Making it count - data science at scale

Moving from single model building to automated model building

equinor


Explore, develop and produce oil and gas
equinor
Operations in more than 30 countries
Deliver wind power to
650,000 British households
Mounting solar panels in Brazil

Our vision for data science at Equinor

Create ML/AI technology that gives Equinor competitive advantages in focus areas.

Avoid model by model tasks (unless we´re exploring), build products.

Our structure supporting our goals and visions

equinor

Core ML Teams

Build ML/NLU/RL technology at scale

Insights & Advanced Analytics

Builds and deliver interactive UX to access & analyse data across the value chain

Knowledge AI

Extracts knowledge from our vast unstructured data and make it available for use

Machine Data ML

Turns machine data into operation optimization capability across the value chain

Self-Learning Robots & Vision

Optimises highly dynamic problems and builds computer vision tools

Business Teams

Work with business areas to implement value add solutions.

Subsurface DS

US DS

IOC

Design Thinking

Data Management | **Data Engineering** | **Software Development** | **Data Platform** | **Procurement**

The data science team

23
by end of October

24%
And growing

5 cities
Stavanger, Bergen, Oslo, Austin, Houston

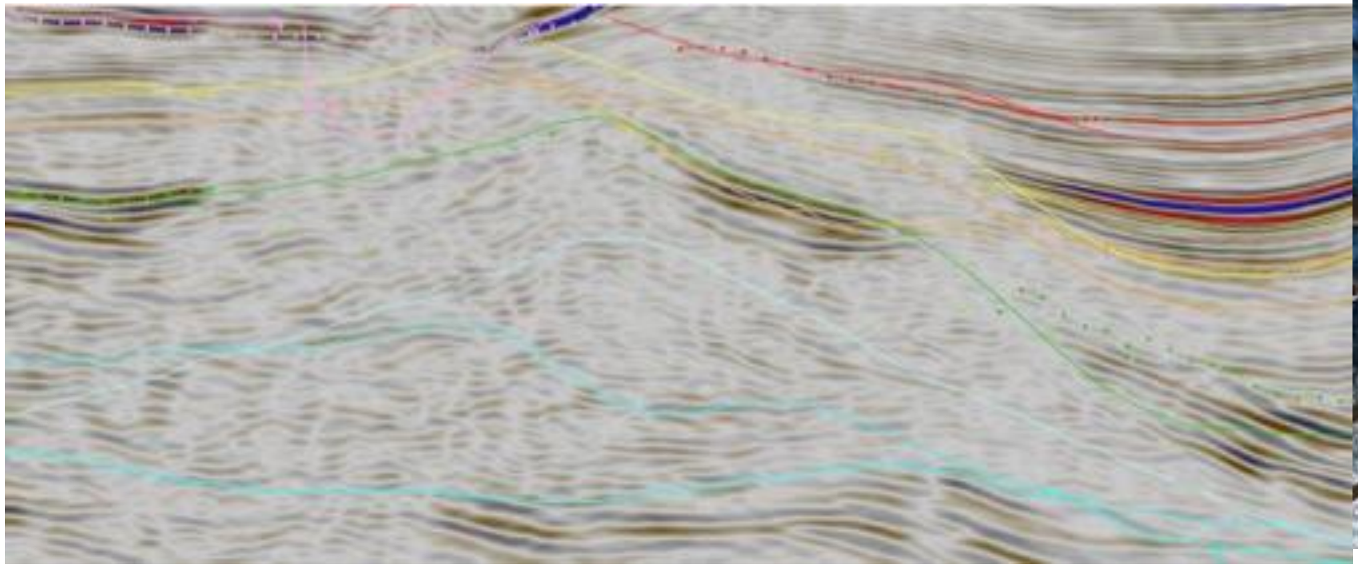
44% PhDs
Mathematics, geoscience, engineering…

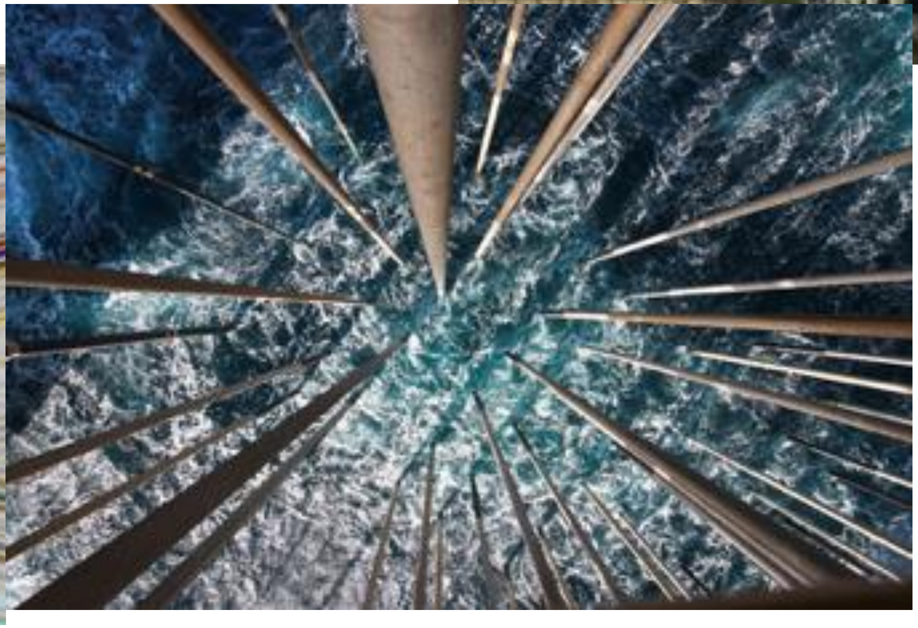
60% External
Great improvement in attracting talent

6 Industries
Consulting, automotive, tech, internet, finance

Equinor operates more than 40 platforms just in Norway

Thousands of potential cases (machines) and tens of thousands (towards of different sensors

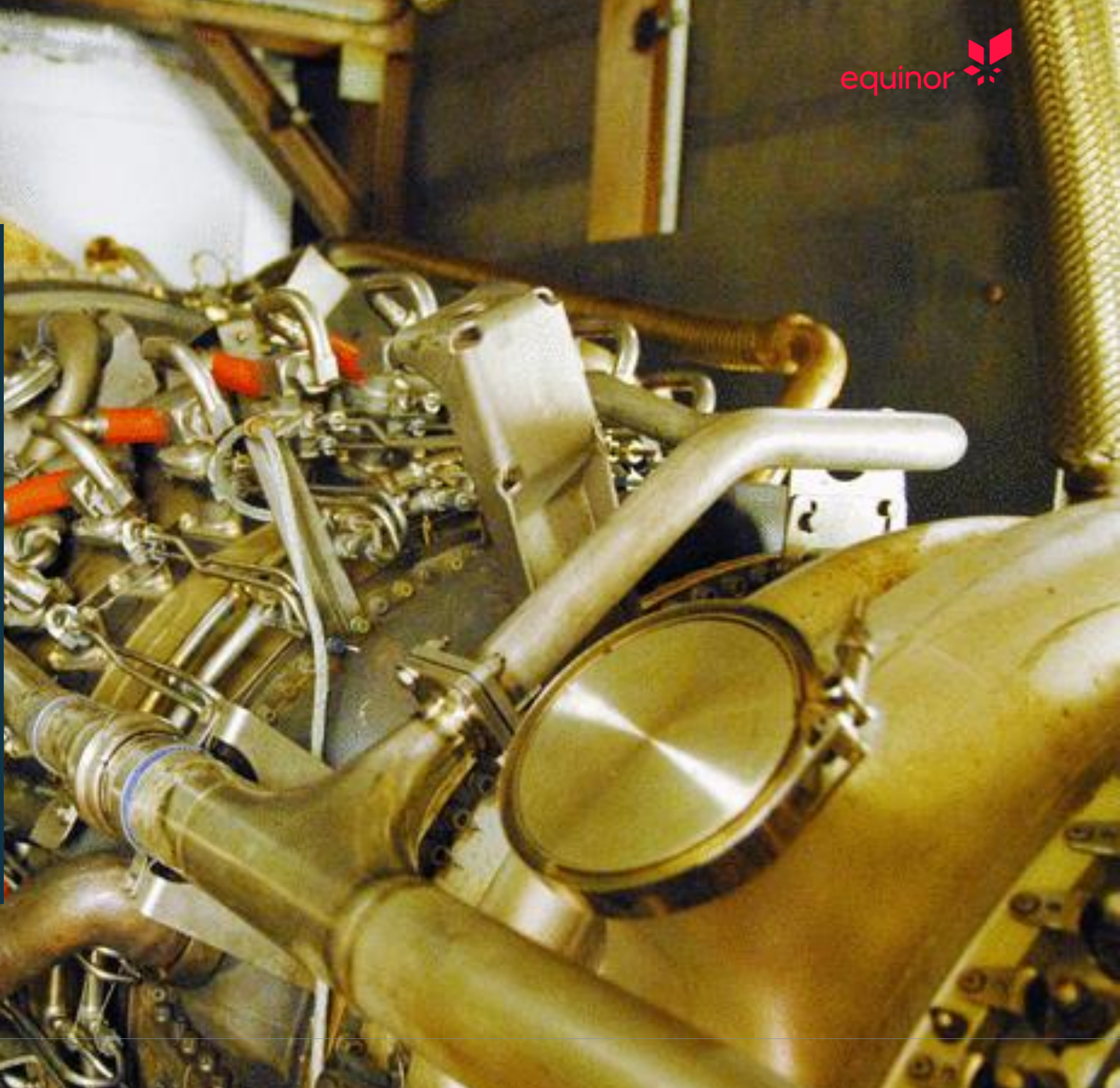




We can't model every machine

One of our cases aim to monitor the *health* of a *machine* and reduce shutdowns.

We are building a general process engineers can apply themselves.



Looking at single machines – but only if re-use is possible

Jupyter notebooks for trying out things.

Not easy to use for scientific studies
Not the tool for solving a problem in a general way

What about the next 120 machines of the same type?

The machine «health» index – from single machines to a reusability

The problem has many instances.

We must be able to solve all of them the same way.

- Different sensors
- Different patterns of operation

equinor

The machine «health» index – from single machines to a reusability

Input:
New machine is (automatically) registered with a barebone of information

Action: Model is built

Action: Model is made available

Action: Predictions are being done automatically

Must be fully automated

Use case driven system development

Data science team collaborate with software developers in **one team**

Mature **use cases** and **develop solution** together.

Building the system

End User

Monitoring Dashboards

Configure/Add Machines

Time Series Data Source

Data pipelines

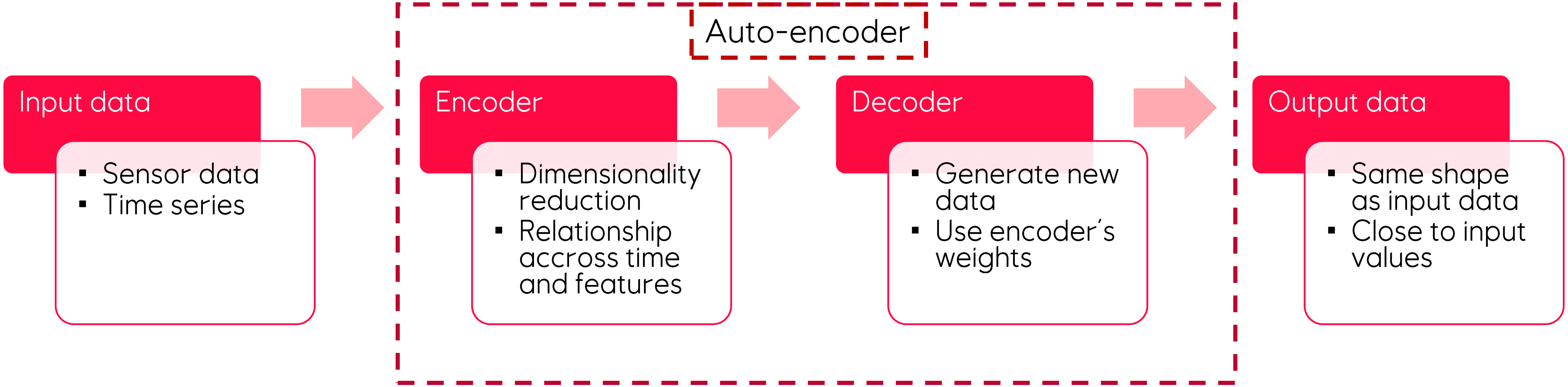
Data scientists and developers

Automated training and deploy

Model APIs

equinor

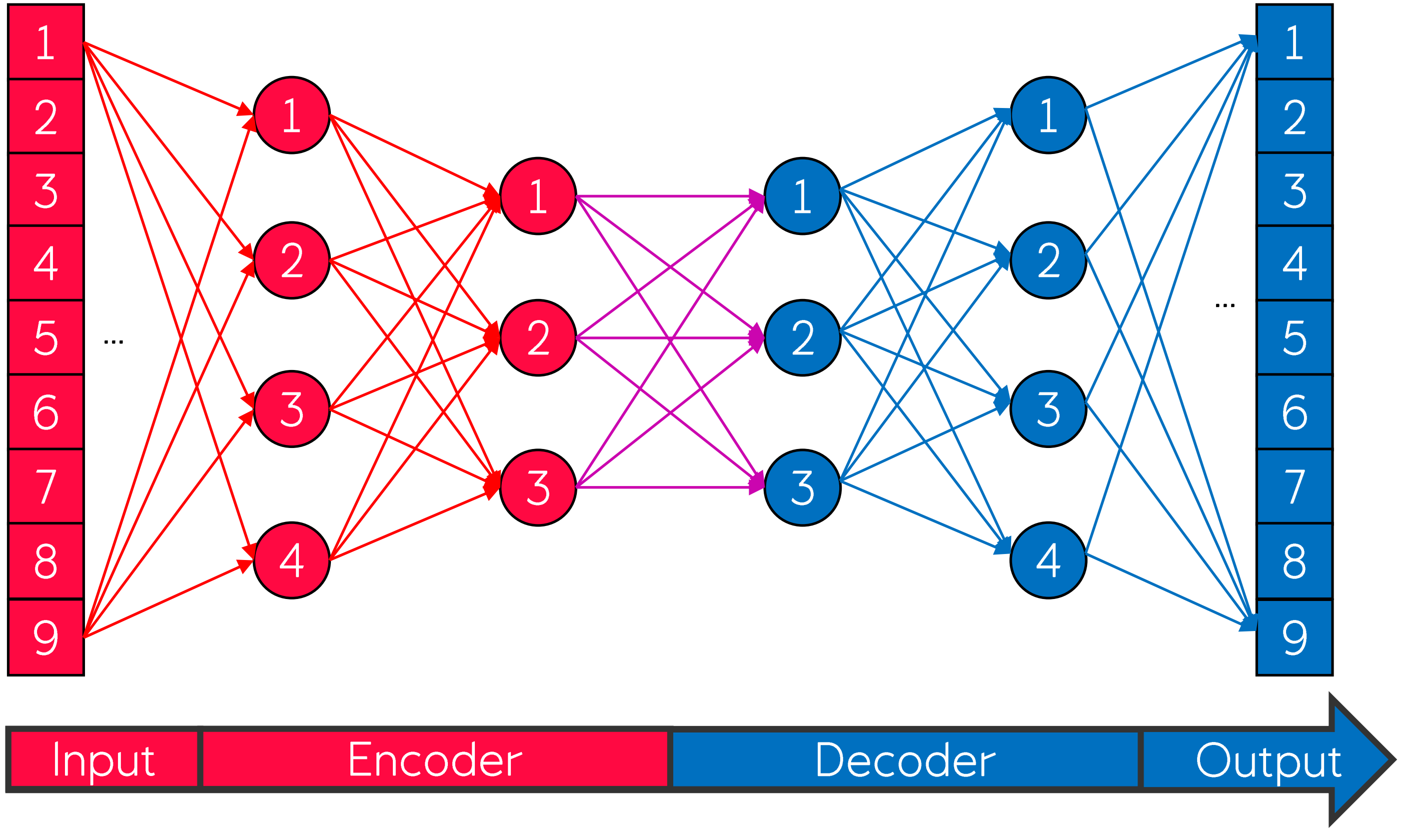
Auto-encoder on sensor time series



The difference between the «Input Data» and the «Output data» indicates normality or anomaly

equinor

The auto-encoder structure: example of feedforward neural network



Summary

To make a difference, we must be able to create products that solve many instances

We need business understanding, analysis and software development skills within one tight group

<u>There are too many cases and problems to be solved one by one</u>