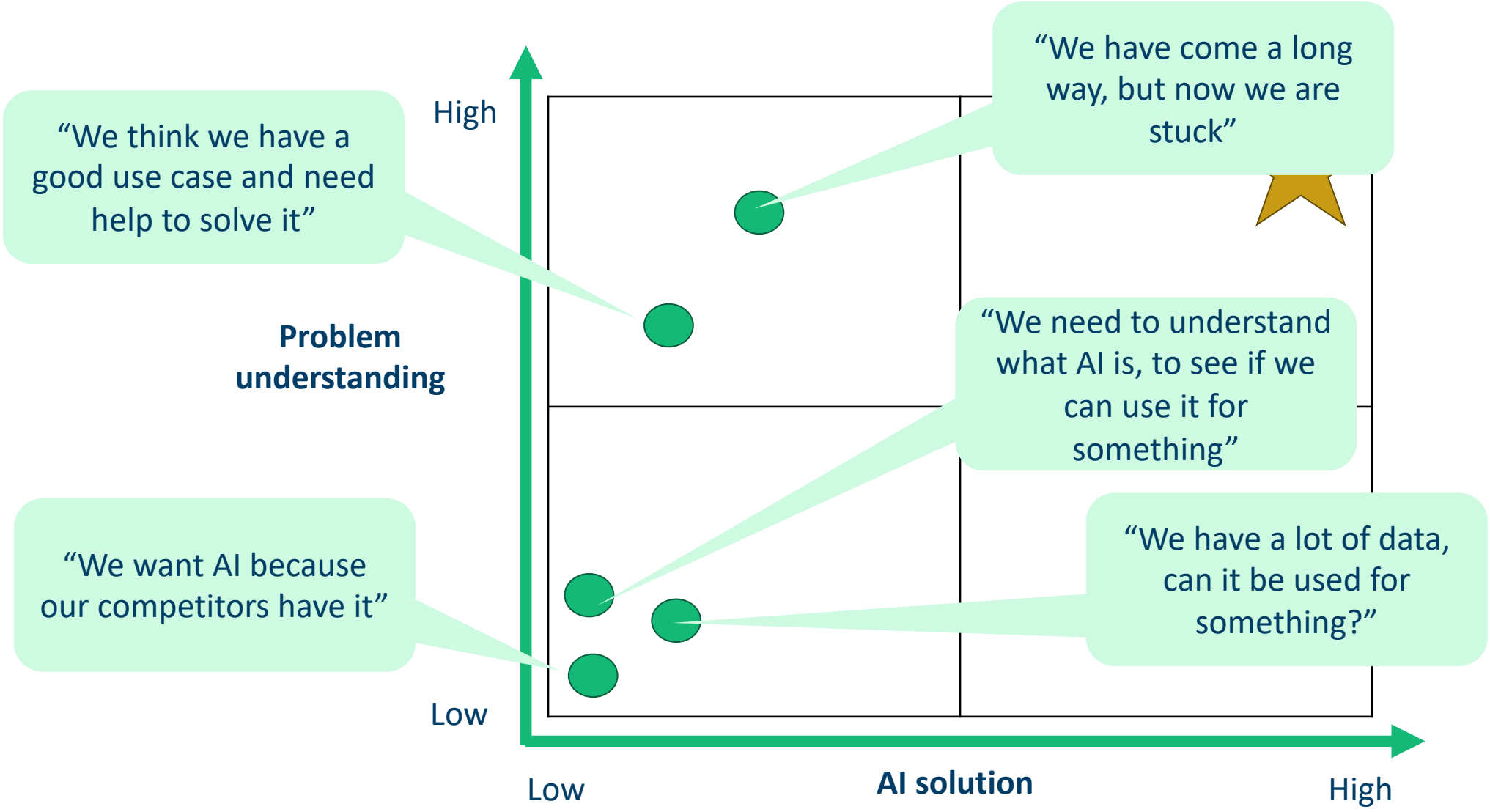
- Old and complex systems
- Limited standardization
- Often poor data quality
- Safety critical systems

# Simplified workstream

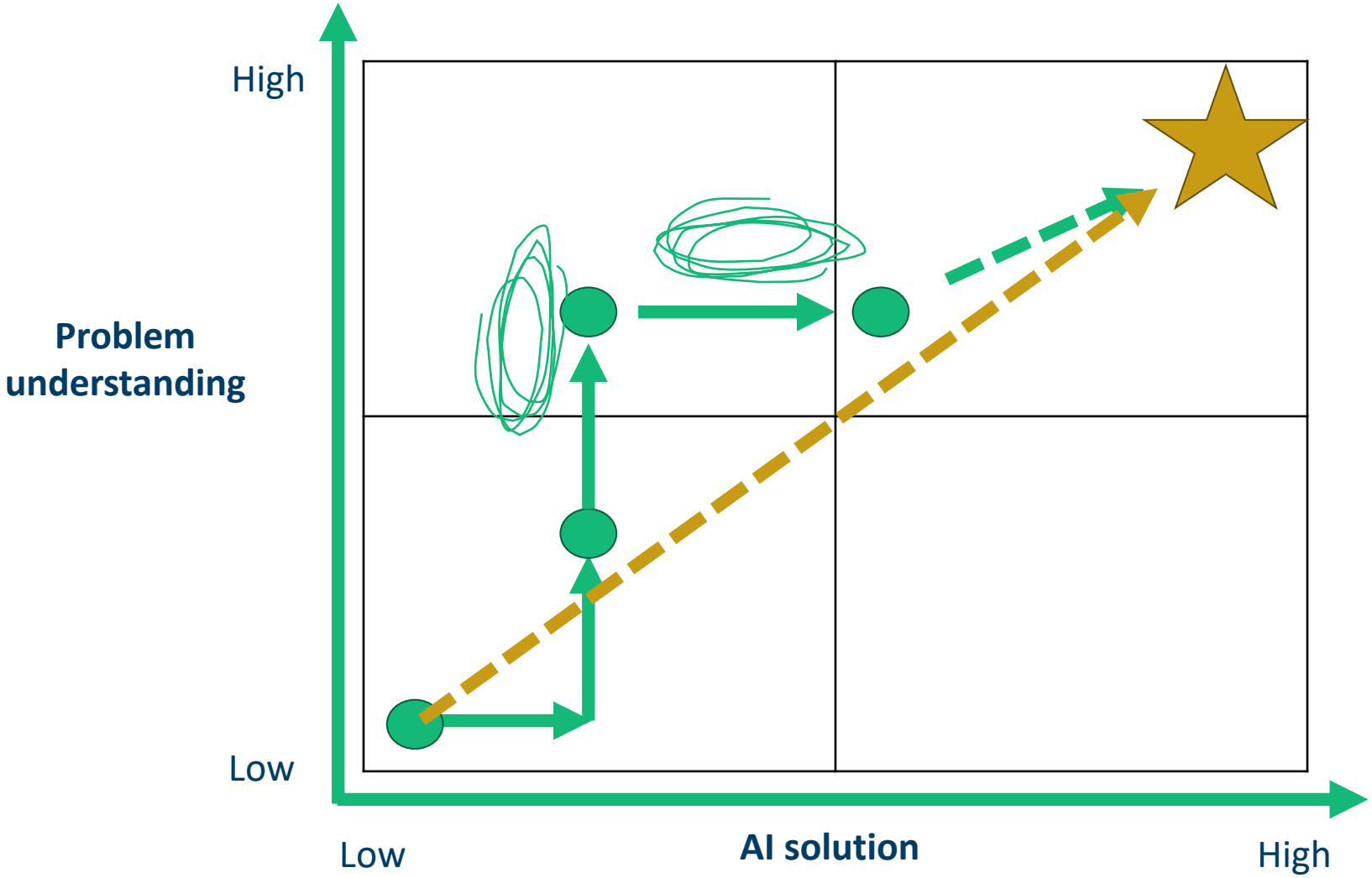
Problem definition → Data → AI modelling → Implementation



# Customers are curious about AI...



# …but it takes time to find the right case and solution



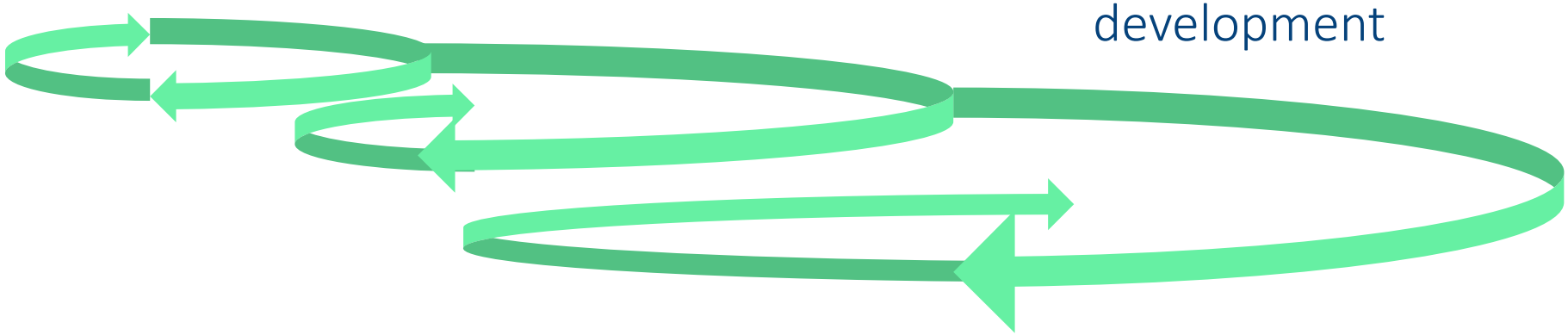
# Iterate problem definition continuously

**Problem definition** → **Data** → **Modelling / AI solution** → **Implementation**

- Business value
- Feasibility

- Data availability
- Quality

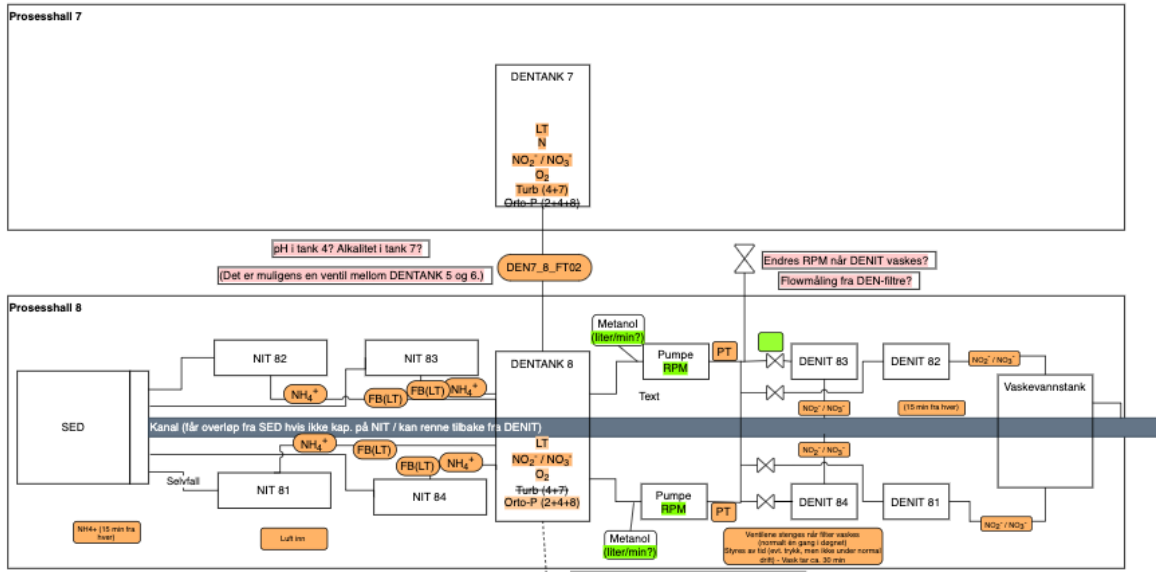
- Method selection
- Model development



# Problem definition

## Example: Waste water treatment

- Made flow charts to get an overview of the process
- Chose a specific first case
- Changed the case definition based on early analysis and better understanding of the process




# Data challenges

## Challenges in getting data:

- Missing data
  - Parameters not measured or recorded
  - Measuring the wrong parameters
  - Not possible to measure
- Messy data, unlabelled or badly labelled
- Data not owned by customer

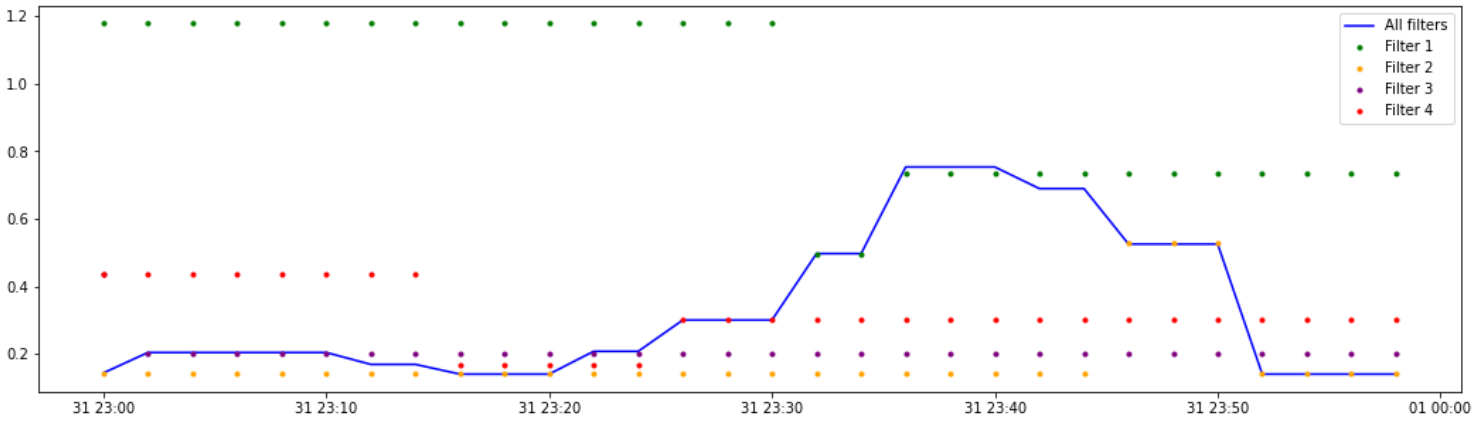
## Challenges when having data:

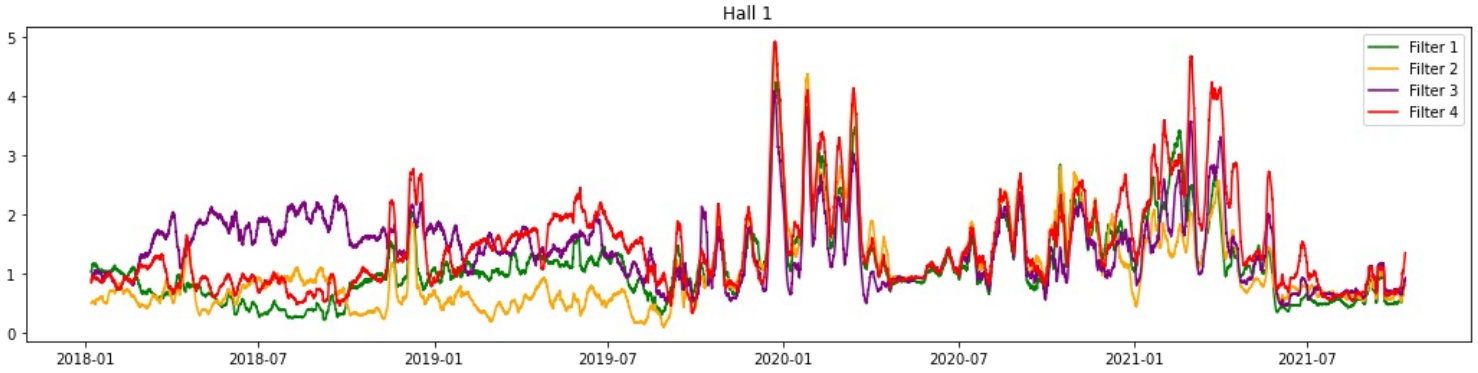
- Relating the data to the process
- Not sufficient data
  - Anomaly detection: enough anomalies?
- Data from model or measurements?
- Extrapolation vs interpolation

# Data challenges

## Example: Waste water treatment

- Output data is actually from four different processes
- Recorded data switches between filters approx. every 15 minutes
- Sampling frequency 10 minutes
- The biological conditions vary



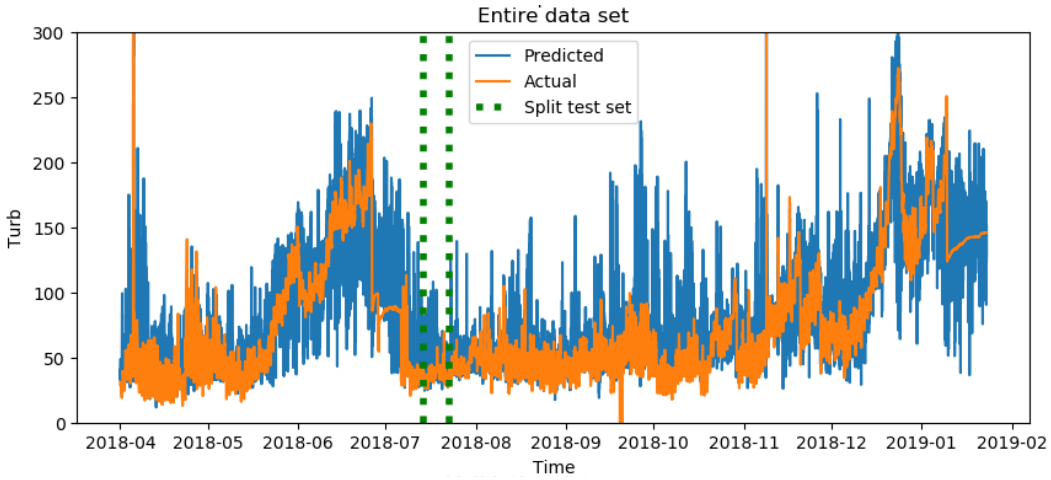
# Data challenges

## Example: Waste water treatment

Estimate solids content in fluid

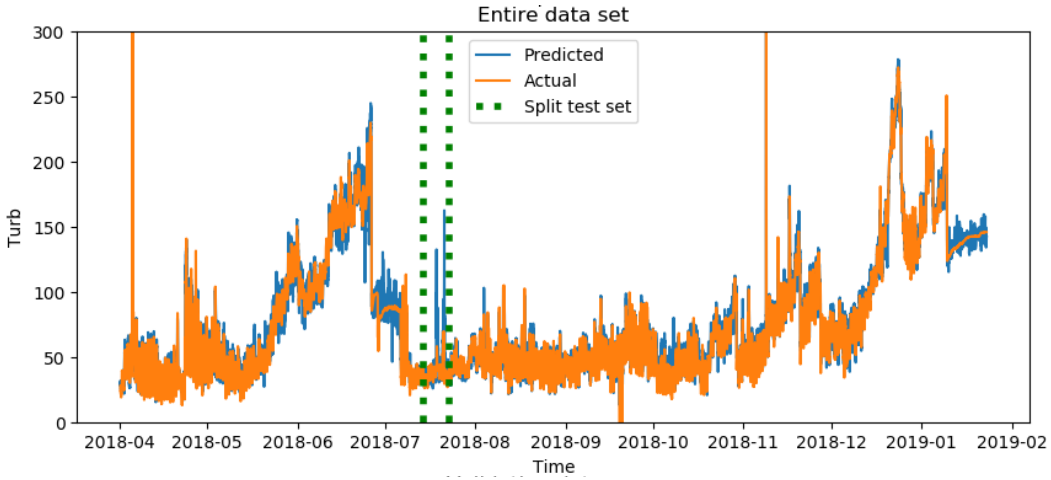
First result:

- A lot of noise in prediction
- Hypothesis: Residence time in the cleaning tanks




Second result:

- Input is from last 8 hours
- Feedback from sewage treatment is important
- Big improvement in prediction

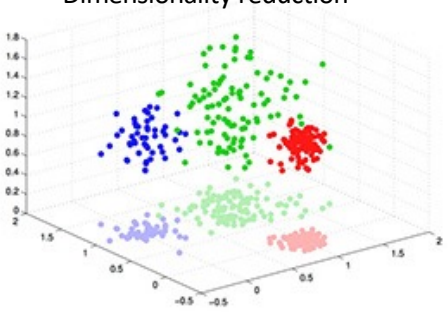



See Moe, Signe,, et al. ”Neural network to analyze wastewater treatment plant with CEPT”

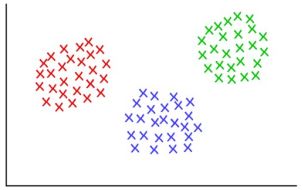
# Appropriate use of ML

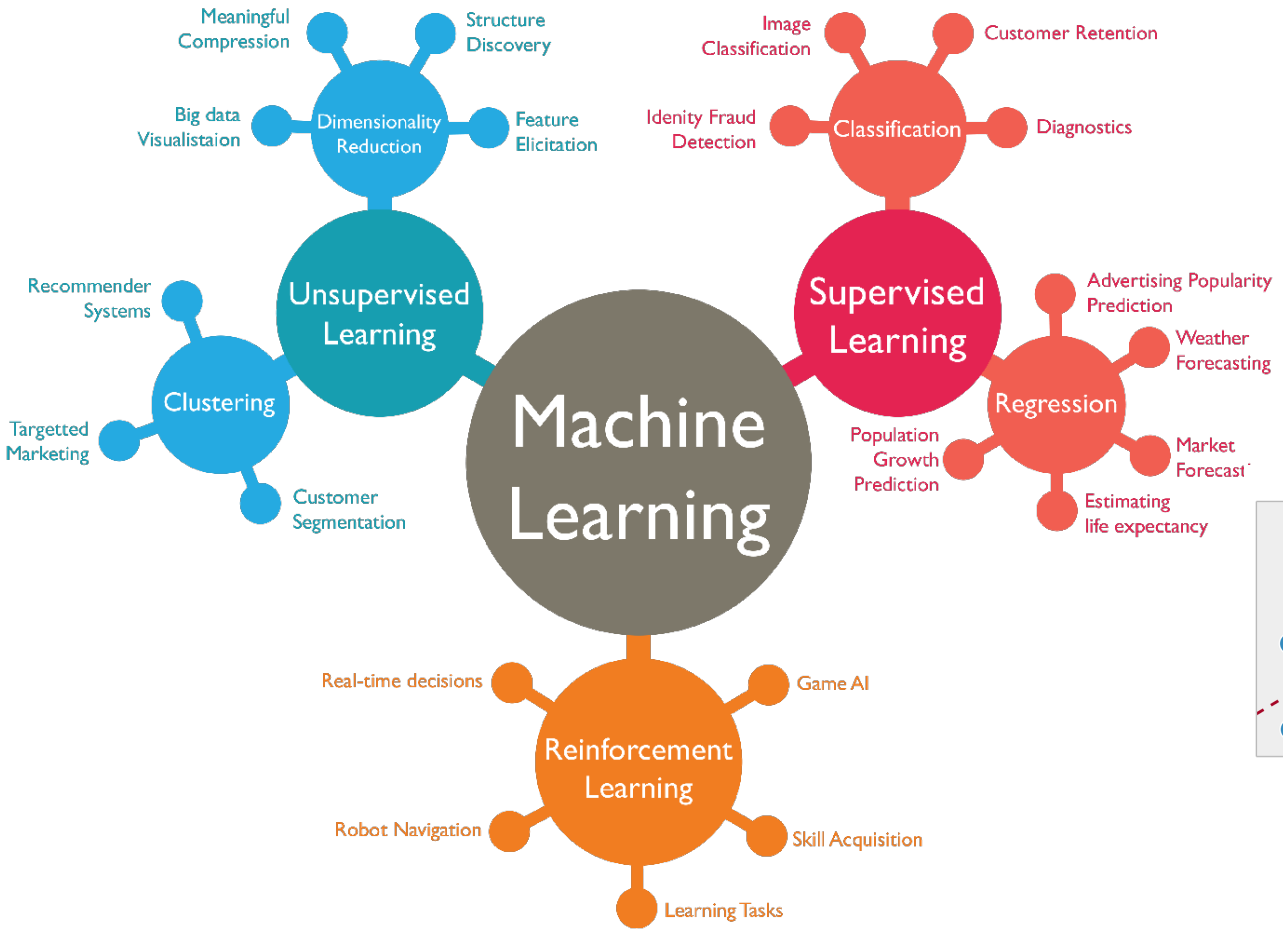
## Select method depending on structure of the problem

Dimensionality reduction

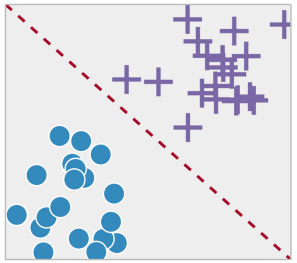


Clustering

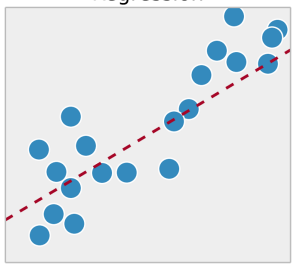





Classification



Regression

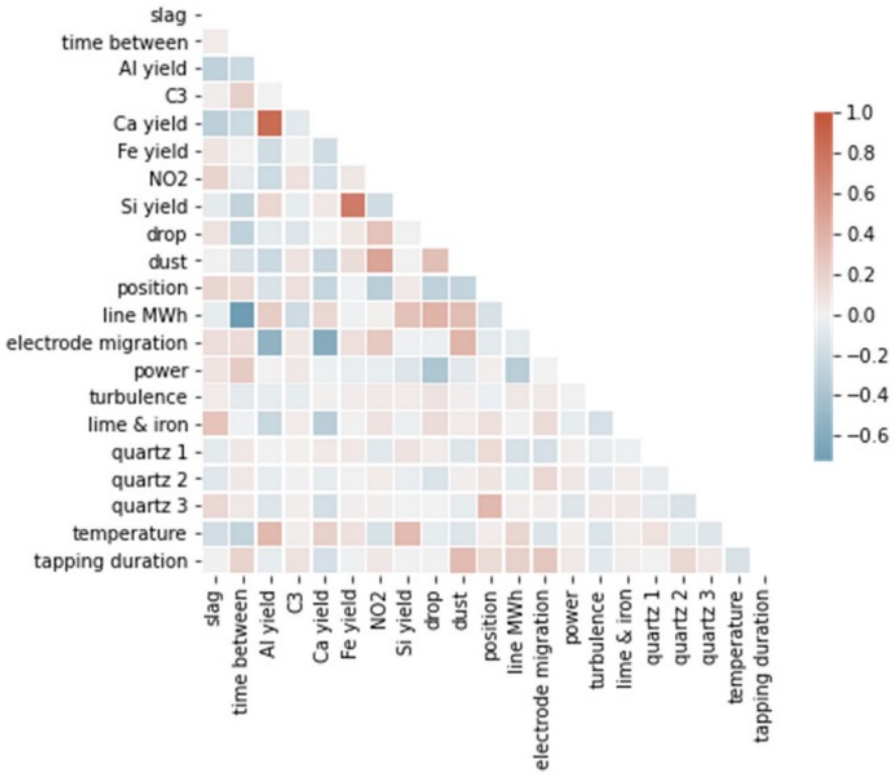


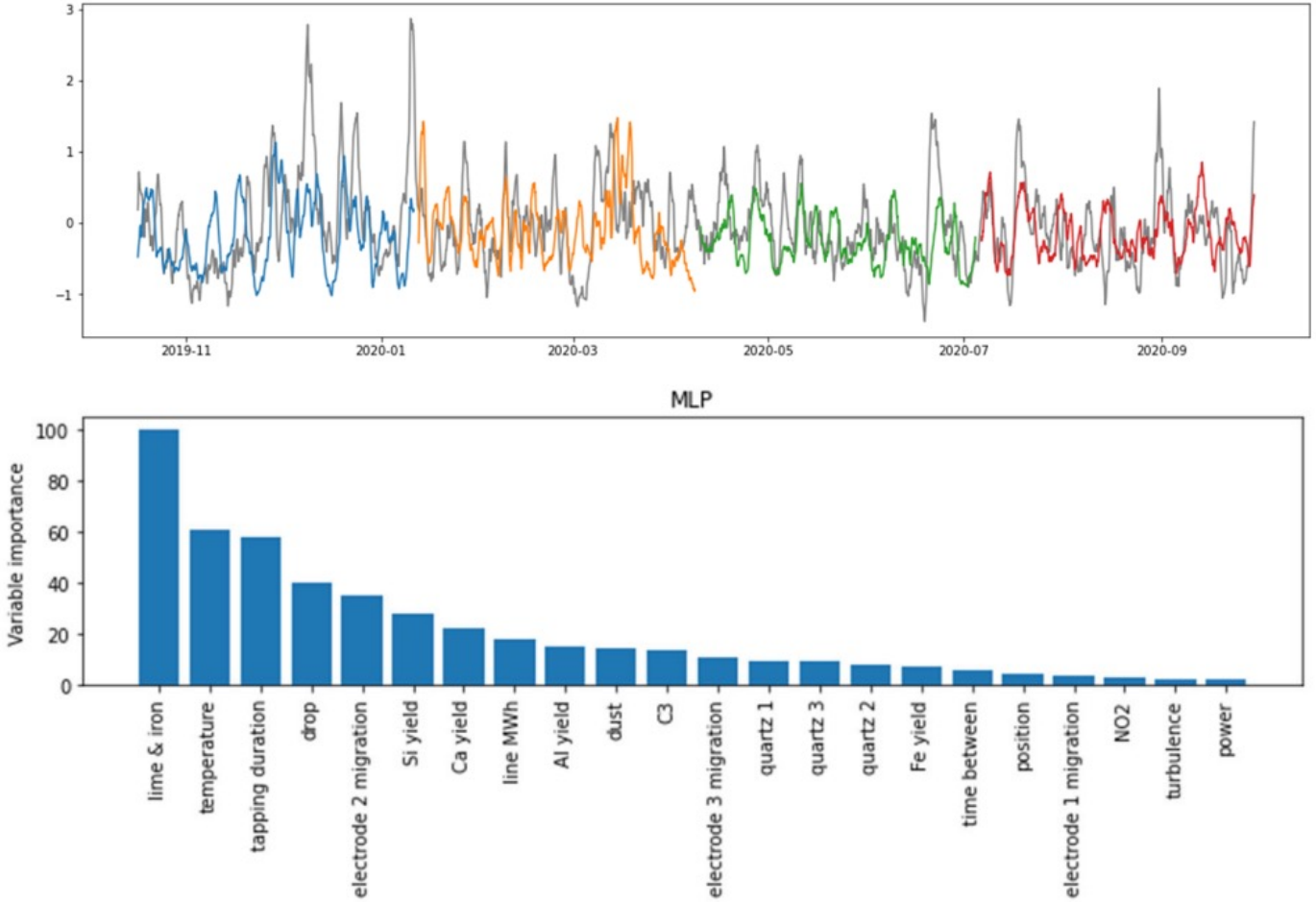
# Appropriate use of ML

- First evaluate whether machine learning is a good, or even feasible, solution
- A data-driven model can never be better than the data
  - Quantity is important
  - Quality is important
- Start with a very simple baseline model. Are we able to beat this with little effort?
  - If not, then we probably don’t have sufficient data

# Appropriate use of ML

## Example: Predicting slag from furnace



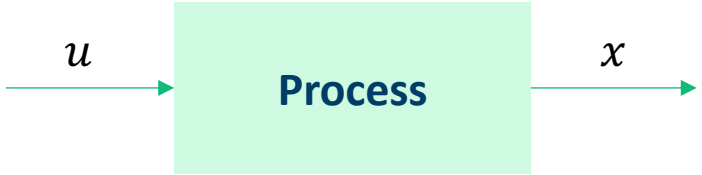




See Linnestad, Kasper, et al. "A Hybrid Digital Twin for Optimal Si-Production." *SSRN 4121131* (2022).

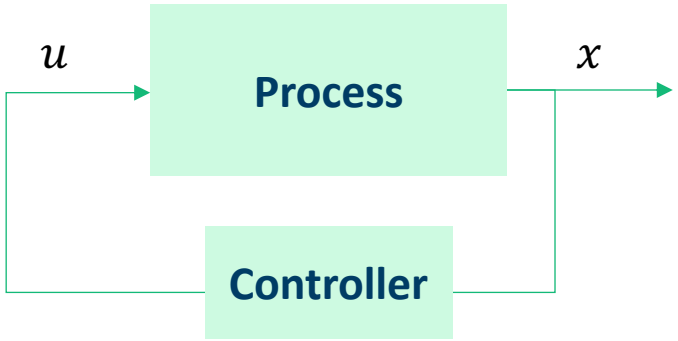
# Learning in a closed loop

- Controlled processes can be challenging to learn
- When we close the loop, we are no longer able to separate the process from the controller
- We are unable to learn how the process would behave without a controller, or with a different controller
- Possible solution: Allow for some disturbance in the process and limited open loop runs

$$\dot{x} = f(x) + g(u)$$




$$\dot{x} = f(x) + g\big(u(x)\big) = h(x)$$



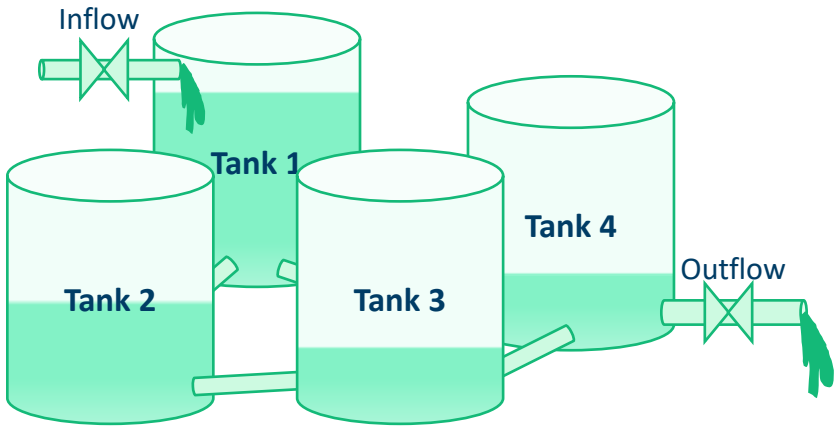
# Hybrid analytics: Physics-informed ML

Combination of known physical/chemical/engineering knowledge, and machine learning

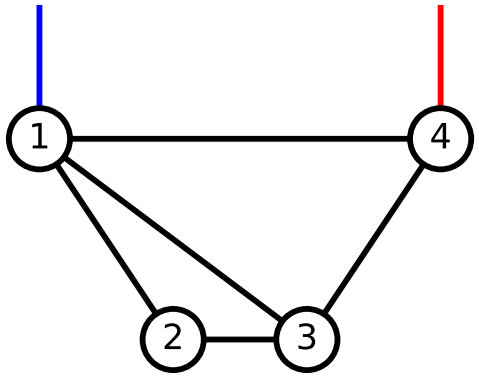
Useful if you:

- Have a model
  - Incomplete model
  - Complete model for parts of the system
- Know structures in the system
  - Mass and energy conservation, dissipation
  - Controlled parameters
  - Connected parts
- Need guaranteed behaviour
  - Stability
  - Constrain to physically reasonable regions

# Hybrid analytics: Physics-informed ML



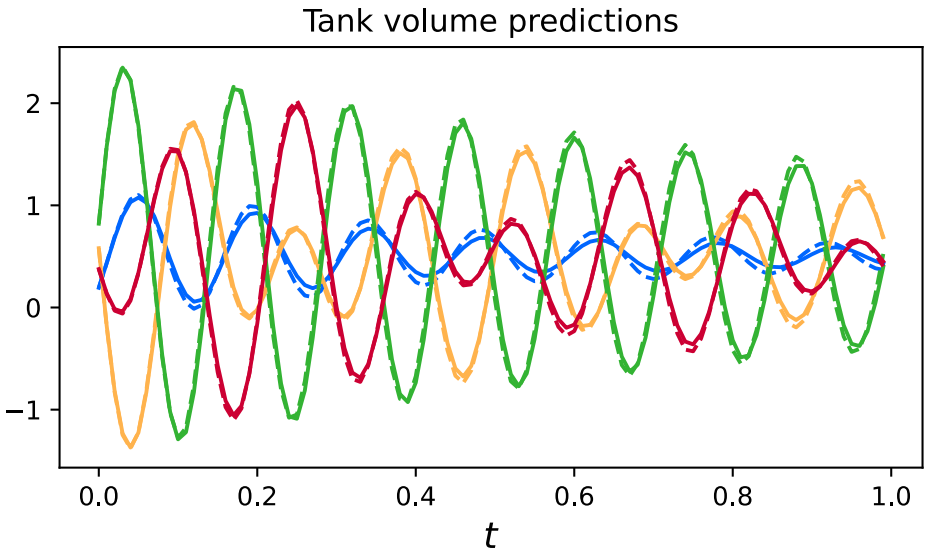

system architecture




laws of physics

$$\frac{d}{dt}\begin{pmatrix}\phi\\ \mu\end{pmatrix} = \begin{pmatrix}-R & B^T\\ -B & 0_{N\times N}\end{pmatrix}\begin{pmatrix}\frac{\partial\mathcal{H}}{\partial\phi}\\ \frac{\partial\mathcal{H}}{\partial\mu}\end{pmatrix} + \begin{pmatrix}F_p\\ F_t\end{pmatrix}$$

+ data




See Eidnes, Sølve, et al. "Port-Hamiltonian Neural Networks with State-Dependent Ports." *arXiv preprint, arXiv:2206.02660* (2022)

# Recommendations

1. Define the problem, and iterate
2. Know your data, and share your knowledge early
   - And your process too
3. Share your data, but also your time
4. Find the appropriate milestones, define specific goals

SINTEF

# Technology for a better society