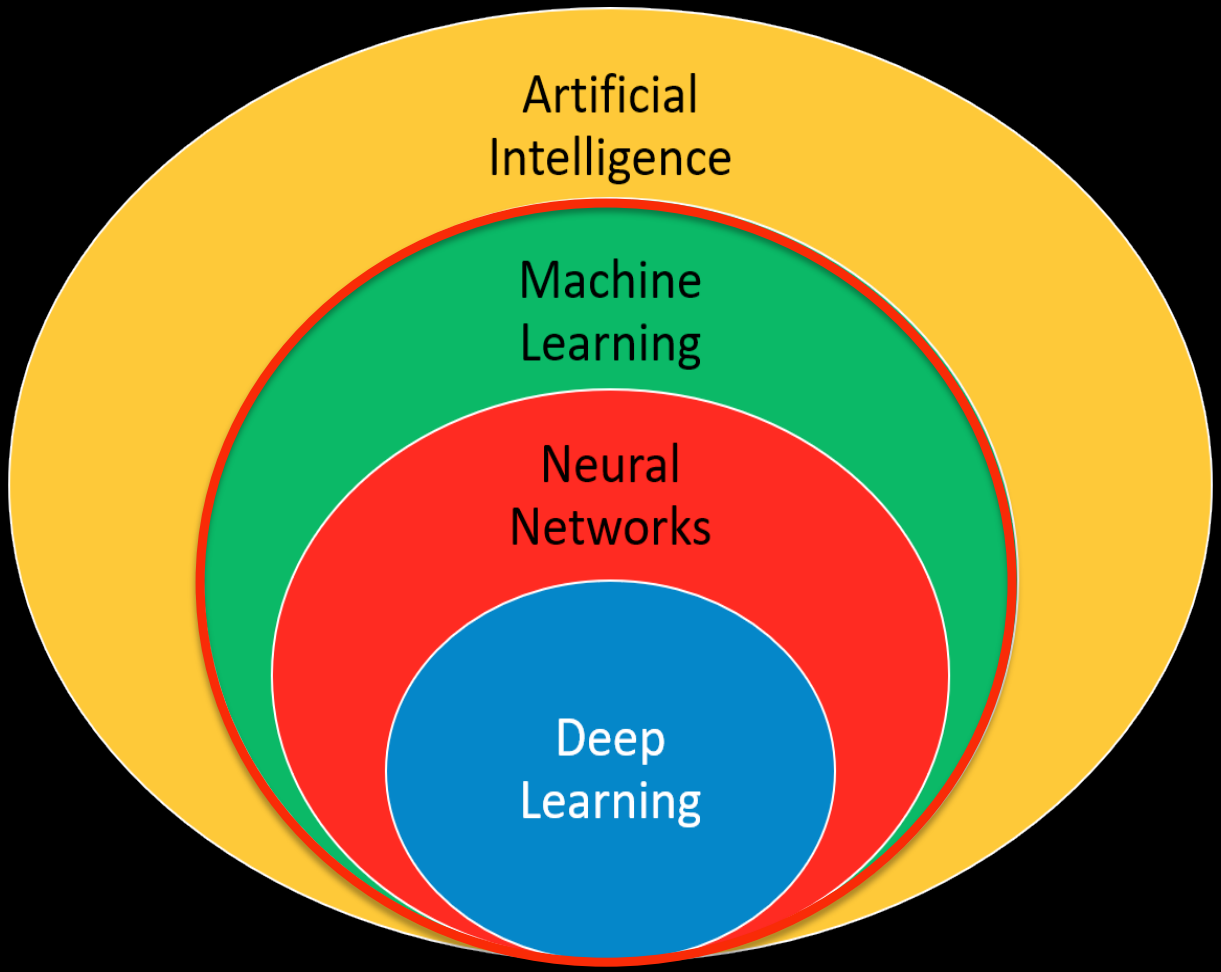

AI engineering
AI Engineering — Meeting New Challenges in System and Software Development of AI-based Systems
CLOSER 2021. 20201-04-30
Ivica Crnkovic, ivica.crnkovic@chalmers.se

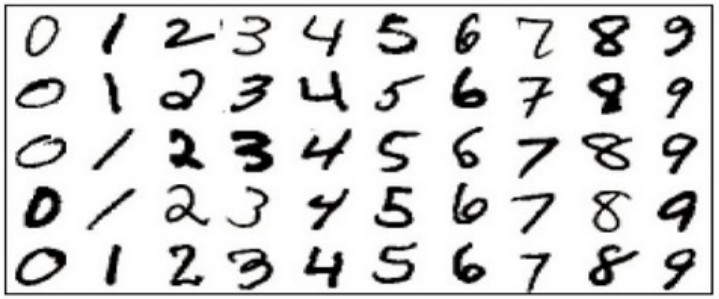

Ivica Crnkovic
Chalmers University of Technology
Professor in Software Engineering
Director of Chalmers AI Research Centre (CHAIR)
ivica.crnkovic@chalmers.se


Artificial Intelligence
Machine Learning
Neural Networks
Deep Learning

Types of Machine Learning

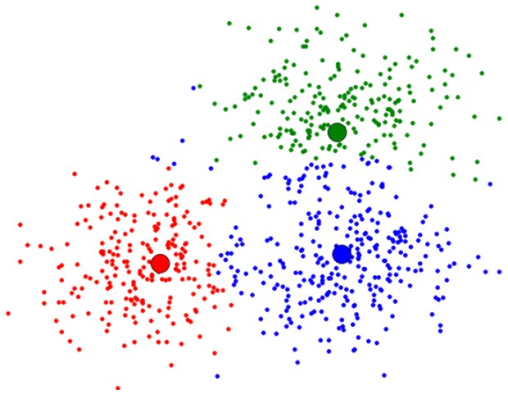
SUPERVISED
95% of ML

Task Driven
predict next value, classification



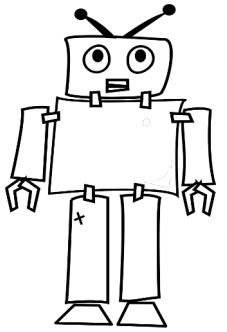
UNSUPERVISED

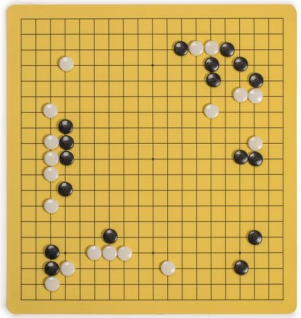
Data Driven
Identify clusters

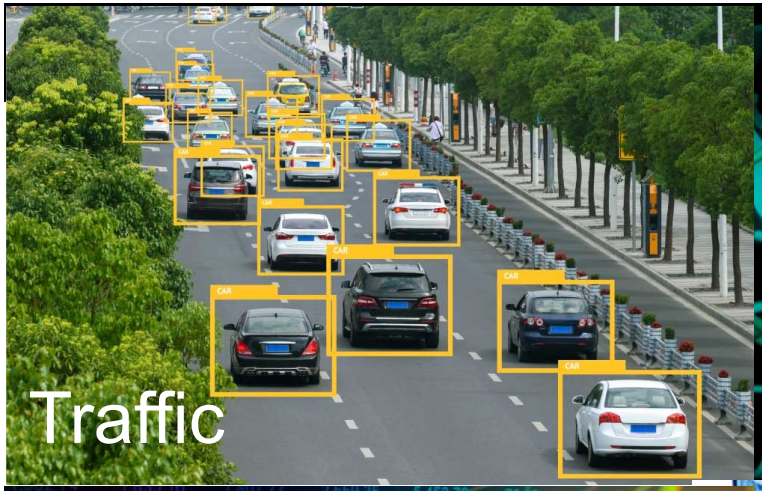


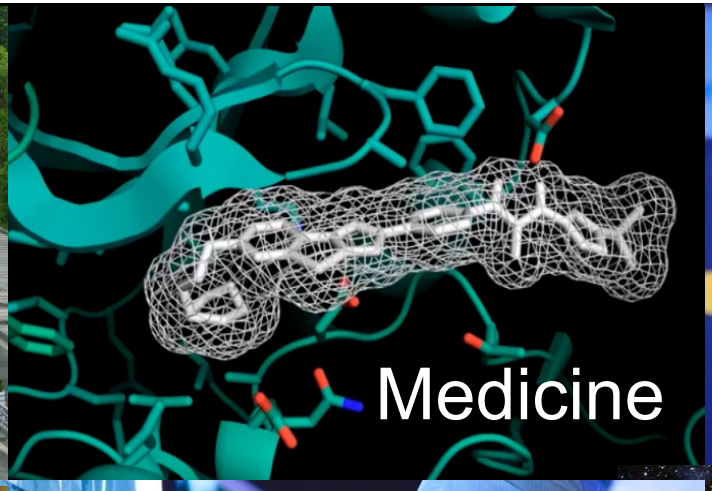
REINFORCEMENT

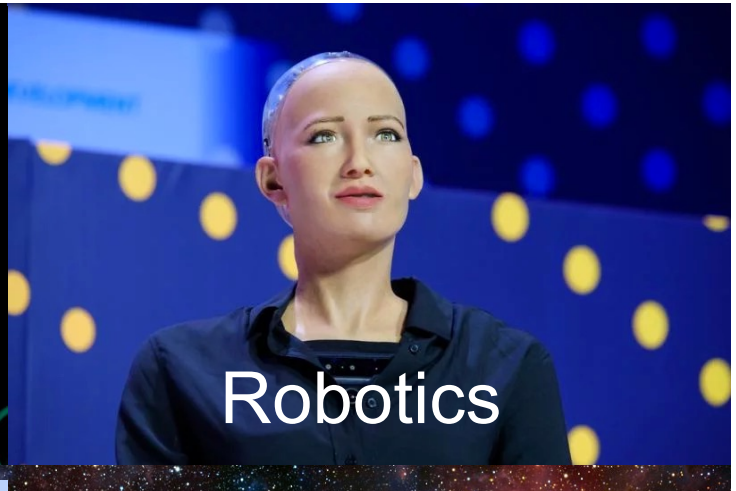
Learn From Mistakes
Trial and Error

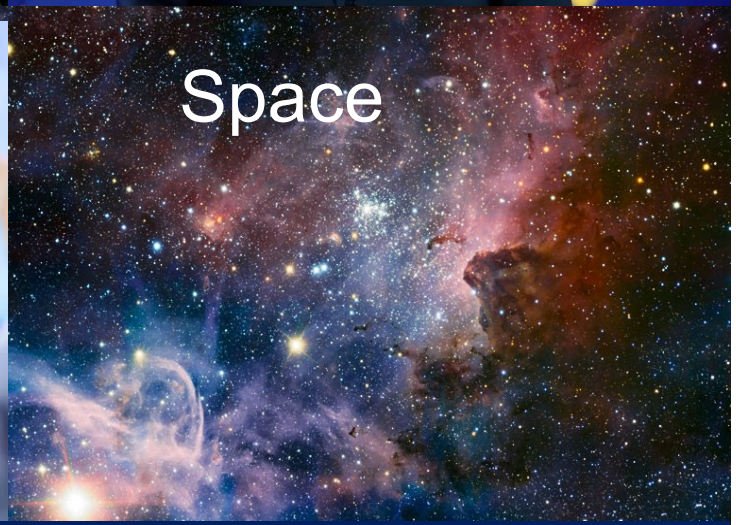












AI success Stories

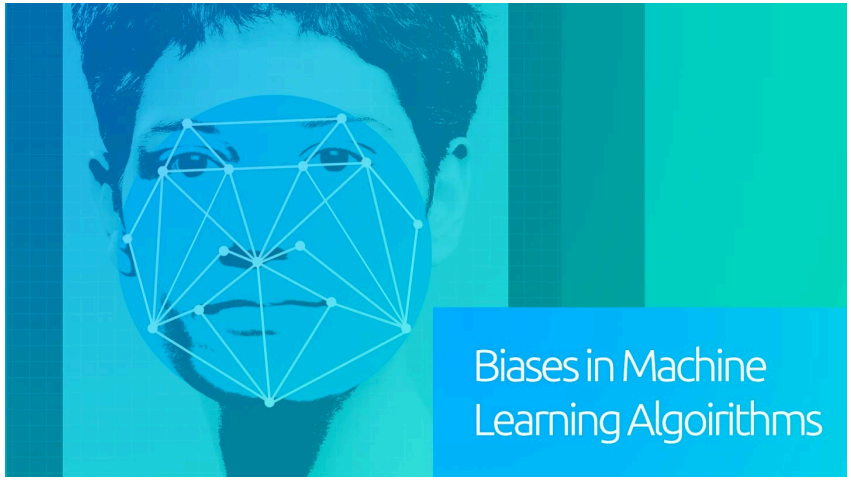
virtually no area that does not (plan to) use AI

AI failures 2018

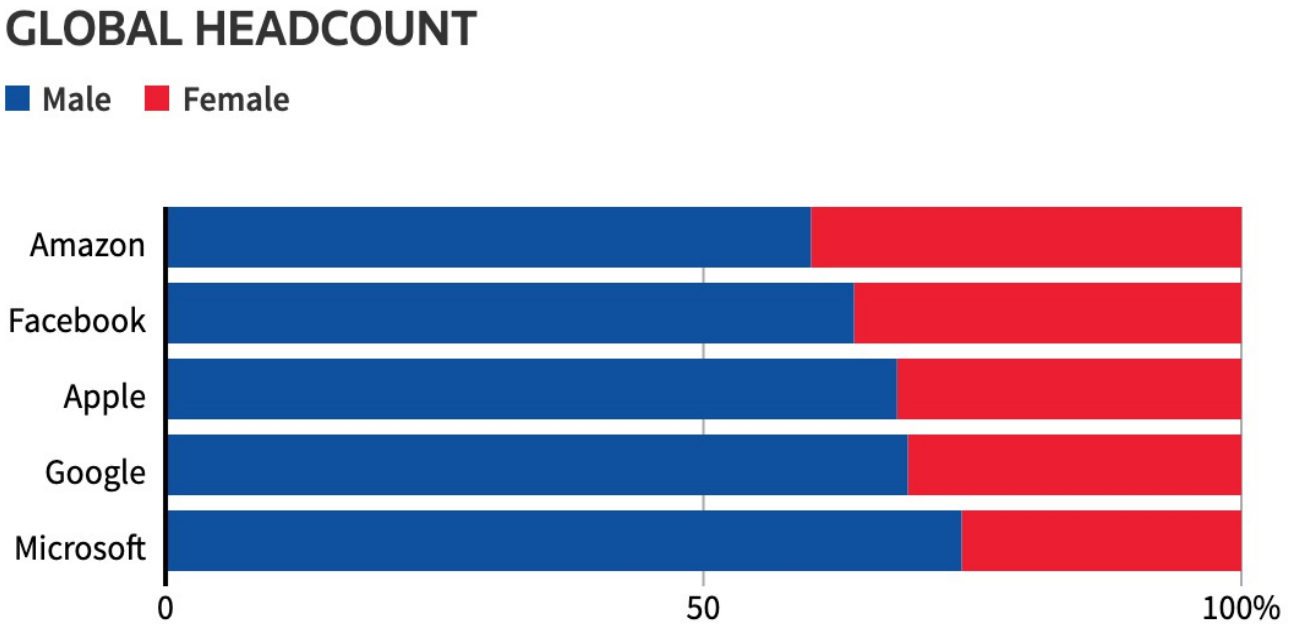
https://medium.com/syncedreview/2018-in-review-10-ai-failures-c18faadf5983

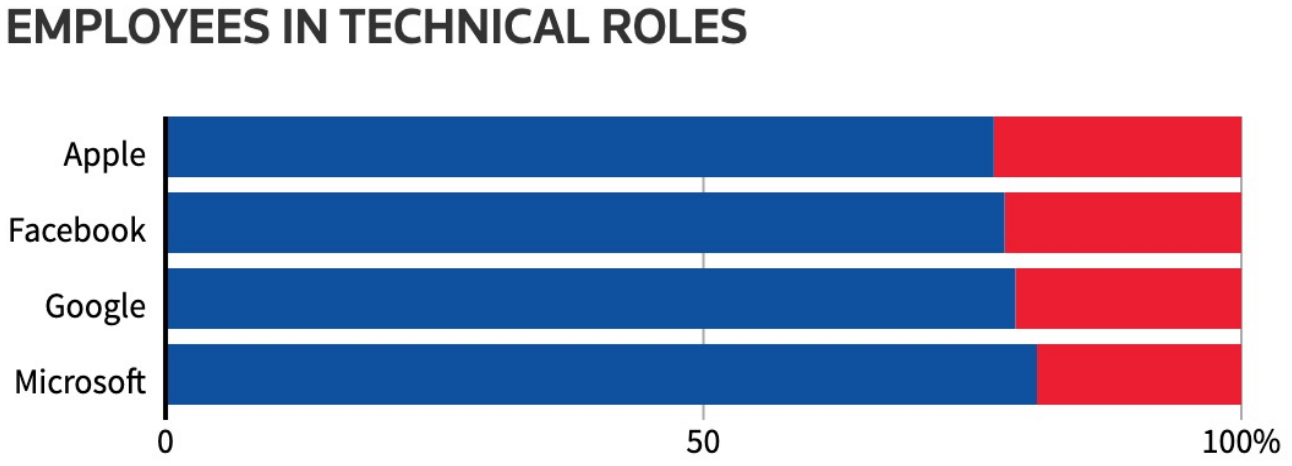
AI failures 2018

Amazon AI recruiting tool is gender biased



https://www.reuters.com/article/us-amazon-com-jobs-automation-ins





AI failures 2018

AI World Cup 2018 - predictions almost all wrong ☺

Why these failures?

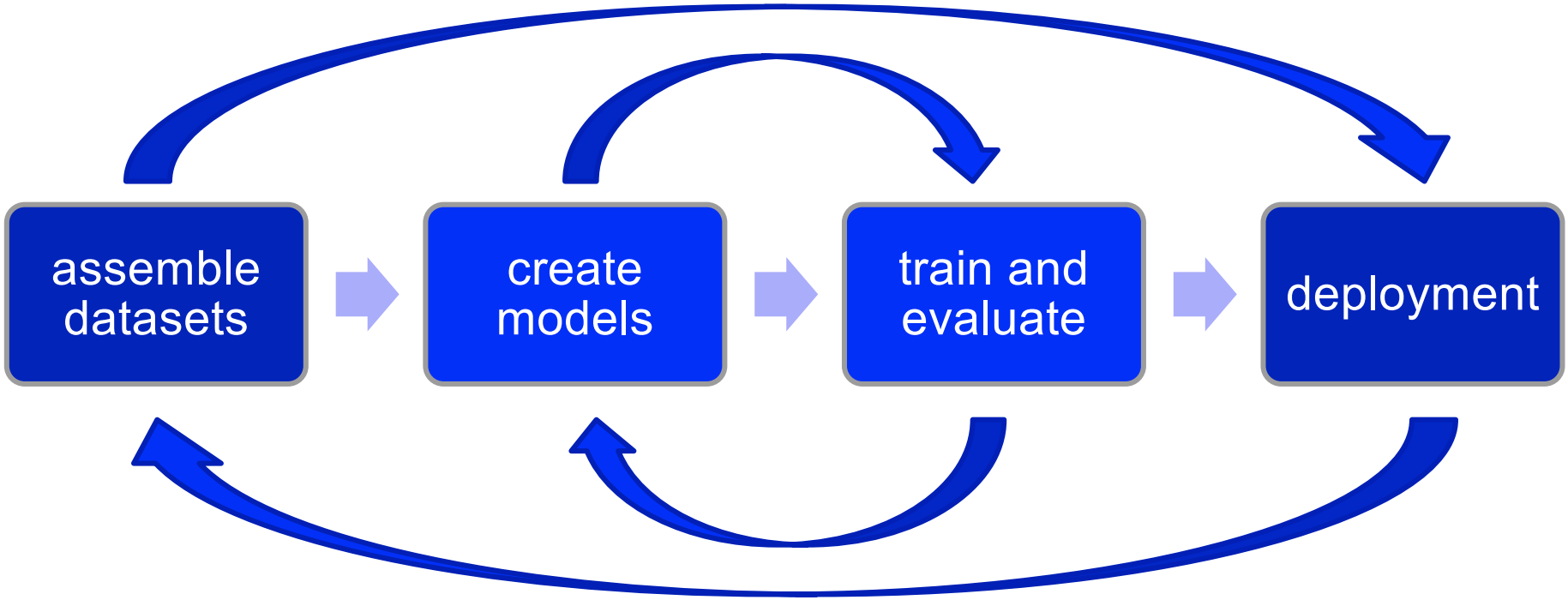
What are the challenges in AI development?

New questions in AI Development

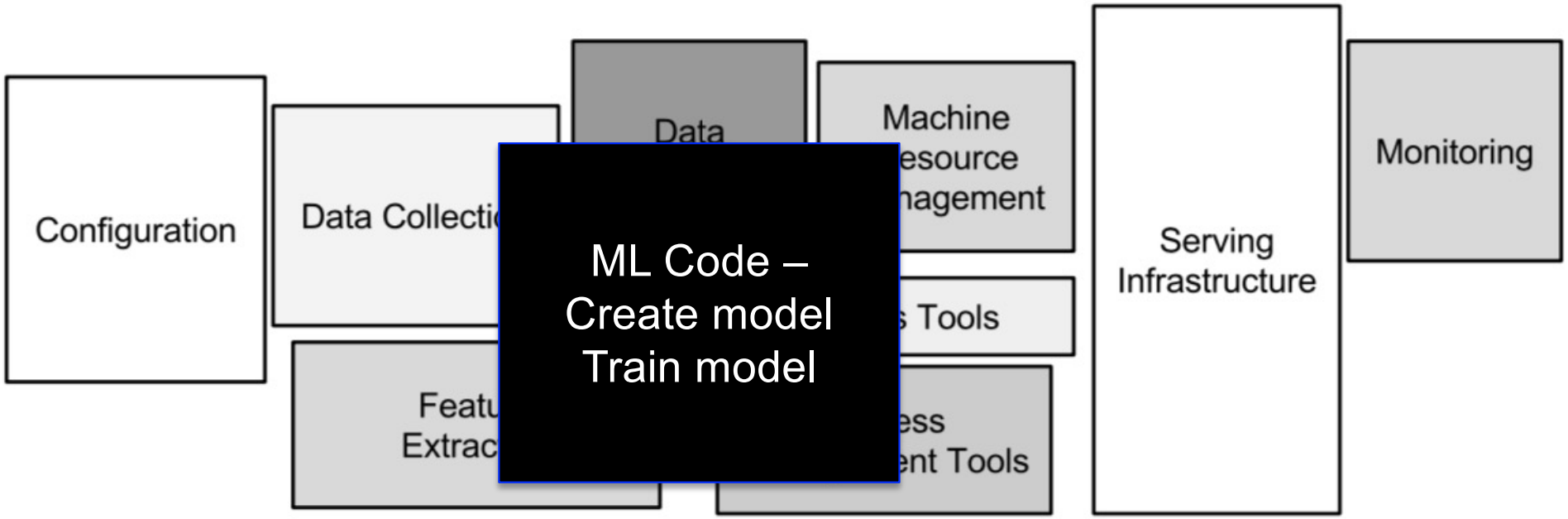
- **Non technical character**
  - Who is owner of data?
  - What are the ethical aspects of using data?
  - What can we allow a machine to decide?
  - How do interpret the results from AI models?

- **Technical nature**
  - How to efficiently collect, store, process, analyse, and present data?
  - How to efficiently build the AI-based systems?
  - How to ensure enough resources (computation, storage, timing)
  - How to ensure dependability/trustworthy of such systems?
  - What system and software architecture are required for AI-systems?
  - **WHAT KIND OF SOFTWARE ENGINEERING SUPPORT IS NEEDED?**

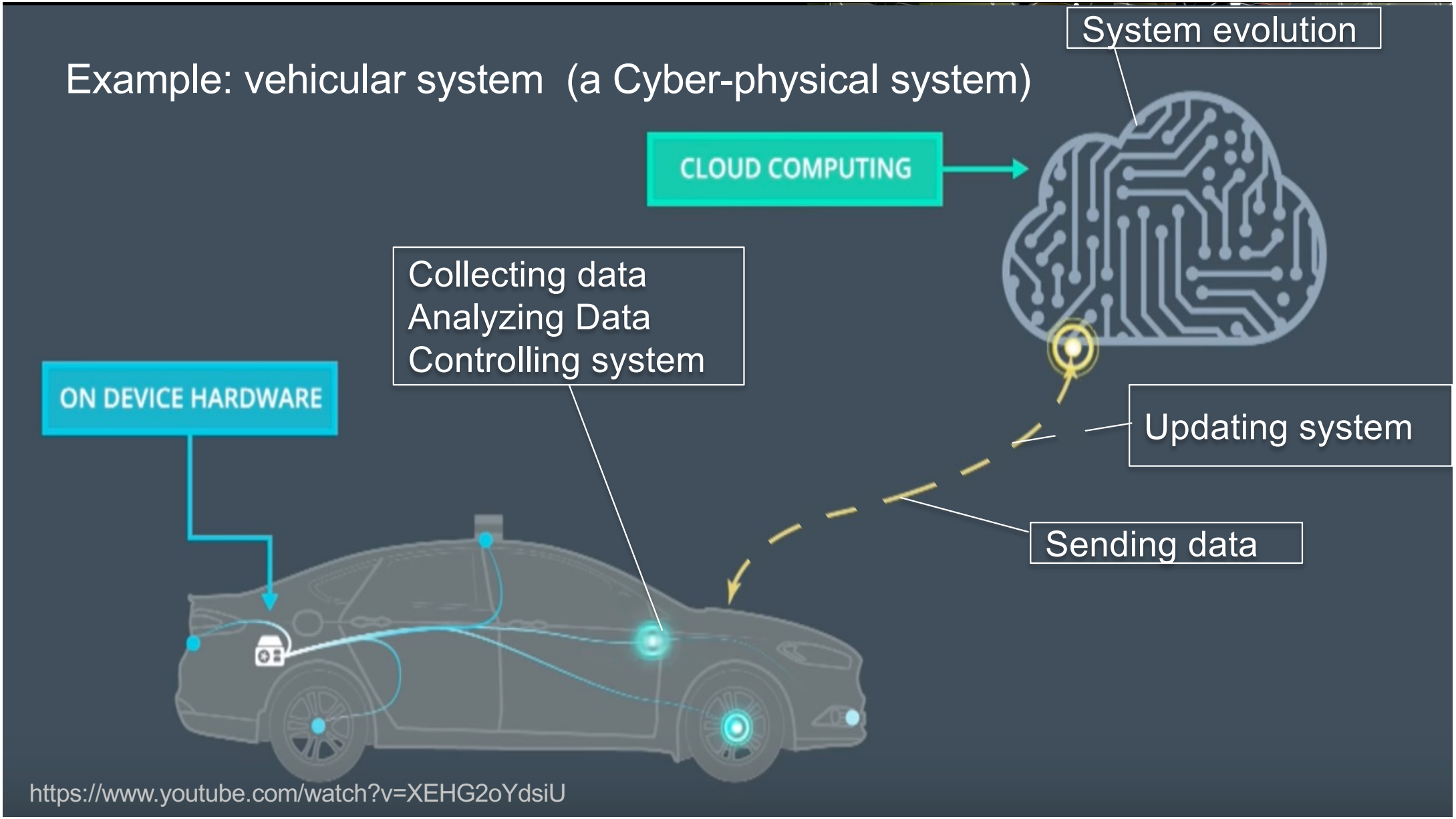
ML Development cycle



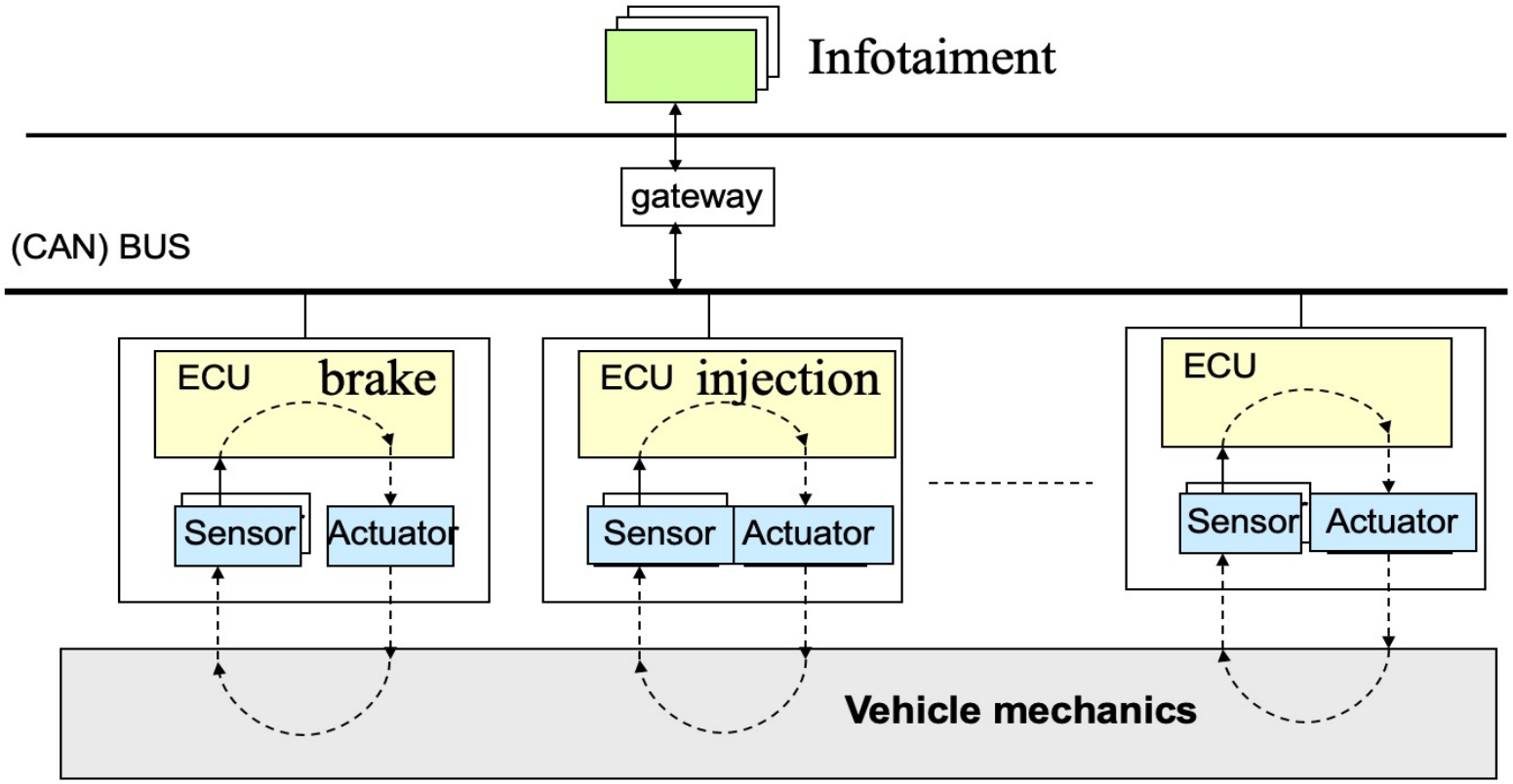
Life cycle



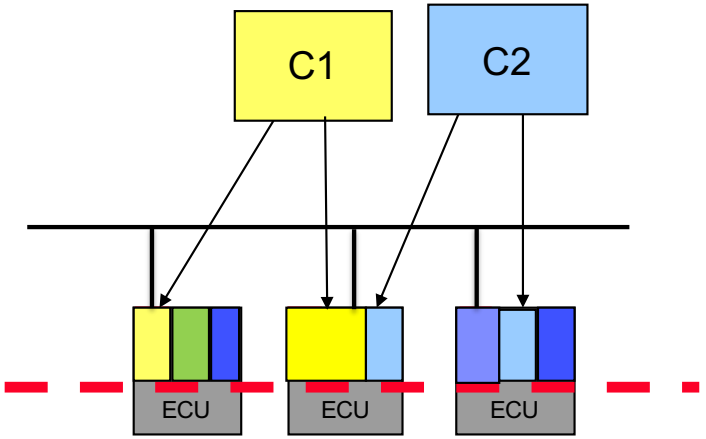
Machine Learning: The High-Interest Credit Card of Technical Debt, *D. Sculley, G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young*


Example: vehicular system (a Cyber-physical system)
System evolution
CLOUD COMPUTING
Collecting data
Analyzing Data
Controlling system
ON DEVICE HARDWARE
Updating system
Sending data
https://www.youtube.com/watch?v=XEHG2oYdsiU

Architecture of a car control system



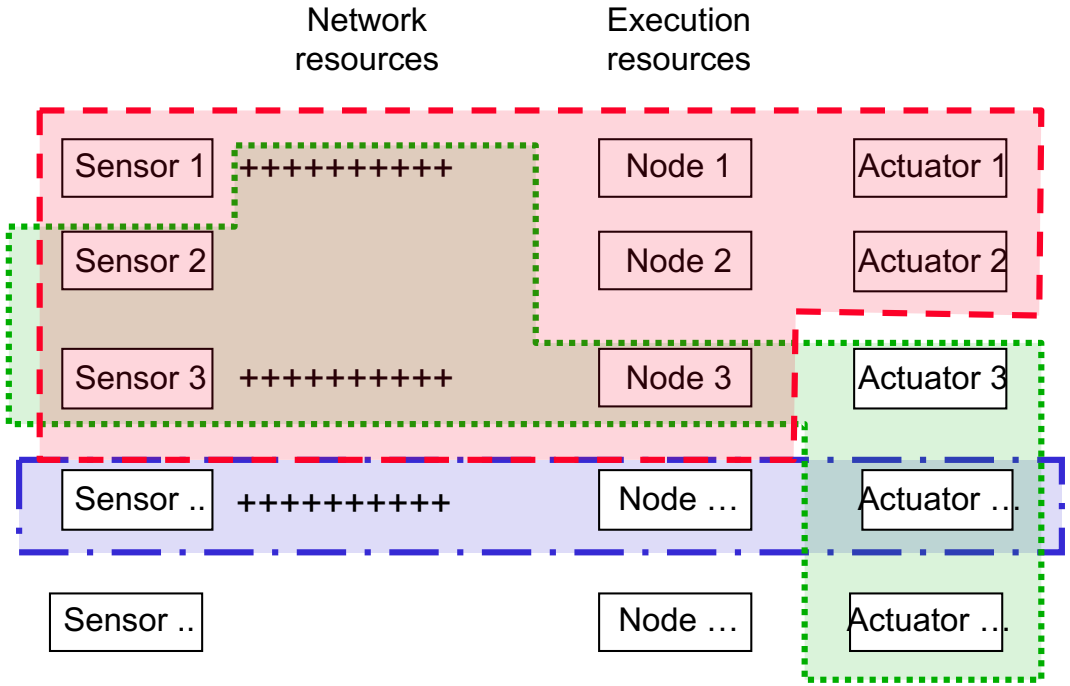
Complex services – distributed components



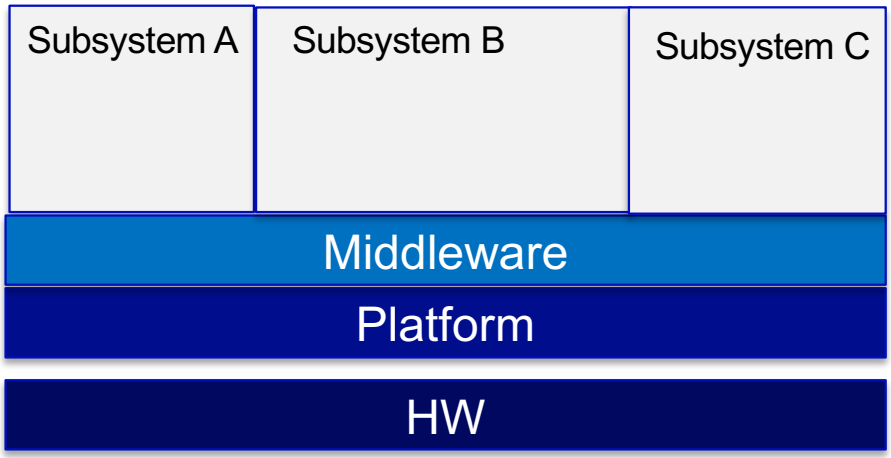
Challenges:

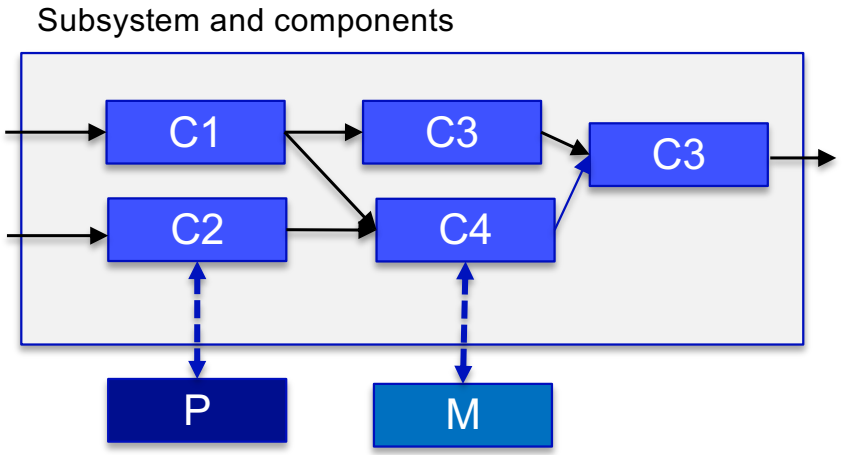
- Real-time requirements
- Shared resources
- Resource constraints

Shared resources



Software architecture

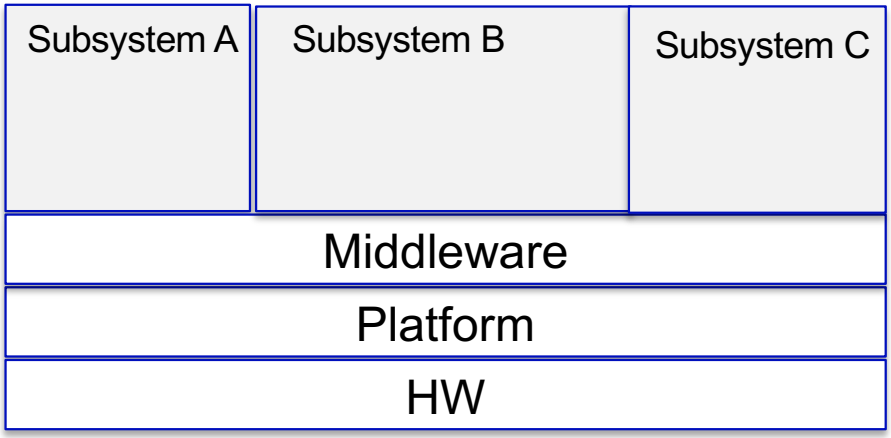


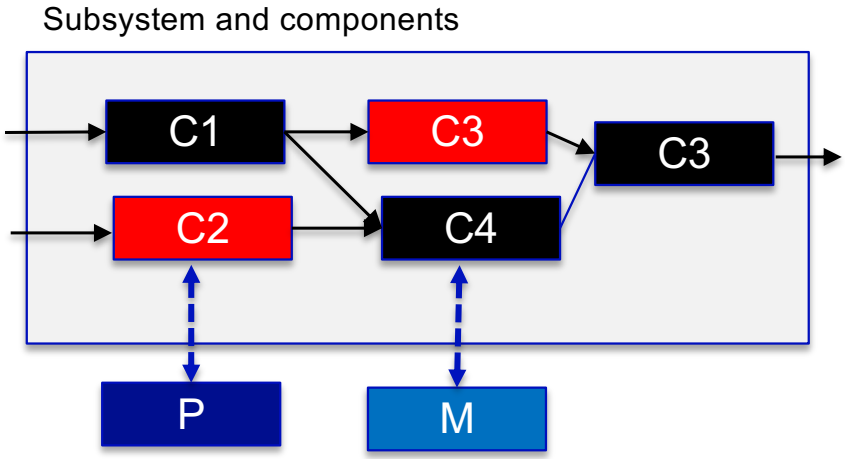


Component-based and service-based approach

- Components
  - Encapsulation of data
  - Encapsulation of functionality
  - Dependency between components defined and controlled

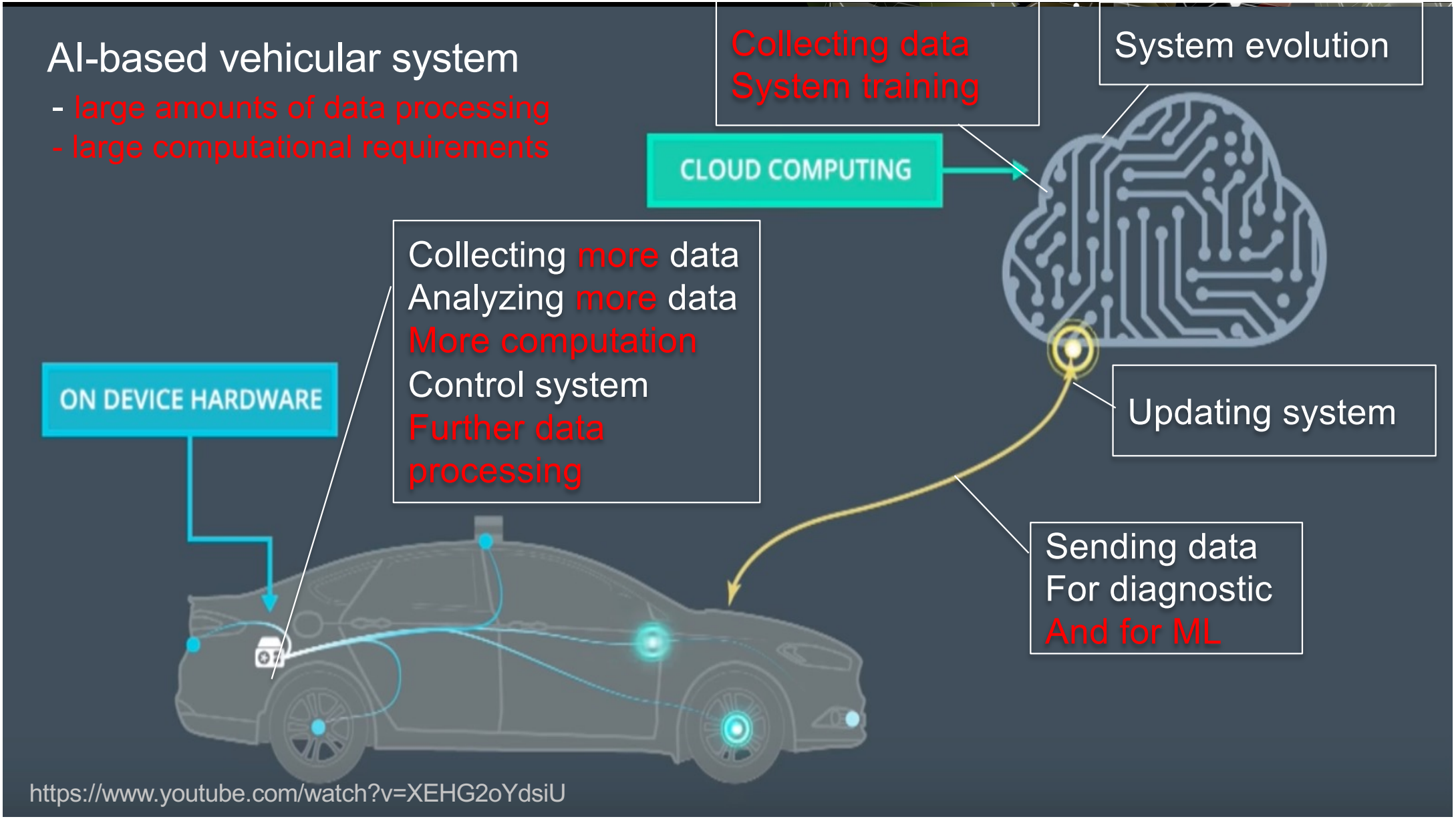
AI-System - system architecture (example)



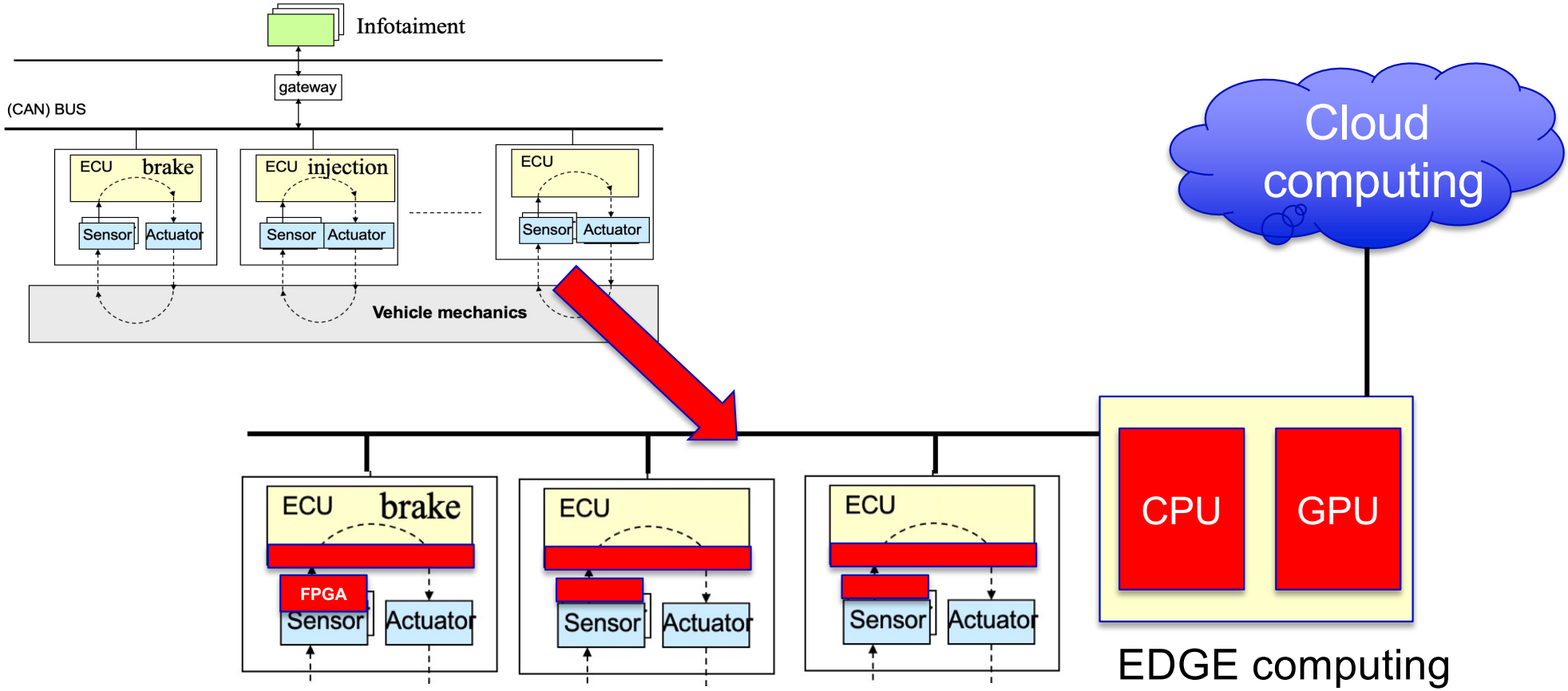


Components – black boxes

- Components
  - Encapsulation of AI-based functionality
  - Dependency between components defined and controlled
  - **What about AI code and data?**


AI-based vehicular system
- large amounts of data processing
- large computational requirements
Collecting data
System training
System evolution
CLOUD COMPUTING
Collecting more data
Analyzing more data
More computation
Control system
Further data processing
ON DEVICE HARDWARE
Updating system
Sending data
For diagnostic
And for ML
https://www.youtube.com/watch?v=XEHG2oYdsiU

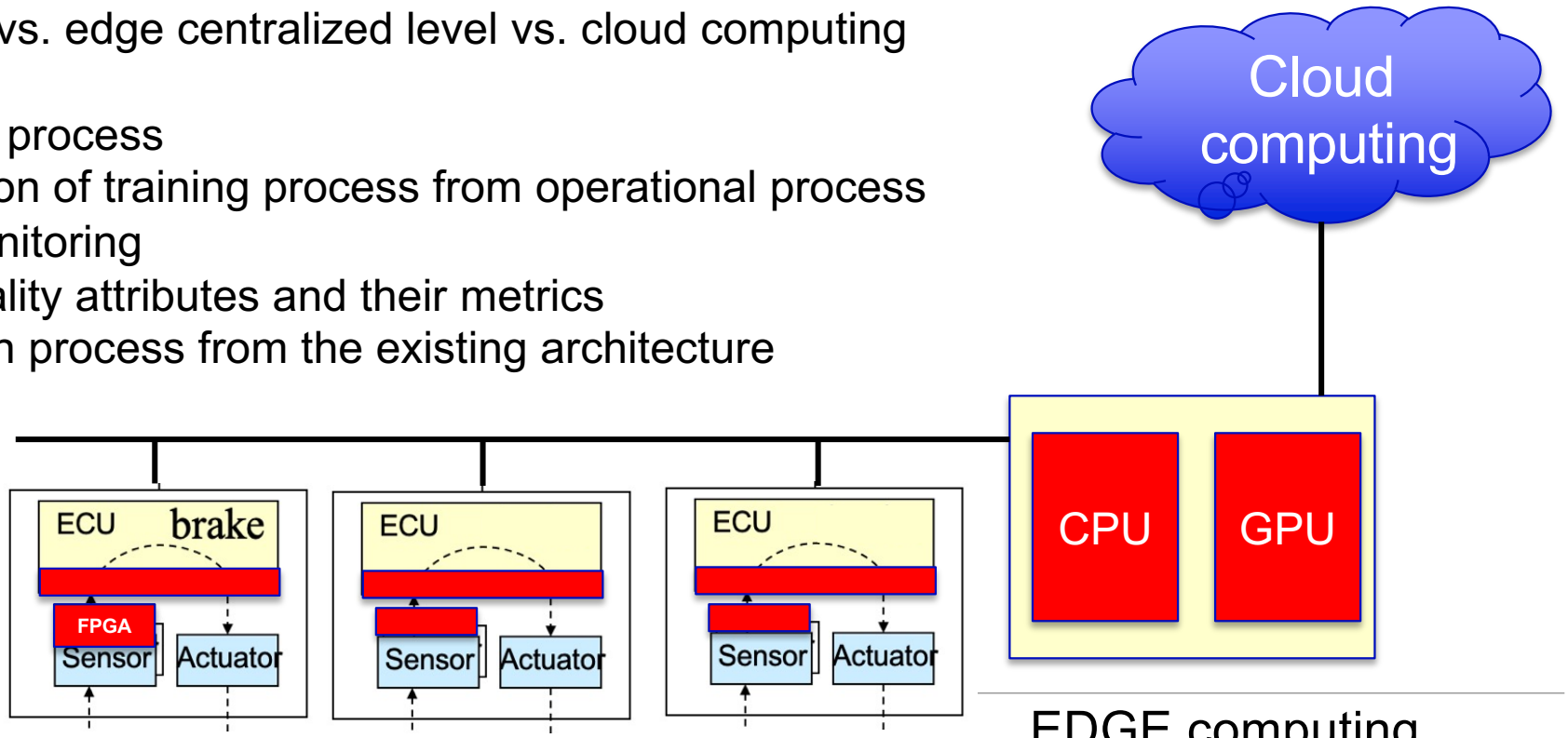
With AI the existing architecture is not feasible



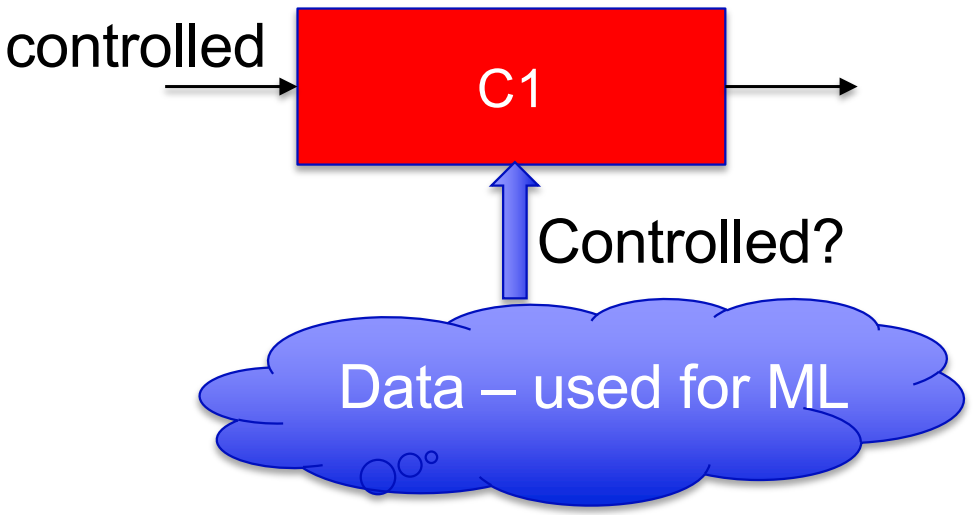
Challenges

- New computation models
- Heterogenous platform
- Edge sensor vs. edge centralized level vs. cloud computing

- Development process
  - separation of training process from operational process
  - New monitoring
  - New quality attributes and their metrics
  - Migration process from the existing architecture



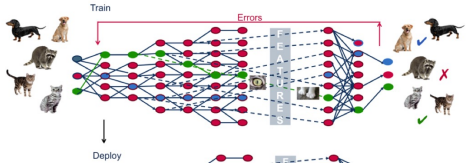
AI- based components – fundamental problems



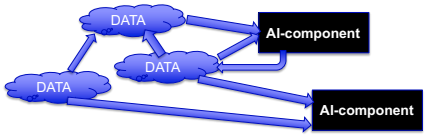
- The results depend not only on the algorithms and controlled data but also on uncontrolled/unknown data
- The AI-based functions are not continuous: small change of data can cause big changes
- **Due to data impact on the model many new challenges appear**

Examples of Data-related challenges

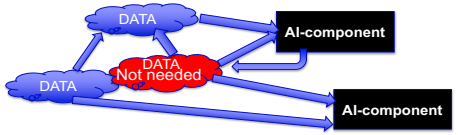
1. Entanglement (Data fusion)



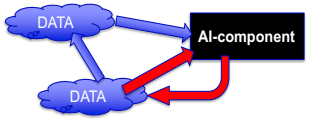
2. Data dependencies
   - Unstable data dependencies



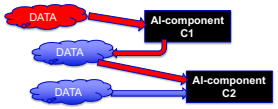
   - Underutilized Data Dependencies



3. Hidden Feedback Loops



4. Undeclared Consumers



5. Correction Cascades

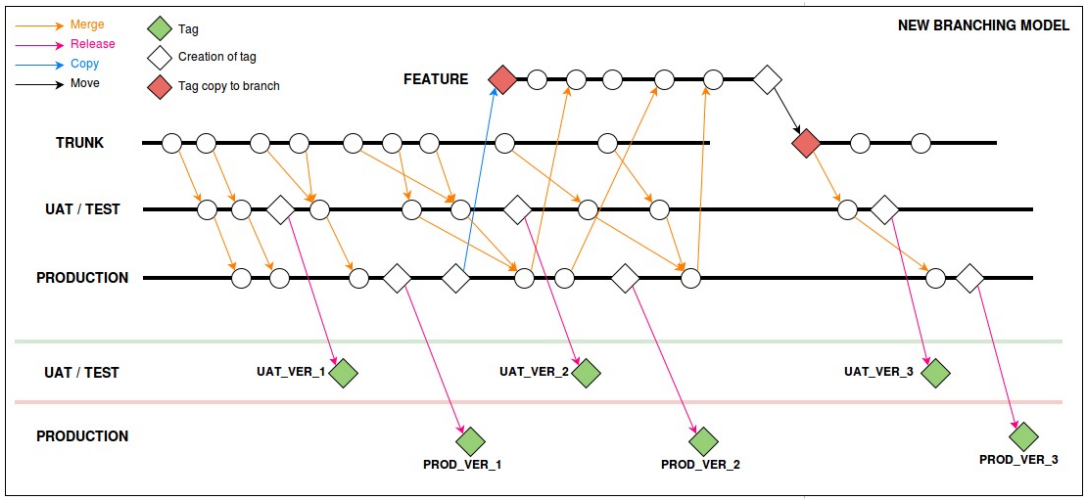
System-design anti-patterns (1)

- **Heterogeneity of data** (different formats, accuracy, semantics) and use of standard ML functions require a lot Glue code
  - *95% of code in AI-based systems is a glue-code (empirical data)*
  - *Requires*
    - *Frequent refactoring of code*
    - *Re-implementing AI models*
- **Pipeline Jungles**
  - ML-friendly format data become a jungle of scrapes, joins, and sampling steps, intermediate files
    - Requires – a close team work of data and domain engineers

System-design anti-patterns (2)

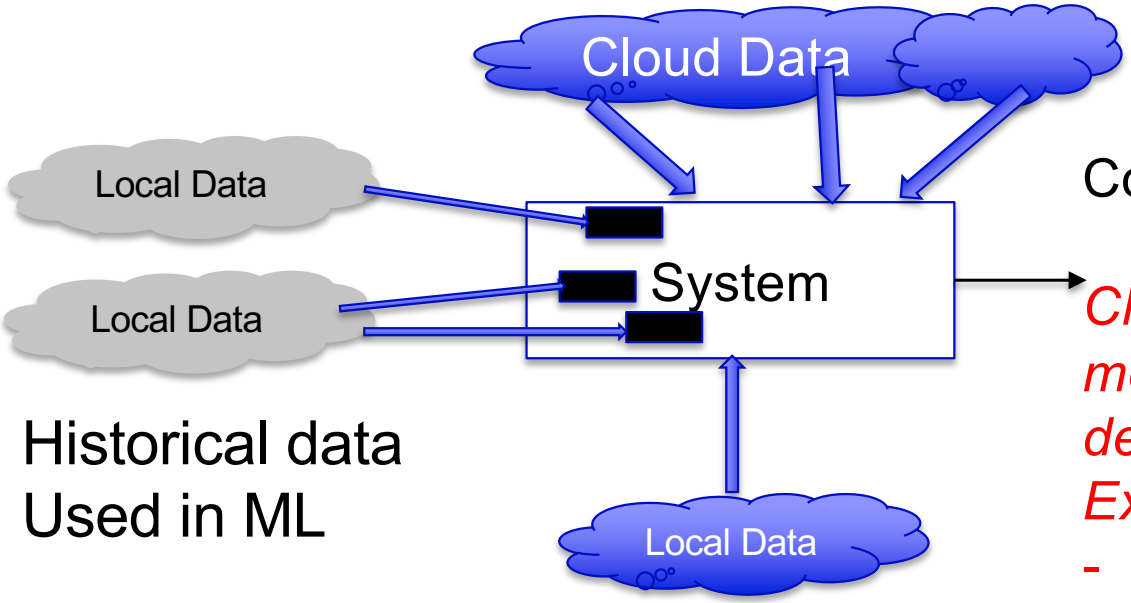
Dead Experimental Code paths

- AI solution requires a lot of experimentation
- A lot of code that will not be used later
- **Problems**
  - **Dead code**



- **Version management – how to preserve useful configuration branches, and remove unnecessary**

Managing Changes in the External World



Historical data
Used in ML

Continuous change of data

Challenge:
models and system behaviour dependent of data
Examples:
- *Threshold changes*
- *Correlation between data*

Requirements: Continuous monitoring of data and system. Continuous test.

Study : Overview of industrial ML systems

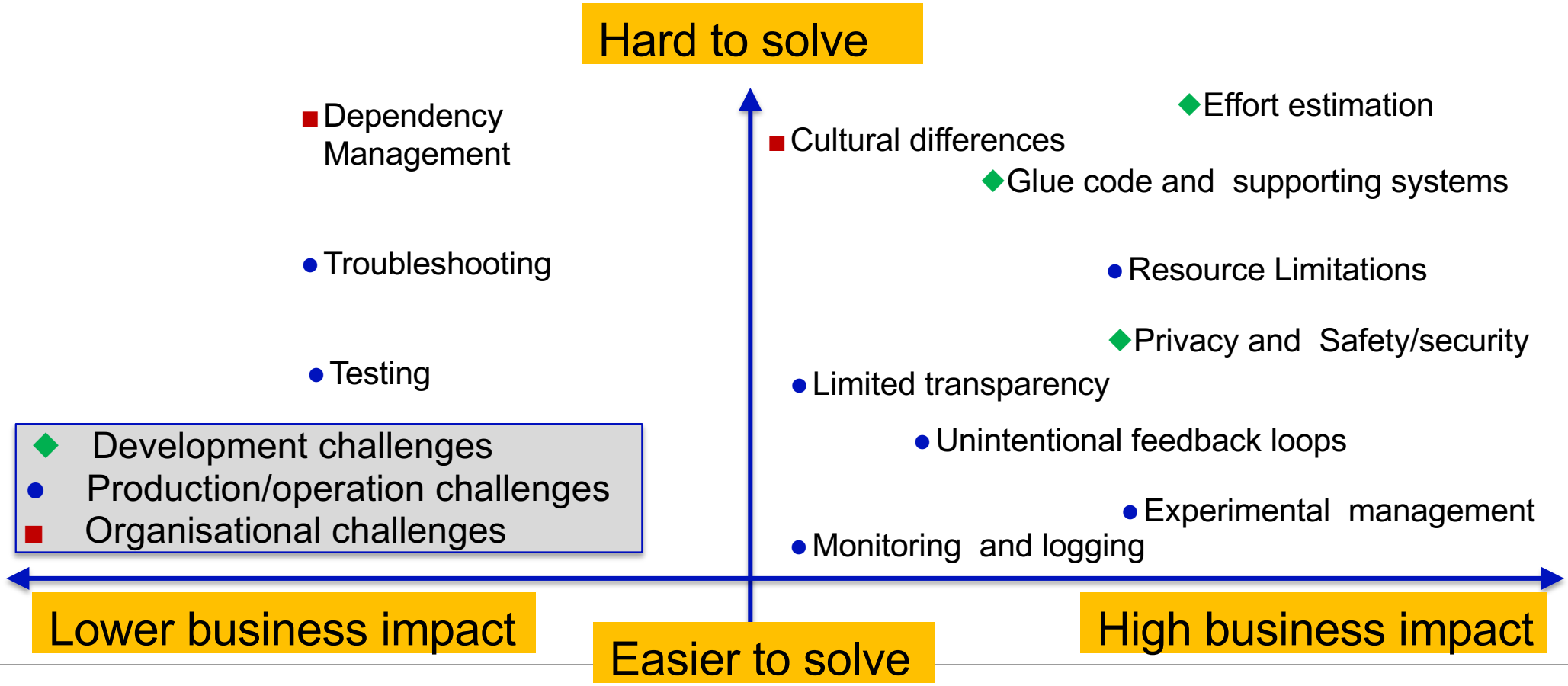
Which challenges are the most difficult in development of AI-based systems?

Case studies

A. Project EST: Real Estate Valuation
B. Project OIL: Predicting Oil and Gas Recovery Potential
C. Project RET: Predicting User Retention
D. Project WEA: Weather Forecasting
E. Project CCF: Credit Card Fraud Detection
F. Project BOT: Poker Bot Identification
G. Project REC: Media Recommendations

Software Engineering Challenges of Deep Learning, Euromicro SEAA 2018, *A. Arpteg, B. Brinne, I. Crnkovic-Friis, J. Bosch*

Study: Engineering challenges of DL



The challenges in evolution of development and use ML components

	Experiment prototyping	Non-critical deployment	Critical deployment	Cascading deployment
Assemble dataset	Issues with problem formulation and specifying desired outcome	Data silos, scarcity of labelled data, imbalanced training set	Limitations in techniques for gathering training data from large-scale, non-stationary data streams	Complex and effects of data dependencies
Create model	Use of non-representative dataset, data drifts	No critical analysis of training data	Difficulties in building highly scalable ML pipeline	Entanglements causing difficulties in isolating improvements
Train and evaluate model	Lack of well-established ground truth	No evaluation of models with business-centric measures	Difficulties in reproducing models, results and debugging DL models	Need of techniques for sliced analysis in final model
Deploy model	No deployment mechanism	Training- serving skew	Adhering to stringent serving requirements e.g., of latency, throughput	Hidden feedback-loops and undeclared consumers of the models

LE Lwakatare, A Raj, J Bosch, HH Olsson, I Crnkovic, A taxonomy of software engineering challenges for machine learning systems: An empirical investigation - International Conference on Agile Software …, 2019

cases

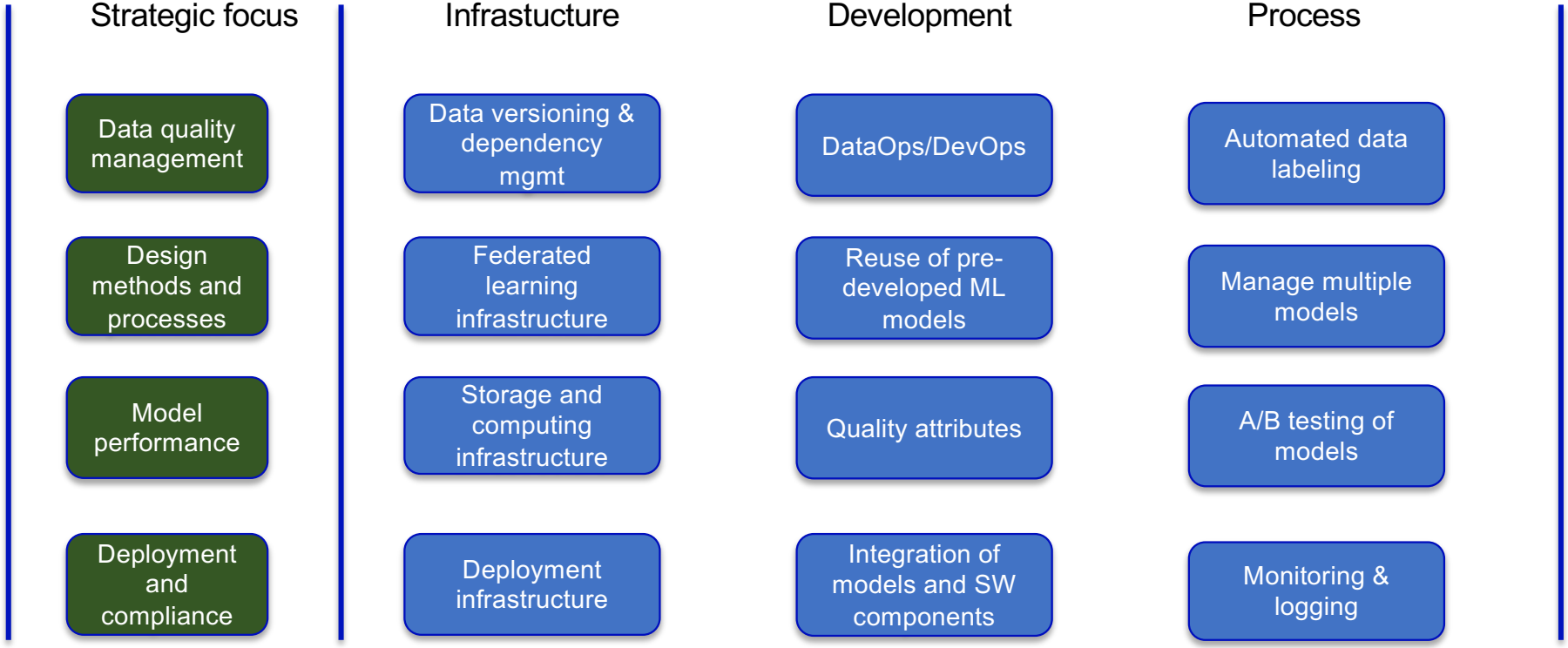
- Real Estate Valuation
- Predicting Oil and Gas Recovery
- Predicting User Retention
- Weather Forecasting
- Credit Card Fraud Detection
- Poker Bot Identification
- Media Recommendations
- Sensor data (automotive)
- Sentiment analysis
- Manufacturing optimization
- Training data annotation
- Failure prediction (telecom)
- OoO reply analysis
- Search engine optimization
- Wind power prediction
- Skin lesion classification

identified problems

		strategic focus
Lack of labelled data Lack of metadata Shortage of diverse samples Heterogeneity in data Data granularity Imbalanced data sets	Data drift Data dependencies Managing categorical data Managing sequences in data Deduplication complexity Data streams for training	data quality management
Experiment management Dependency management Unintended feedback loops Effort estimation Cultural differences Specifying desired outcome	Lack of modularity Sharing and tracking techn. Reproducibility of results Data extraction methods Tooling	design methods and processes
Overfitting Scalable ML pipeline Quality attributes Statistical Understanding	Limited transparency Training/serving skew Sliced analysis of final model	model performance
Monitoring and Logging Testing Troubleshooting Data sources and distribution Glue code and support	Privacy and data safety Data silos Data storage Resource limitations	deployment & compliance

J Bosch, H Holmström Olsson, I Crnkovic: Engineering AI Systems A Research Agenda (arXiv:2001.07522, 2020 preprint)

AI Engineering challenges

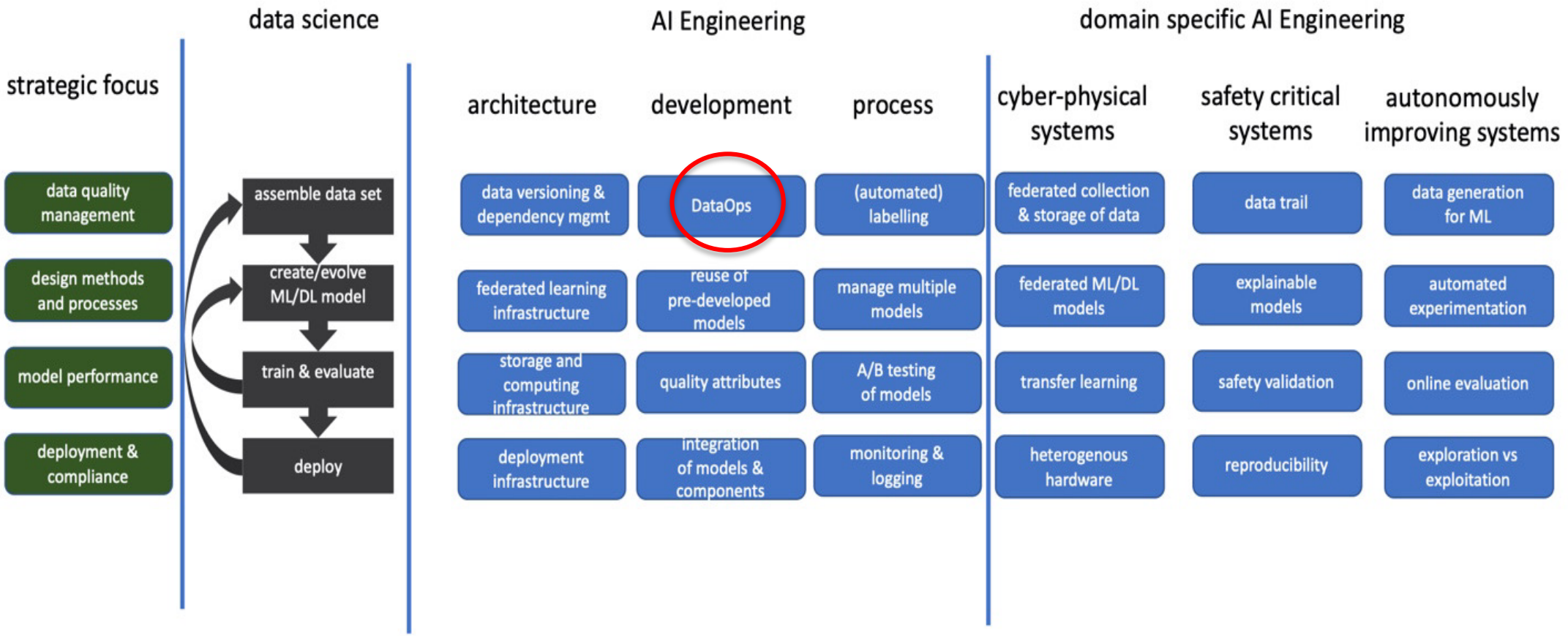


J Bosch, H Holmström Olsson, I Crnkovic: Engineering AI Systems A Research Agenda (arXiv:2001.07522, 2020 preprint)

Domain-specific AI Engineering challenges

Strategic focus	Cyber-physical systems	Safety-critical systems	Autonomously Improving systems
Data quality management	Federated collection & storage of data	Data trail	Data generation for ML
Design methods and processes	Federated ML models	Explainable models	Automated experimentation
Model performance	Transfer learning	Safety validation	Online evaluation
Deployment and compliance	Heterogenous HW platforms	reproducibility	Exploration vs. exploratiom

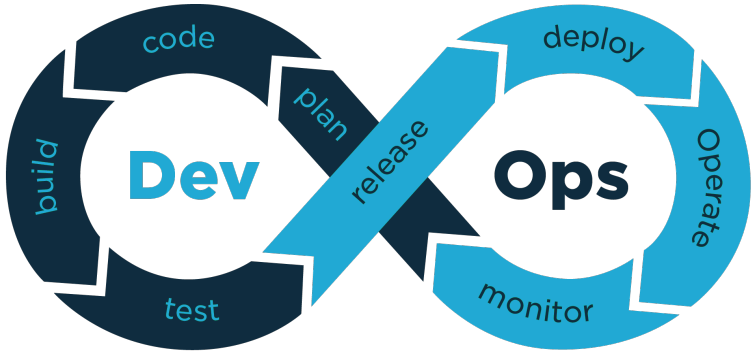
J Bosch, H Holmström Olsson, I Crnkovic: Engineering AI Systems A Research Agenda, arXiv:2001.07522, 2020 preprint)



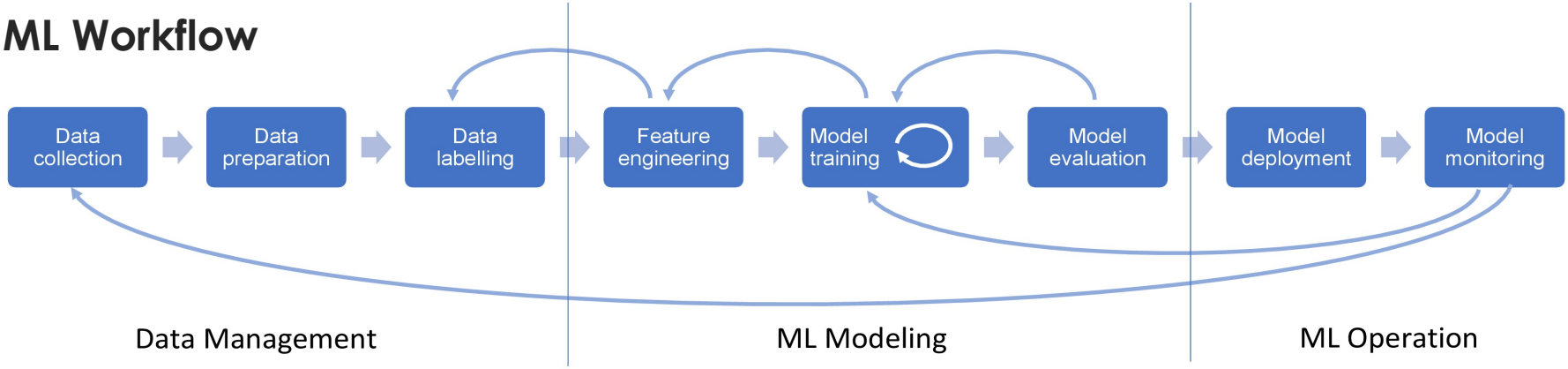
J Bosch, H Holmström Olsson, I Crnkovic: Engineering AI Systems A Research Agenda, arXiv:2001.07522, 2020 preprint)

Example – Process integration

DevOps

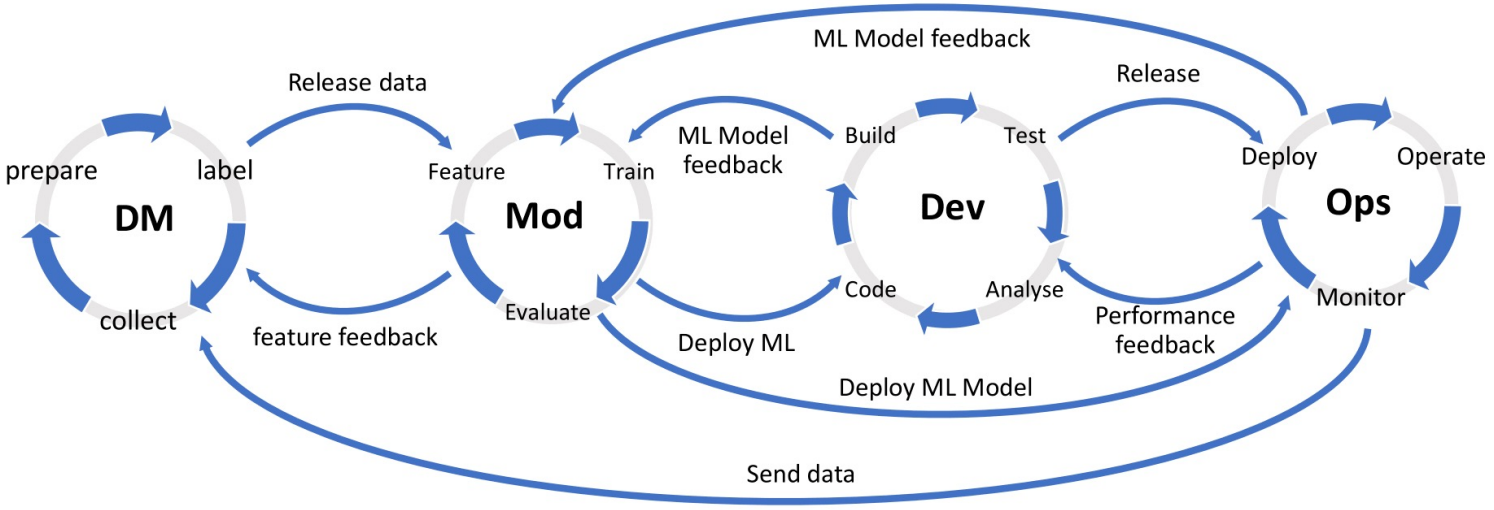


ML Workflow



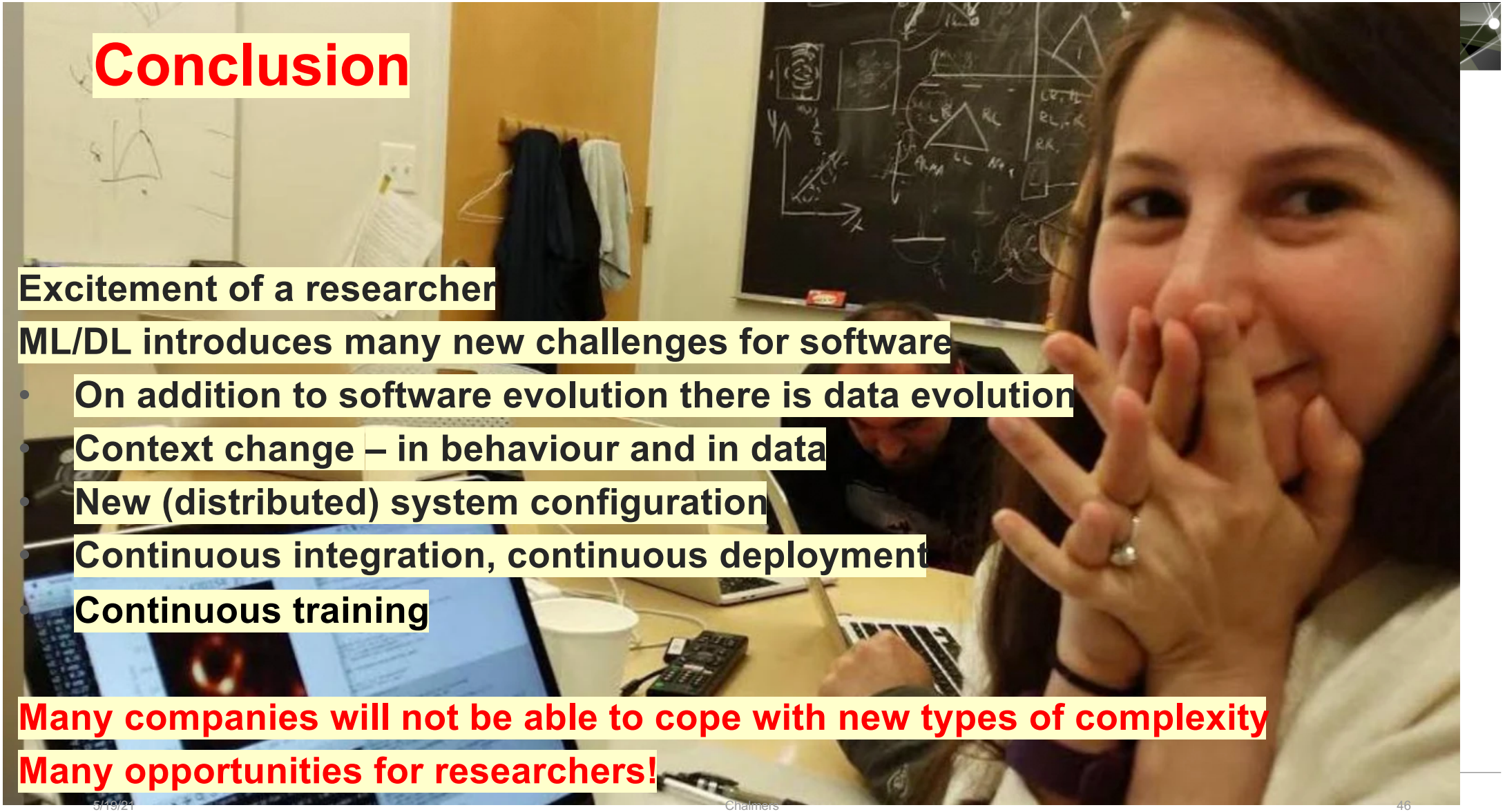
Lwakatare, Crnkovic & Bosch. DevOps for AI – Challenges in Development of AI-enabled applications, SoftCom 2020

Conceptual ML Workflow and DevOps Process Integration



DM – Data Management
Mod – ML Modeling
Dev – Development
Ops – System Operation

Conclusion



Excitement of a researcher

ML/DL introduces many new challenges for software

- On addition to software evolution there is data evolution
- Context change – in behaviour and in data
- New (distributed) system configuration
- Continuous integration, continuous deployment
- Continuous training

Many companies will not be able to cope with new types of complexity

Many opportunities for researchers!

CHALMERS

UNIVERSITY OF TECHNOLOGY