University of Amsterdam
Multi-scale Networked Systems Research Group

Data-Centric Analysis of Complex Industrial Systems

Uraz Odyurt
2022-06-29

Why data-centric solutions?

- Modern systems in many domains are **data-rich ecosystems**
  => Lots of **sensors**

- Computerisation
  => Everything is a **computing platform**
  => Strong presence of **software**

- Systems are continuously evolving during their lifecycle

- Dealing with **non-determinism**
  => A common trait of CPS

There are challenges …

- Data scarcity
  => Sometimes there is not enough data
  => Not of the kind we need
  => Gaps in data streams

- Data deluge
  => Because of: Transfer limitations, processing limitations, latency

- Knowledge incorporation (more on this at the end …)

- How to generalise?
  => Oftentimes solutions are use-case specific

Example use-case:
Anomaly detection/identification

for semiconductor photolithography machines


20
10
0
x100
min-1
60
70
80
km/h
60
40
20
0
140
160
180
200
ABS

iDAPT Project

- *Interactive DSL for Composable EFB Adaptation using Bi-simulation and Extrinsic Coordination*
- National project funded by NWO
- Main project user: ASML Netherlands B.V.
- Other users: TNO, Thales, Radboud University Nijmegen

Robustness

- Things go wrong, no matter how good the design (**design is not static**)

- This is about **detecting** that something is going wrong
  => Unwanted behaviour
  => Light turns on …

- This is also about **distinguishing** between different unwanted behaviour
  => Different things can go wrong in a complex system
  => Different lights, different modes

- If we know about it, we can **fix** it, or reduce its effects
  => Increased robustness

This is exactly what the solution is about!

Robustness: Anomaly detection/identification

- Anomaly: A **readily** detectable deviation in system's normal behaviour
  => A symptom
- Anomaly detection
  => Behaviour is not as intended
- Anomaly identification
  => What sort of trouble are we talking about?
  => Which part? (subsystem)
  => How bad? (severity)
- Predicting anomalous behaviour

It is all about normal behaviour vs anomalous behaviour.

Industrial CPS and phases

- Industrial CPS are **purpose-built**
  => A limited domain of activities and tasks

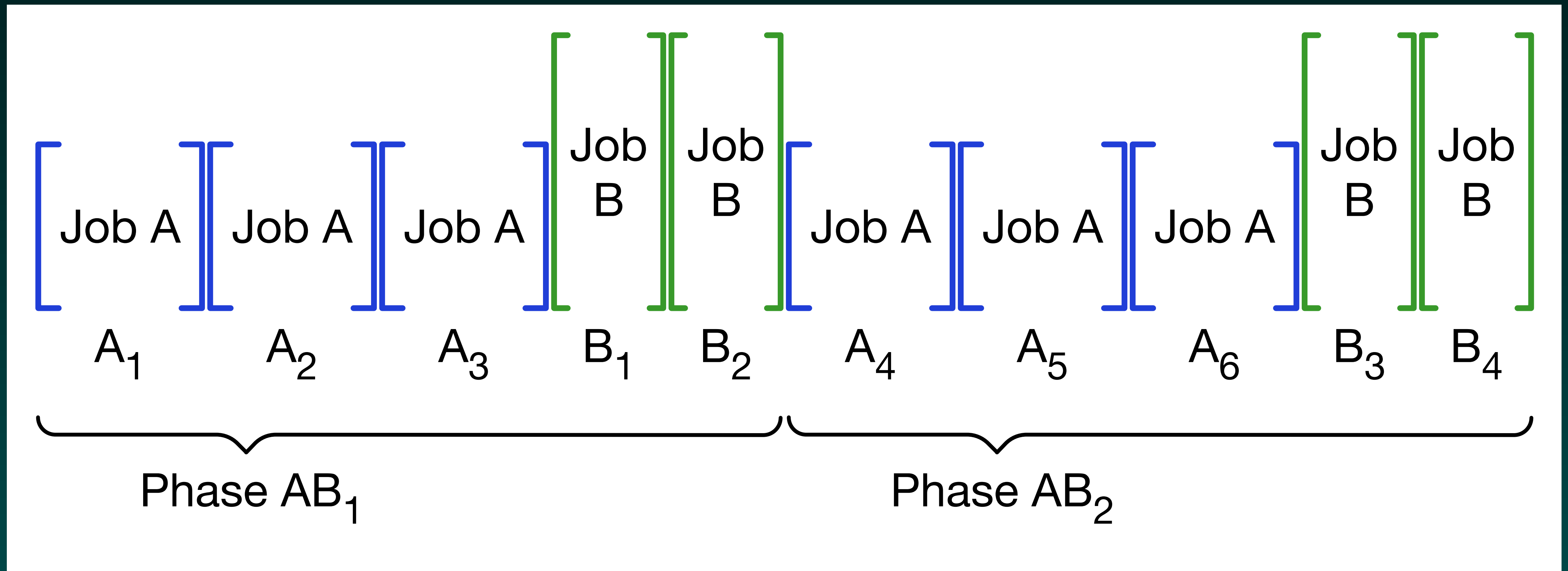
- We want to exploit this repetitiveness for behavioural monitoring

- Execution phases => **Units of execution**

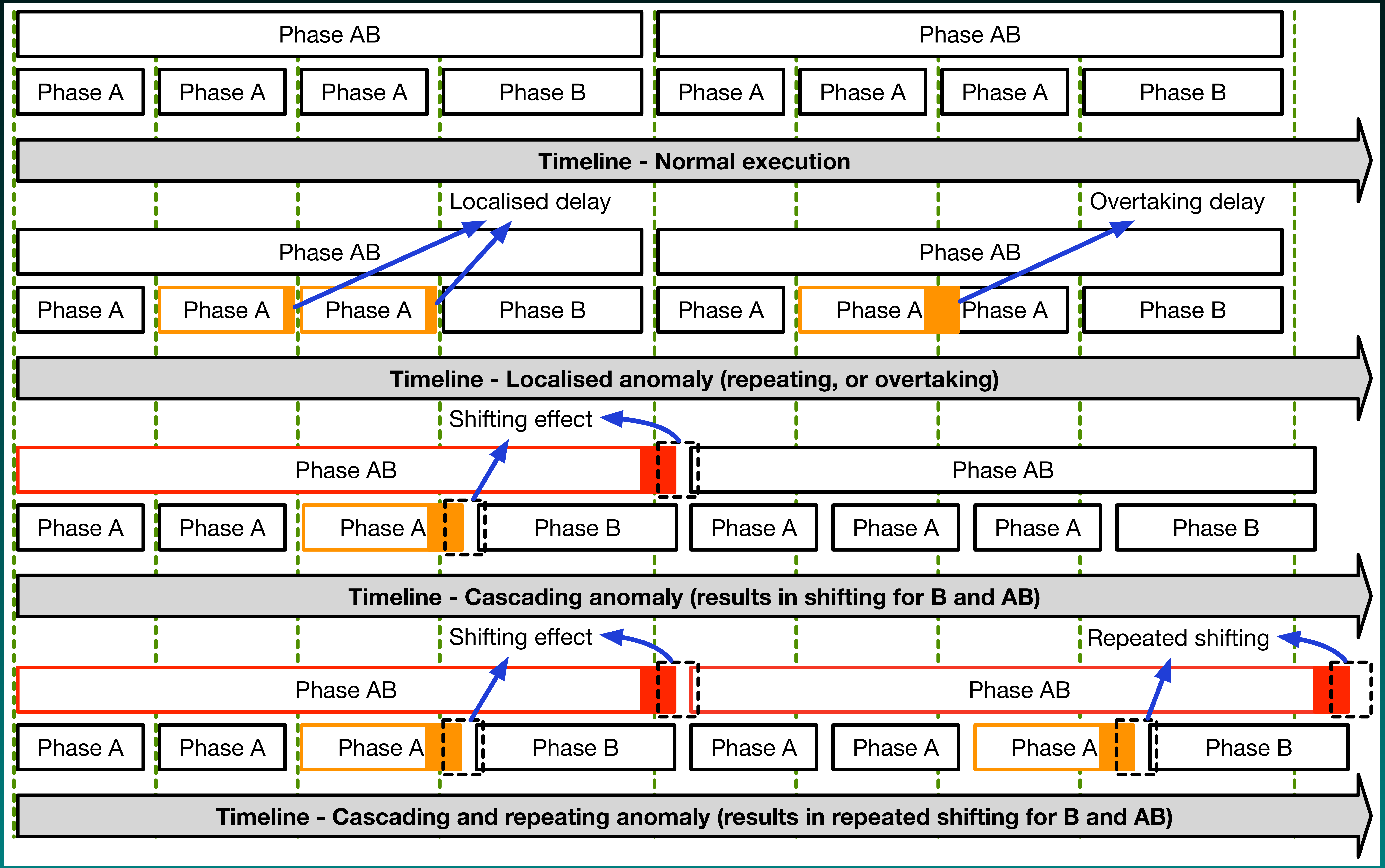
  ➡ Atomic phases: Smallest **repetitive** unit of execution behaviour

  ➡ Combo phases: **Repetitive** combinations of a collection of atomic phases

- Observation and analysis needs will determine phase granularity



Anomalies and their effects



Our subject: Industrial CPS

- We are not dealing with cars, or engines
- We are dealing with industrial machinery
  => But, a specific breed, controlled by computers
- Characteristics of industrial CPS
  => Bunch of computers working together, collectively!
  => Different types of computers, heterogeneous
  => Interaction with the environment
  => Software, software, software, software, software, …

Industrial Cyber-Physical Systems (CPS) are highly repetitive systems

Semiconductor photolithography machines

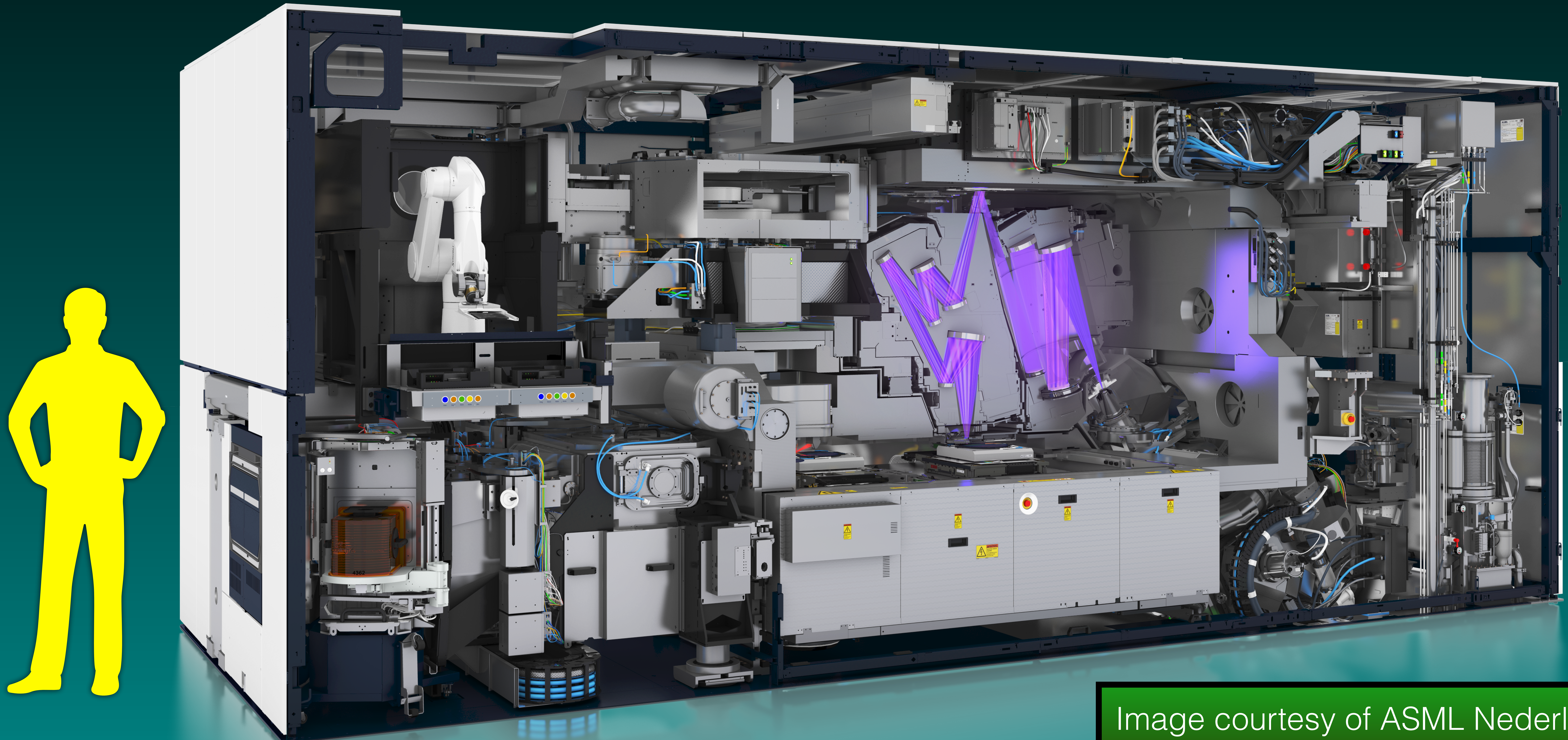
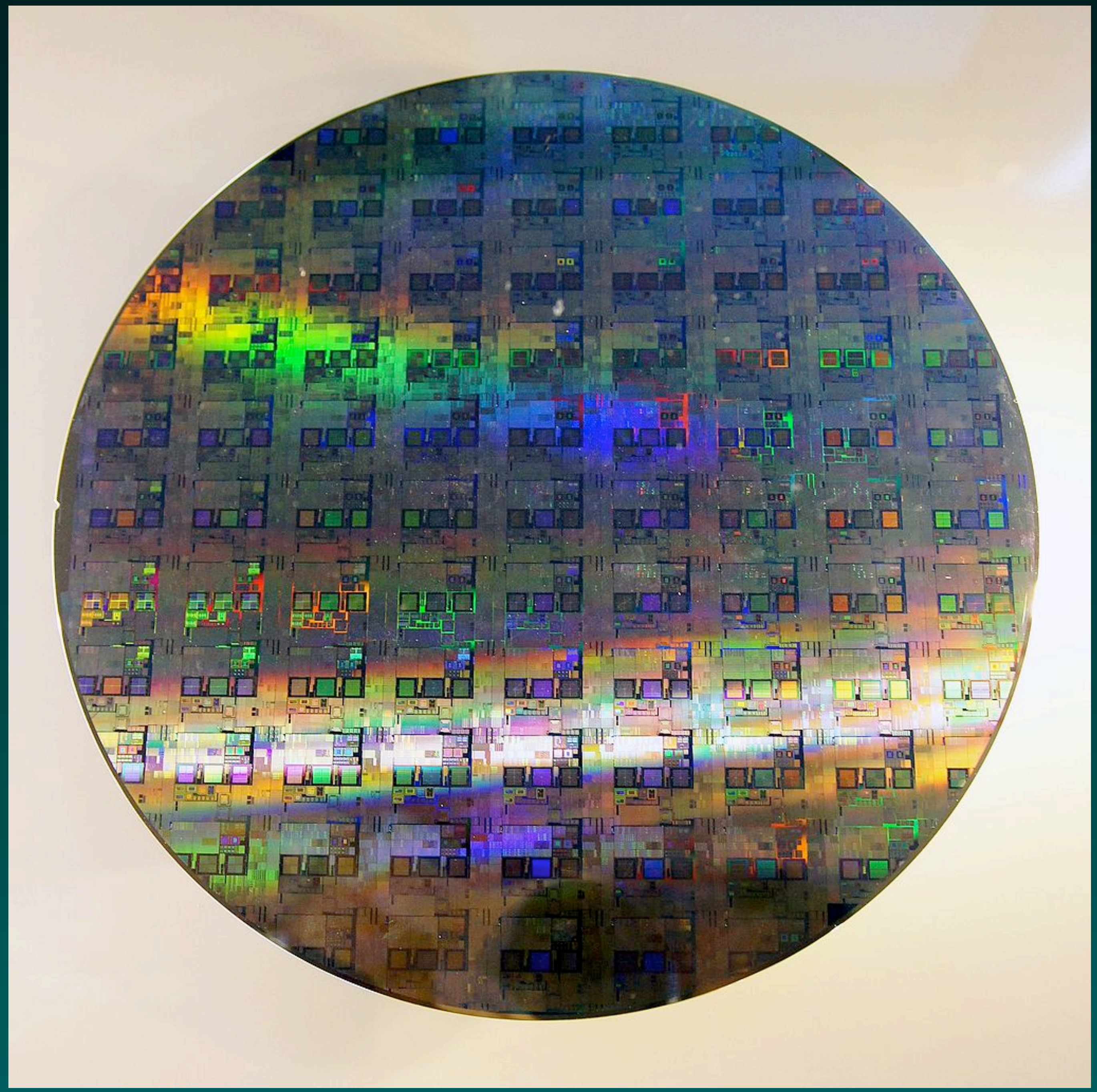


Image courtesy of ASML Nederlands B.V.

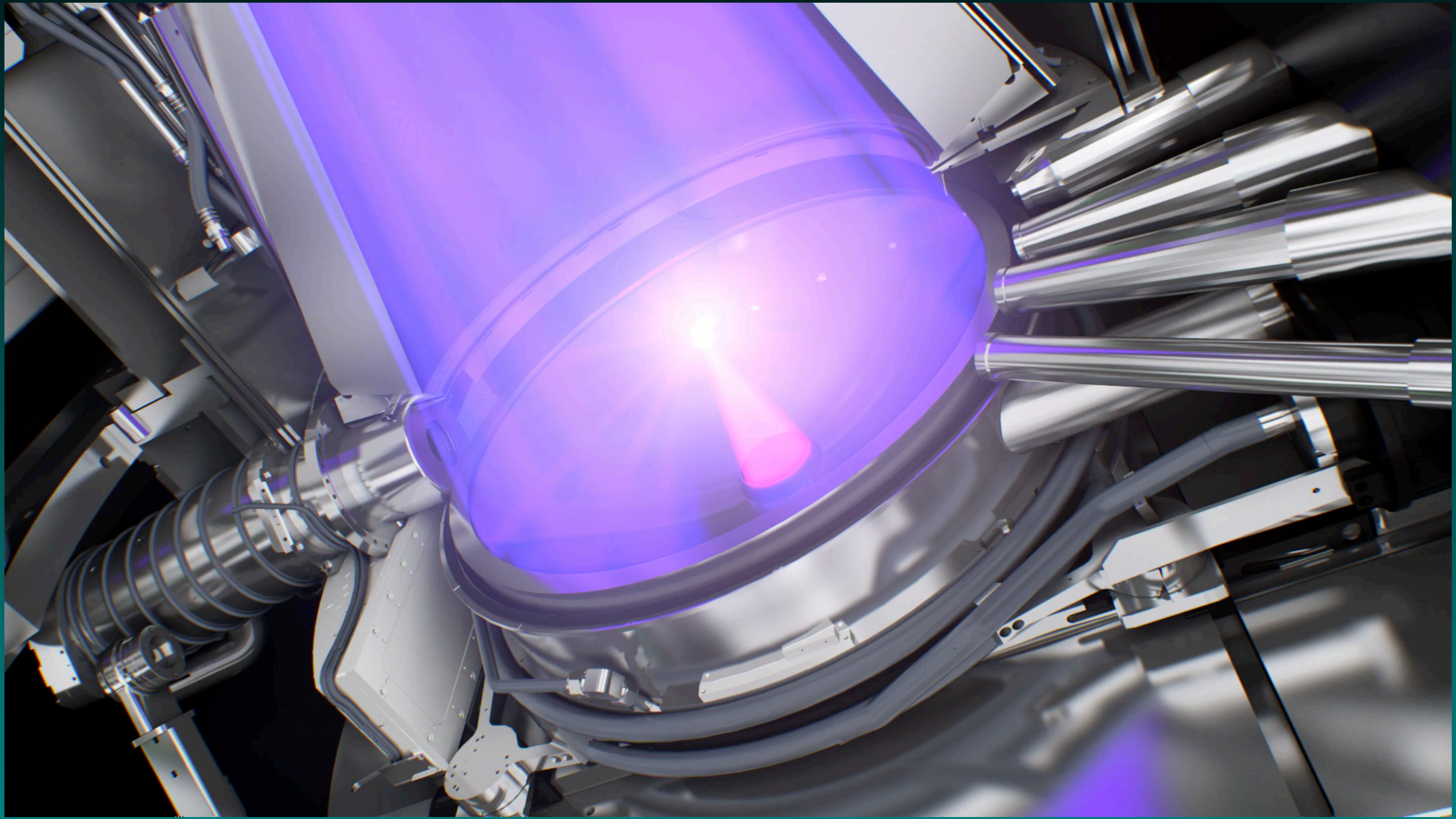
Semiconductor photolithography machines

- Very similar to photography
- Involves light (EUV) and therefore, a **light source**
- Involves film (**wafer**) with **photosensitive material**
- Involves patterns to be applied (**reticle**)
- Involves chemical developers
- Has to be done at scale, otherwise a smartphone will cost €10000 …



Photolithography

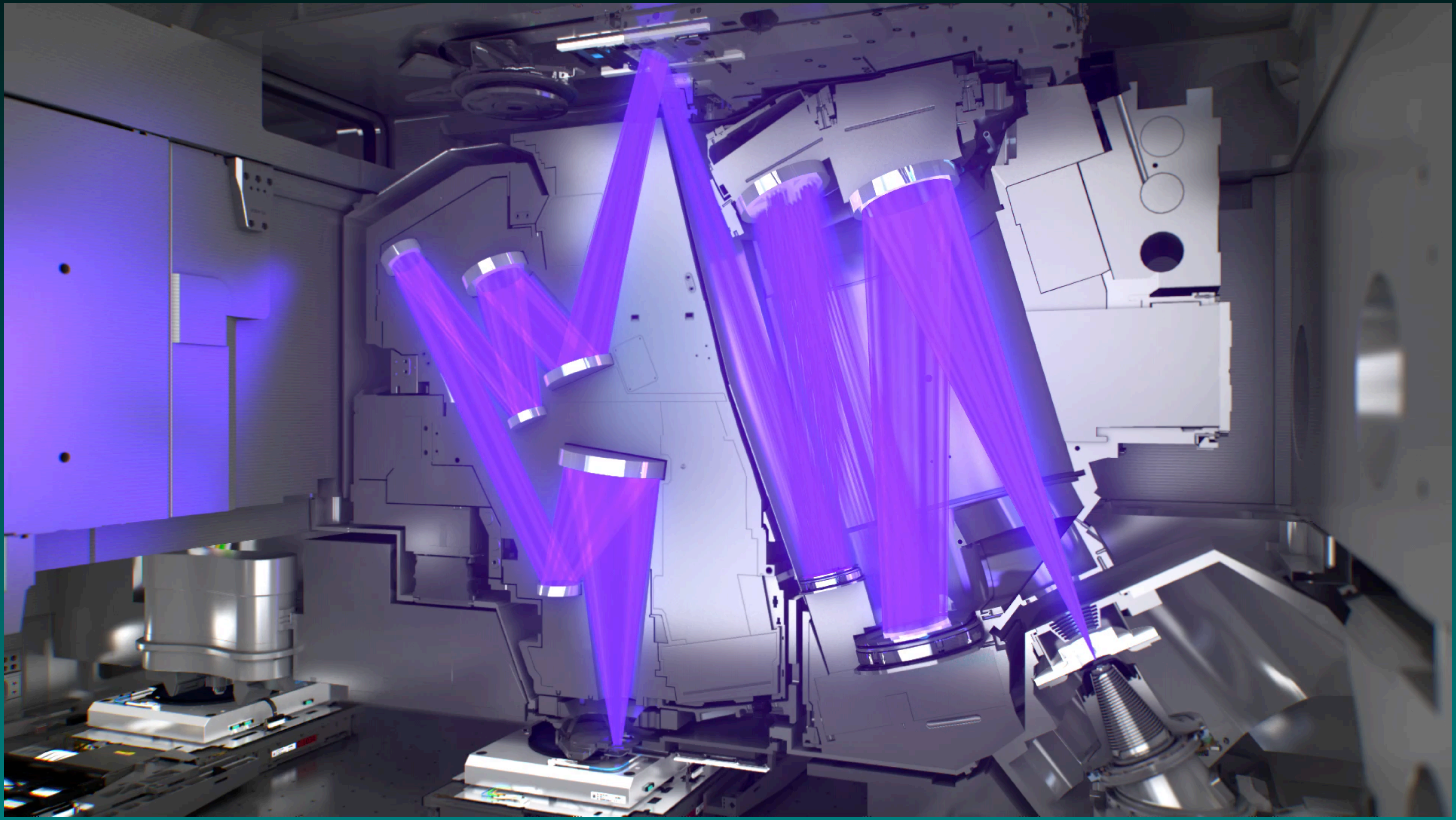
EUV generation - The light source



Video courtesy of
ASML Nederlands B.V.

Photolithography

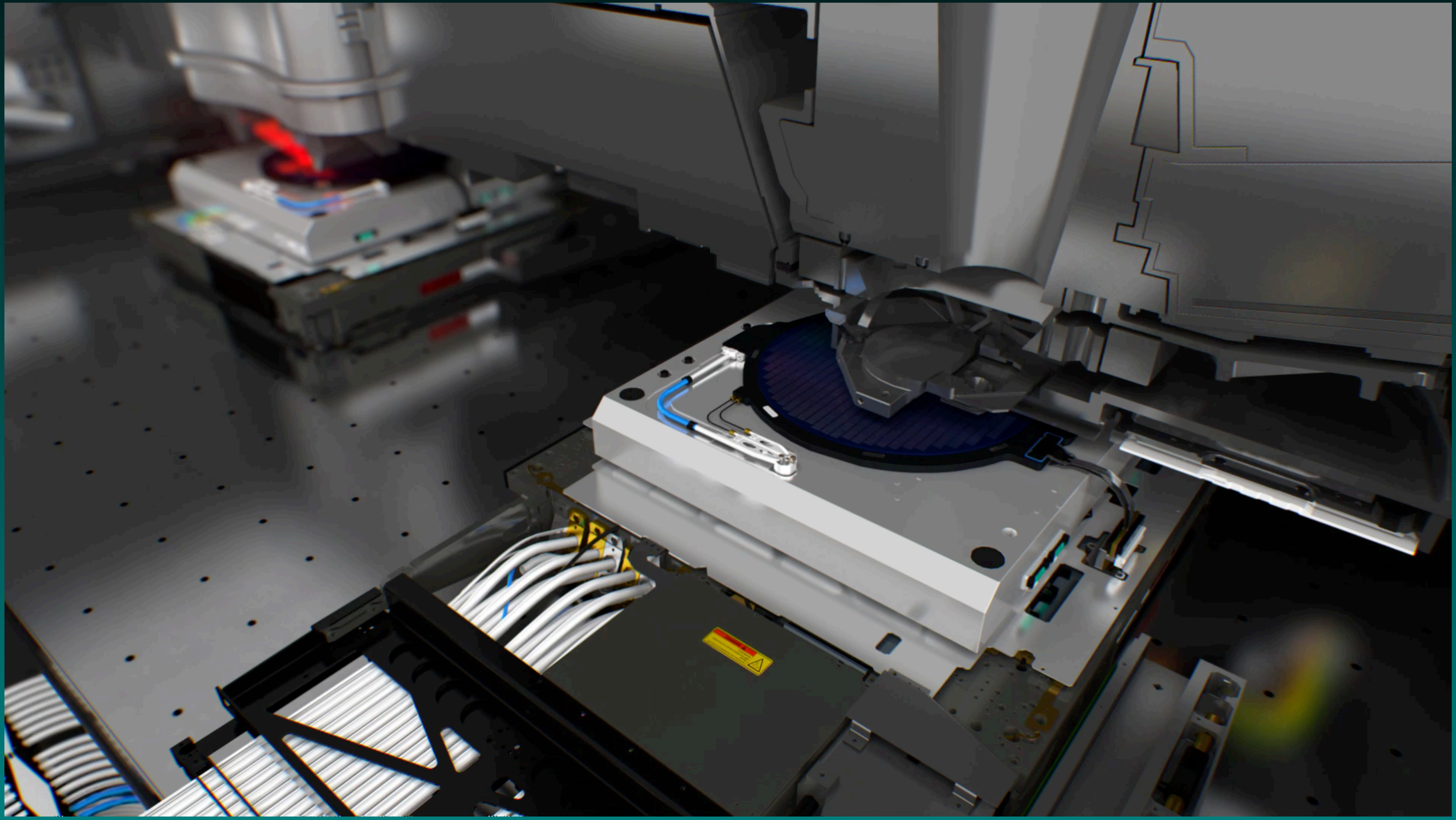
Light path and patterns



Video courtesy of
ASML Nederlands B.V.

Photolithography

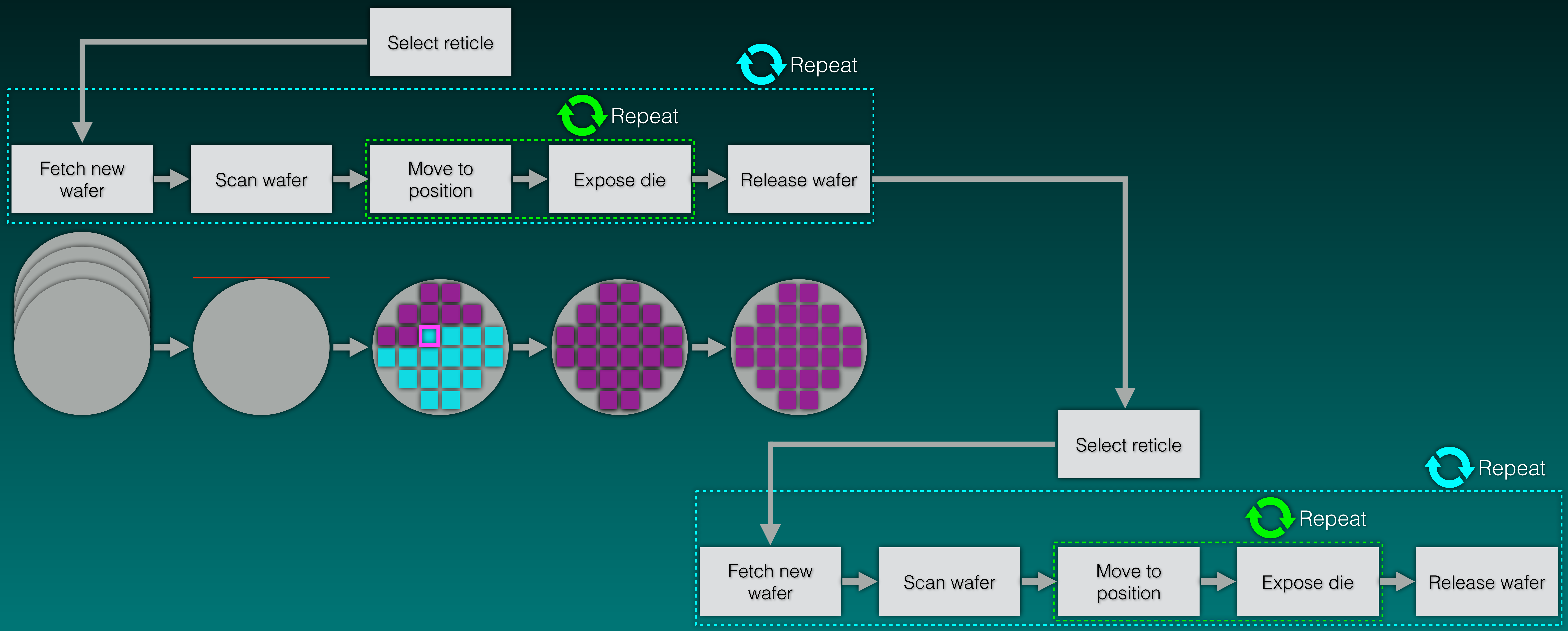
Stepper unit exposing photosensitive material



Video courtesy of
ASML Nederlands B.V.

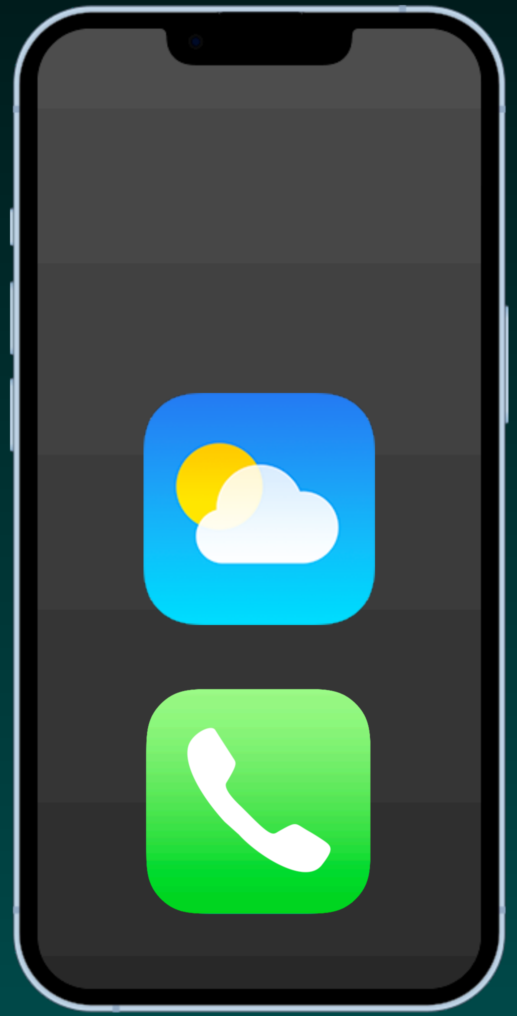
Characteristics of industrial CPS

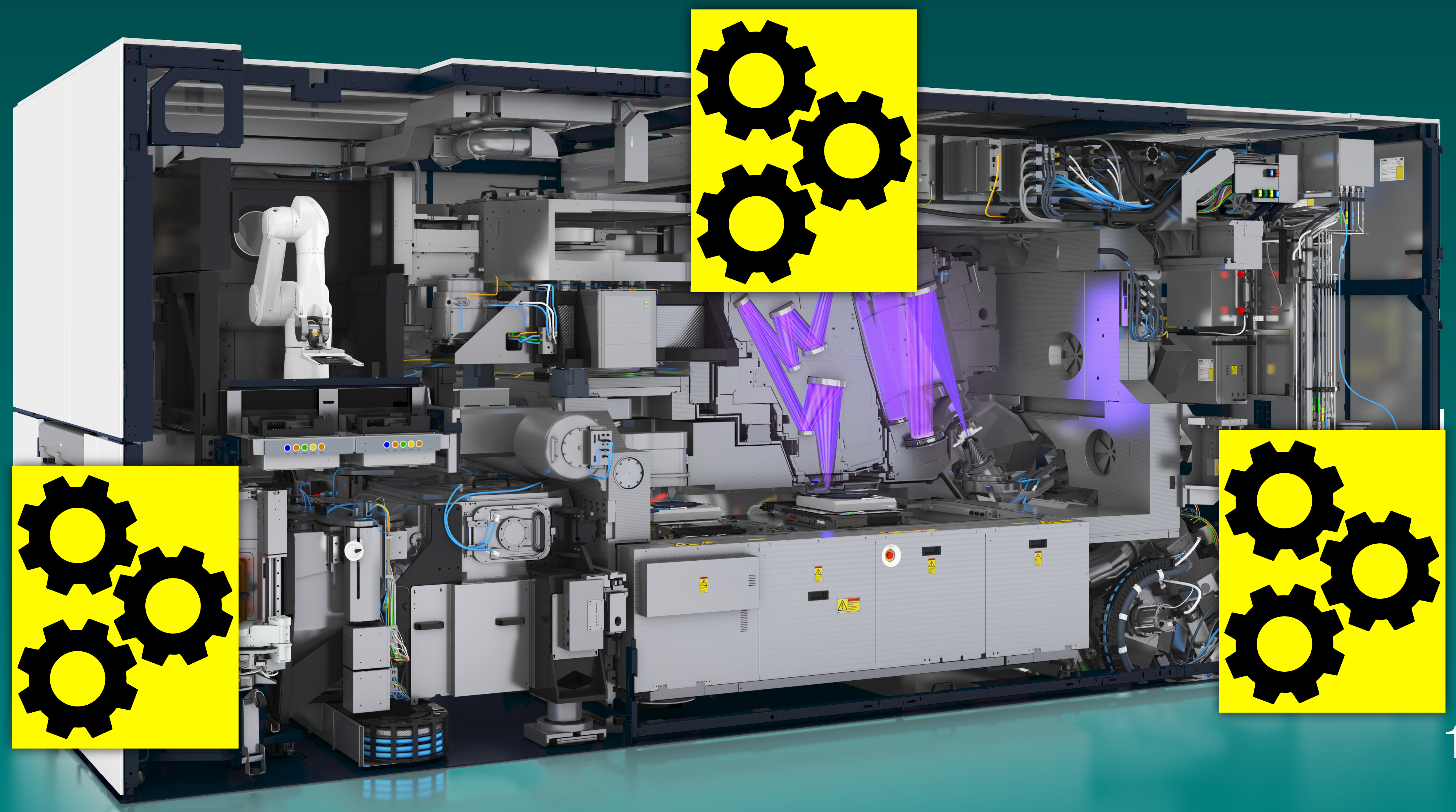
Repetition in a semiconductor photolithography machine



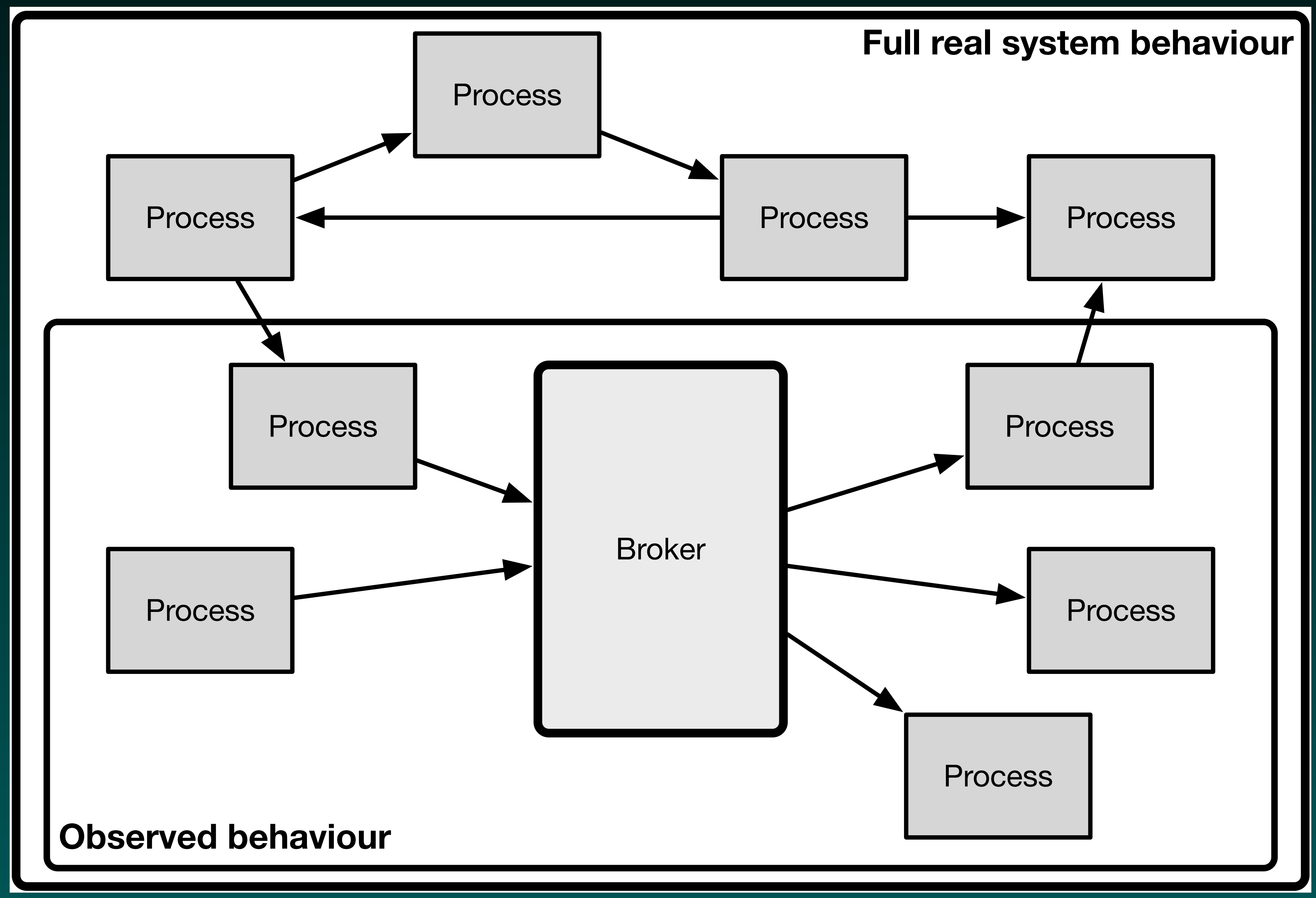
Software is taking over

- Everything is being computerised, industrial CPS included
- Computers are platforms for software
  => Software can be added or removed
  => New or extended software, extended functionality
- Software is a big source of data
  => We can take advantage of sensors and collect
- Software is getting too large
  => Complexity, extremely costly to get it right at design time





Minimal, but extensive enough data

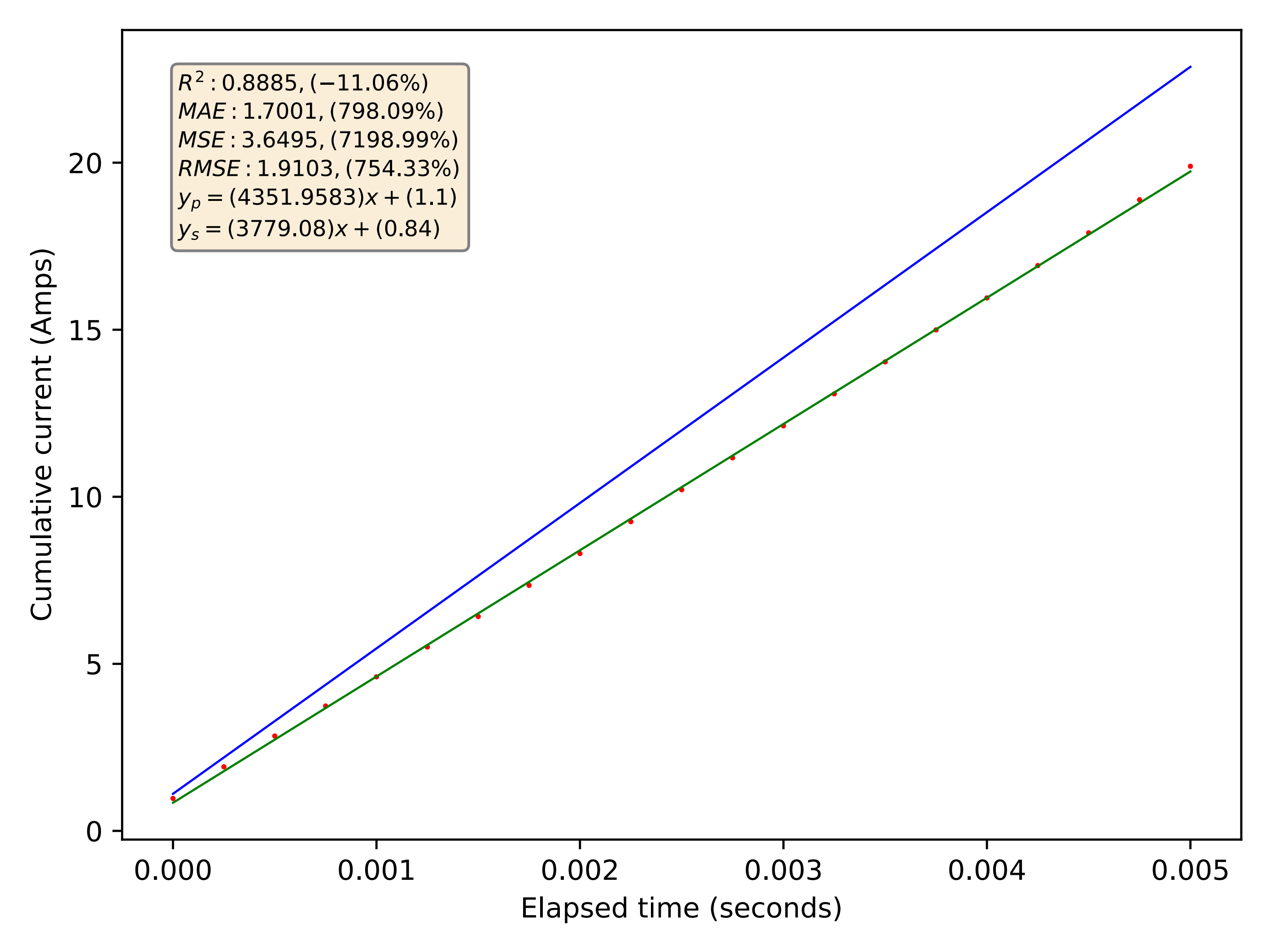


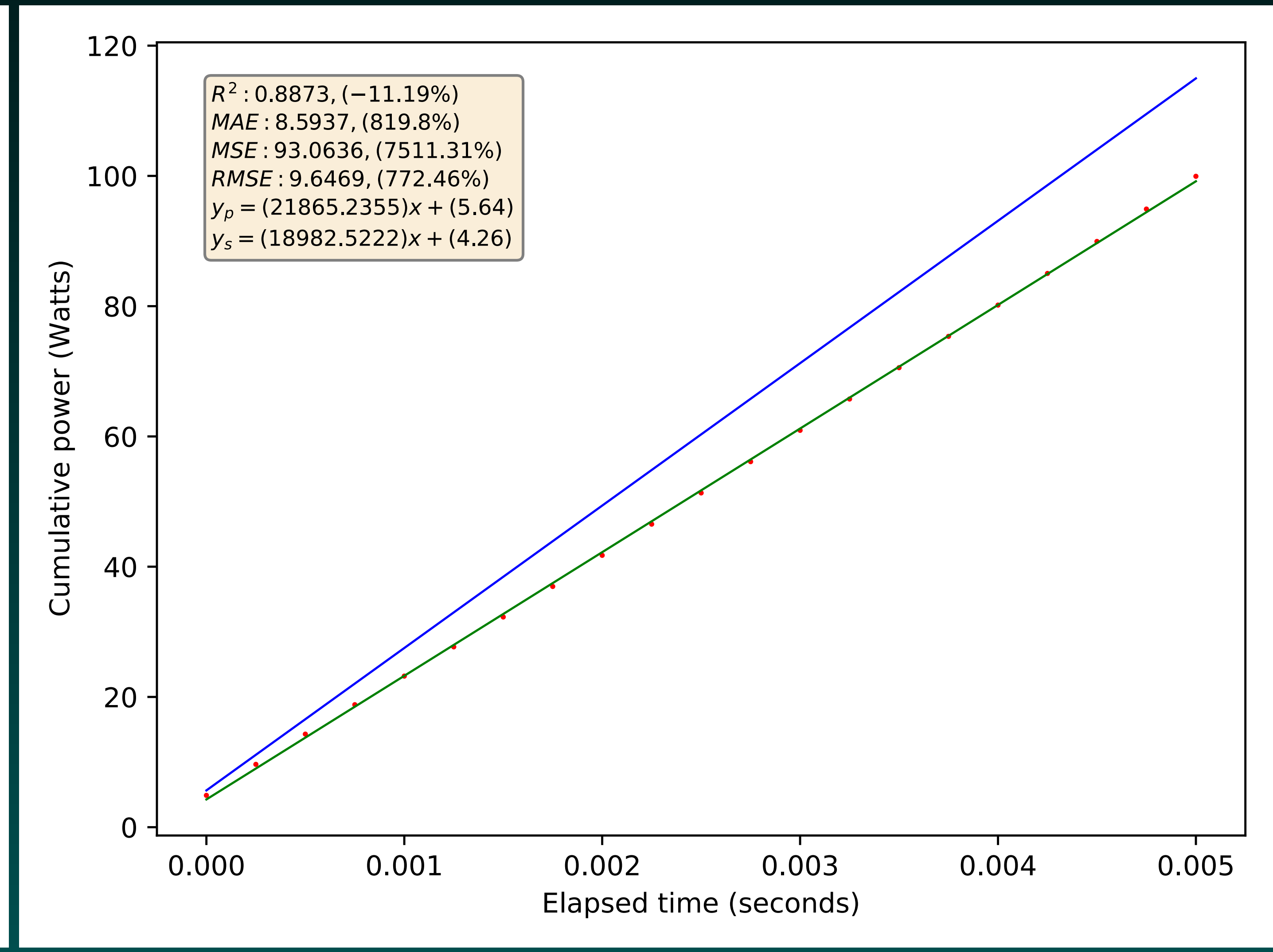
- We look into efficient collection of data
  => Enough to understand what is going on (understand the behaviour)
  => Be efficient, not too much data

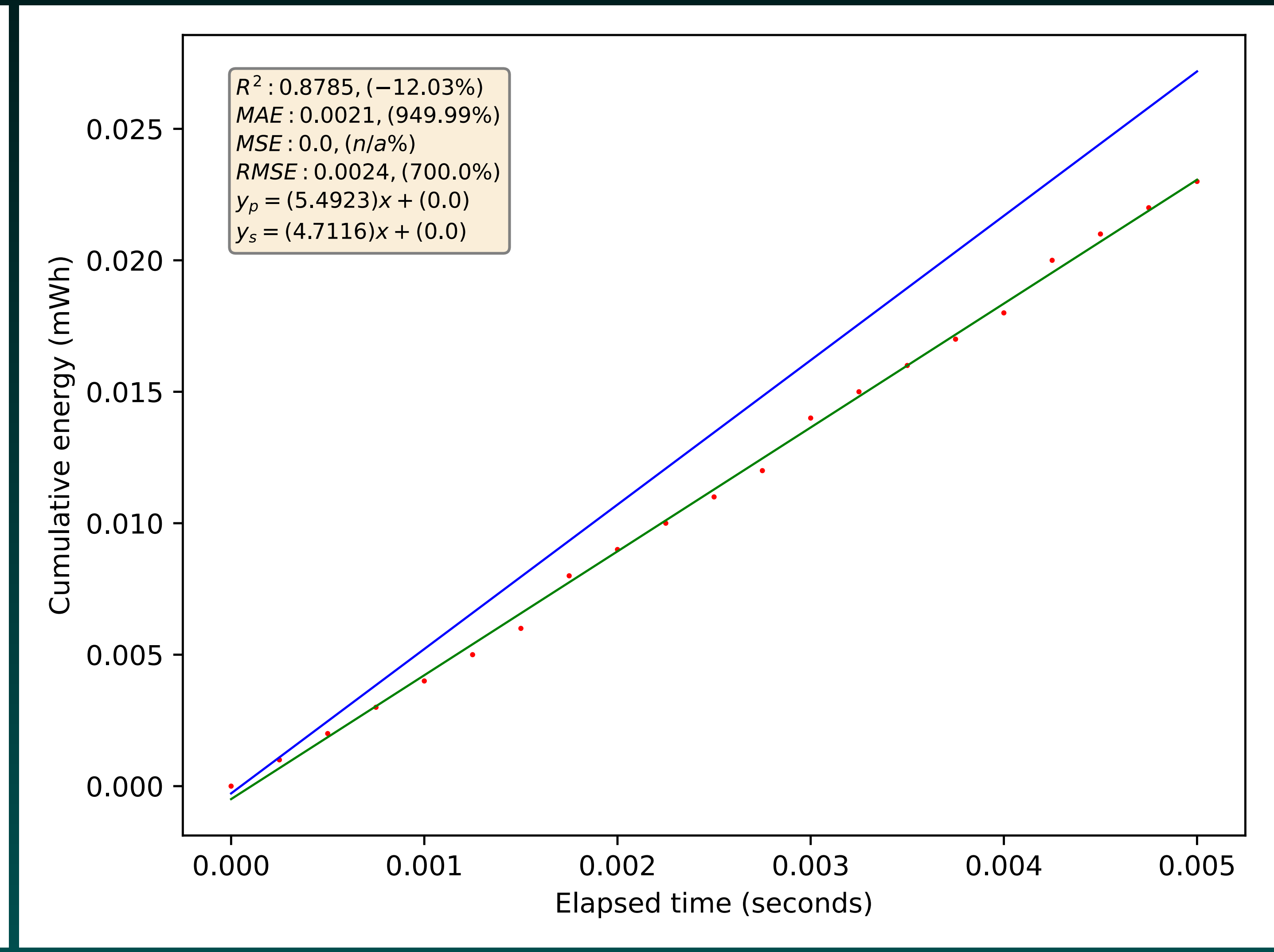
- Process the data in different ways
  => Ability to generate **fingerprints** for the behaviour
  => Ability to compare different fingerprints
  => Take advantage of Artificial Intelligence (AI)

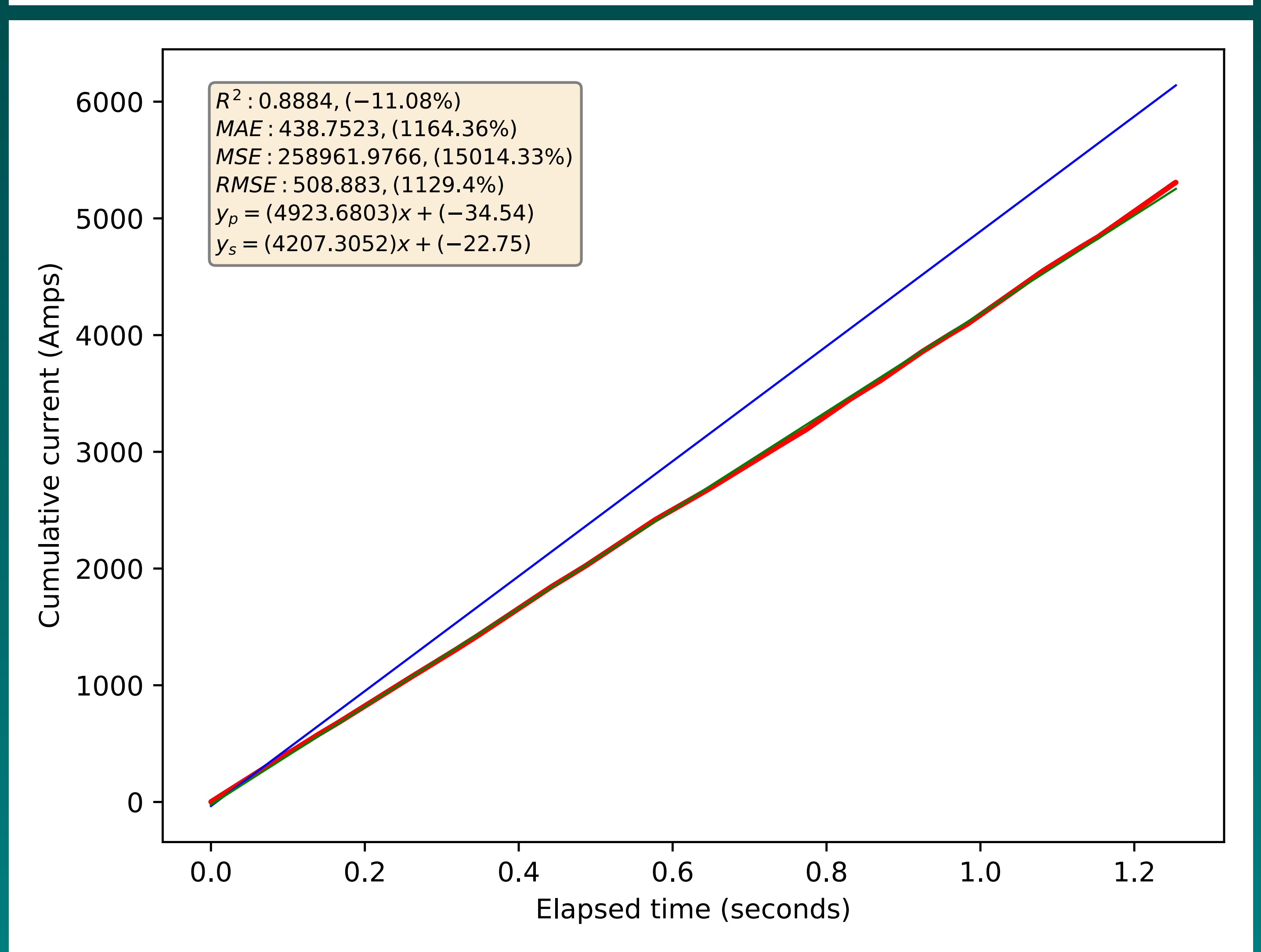
It is all about data and what we can learn from it.

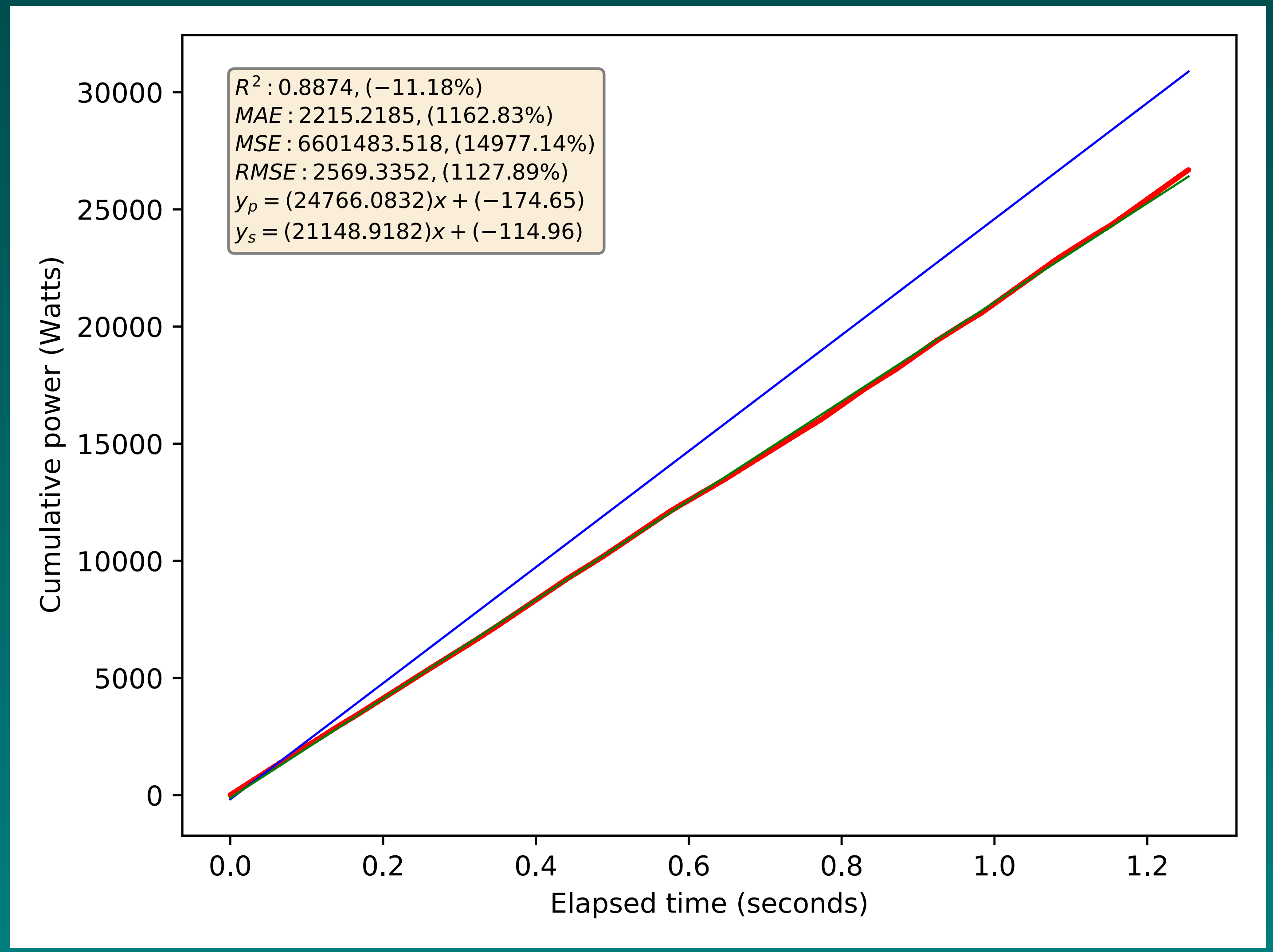
How about those fingerprints?

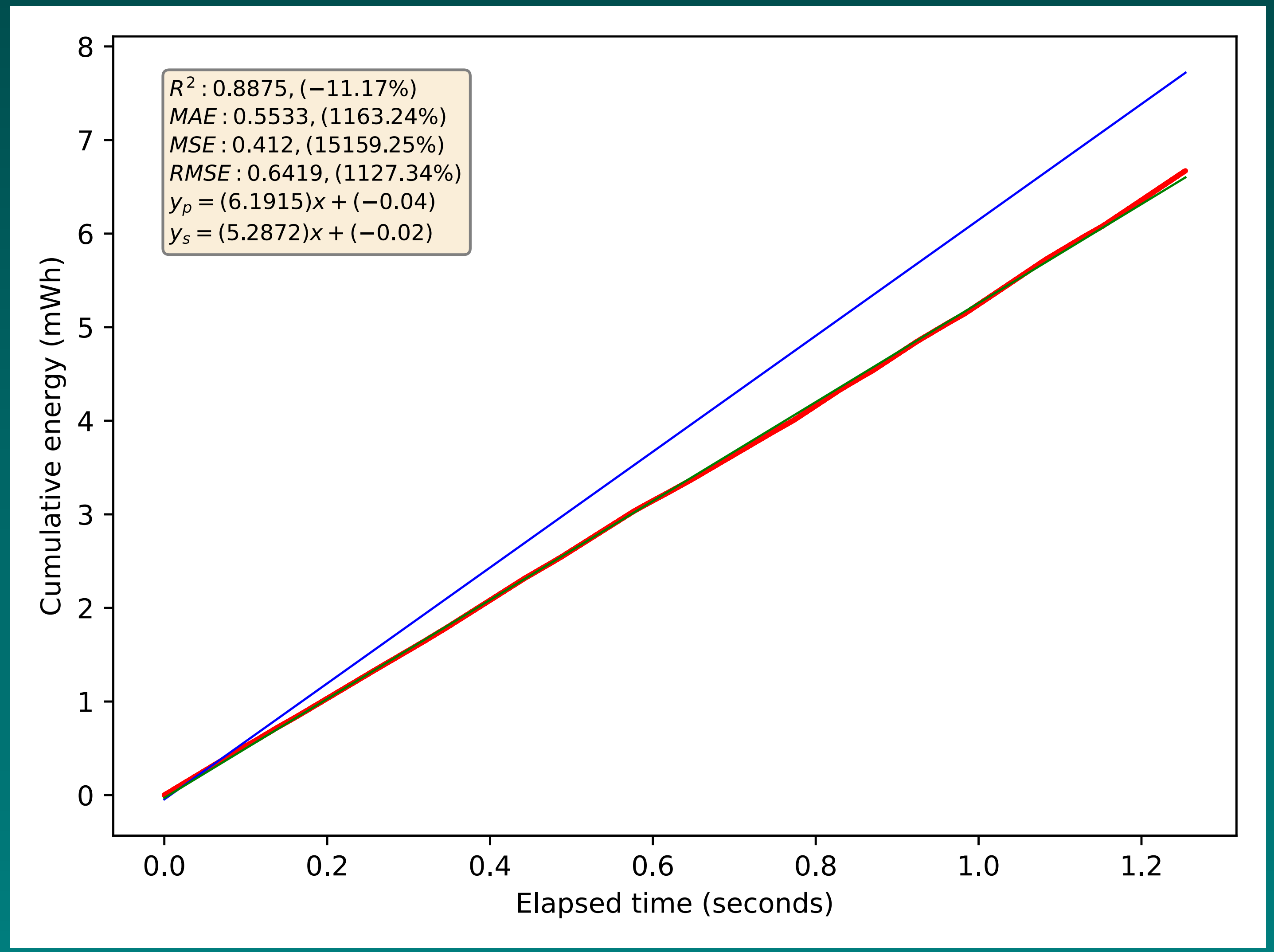










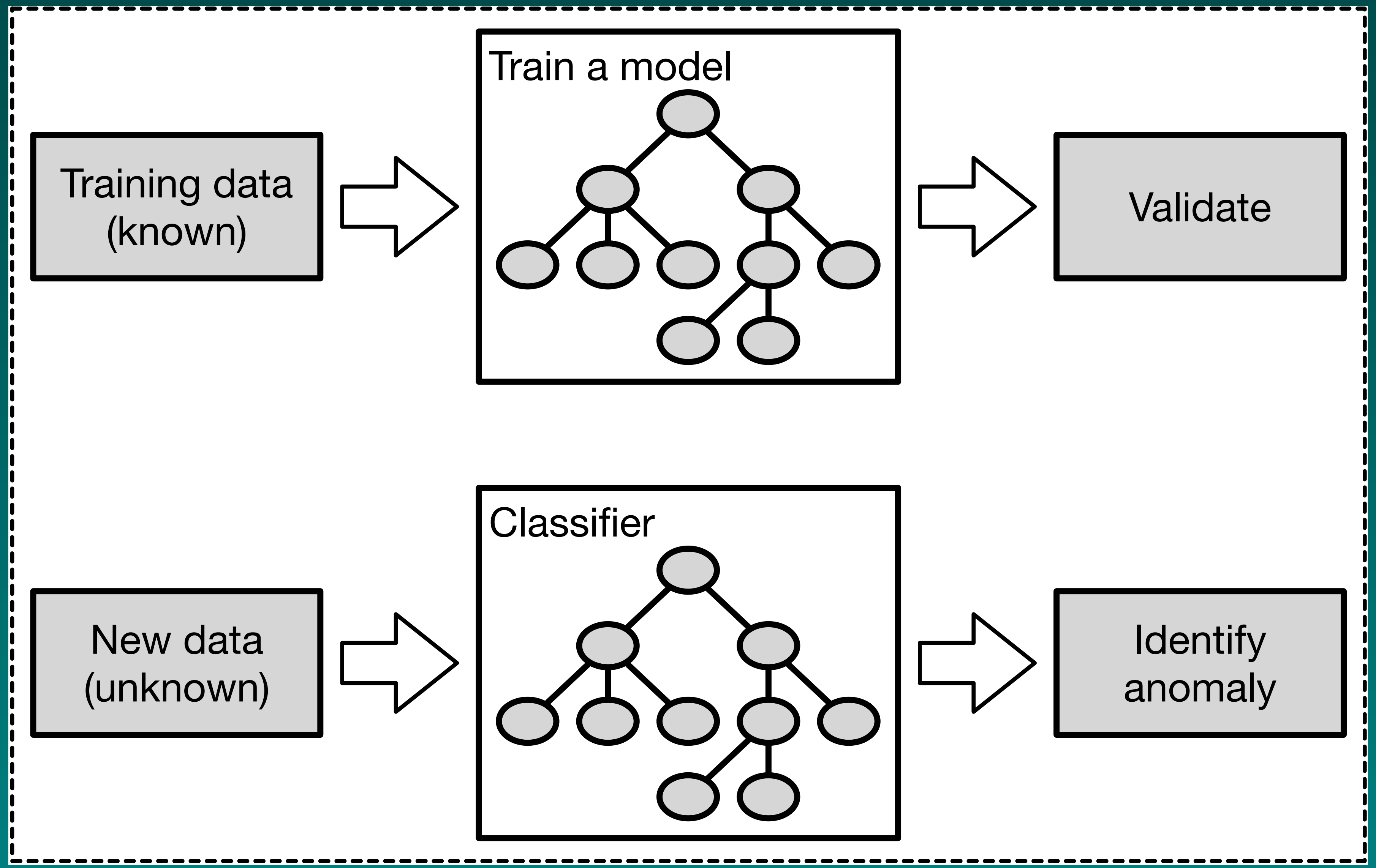


Where there is data, there is AI

- AI models are trained with known data and can react to unknown data
  => Based on what they have learned

- Specifically for us: Classifiers
  => Use known data to train the model
  => Teach it different behavioural categories
  => Use the model to classify unknown behaviour

- There are different types of classifiers
  => Traditional models, e.g., Decision Trees
  => And deep learning models, specifically, Convolutional Neural Networks (CNN)



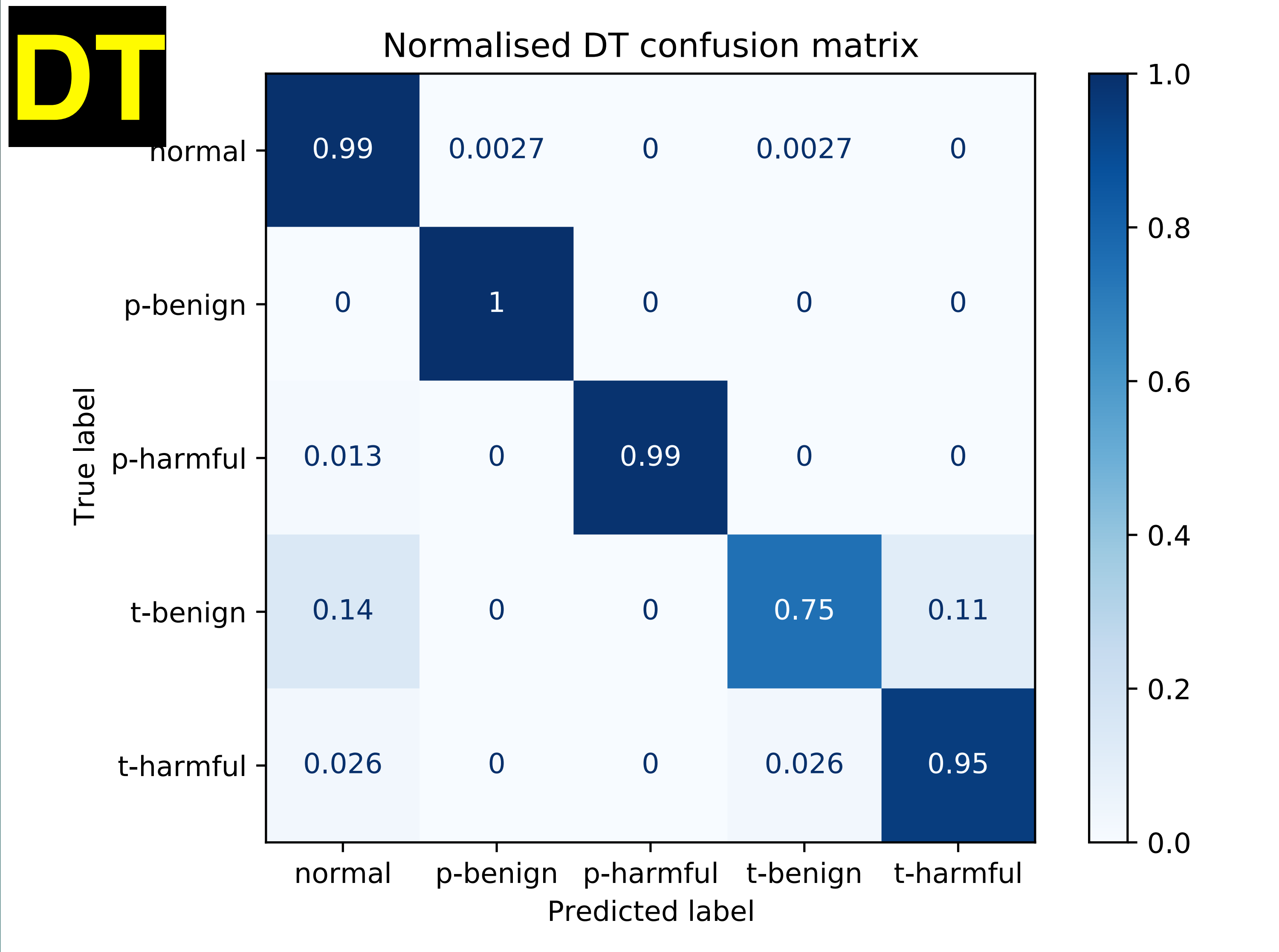
Where there is data, there is AI

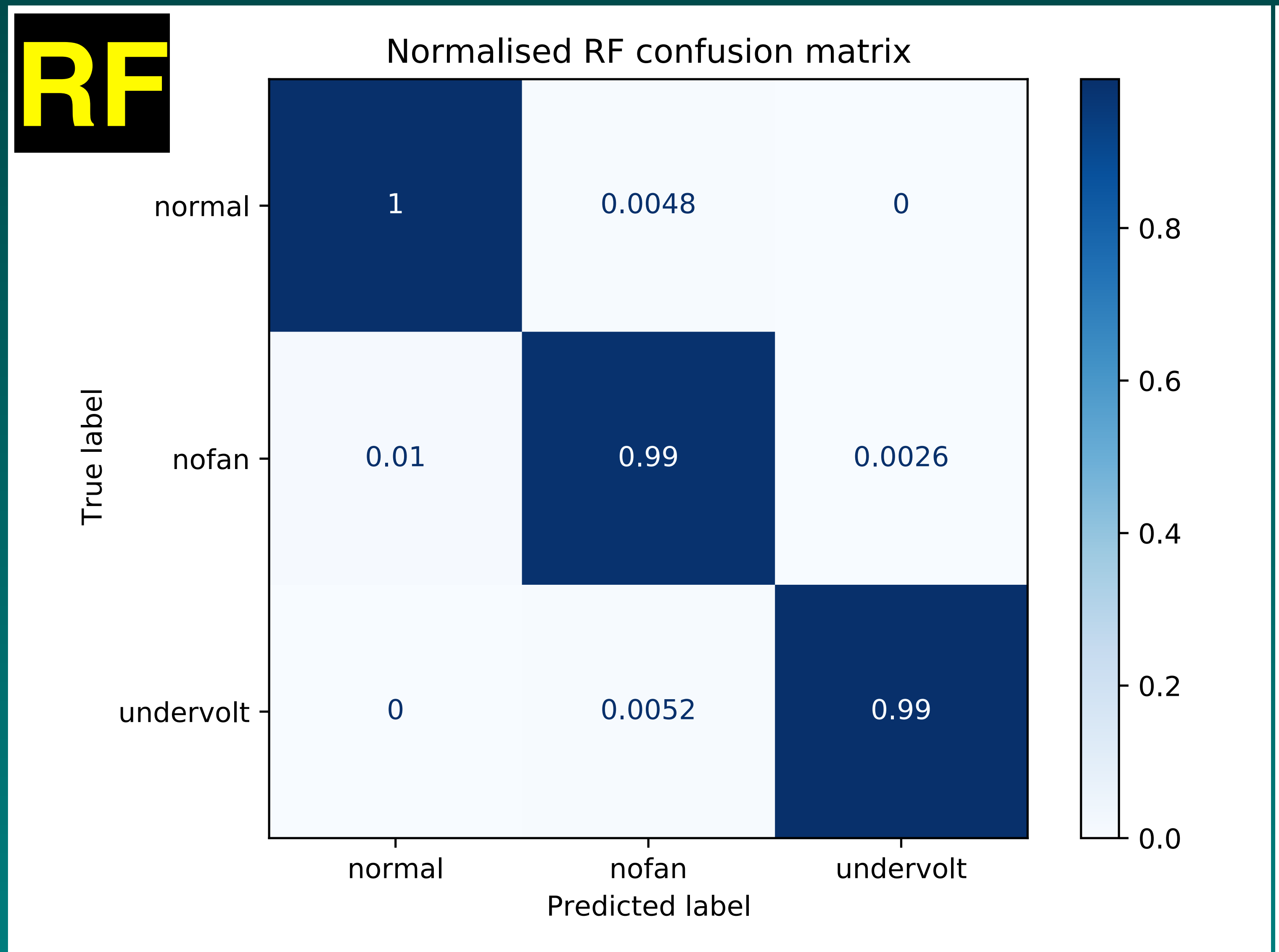
- How good is a classifier?
  => This is evaluated by the accuracy of its predictions
  => Classifiers learn from known data
  => Learning performance varies

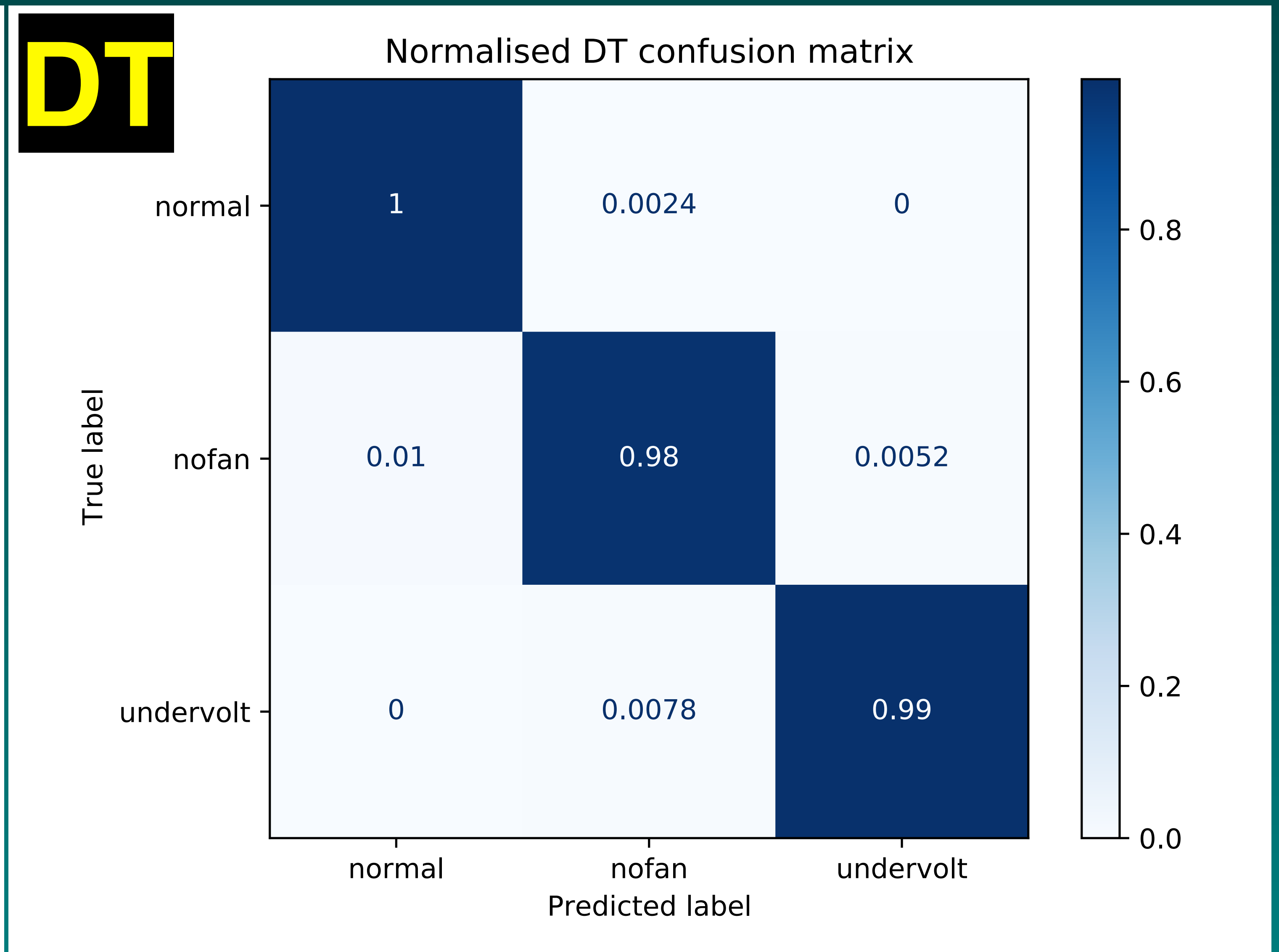
- Correct classification rates:
  => Solution using traditional ML: **99.23%**
  => Solution using DL with CNN: **94.85%**
  (with room to improve …)

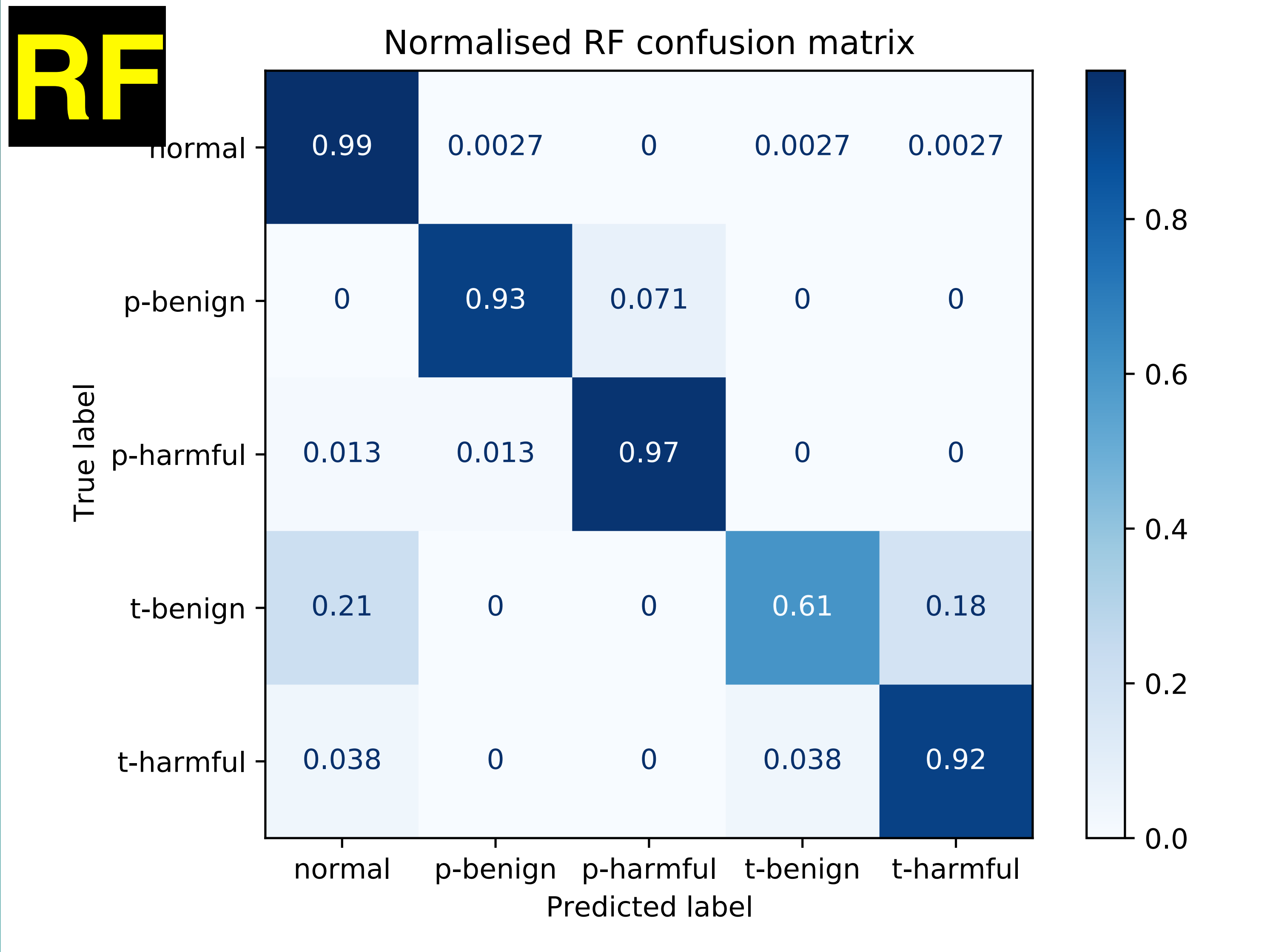
Where there is data, there is AI

- How good is a classifier?
  => Evaluated by the accuracy of predictions
  => Classifiers learn from data
  Known data for supervised …
  => Learning performance varies





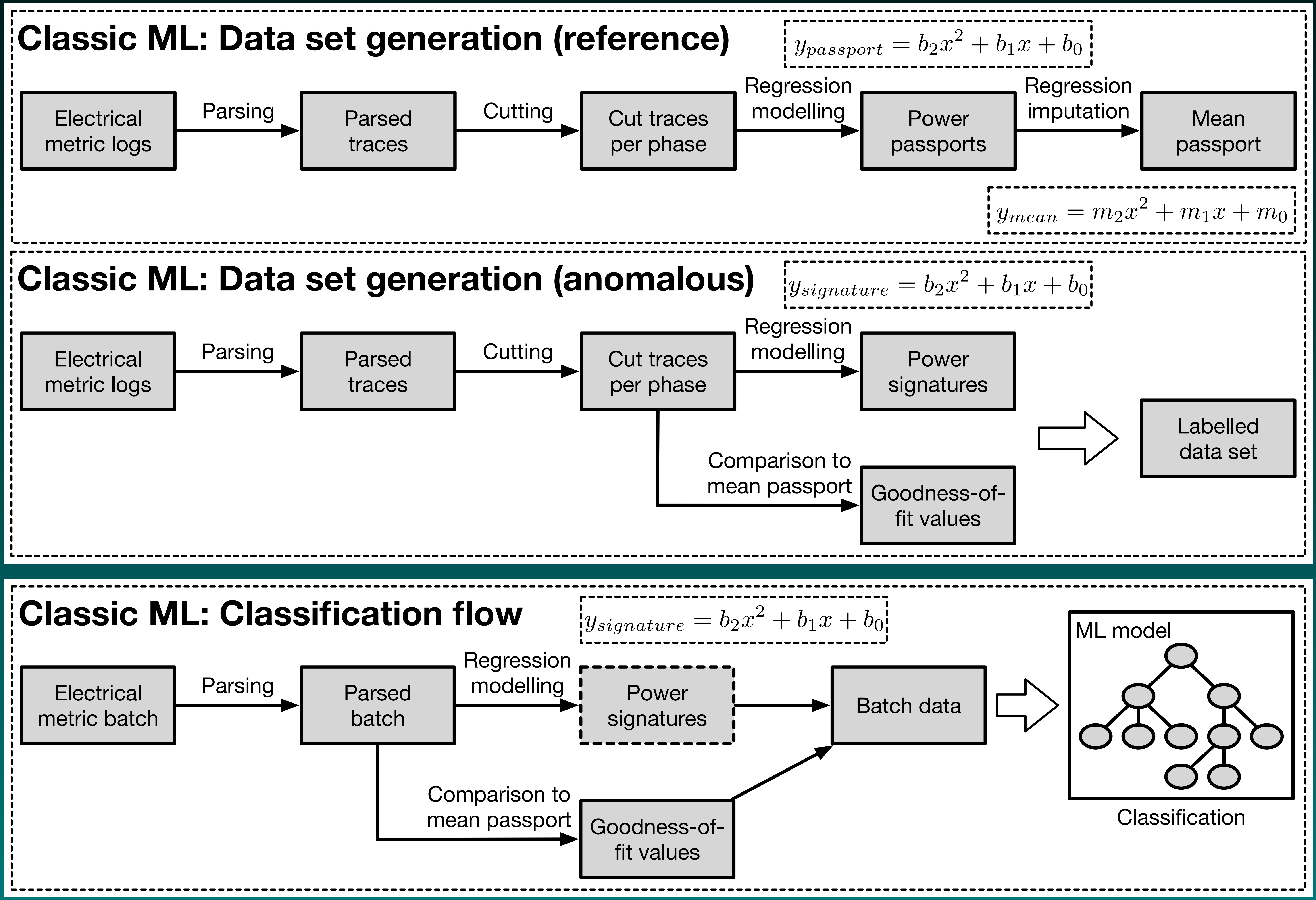




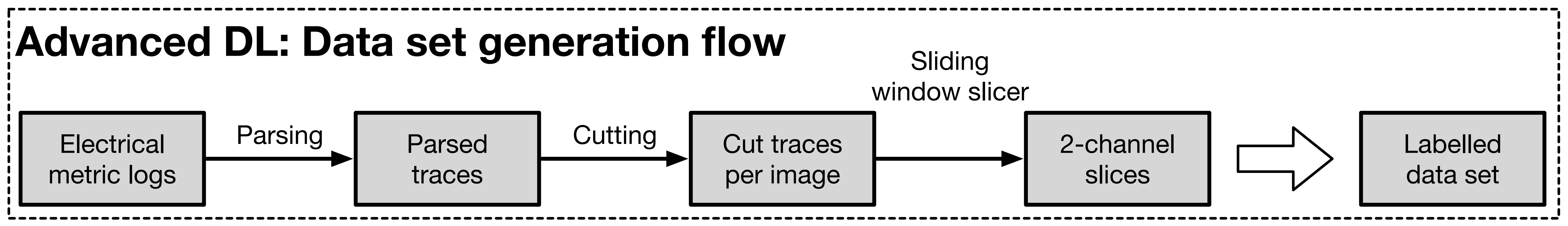
Complete solution 1

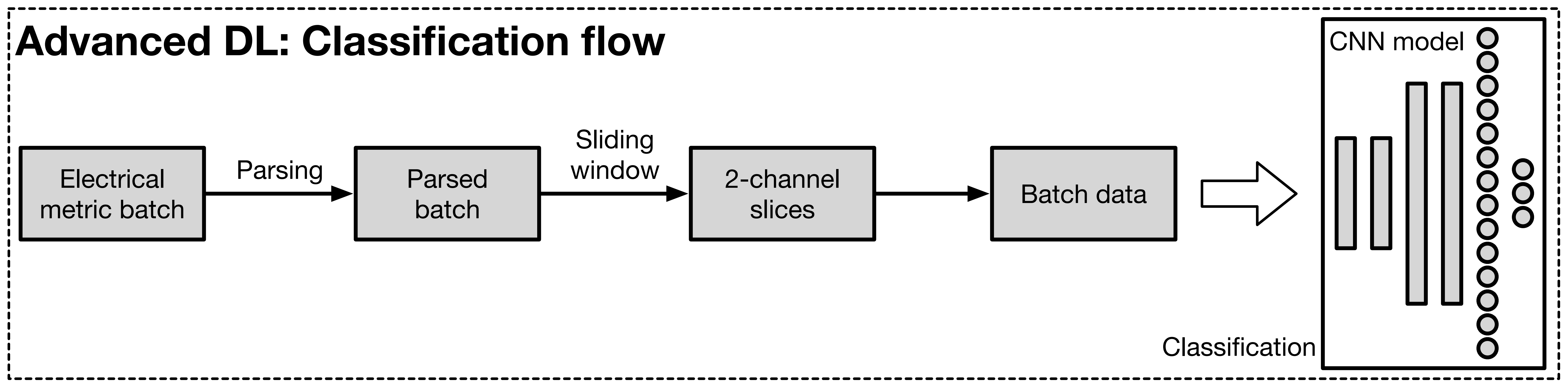
- The solution combines fingerprinting techniques with AI techniques

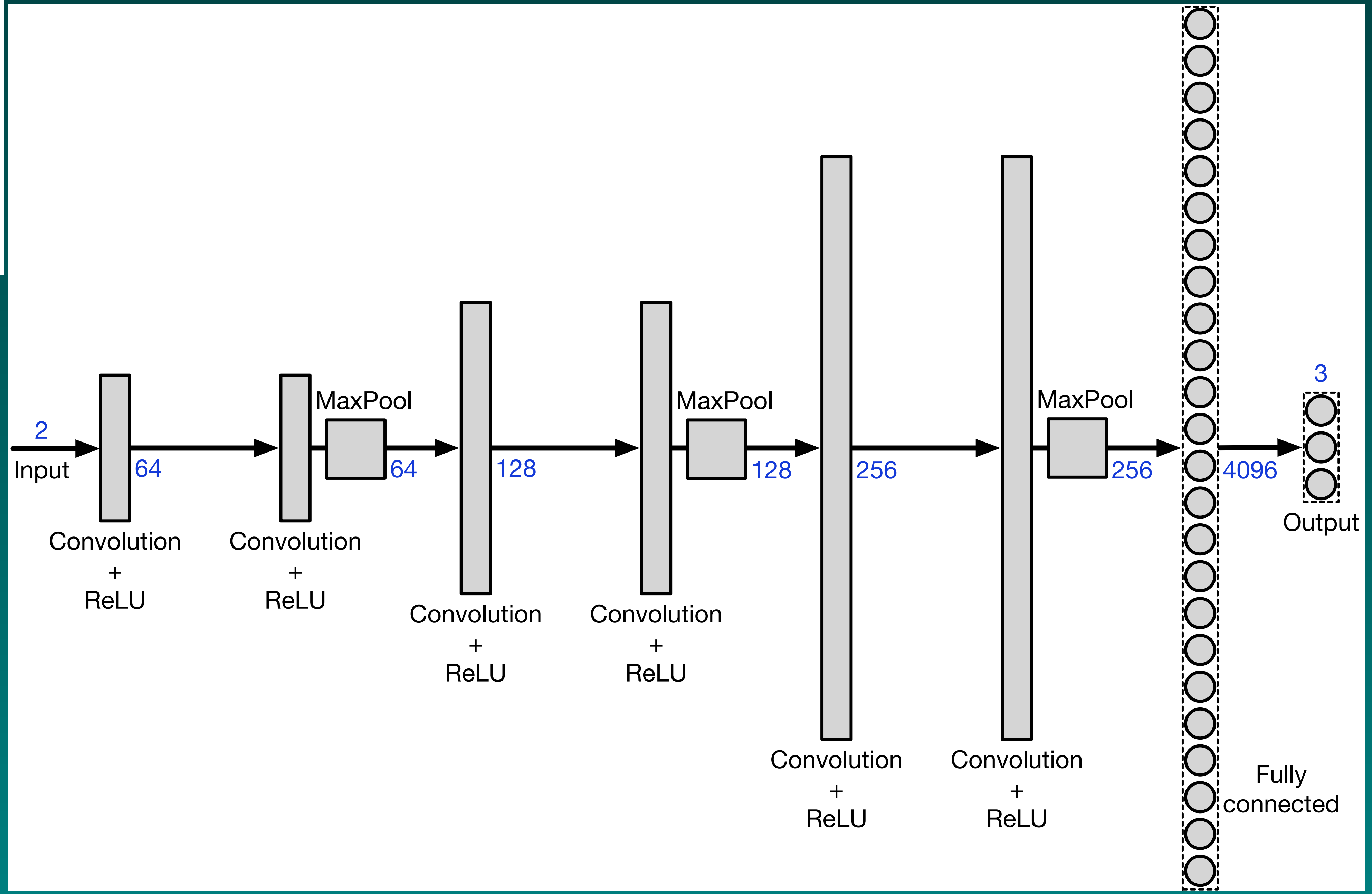
- Different solutions addressing different information positions



Complete solution 2







Comparison

Classic ML vs Advanced DL

- Classic ML advantages:
  => Exceptional accuracy -> **99.23%**
  => Very fast training and inference (after preprocessing)
  => Explainable output (take a look at the DT and backtrack)

- Advanced DL:
  => Good accuracy -> **94.85%** (can be even better)
  => Minimal preprocessing (just proper formatting)
  => Can be relatively quickly put together

To be compared:

- Training speed, classification speed, accuracy, overhead
- Reduction in feature engineering, CNN model design effort

Incorporation of knowledge

- Different information positions will define:
  => Solution specifics
  => Type of model
  => Achievable performance
  (model performance, speed, …)

- Two types of knowledge
  => **Readily** available
  => **Extracted** from data

- Can we develop a framework? Generalise?

		Knowledge position		
		White	**Grey**	**Black**
Data position	**White**	White-White (WW)	White-Grey (WG)	White-Black (WB)
	Grey	Grey-White (GW)	Grey-Grey (GG)	Grey-Black (GB)
	Black	Black-White (BW)	Black-Grey (BG)	Black-Black (BB)

Science does not happen in vacuum

- **UvA:**
  prof. dr. ir. Cees de Laat
  prof. dr. Andy D. Pimentel
  dr. Hugo Meyer
  Simon Polstra
  Julius Roeder
  Dolly Sapra

- **ASML:**
  dr. Evangelos Paradas
  dr. Ignacio Gonzalez Alonso

And many more who
guided us,
supported us with bureaucratic matters,
questioned and criticised our work,
or shared their opinion.

Thank you!