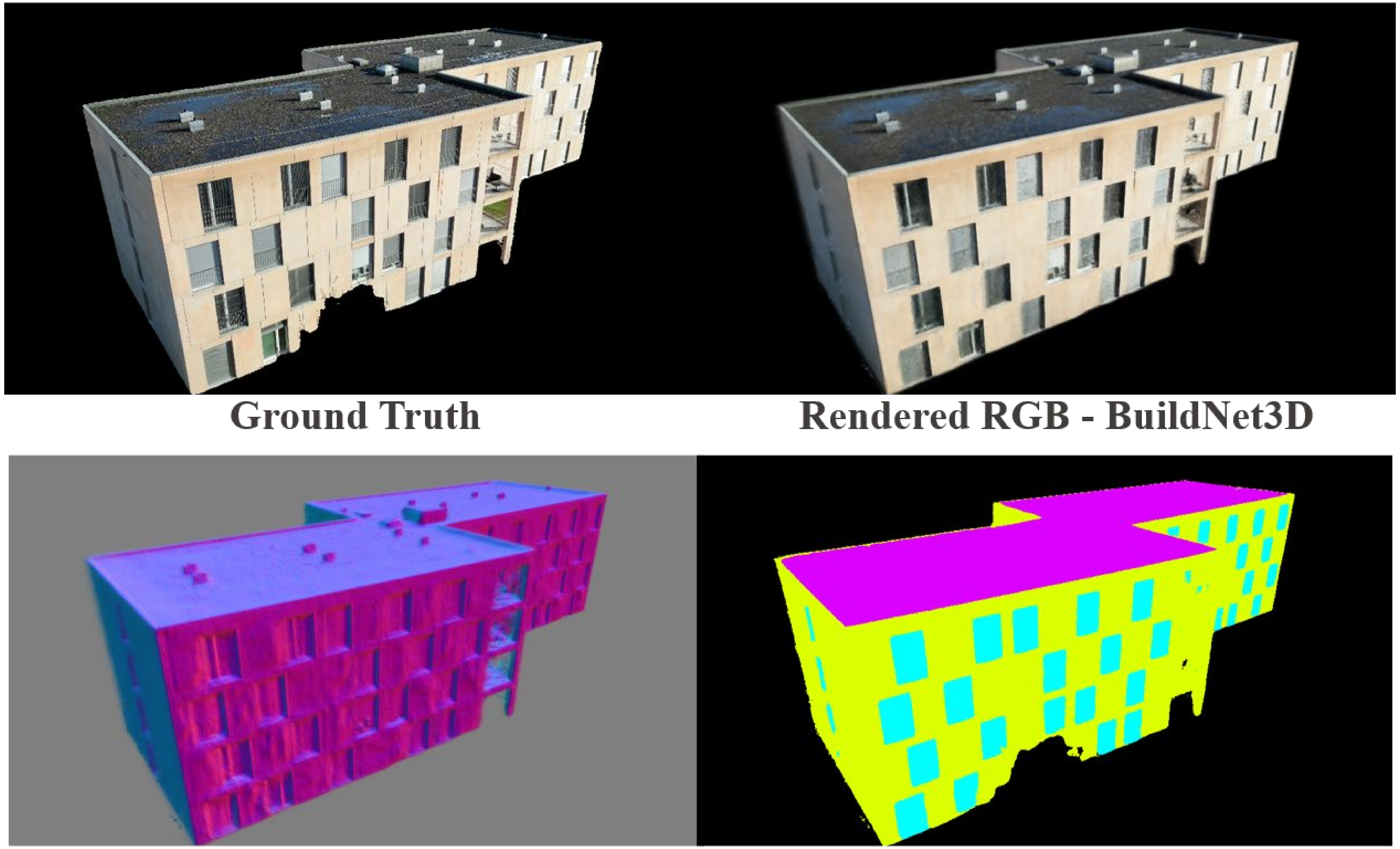
EPFL

Olga Fink / Alexandre Alahi

Introduction to Machine Learning

Estimating Building Envelope Characteristics

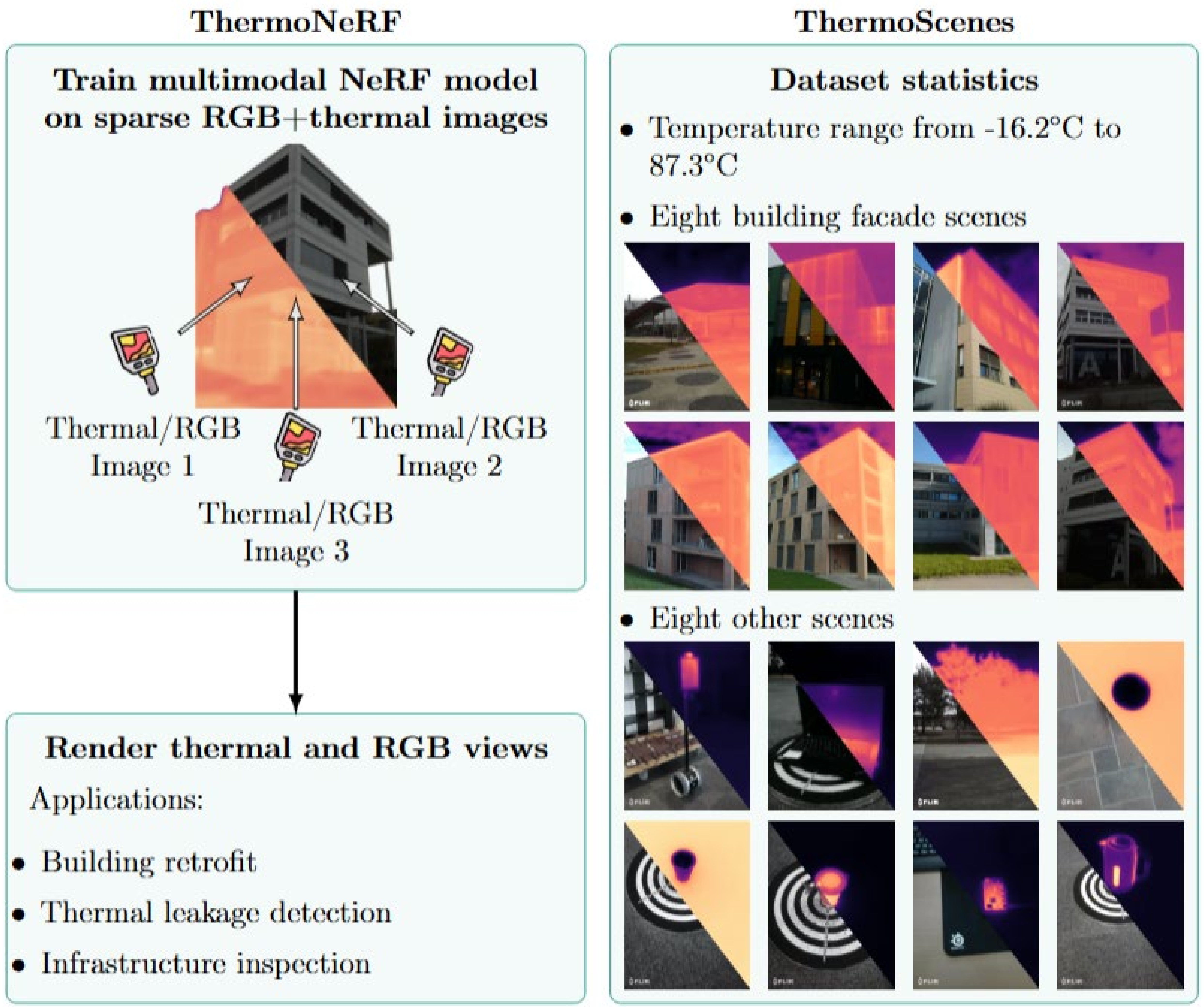


Ground Truth | Rendered RGB - BuildNet3D

Rendered Normal - BuildNet3D | Rendered Semantics- BuildNet3D

Xu, C., Mielle, M., Laborde, A., Waseem, A., Forest, F. and Fink, O., 2025. Exploiting Semantic Scene Reconstruction for Estimating Building Envelope Characteristics. *Building and Environment*

Thermal 3D reconstructions from sparse images

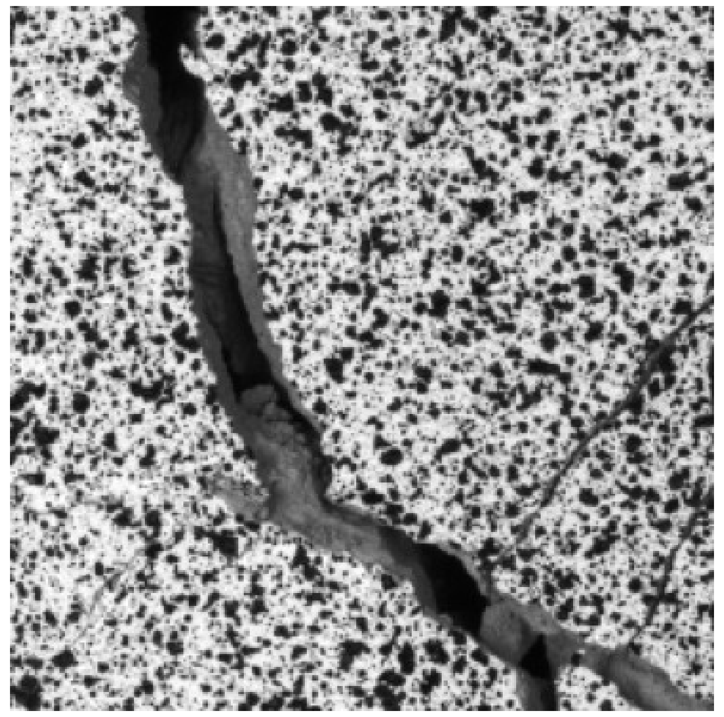


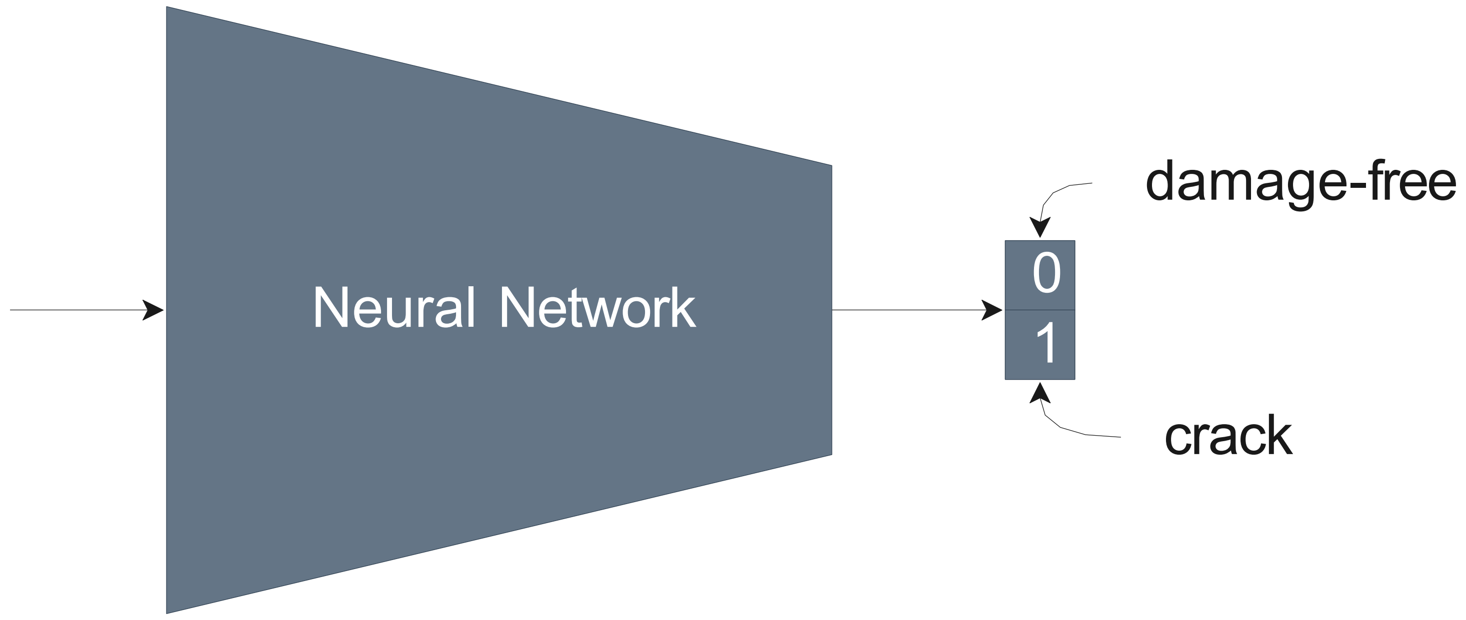
Hassan, M., Forest, F., Fink, O. and Mielle, M., 2024. ThermoNeRF: Multimodal Neural Radiance Fields for Thermal Novel View Synthesis. *arXiv preprint arXiv:2403.12154*.

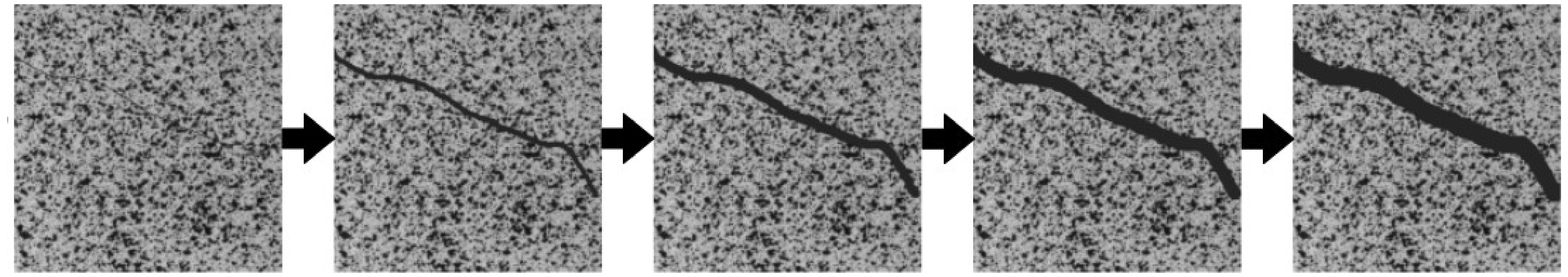
Machine learning for image-based crack detection

Data-driven approaches based on supervised deep learning have demonstrated excellent performance in detection of cracks in images, but they require **large annotated datasets** for training.

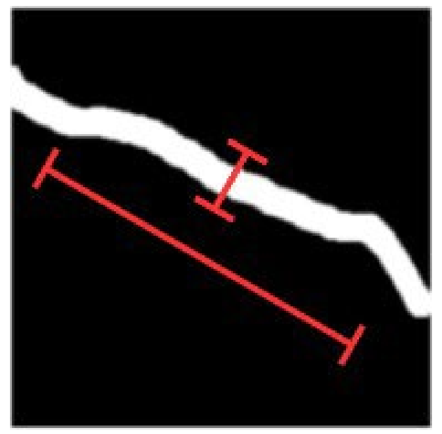
Classificationtask:







Severity quantification and monitoring is crucial for timely decision-making.

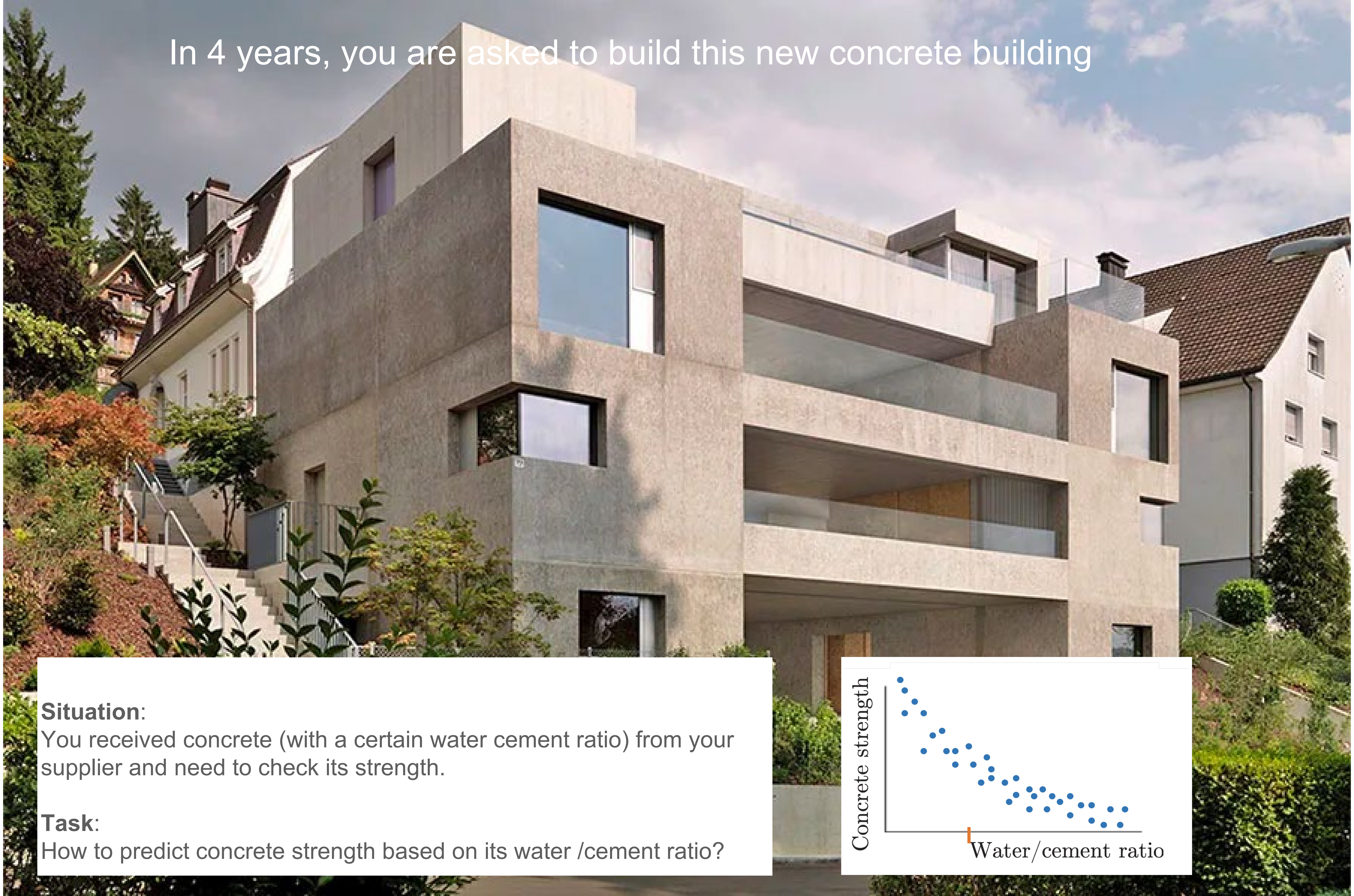
What we need:  → Severity metrics

- Does not allow severity quantification
- Fast and easy image-level annotation (1 bit)

Forest, F., Porta, H., Tuia, D., & Fink, O. (2024). From classification to segmentation with explainable AI: A study on crack detection and growth monitoring. Automation in Construction, 165, 105497.

Olga Fink / Alexandre Alahi

In 4 years, you are asked to build this new concrete building



Situation:
You received concrete (with a certain water cement ratio) from your supplier and need to check its strength.

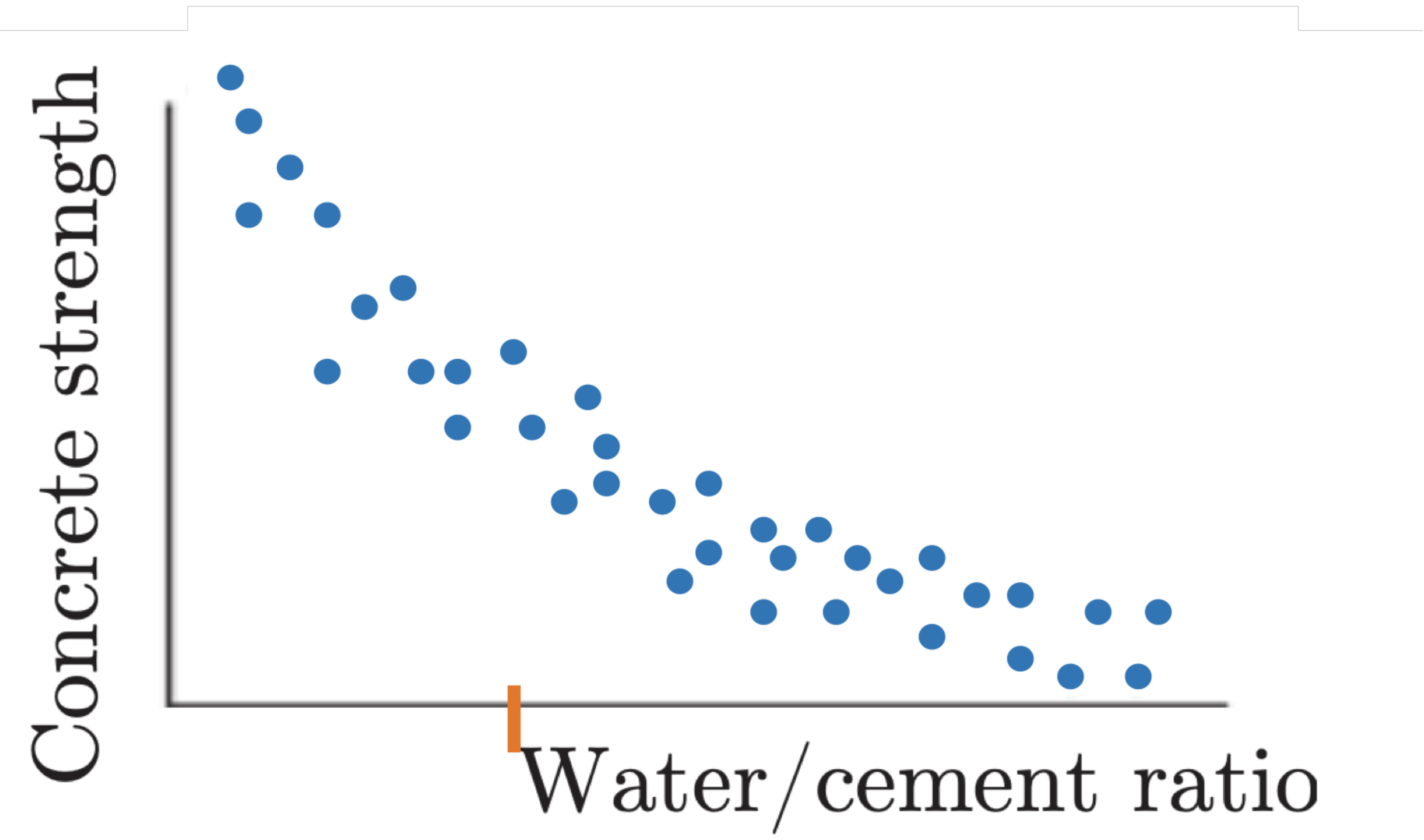
Task:
How to predict concrete strength based on its water /cement ratio?

Mathematically,

Let:

X = Water/cement ratio
Y = Concrete strength

Can I find a function **f** of parameters **W** such that:
Y = f(X,**W**)
Attempt1 : f(X,**W**) = $X*W_1 + W_2$



Attempt2 : f(X,**W**) = $X^2*W_1 + X*W_2 + W_3$

W is a vector commonly referred to as unknown parameters

The goal of this class is to learn how to find a good function **f** and its parameters **W**!

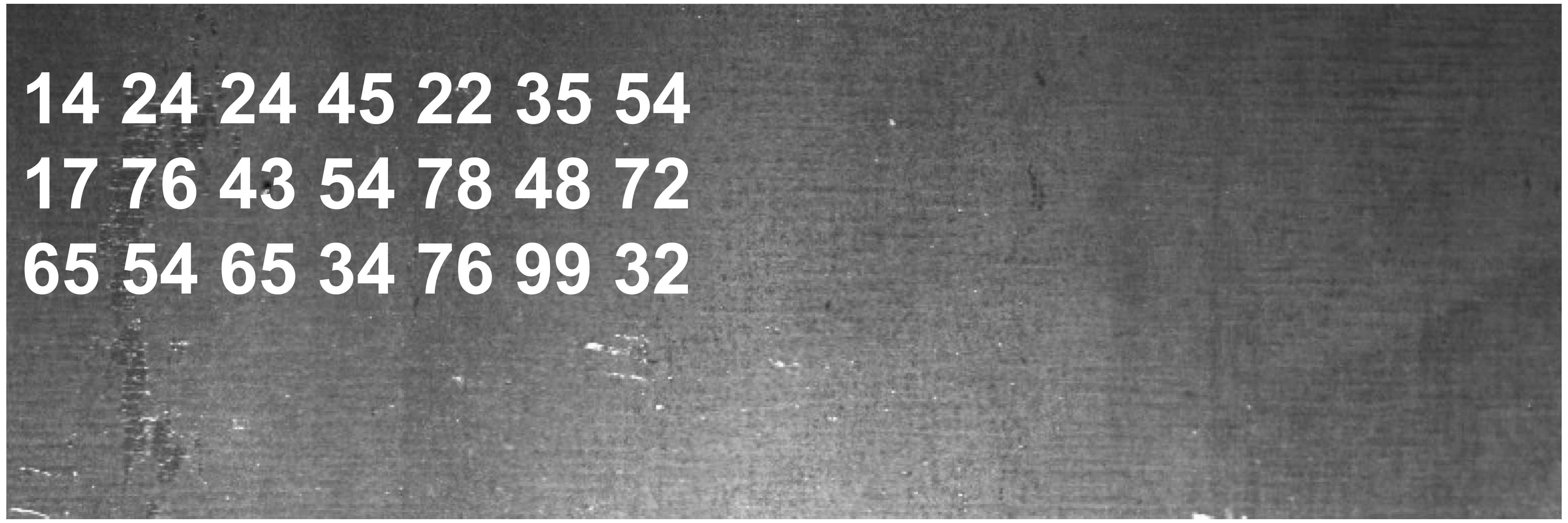
You got promoted to be the director of a Steel Factory!

Situation:
Your goal is to make the production of steel sheet more efficient by identify defects.

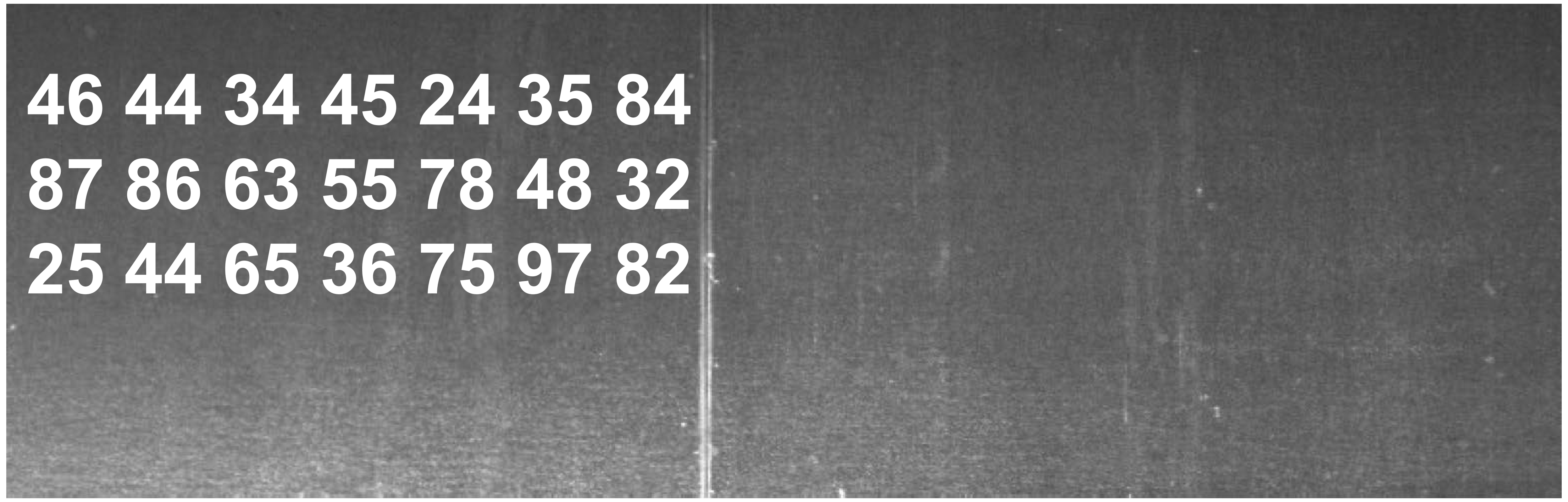
Task:
How to detect and classify defects from images of steel sheets?

Mathematically:
Let,
X = Image (matrix of numbers)
Y = No Defect (0), or Defect (1)

Can I find a function f of parameters **W** such that:
Y = f(**X**,**W**)



No Defect



Defect


FAKE

Welcome in civil-226

Outline

- Administrative
- AI revolution
- Data revolution
- ML revolution

Artificial Intelligence
AI

Machine Learning
ML

Olga Fink / Alexandre Alahi

Administrative

Administrative
Teacher Staff

Lecturers

Olga Fink

Alexandre Alahi

TAs

Yasaman Haghighi

Leandro Von Krannichfeldt

Bastien Van Delft

+

Amelie Menoud

Pedro Rodrigues Da Costa Savini

Administrative
Grading

This course will be graded as follows:

- **Exercises**: Not graded
  - Every week - To practice the techniques learned in class.
- **Graded Homework (individual)**: 10%
  - Similar to a weekly exercise, except it will be graded.
- **Projects (in groups)**: 40%
  - Coding project to solve a problem using ML.
- **Final Exam**: 50%
  - Standard final exam.

Administrative
Requirements

Domains covered by ML

- Linear algebra
- Calculus
- Probability & statistics
- Programming (in Python)

Course Requirements:

- MATH-111
- CS-119(h)

Lectures

- Start with a section specific real-world example

- Real-time demo (from any domain)
  - e.g., https://playground.tensorflow.org

- Have references to on-line resources/animations
  - e.g., https://www.deeplearning.ai/ai-notes/optimization/

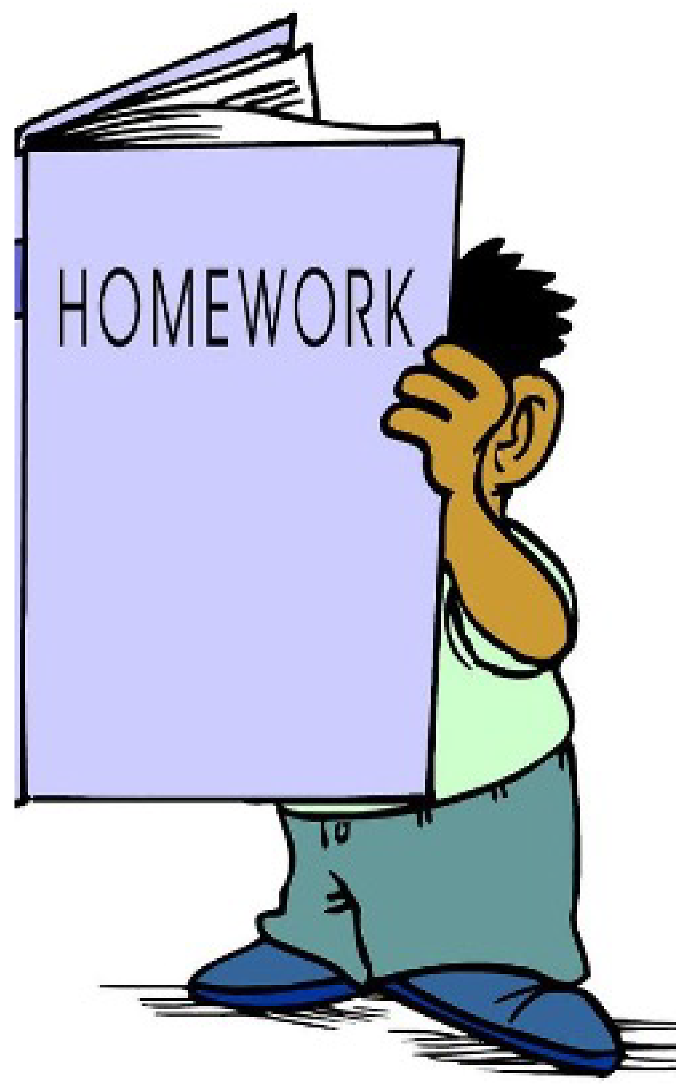
- Guest speaker from industry

Exercices

Olga Fink / Alexandre Alahi

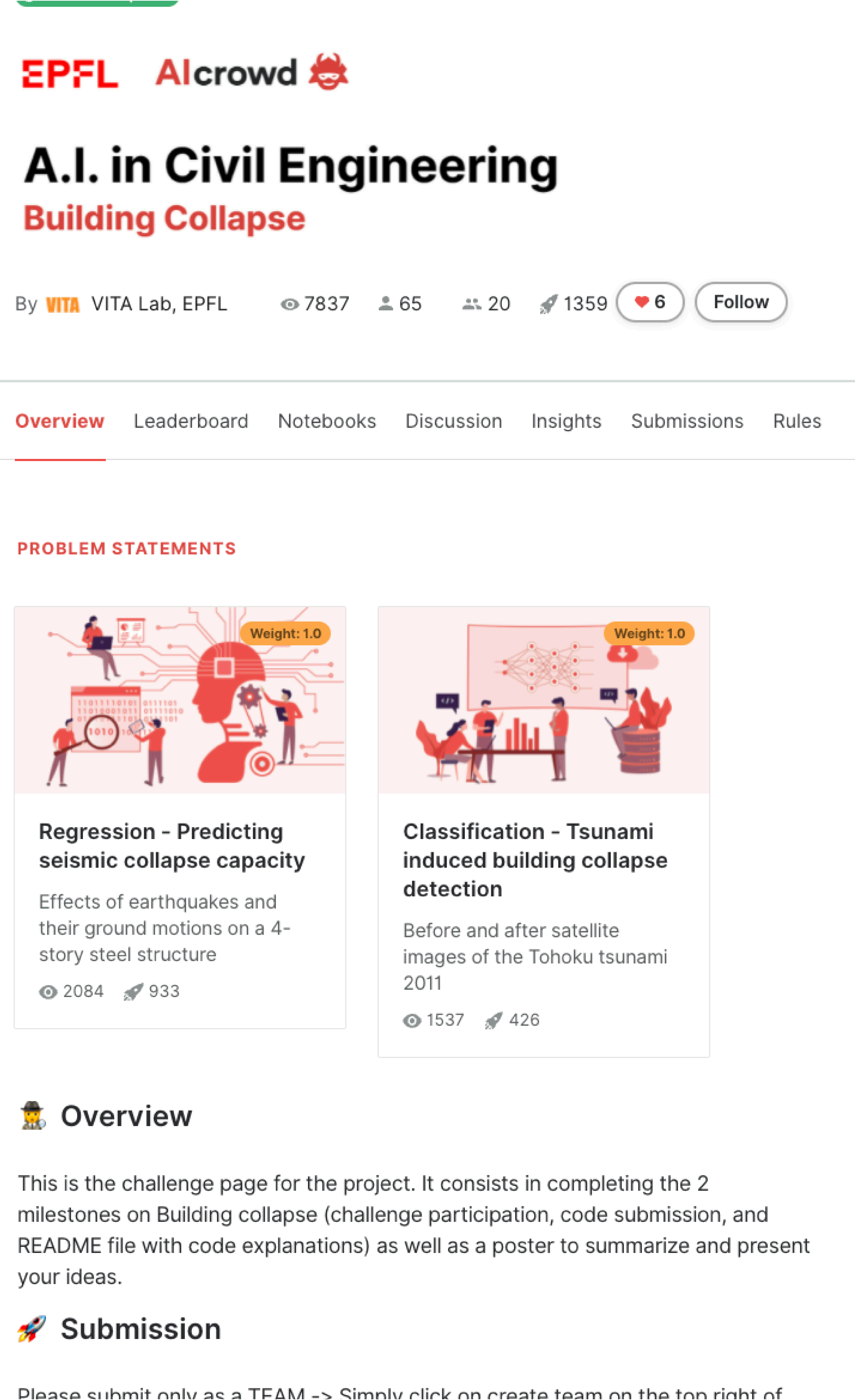
- **Notebooks**
  - Visual, direct feedback
  - EPFL NOTO
  - 3rd party tools (colab)
- Generic

https://github.com/vita-epfl/IntroML-2025



Project

- AIcrowd
  - Leaderboard
  - https://www.aicrowd.com/challenges/a-i-in-civil-engineering-building-collapse
- Poster session
- Awards:
  - 1st & 2nd in the leaderboard
  - Best poster



Optional

Administrative

Additional resources

Books

- K.P. Murphy, Probabilistic Machine Learning: An Introduction, 2021
- M.P. Deisenroth, A.A. Faisal, C.S. Ong, Mathematics for Machine Learning, 2020
- J. VanderPlas, Python Data Science Handbook, 2016

Online courses

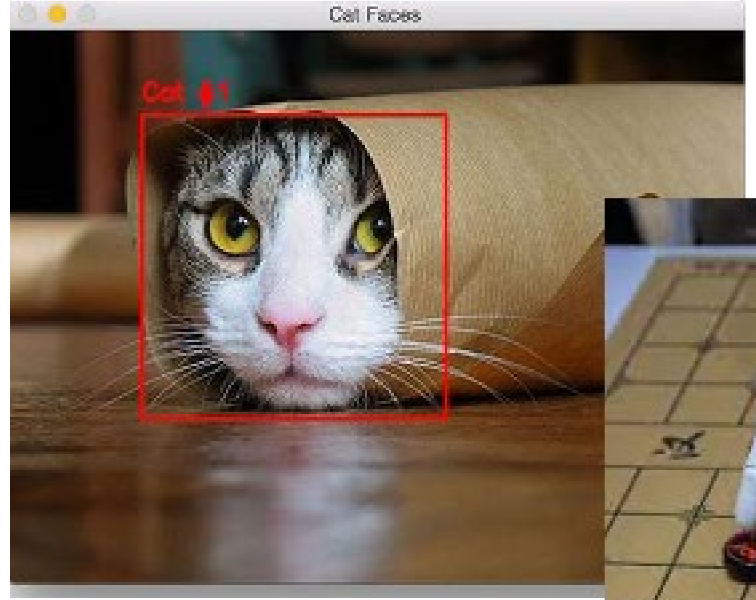
- Machine Learning Course by Andrew Ng (Stanford)
- Deep Learning Specialization by Andrew Ng (deep learning.ai)

Other

- Google Machine Learning Glossary

AI revolution

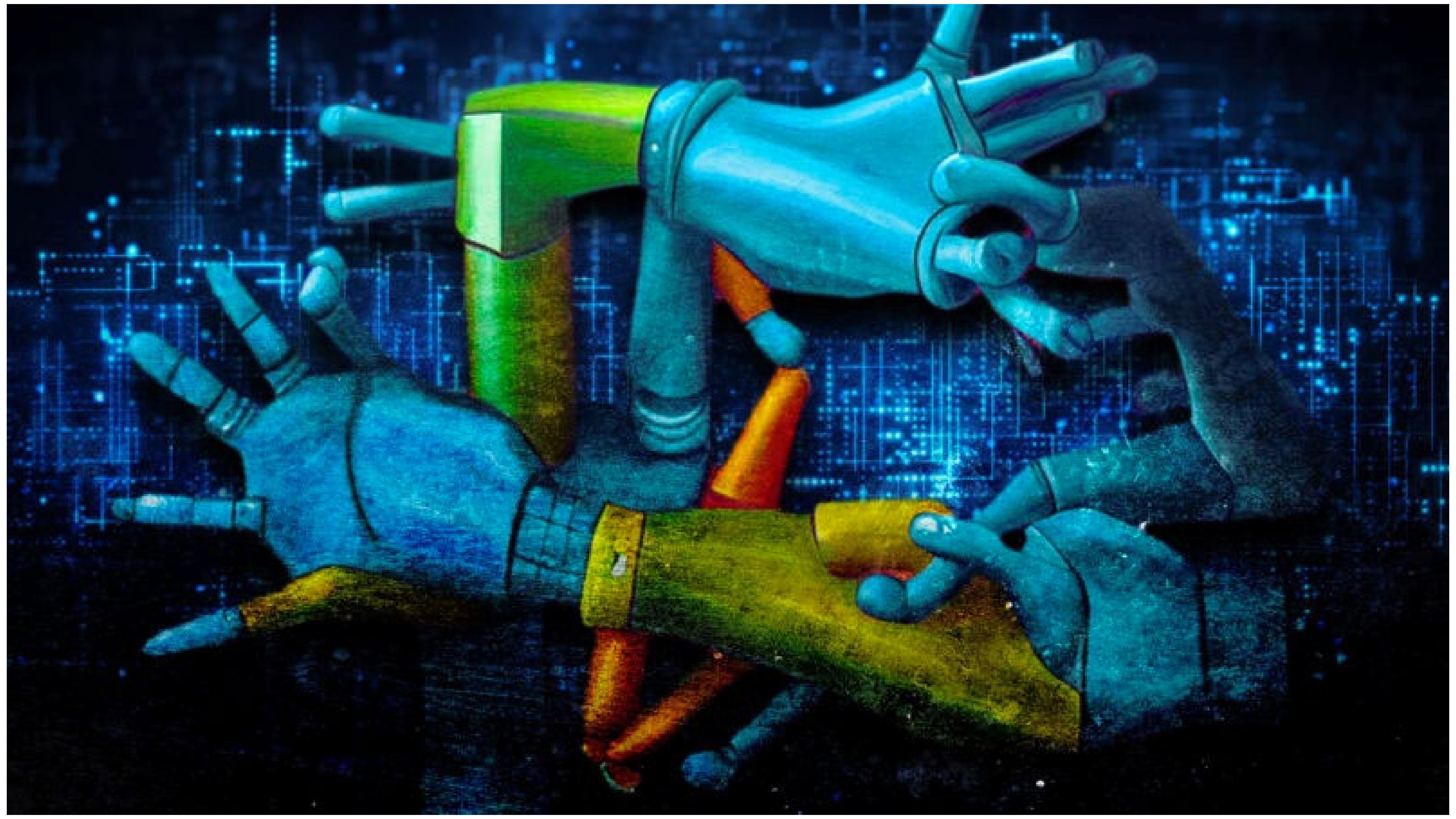
AI revolution



"A teddy bear on a skateboard in Times Square"

Generative AI

“a pair of robot hands holding pencils drawing a pair of human hands, oil painting, colorful,”



“a shark that looks like a car that looks like a face”

AI Takes the Nobel Prize

What is AI?

What is AI?

AI is intelligence demonstrated by machines…

What is AI?

Olga Fink / Alexandre Alahi

What is Intelligence? Ability to accomplish complex goals

Credits: reach.ch

What is AI?

"any device/system/software that **perceives** its environment and **takes actions** that maximise success for its goal"

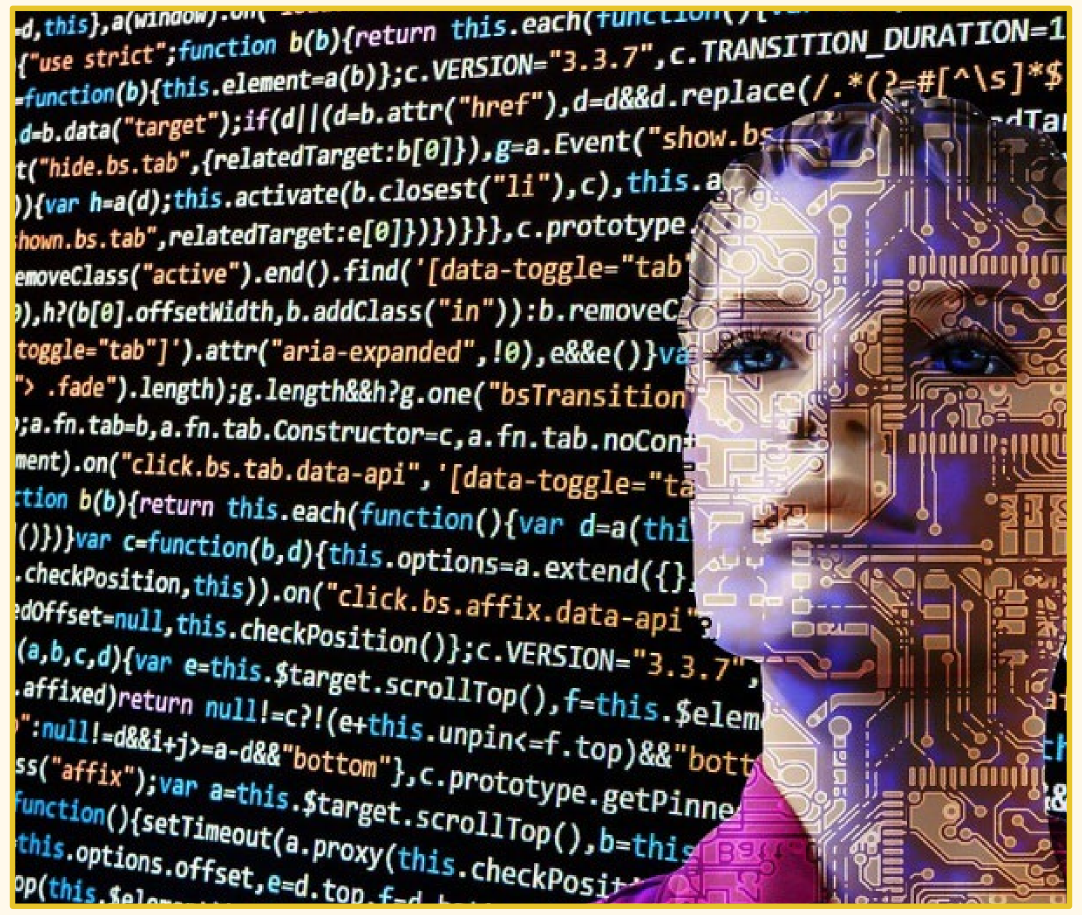
EPFL

What is artificial intelligence?

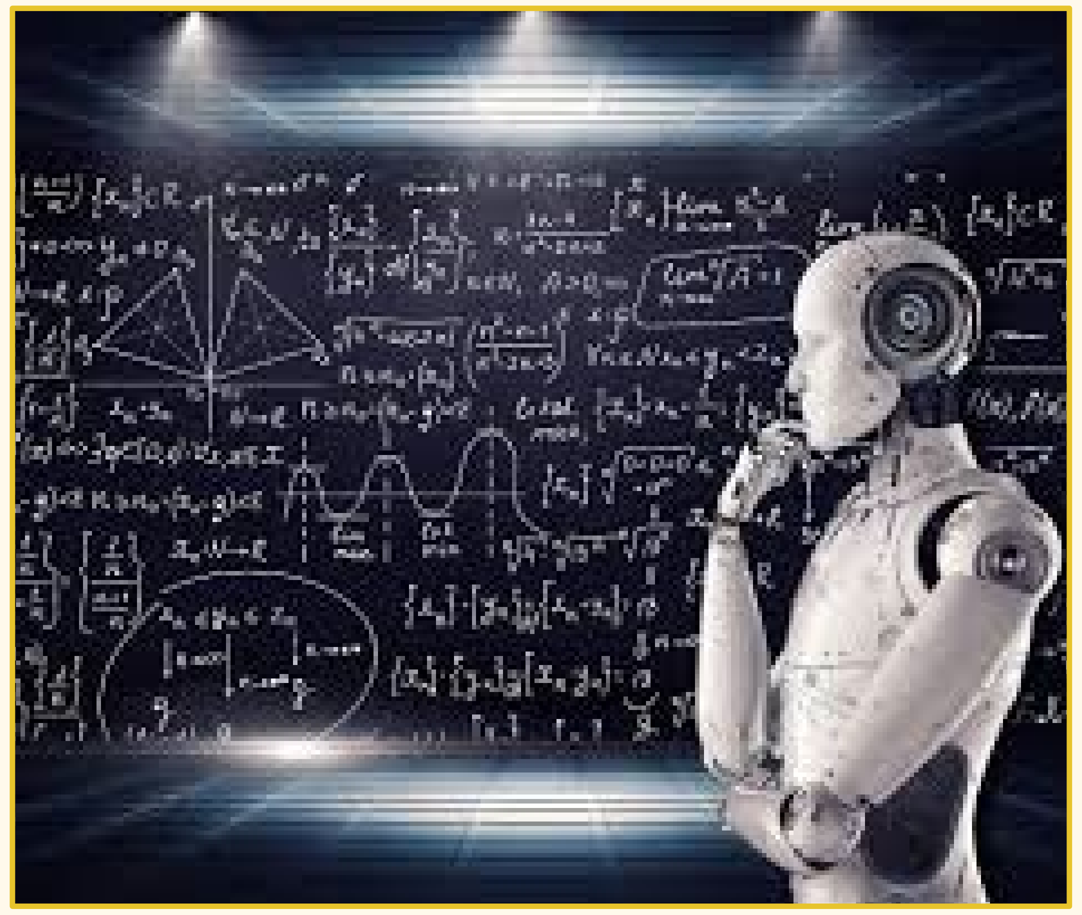
Artificial intelligence (AI) refers to the ability of a computer or machine to perform tasks that would normally require human intelligence, such as learning, problem solving, decision making, and language understanding.

In which areas does AI excel today?

Artificial Narrow Intelligence



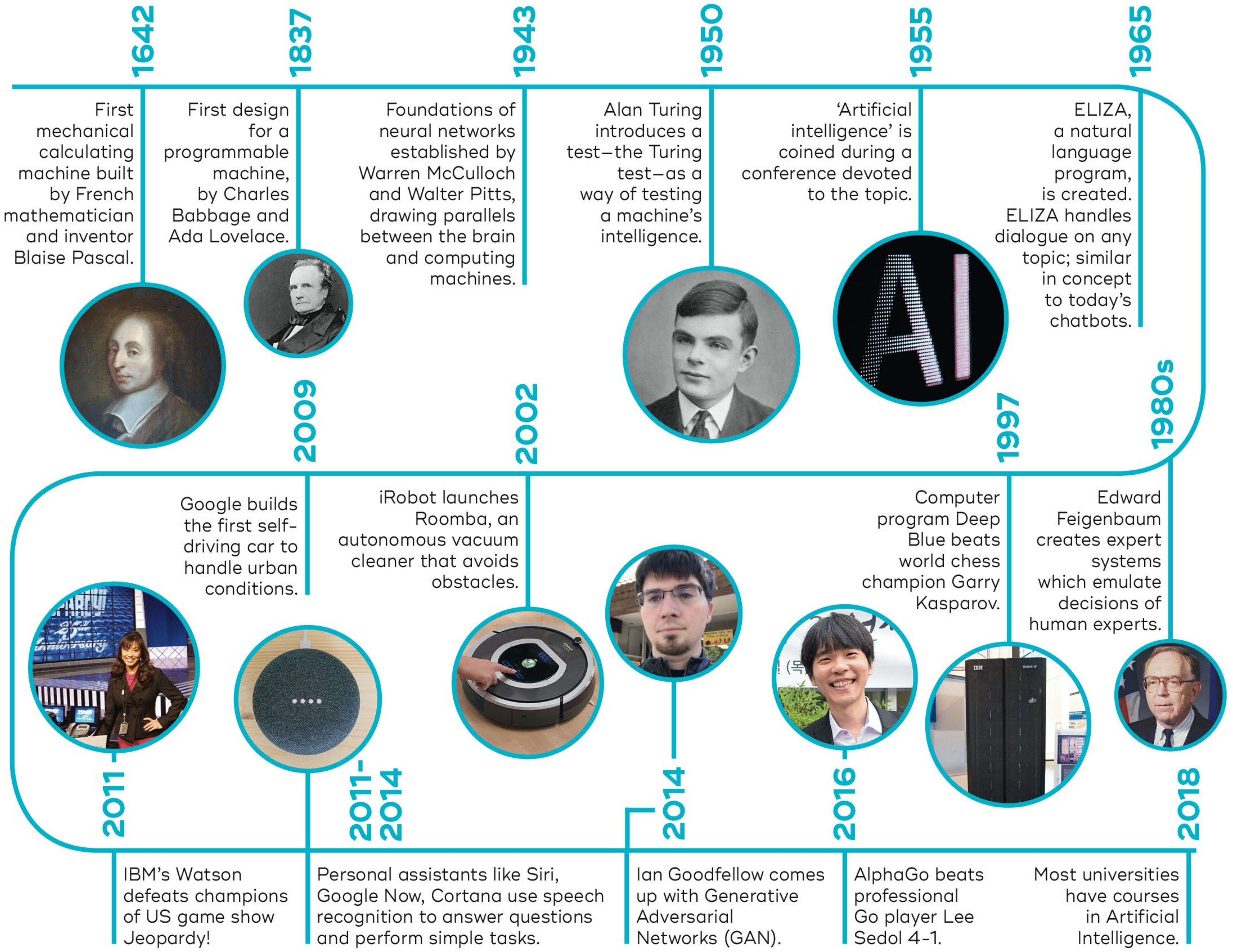
Artificial General Intelligence



Artificial Superintelligence

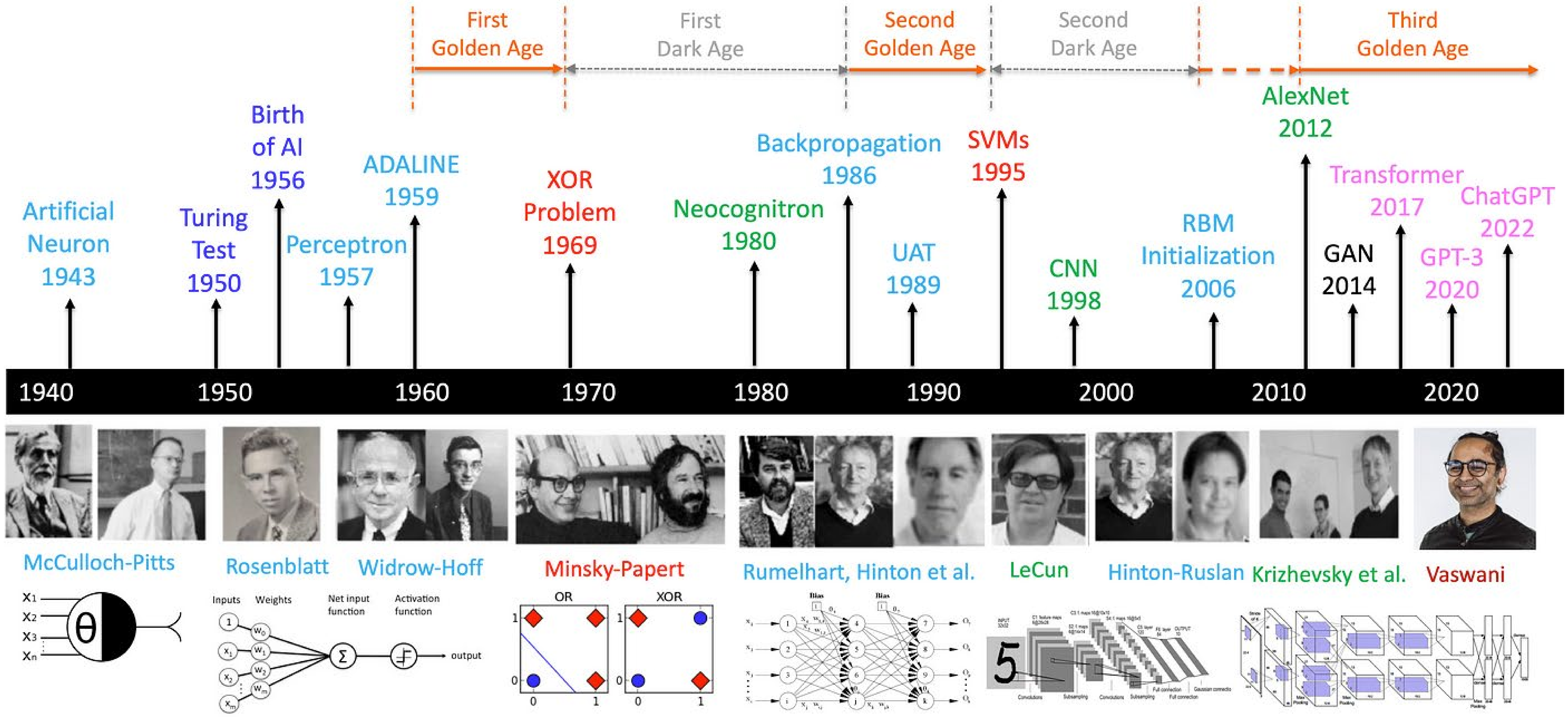
History of AI

History of AI



Olga Fink / Alexandre Alahi

History of AI: a bit different perspective



Olga Fink / Alexandre Alahi

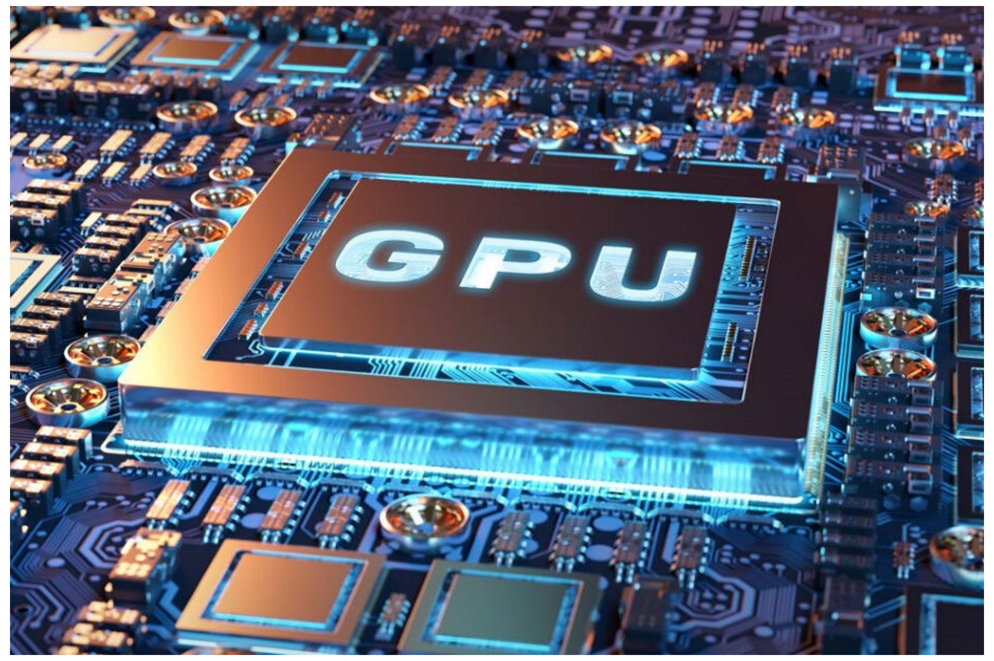
History of AI
Interplay between Data and algorithms

Most AI breakthroughs used old algorithms but recent data sets.

Year	Breakthroughs in AI	Datasets (First Available)	Algorithms (First Proposed)
1994	Human-level spontaneous speech recognition	Spoken Wall Street Journal articles and other texts (1991)	Hidden Markov Model (1984)
1997	IBM Deep Blue defeated Garry Kasparov	700,000 Grandmaster chess games, aka "The Extended Book" (1991)	Negascout planning algorithm (1983)
2005	Google's Arabic- and Chinese-to-English translation	1.8 trillion tokens from Google Web and News pages (collected in 2005)	Statistical machine translation algorithm (1988)
2011	IBM Watson became the world Jeopardy! champion	8.6 million documents from Wikipedia, Wiktionary, Wikiquote, and Project Gutenberg (updated in 2010)	Mixture-of-Experts algorithm (1991)
2014	Google's GoogLeNet object classification at near-human performance	ImageNet corpus of 1.5 million labeled images and 1,000 object categories (2010)	Convolution neural network algorithm (1989)
2015	Google's Deepmind achieved human parity in playing 29 Atari games by learning general control from video	Arcade Learning Environment dataset of over 50 Atari games (2013)	Q-learning algorithm (1992)
Average No. of Years to Breakthrough:		3 years	18 years

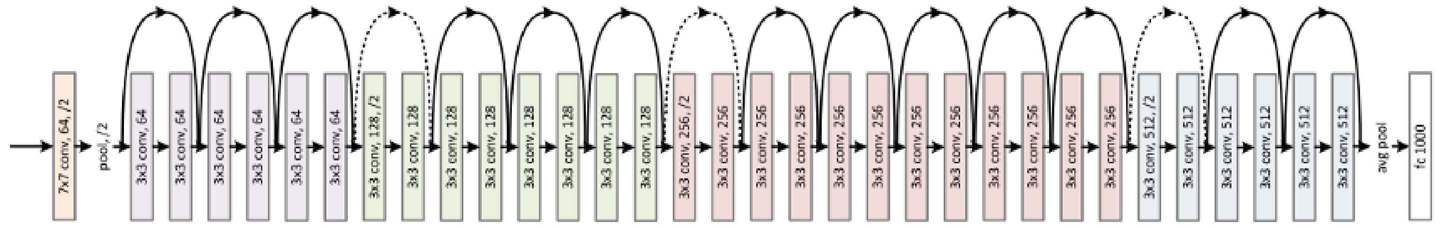
Why is everyone talking about AI now?

Some pillars of AI aren't recent. So why is everyone excited about it now?



IMAGENET

PyTorch



Hardware

Progress in hardware such as Graphical Processing Unit (GPU)

Data

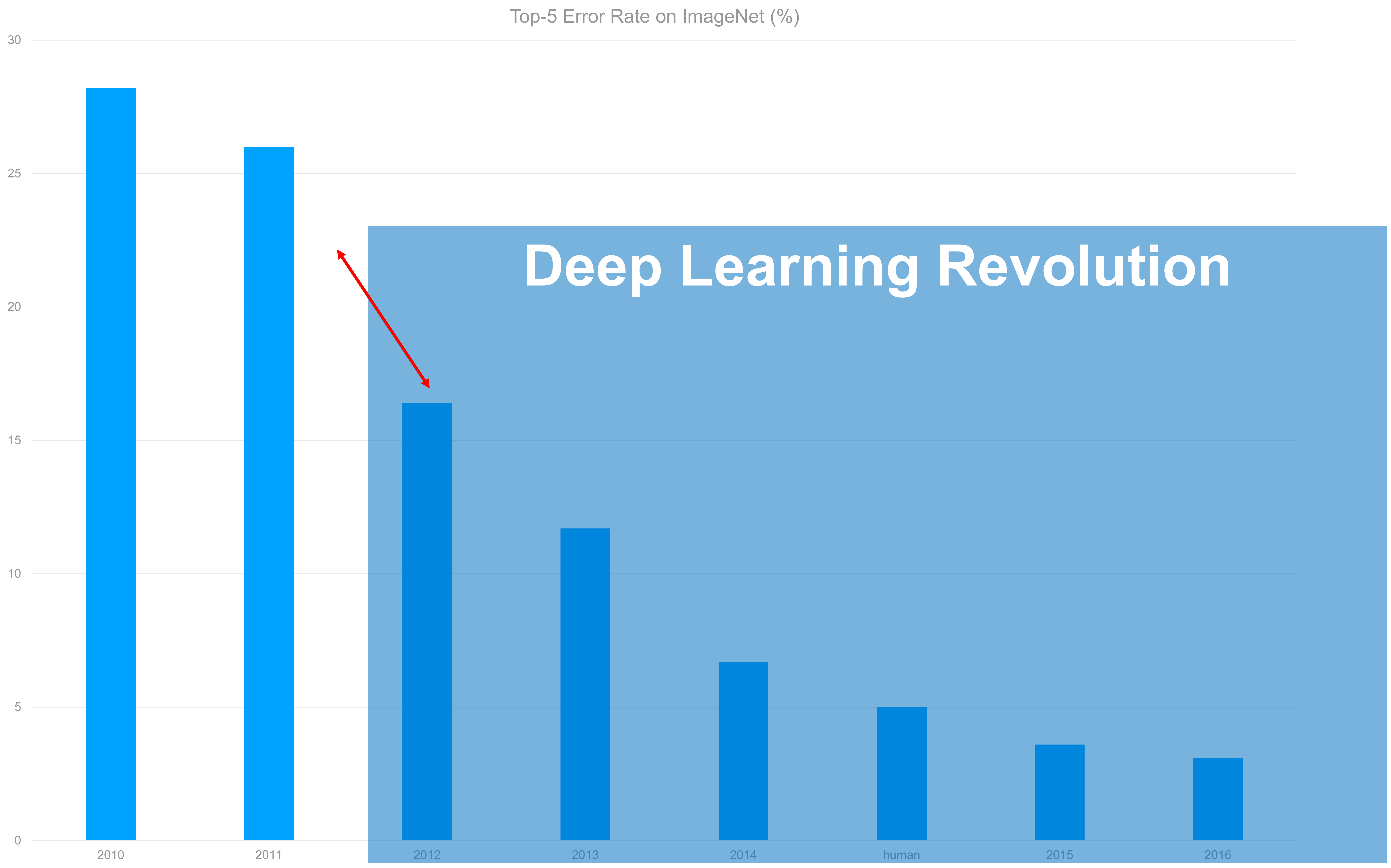
Massive amount of labeled data

Open Science/Software

Shared models/frameworks

Machine Learning is the tool that leverages all these!

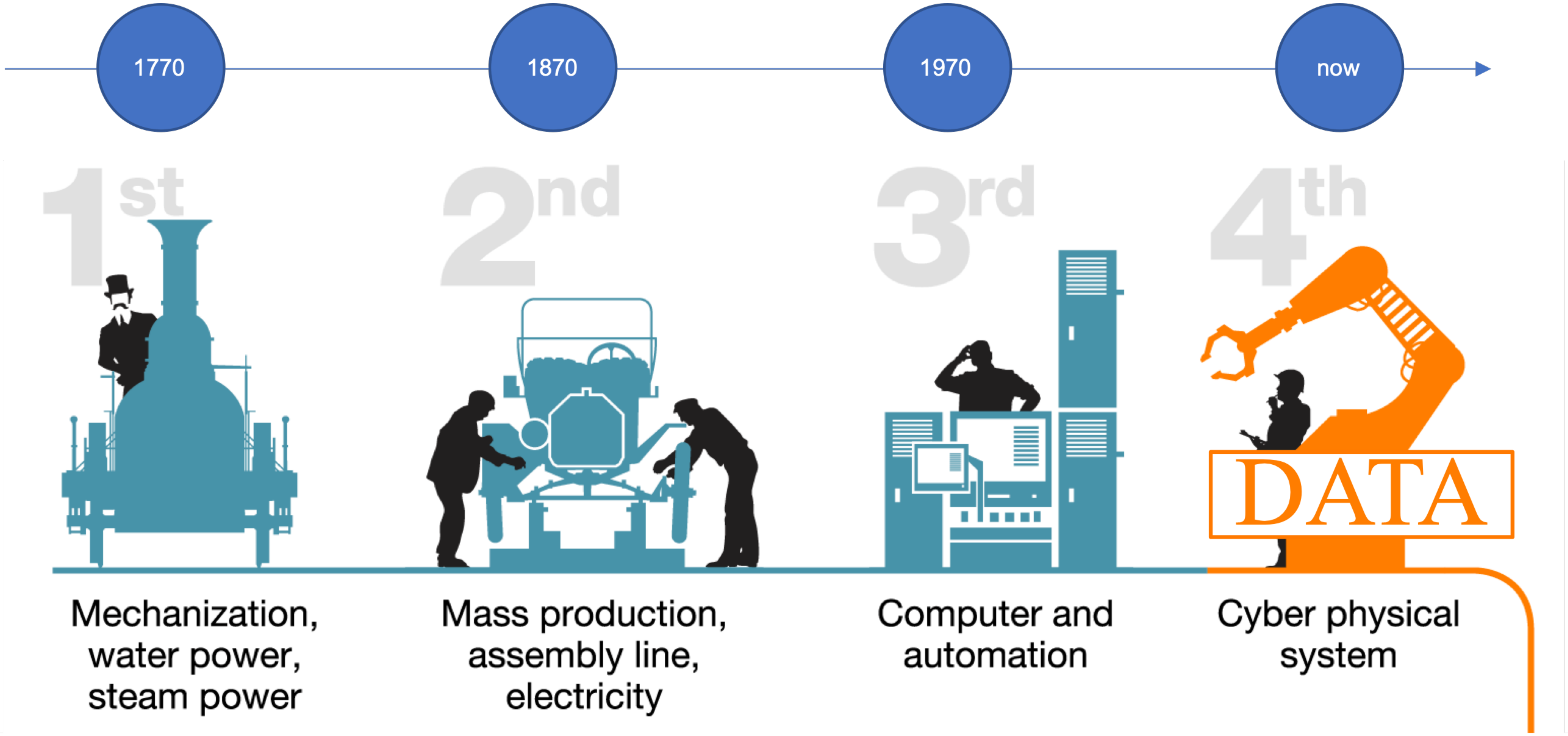
Power of depth



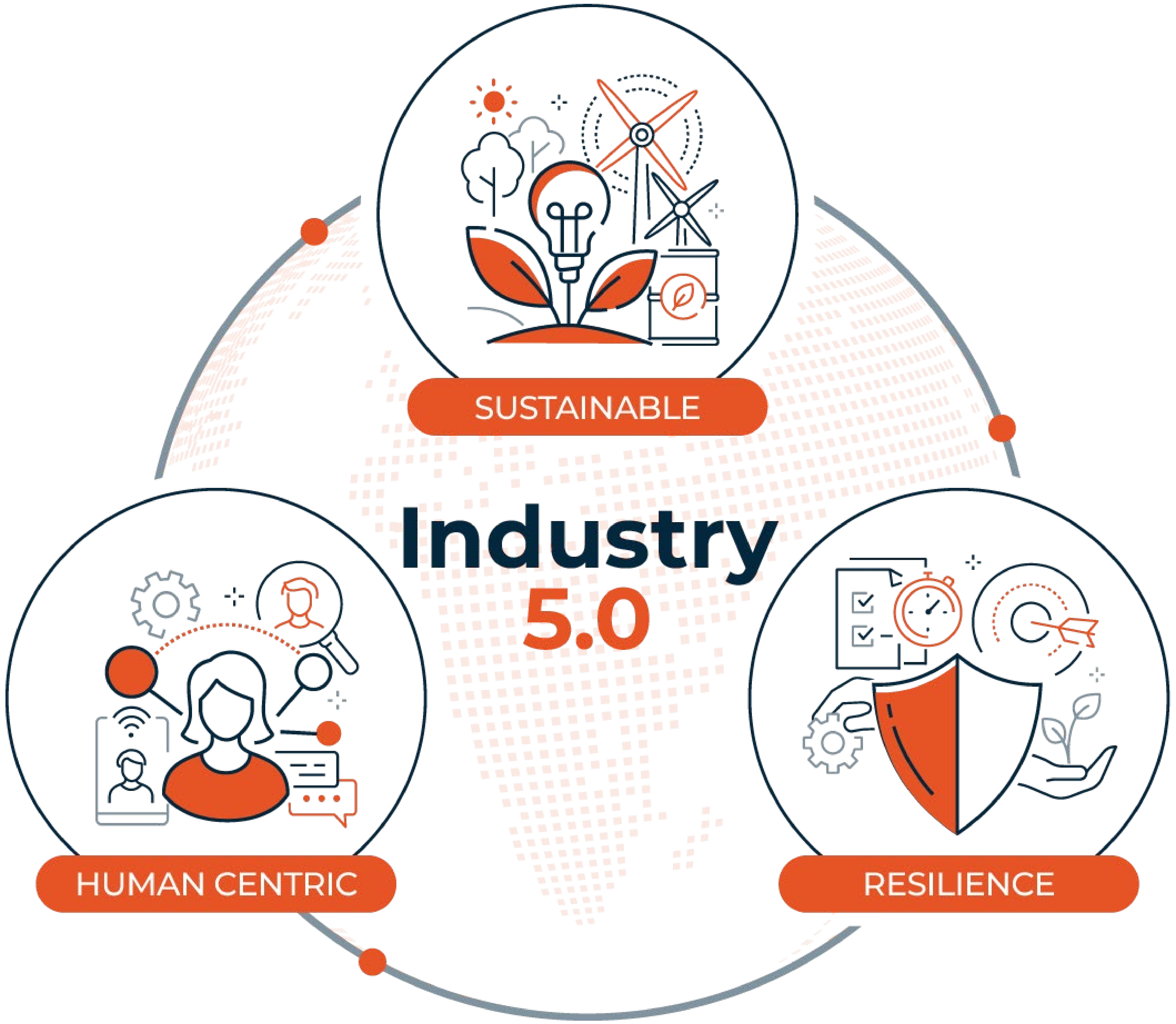
Data revolution

Data revolution

4th industrial revolution

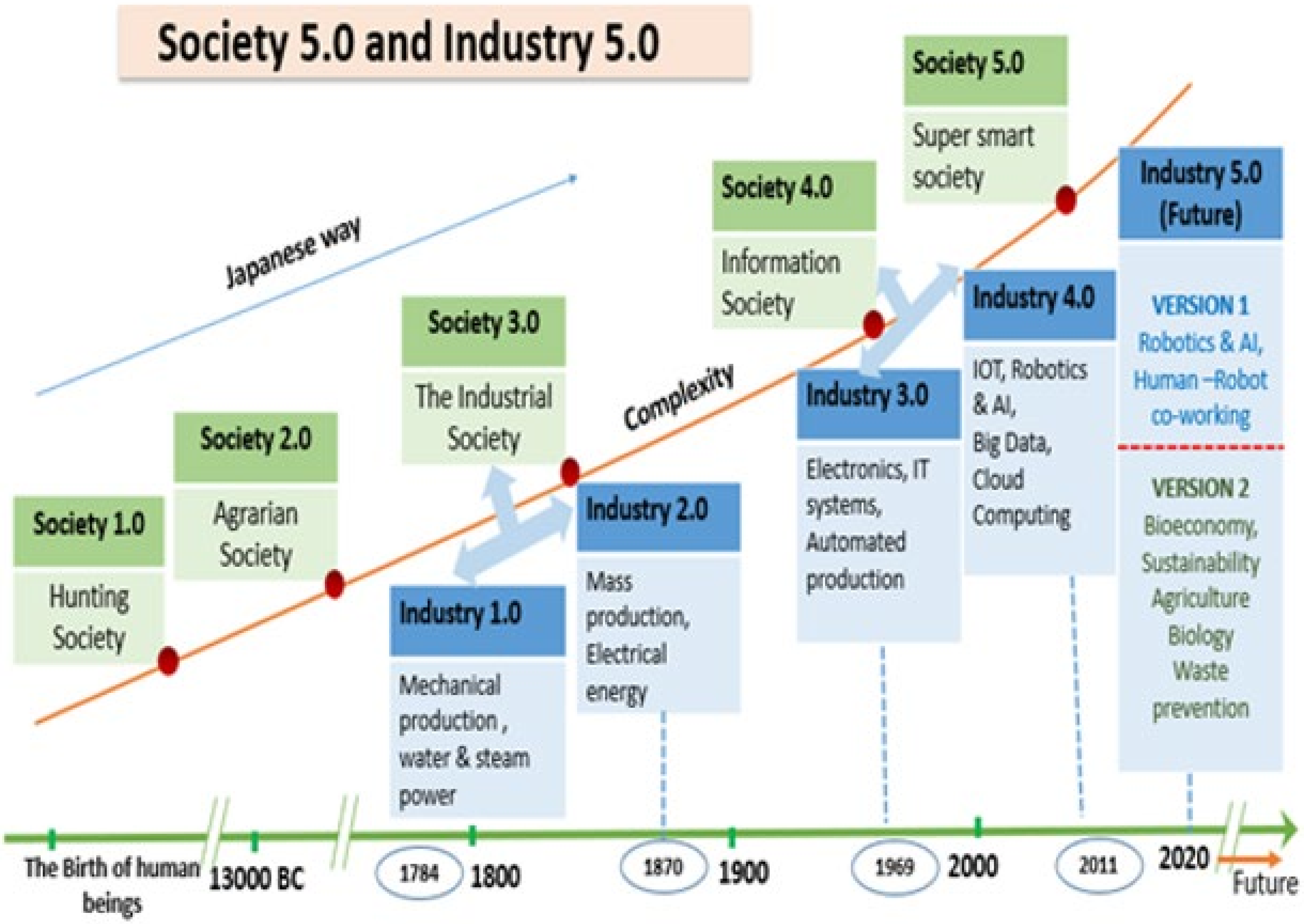


From Industry 4.0 to Industry 5.0



Source: Momenta Partners

A new industrial/societal revolution?



Source: SCHUMPETER CIRCLE

Data revolution
Amount of data

… and Machine Learning is a tool to learn from Big Data.

ML revolution

What is ML? Is it AI?

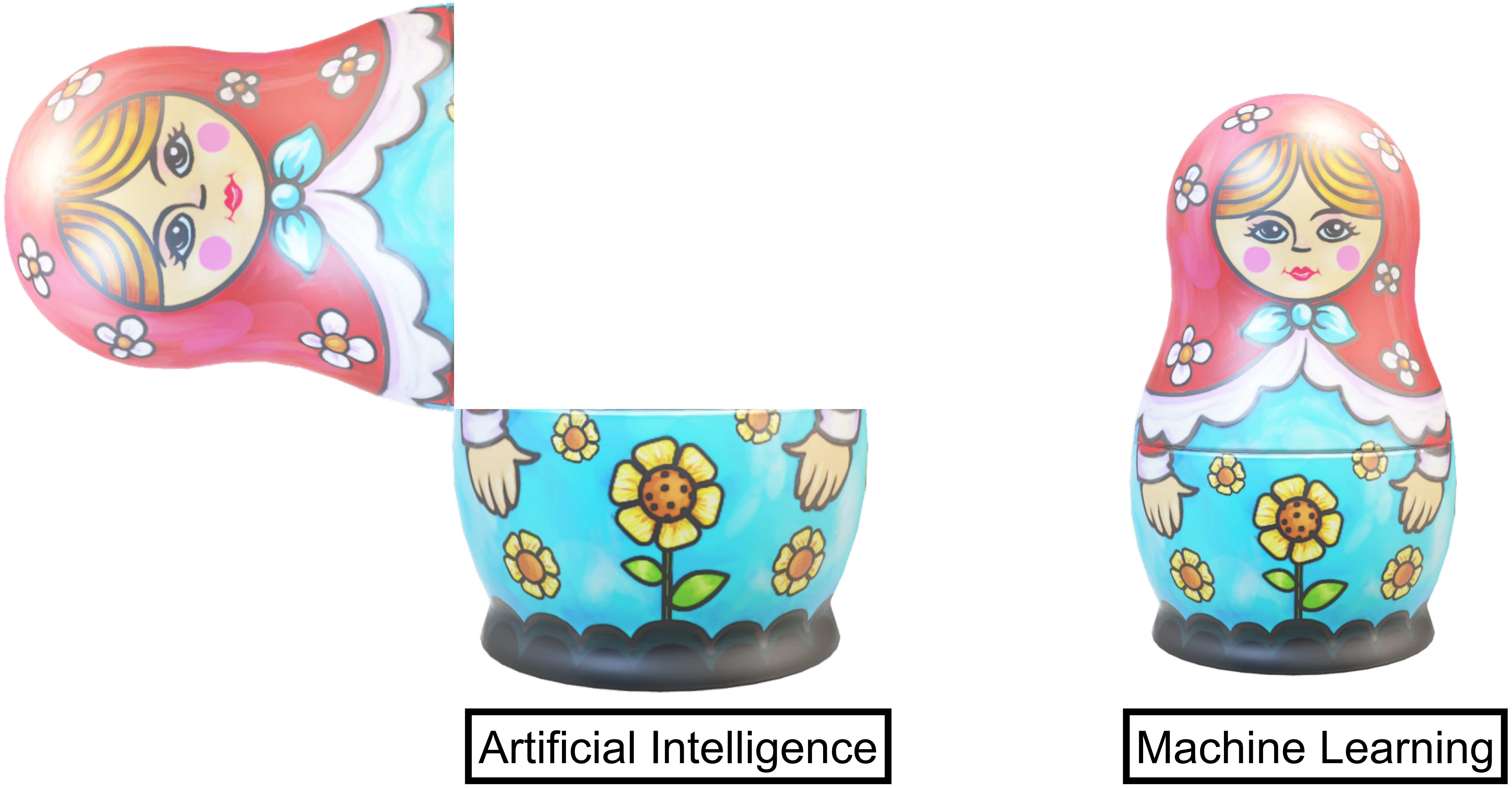
AI vs ML

Artificial Intelligence

Olga Fink / Alexandre Alahi

What is ML? Is it AI?

AI vs ML



What is ML? Is it AI?
AI vs ML

Artificial Intelligence

Machine Learning

What is the difference between AI and machine learning?

Machine learning is considered as a part of AI

Other methods for achieving AI include

- rule-based systems,
- evolutionary computation
- expert systems…

What is ML?

Olga Fink / Alexandre Alahi

- Instead of listing out the rules, let machines automatically learn how input **data** is correlated with a given task/objective/outcome/output.

- Let machines learns from experience/examples/data/feedbacks…

Machine Learning

Machine learning research is part of research on artificial intelligence, seeking to provide knowledge to computers through data, observations and interacting with the world. That acquired knowledge allows computers to correctly generalize to new settings.

Defintion by Yoshua Bengio, Université de Montréal

Olga Fink / Alexandre Alahi

Examples of ML

Is it a ML example?

Video credit: Tesla
https://www.tesla.com/autopilot

Examples of ML

Research example - GAN

The following people don't exist… They are generated by an AI

Try yourself at: https://thispersondoesnotexist.com

Examples of ML
Research example - GAN

GAN can do much more.

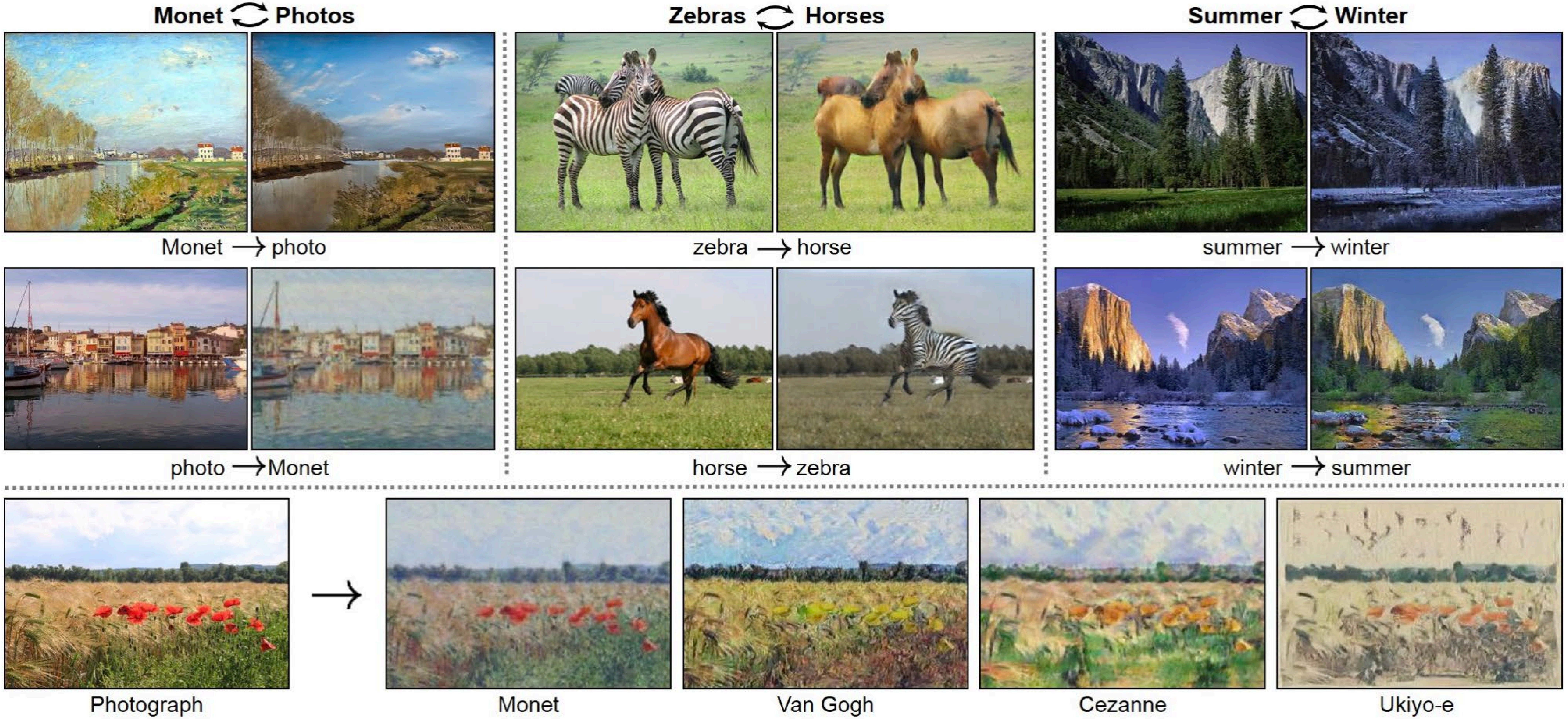
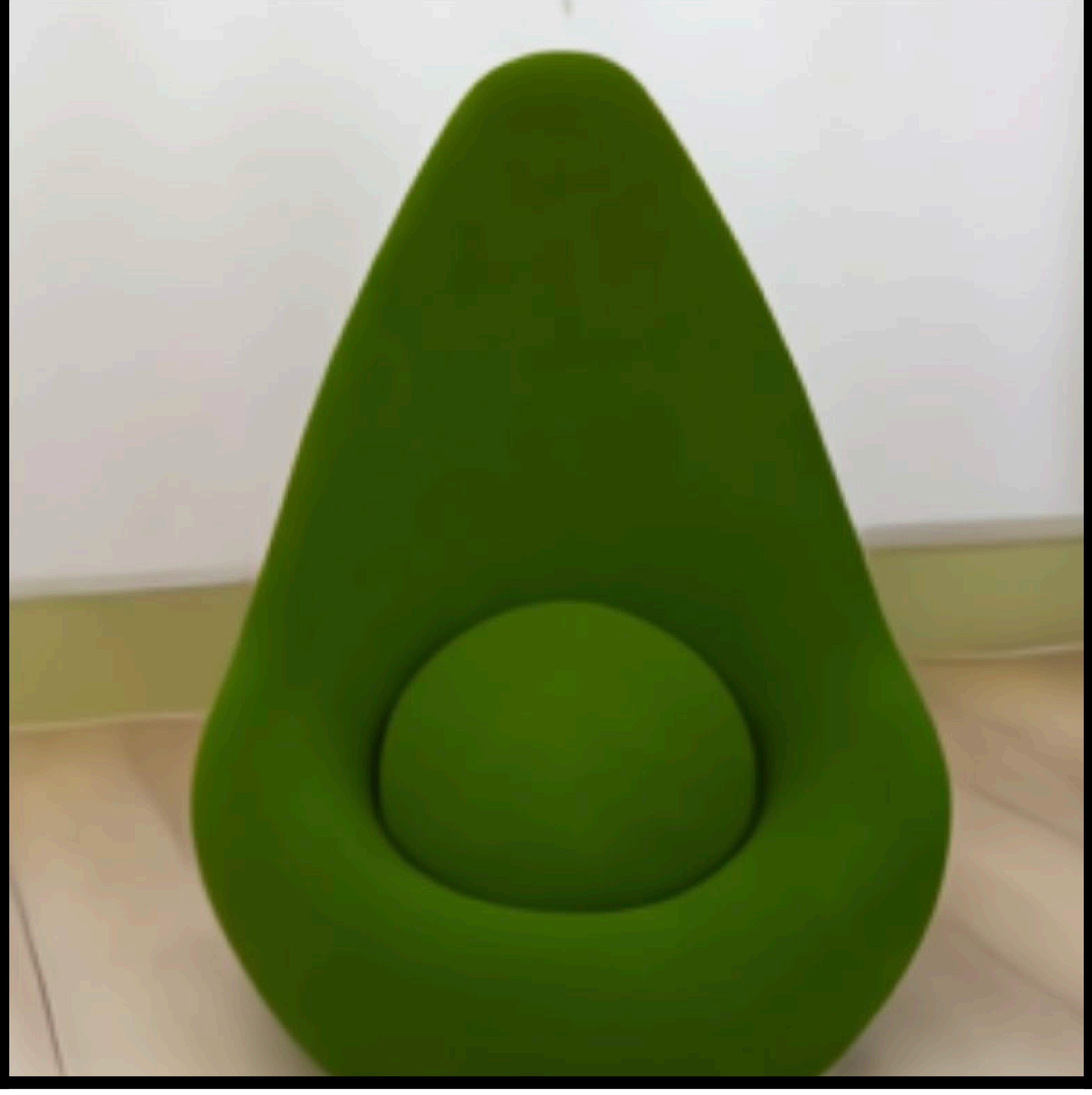
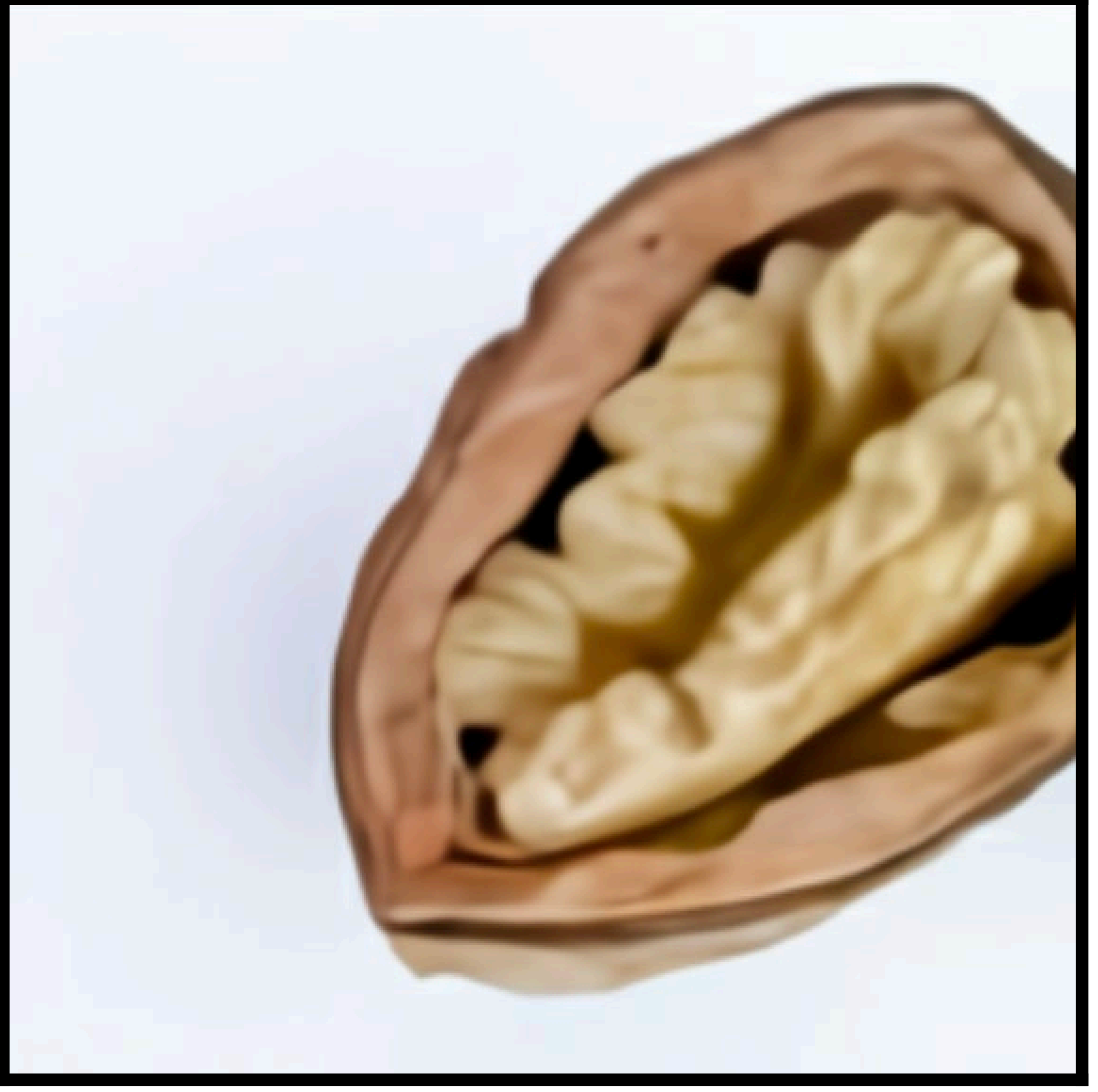


Image credit: Zhu et al., [Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks](), 2017

Examples of ML
Research example - Text to image

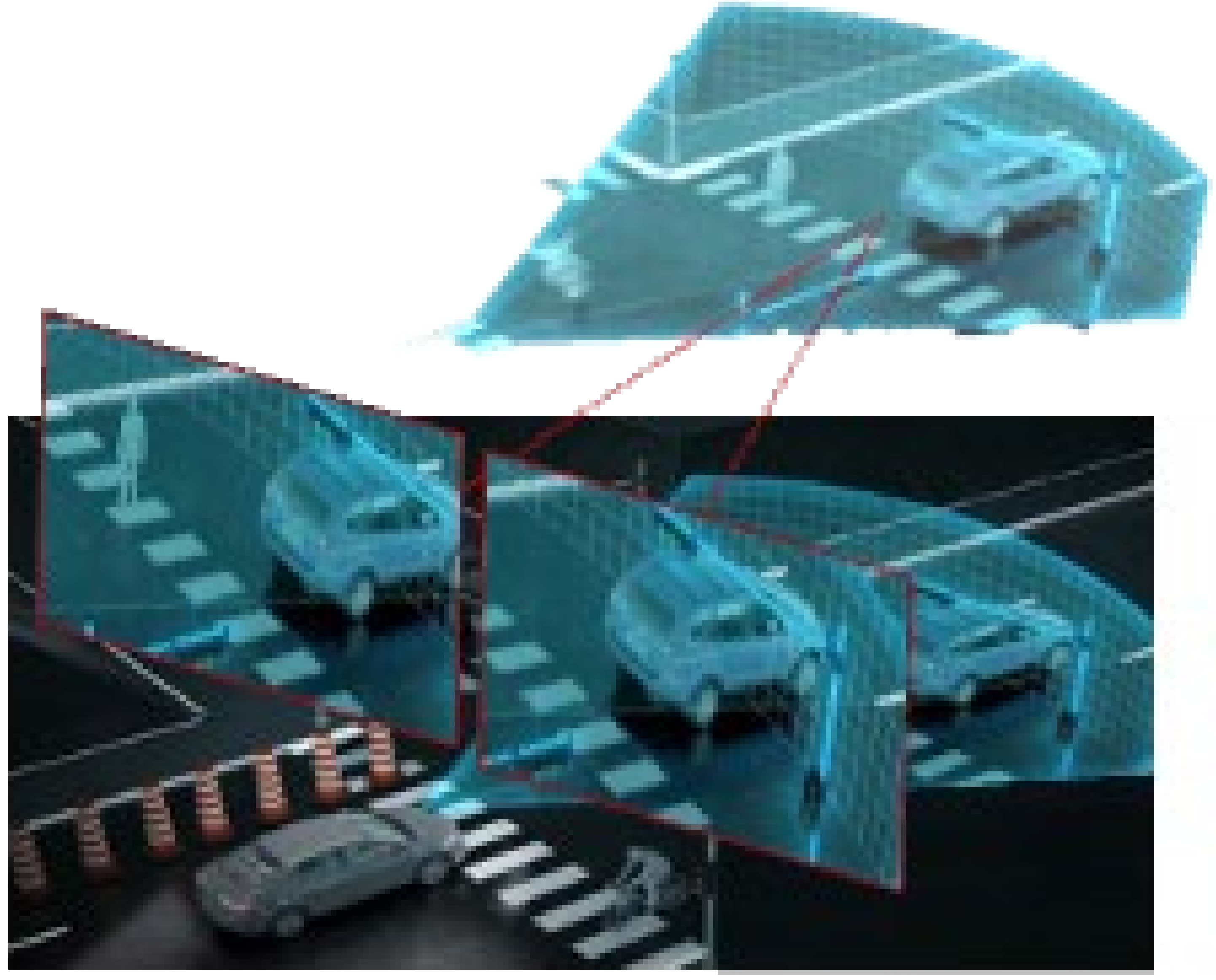
The following image were created from scratch based on the input text.

Try yourself at: https://openai.com/blog/dall-e/

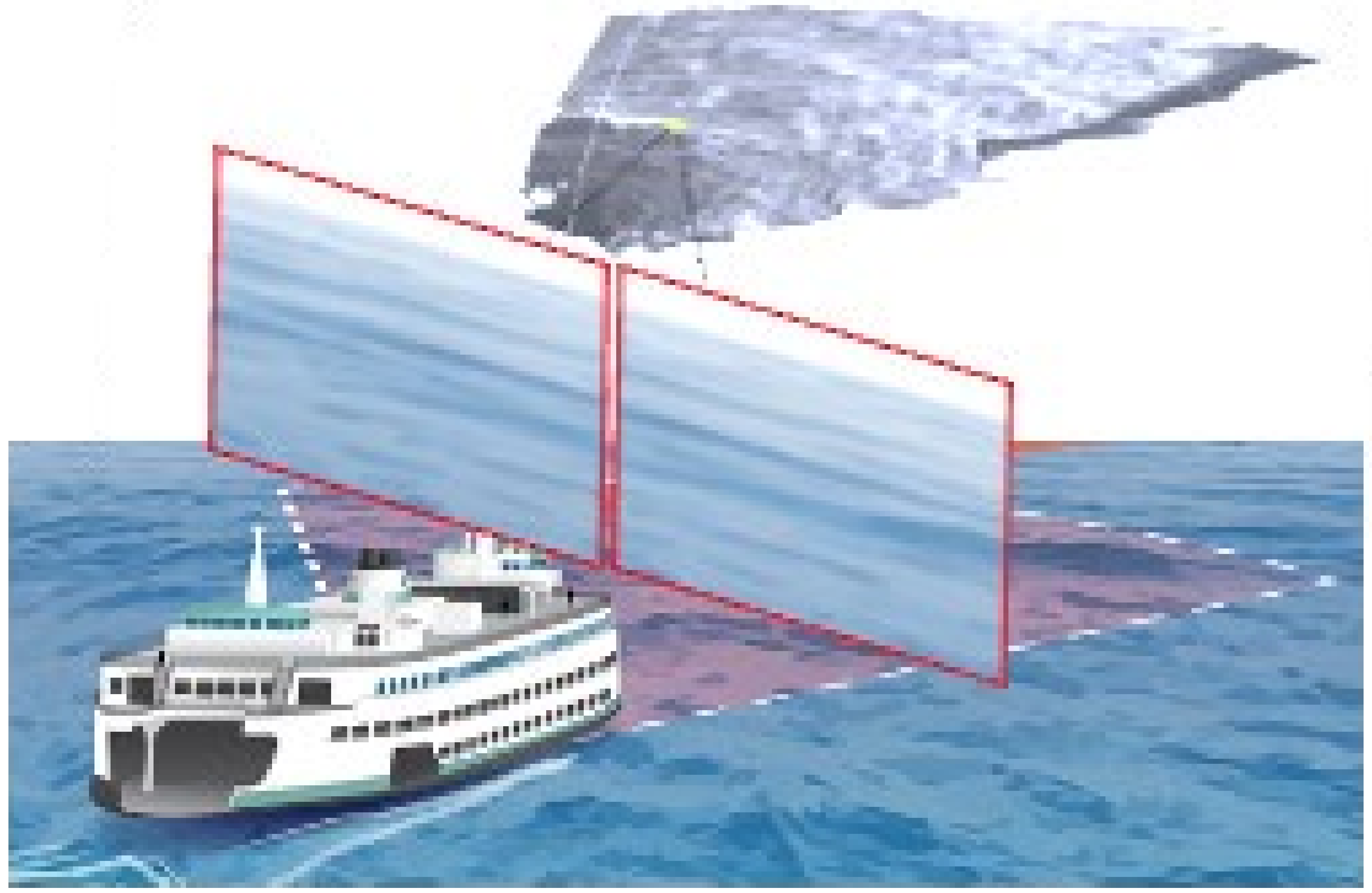
Input:	"An arm chair in the shape of an avocado"	"A collection of glasses is sitting on a table"	"A cross-section view of a walnut"
Output:			

Within EPFL…

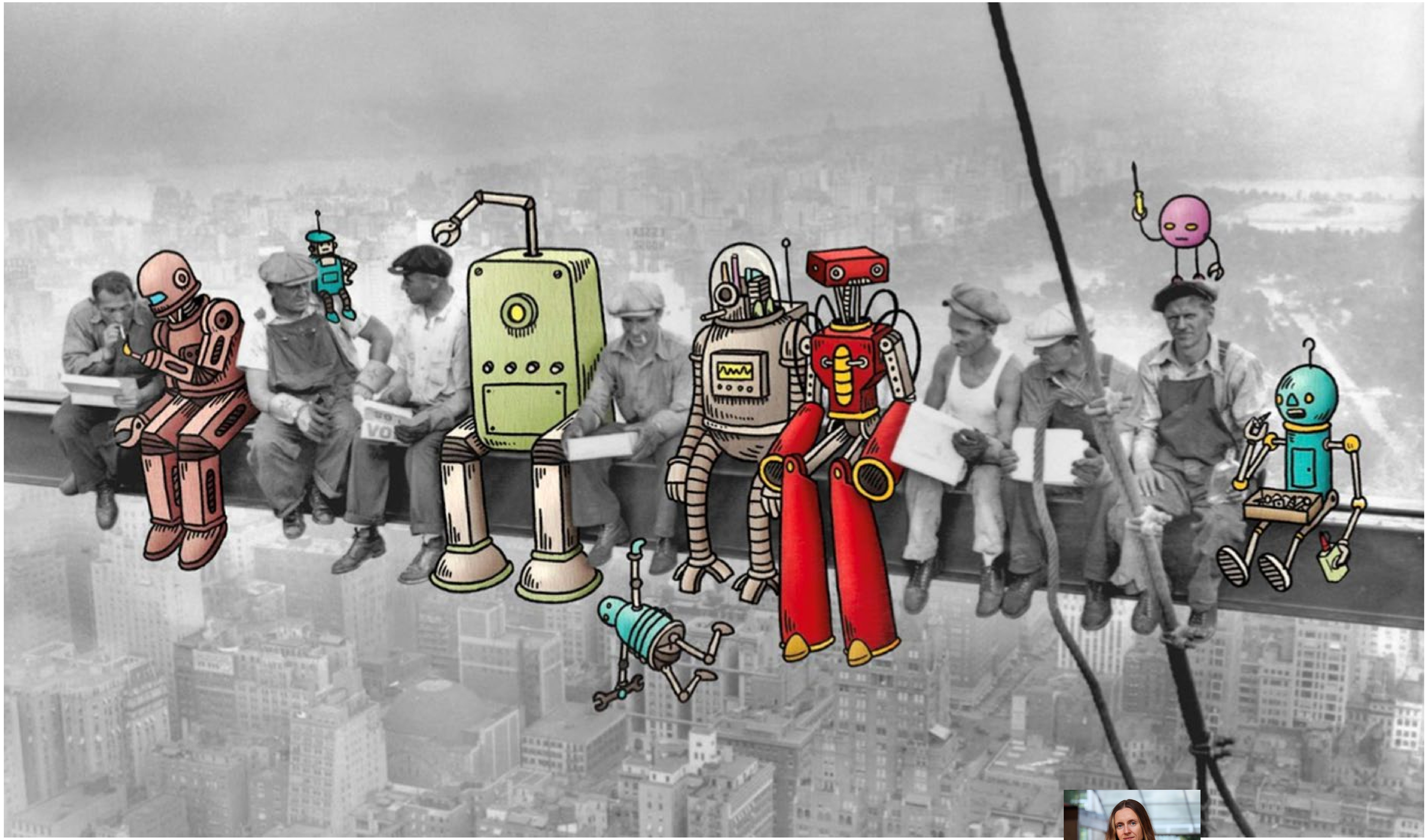
- 3D reconstruction of traffic participants



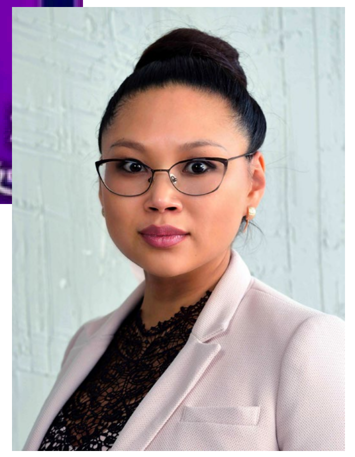
- 3D reconstruction of ocean waves



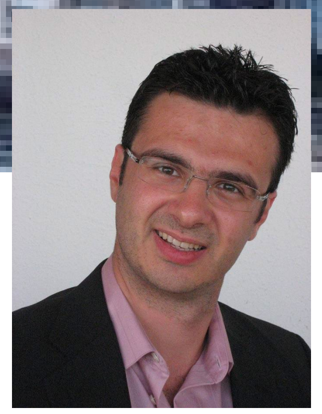
EERL lab



CRCL lab



ICE lab



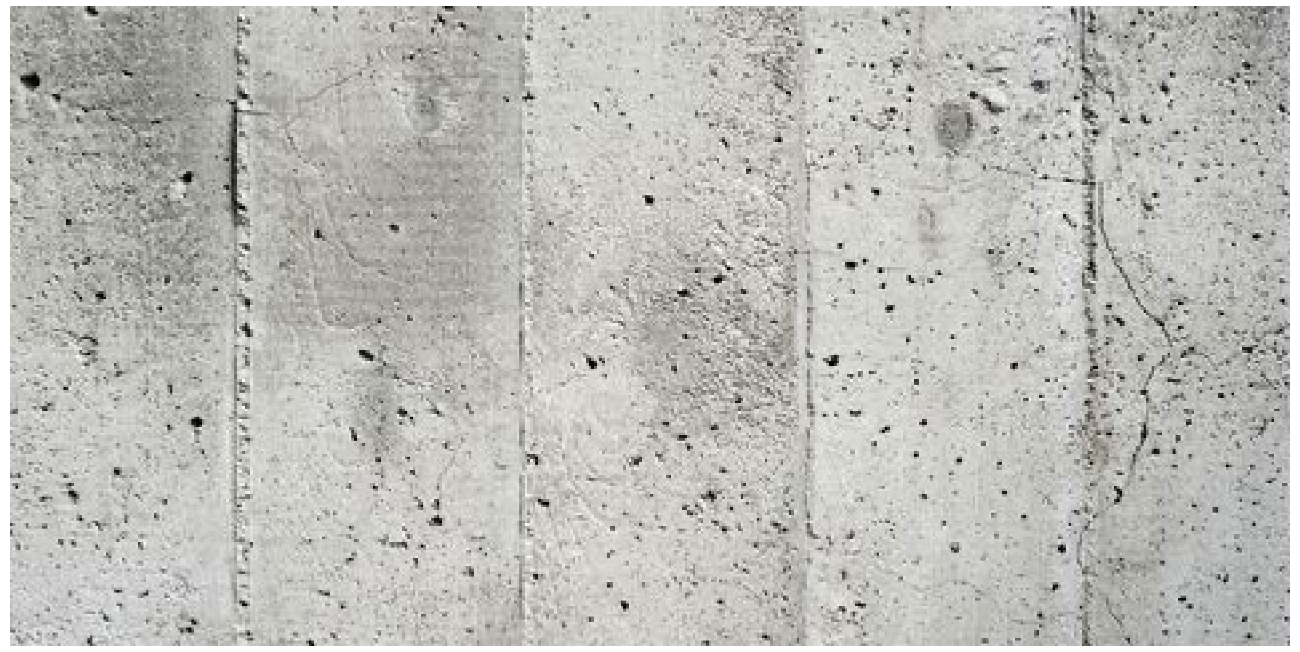
Luts lab

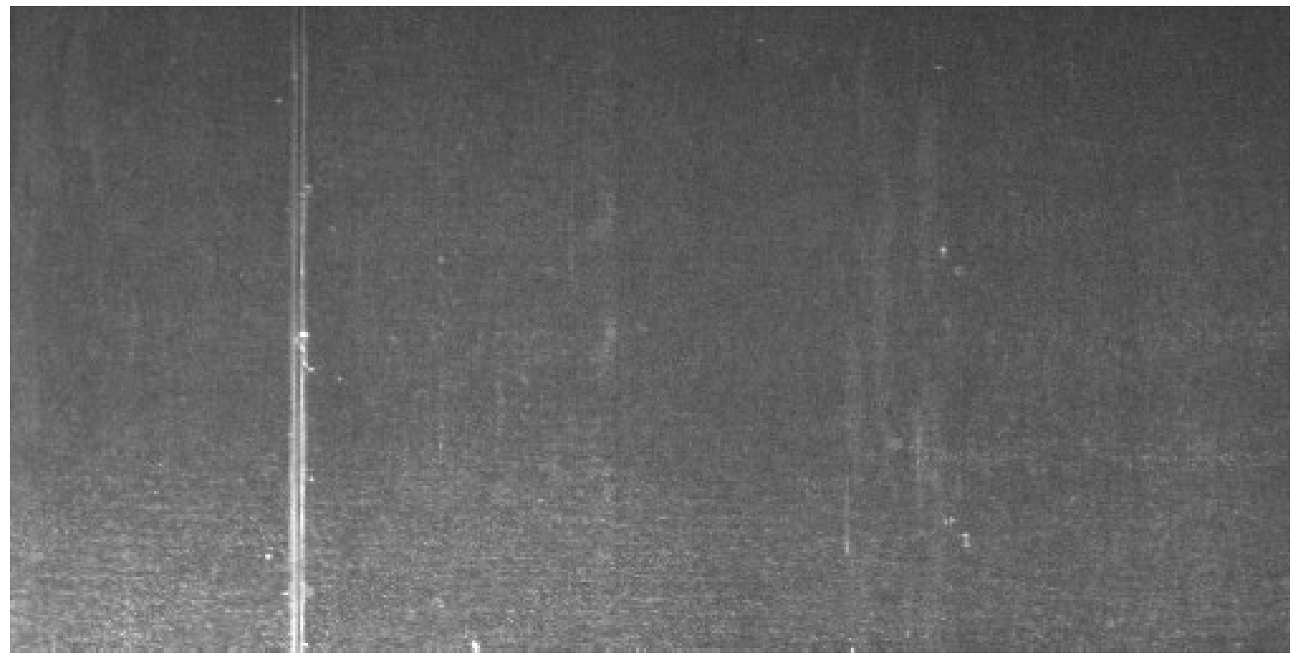
To recap

ML for civil engineers

In the following weeks, you will learn how to

- Predict concrete strength
- Determine building damage following an earthquake
- Predict slope stability
- Detect steel defects
- Predict hot water usage
- Predict the number of vehicles on a bridge
- Predict traffic flows
- …





Olga Fink / Alexandre Alahi

EPFL

Typical steps to follow in a machine learning (ML) project (1/3)

1. Problem Definition and Understanding
   - Clearly articulate the problem you aim to solve.
   - Identify the objectives and success criteria.
   - Ask the right questions: Identify the questions your data can answer to solve the business problem.
   - Understand the data: Learn the context, limitations, and opportunities within the available data.
2. Identify the Value:
   - Determine the potential value and impact of solving the problem (both from business and from the scientific perspective)
   - Assess how the solution will benefit stakeholders and align with business goals.
3. Collect Data:
   - Gather relevant data from various sources.
   - Make sure that your data is representative for the test data (application data) → be aware of the variability of the operating conditions (→ domain shift)
   - Ensure data quality and completeness.
   - Acquire labels for your data if possible (ensure the quality of labels)
4. Explore and Understand the Data:
   - Perform exploratory data analysis.
   - Visualize data to understand patterns and relationships.
   - Understand the data distribution (Explore the statistical properties, distributions, and trends in the data)
   - Identify data quality issues (Look for missing values, outliers, and inconsistencies that may affect analysis).
   - Generate hypotheses (Form hypotheses based on patterns observed during data exploration).

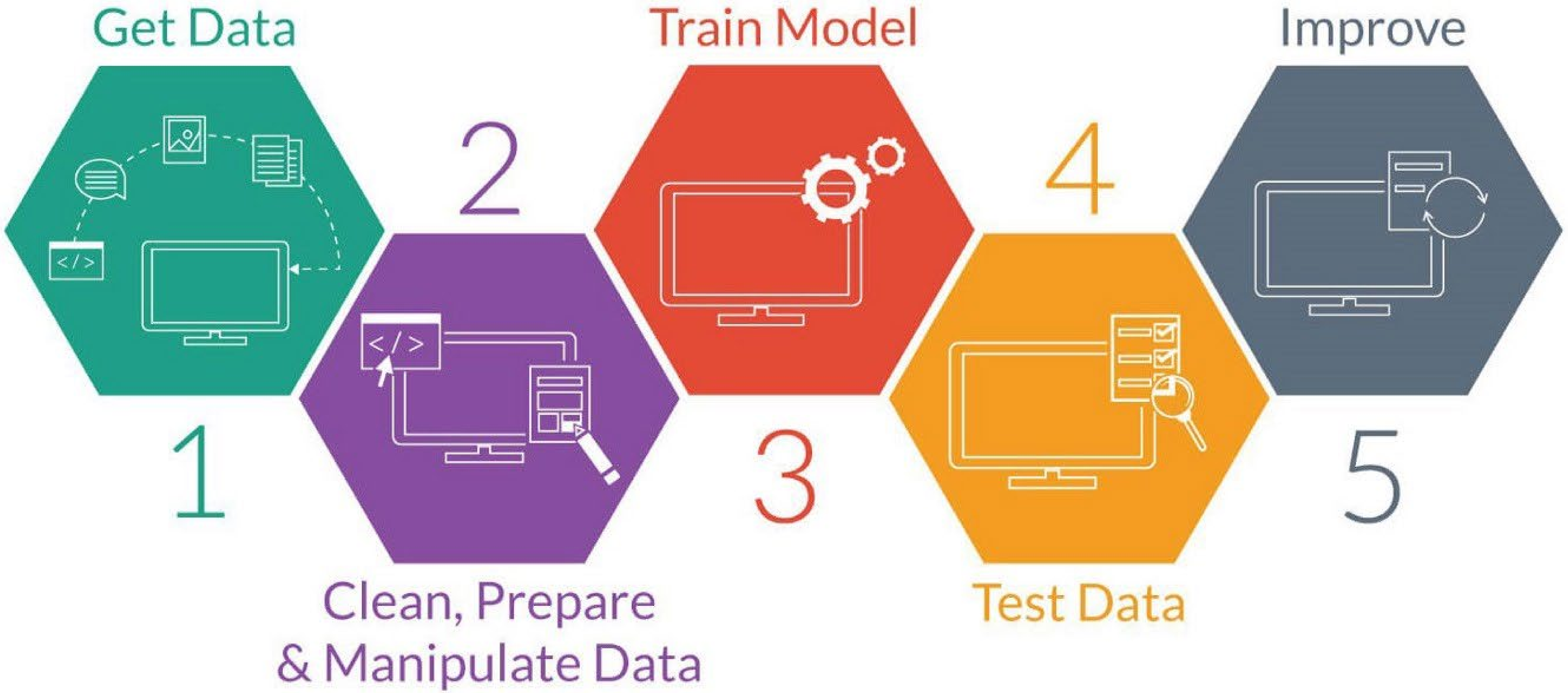
Typical steps to follow in a machine learning (ML) project (2/3)

5. Prepare Data:
   - Clean the data (handle missing values, outliers).
   - Feature engineering (create meaningful features, transform existing ones).
   - Split data into training, validation, and test sets.
   - Scaling and normalization (Prepare the data for modeling by applying appropriate scaling techniques)
   - Address class imbalances: If needed, use techniques like oversampling, undersampling, or synthetic data generation to balance the dataset.
6. Select and Train Models:
   - Select appropriate models (Choose models based on the problem type (regression, classification, clustering, etc.) and data characteristics)
   - Split the dataset (Use train-test splits (or cross-validation) to ensure your model's generalization to unseen data)
   - Baseline model (Start with a simple model as a baseline to compare more complex models).
   - Tune Hyperparameters (Optimize model hyperparameters for better performance).
   - Iterate and improve (Experiment with different models and tuning hyperparameters)
   - Avoid overfitting (Implement regularization techniques or apply cross-validation to avoid overfitting the training data).
7. Evaluate Models:
   - Choose appropriate metrics (Select evaluation metrics that are relevant to the problem (e.g., accuracy, precision, recall, F1-score, RMSE))
   - Validate on unseen data (Always validate your model on a test set or through cross-validation to assess its performance)
   - Compare models: Compare different models based on performance metrics and business value.

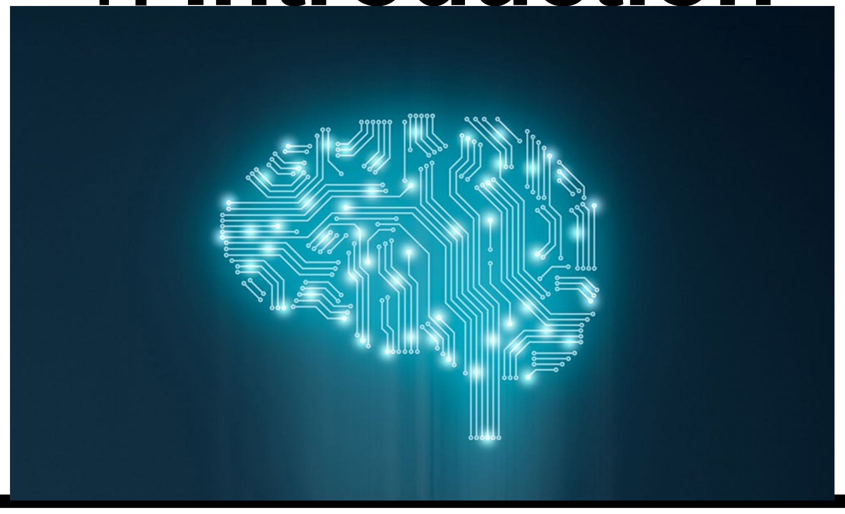
Typical steps to follow in a machine learning (ML) project (3/3)

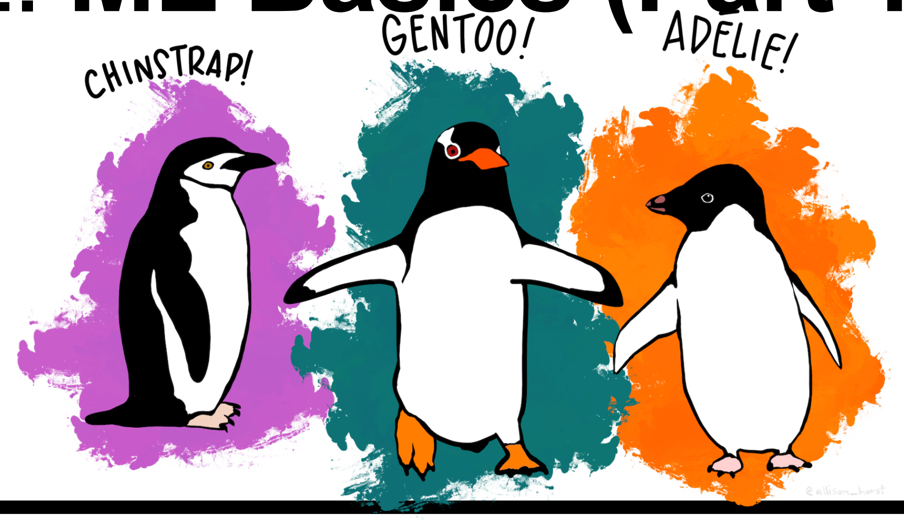
8. Model Interpretation
   - What do the results mean?
   - Explainability (Ensure that you understand how your model works, and use XAI techniques such as SHAP values or feature importance to explain the results).
   - Check assumptions (Ensure that the model's assumptions hold and are consistent with the problem and data context)
9. Test Model:
   - Evaluate the final model on the test dataset to gauge its real-world performance.
10. Deploy Model:
    - Integrate the model into a production environment.
11. Monitor and Maintain:
    - Continuously monitor model performance.
    - Update the model as needed based on new data and changing conditions.
    - Monitor how your model is actually used
12. Documentation and Communication
    - Document thoroughly (Keep detailed notes on every step, including data sources, preprocessing steps, model choices, evaluation, and deployment procedures)
    - Communicate results (Present findings to stakeholders in a clear, actionable manner. Tailor communication based on the audience (e.g., technical team vs. business executives).)

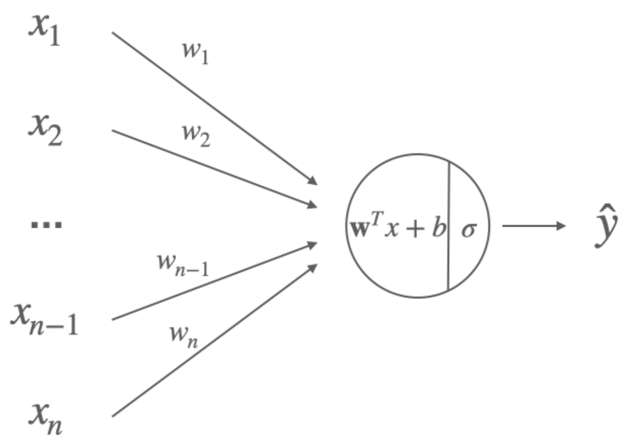
What is ML?
The process



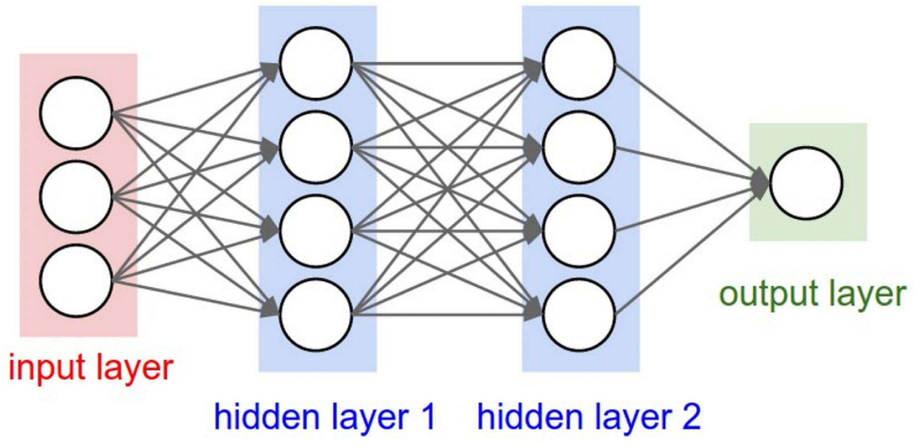
Road map
Lectures

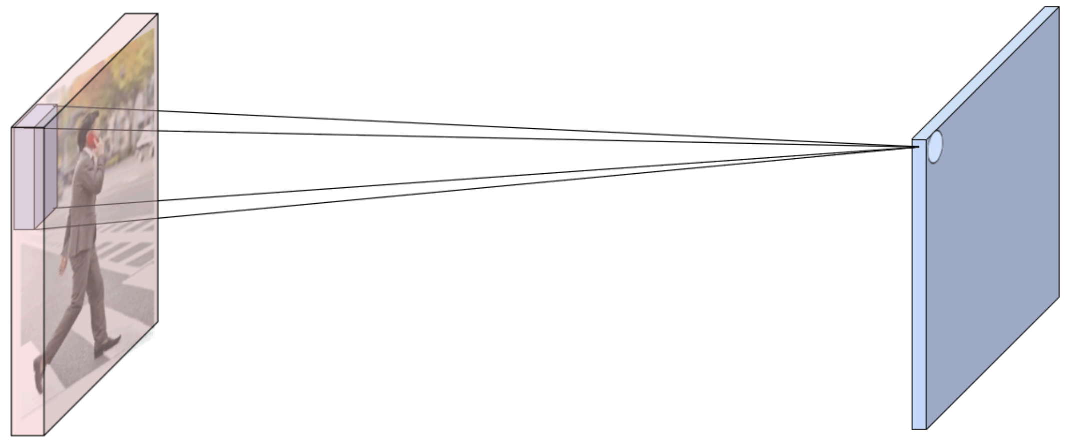
1. Introduction


2. ML Basics (Part 1)


3. ML Basics (Part 2)


4. Role of Input


5. Deep Learning (Part 1)


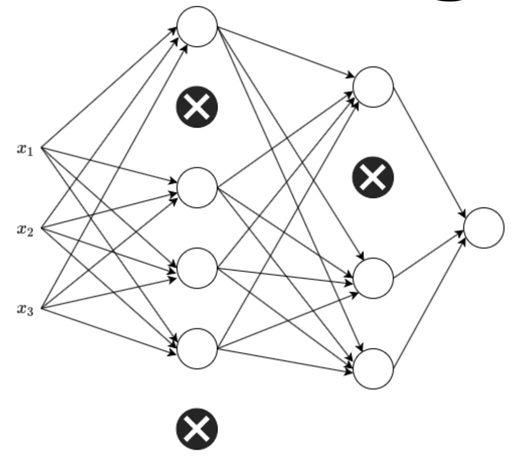
6. Deep Learning (Part 2)


Olga Fink / Alexandre Alahi

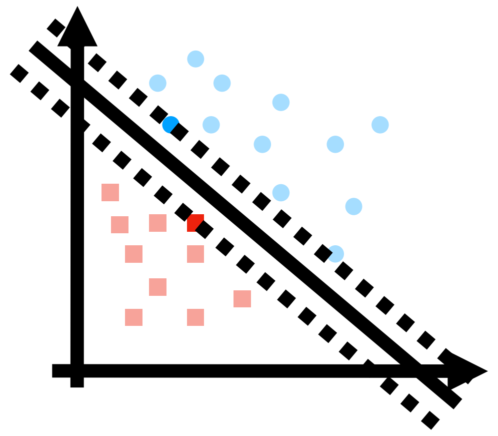
Road map

Lectures

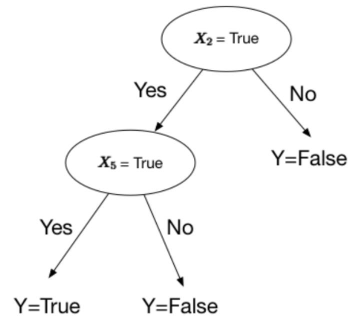
7. Deep Learning (Part 3)



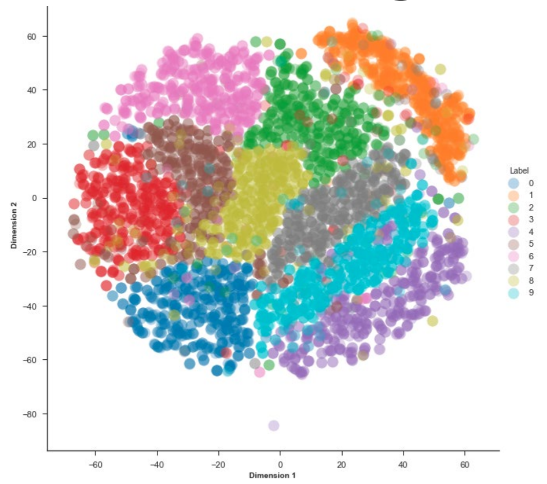
8. Support Vector Machine



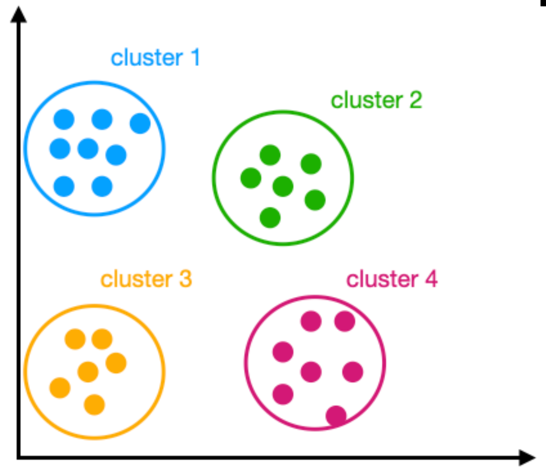
9. Other Supervised Learning Models



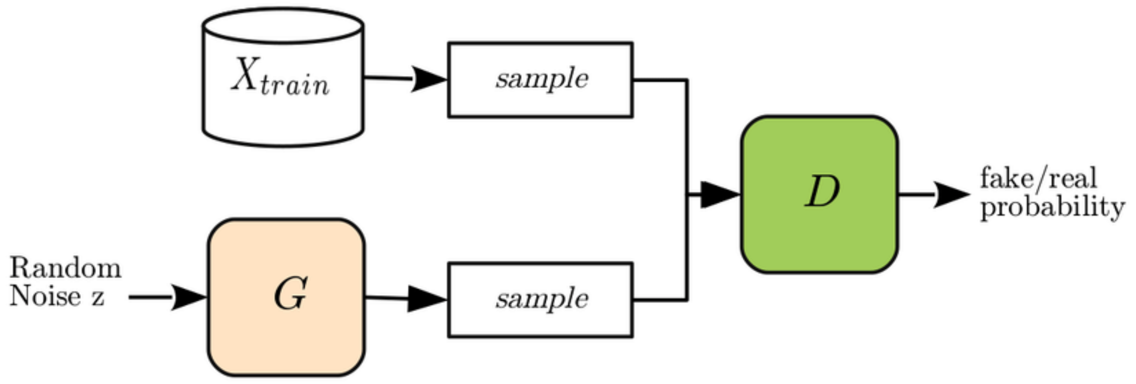
10. Dimensionality Reduction



11. Clustering



12. Other Learning Examples



At the end of the course

You will:

- Know the most common ML techniques
- Know how to implement ML techniques in Python using popular data science libraries
- Have all the tools needed to start a ML project on your own
- Be able to read ML literature

Olga Fink / Alexandre Alahi

Research

Papers on ML in Civil Engineering:

- Chang, Lin, Chang, Applications of neural network models for structural health monitoring based on derived modal properties, 2018
- Dong, Chen, Guan, Cost Index Predictions for Construction Engineering Based on LSTM Neural Networks, 2020
- Huang, Huang, Lyu, An Improved KNN-Based Slope Stability Prediction Model, 2020
- Silva, Moita, Arruda, Machine learning techniques to predict the compressive strength of concrete, 2020
- Feng et al., Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach, 2020
- Li, Zhao, Image-Based Concrete Crack Detection Using Convolutional Neural Network and Exhaustive Search Technique, 2019
- Gao, Kong, Mosalam, Deep leaf-bootstrapping generative adversarial network for structural image data augmentation, 2019
- …

Optional

Resources
To go further

- « What is Artificial Intelligence Exactly? » *by Dagogo Altraide*: https://www.youtube.com/watch?v=kWmX3pd1f10&list=