Deep Learning a Primer with future Outlooks

Prof . Simone Calderara

AImagelab Computer Vision, Pattern Recognition & Machine Learning

Dief University Of Modena and Reggio Emilia

Contact: simone.calderara@unimore.it

Artificial Intelligence

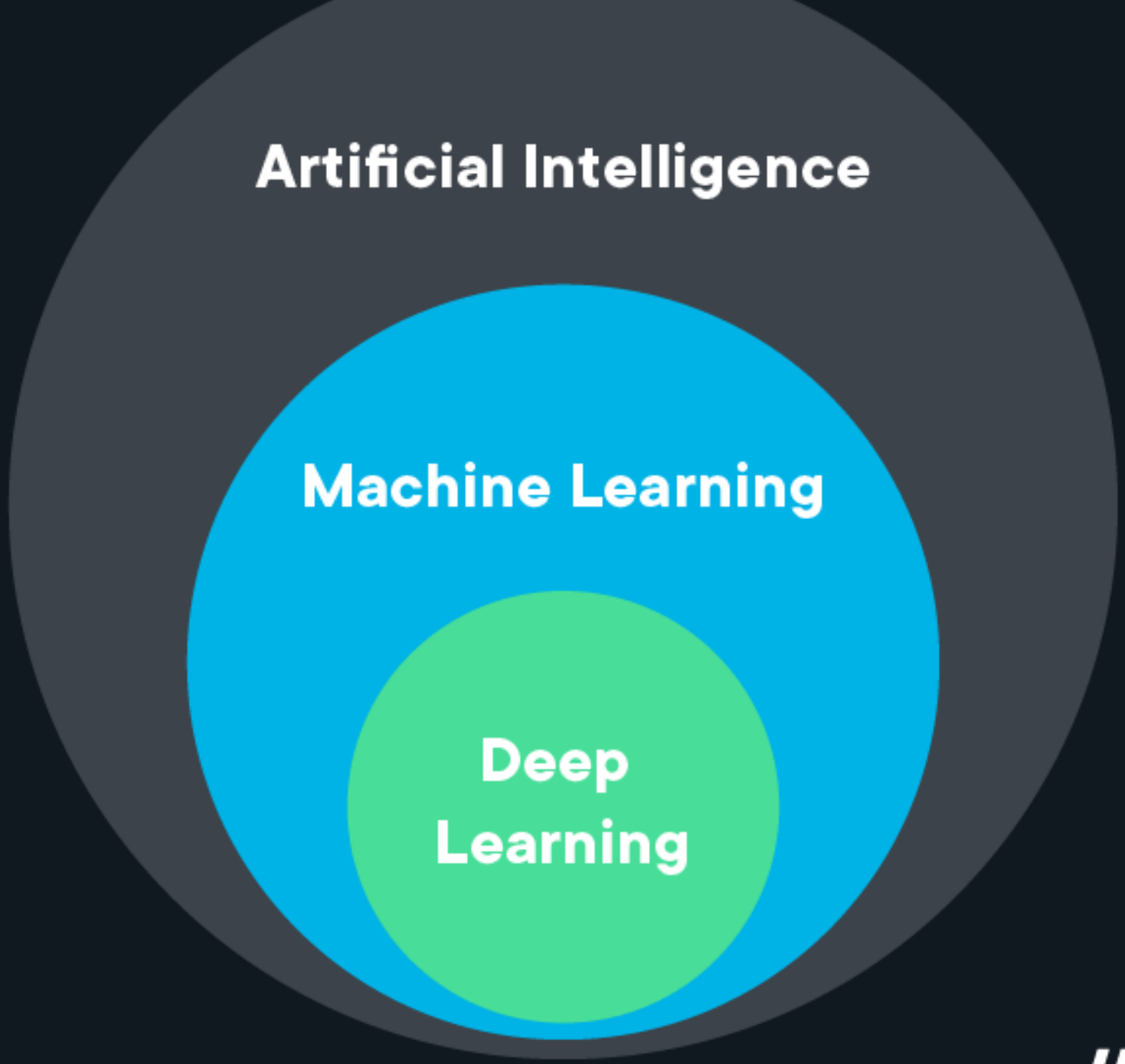
A science devoted to making machines think and act like humans.

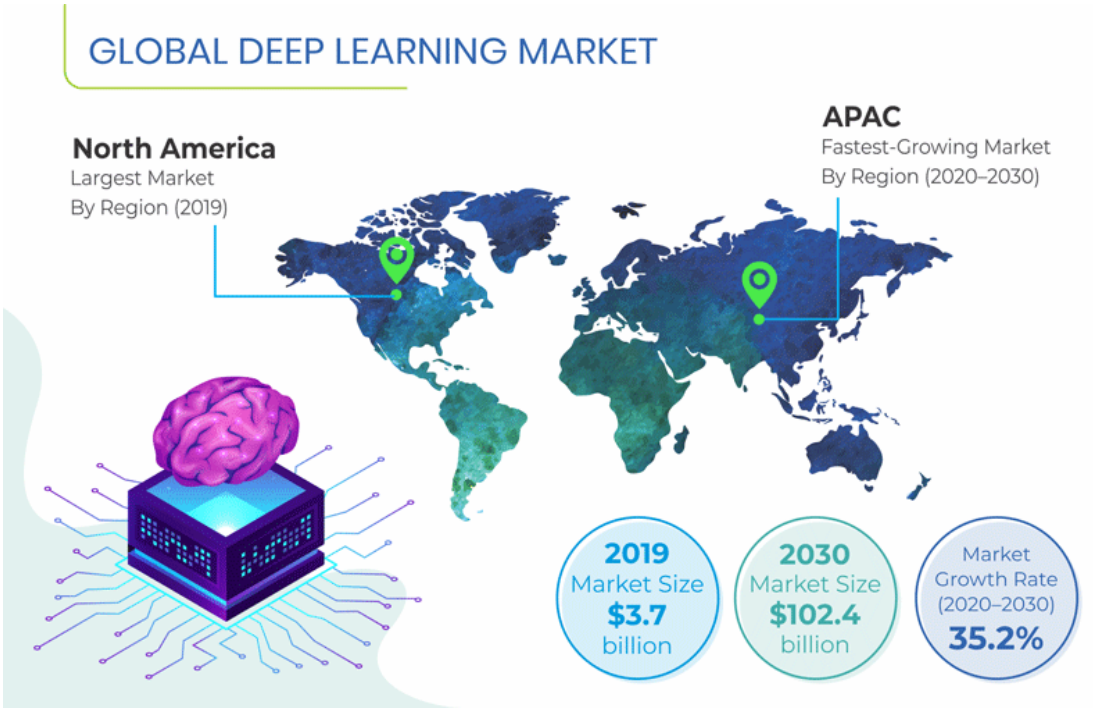
Machine Learning

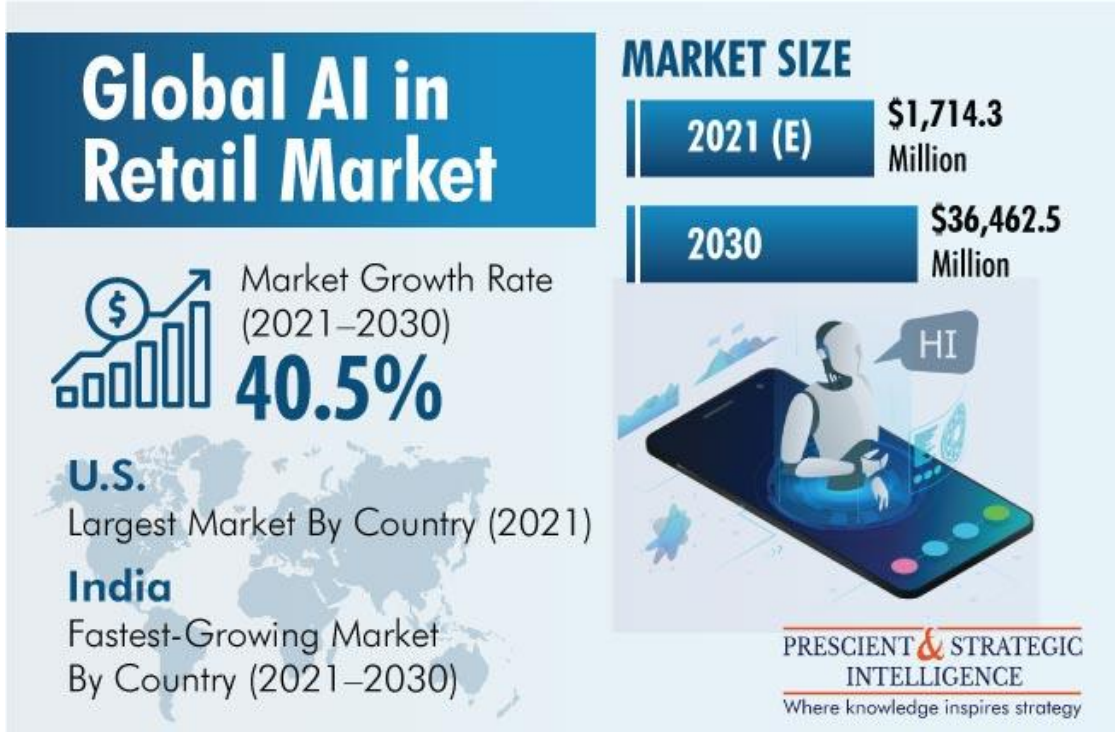
Focuses on enabling computers to perform tasks without explicit programming.

Deep Learning

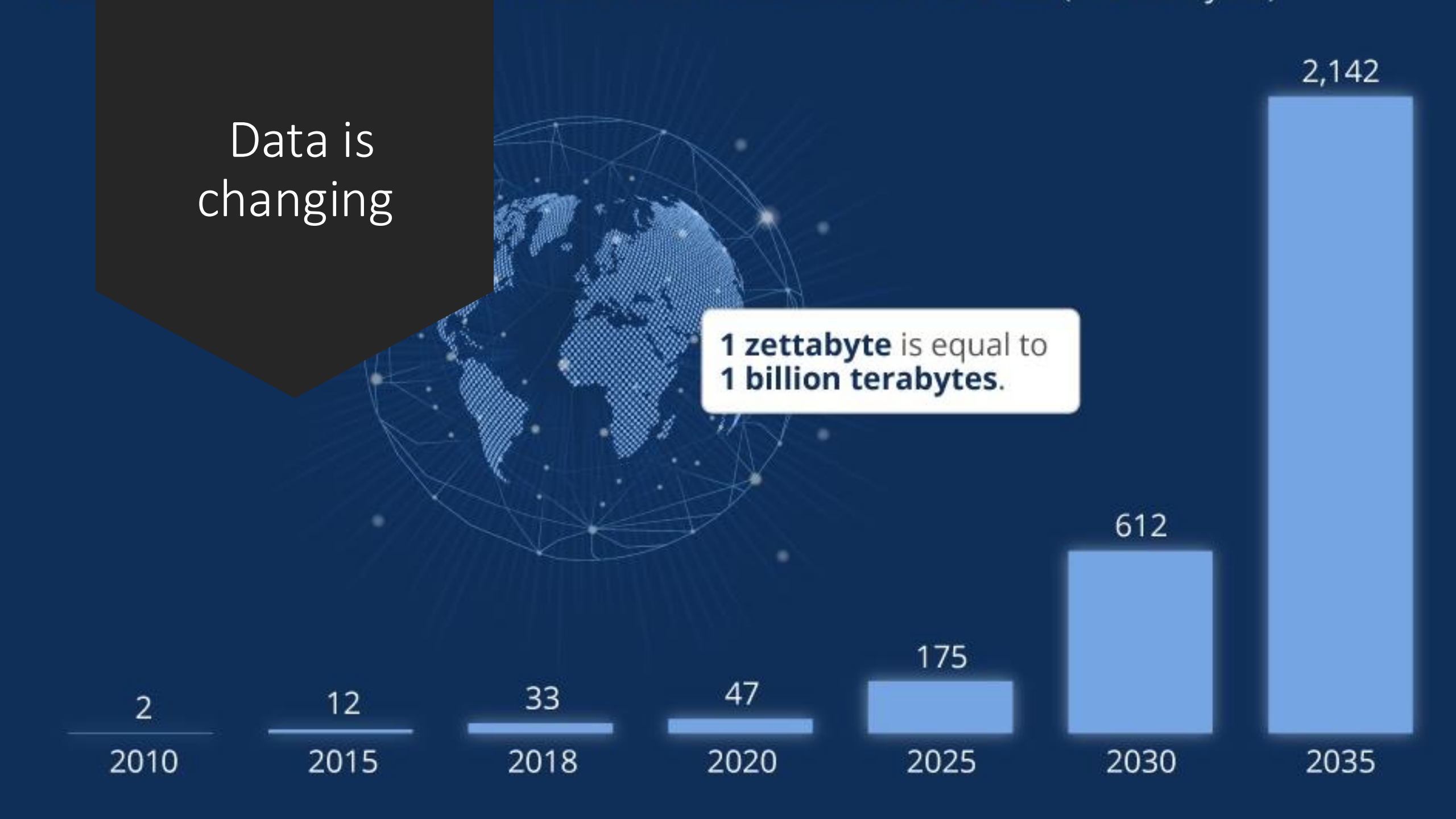
A subset of machine learning based on artificial neural networks.




GLOBAL DEEP LEARNING MARKET
North America
Largest Market
By Region (2019)
APAC
Fastest-Growing Market
By Region (2020–2030)
2019
Market Size
$3.7
billion
2030
Market Size
$102.4
billion
Market
Growth Rate
(2020–2030)
35.2%


Global AI in Retail Market
Market Growth Rate
(2021–2030)
40.5%
U.S.
Largest Market By Country (2021)
India
Fastest-Growing Market
By Country (2021–2030)
MARKET SIZE
2021 (E)
$1,714.3
Million
2030
$36,462.5
Million
HI
PRESCIENT & STRATEGIC
INTELLIGENCE
Where knowledge inspires strategy

Data is changing



A DAY IN DATA

The exponential growth of data is undisputed, but the numbers behind this explosion - fuelled by internet of things and the use of connected devcies – are hard to comprehend, particularly when looked at in the context of one day

500m

tweets are sent every day

Twitter

4PB

of data created by Facebook, including

350m photos

100m hours of video watch time

Facebook Research

294bn

billion emails are sent

Radicati Group

320bn

emails to be sent each day by 2021

306bn

emails to be sent each day by 2020

3.9bn

people use emails

4TB

of data produced by a connected car

Intel

DEMYSTIFIYING DATA UNITS

From the more familiar 'bit' or 'megabyte', larger units of measurement are more frequently being used to explain the masses of data

Unit		Value	Size
b	bit	0 or 1	1/8 of a byte
B	byte	8 bits	1 byte
KB	kilobyte	1,000 bytes	1,000 bytes
MB	megabyte	$1,000^2$ bytes	1,000,000 bytes
GB	gigabyte	$1,000^3$ bytes	1,000,000,000 bytes
TB	terabyte	$1,000^4$ bytes	1,000,000,000,000 bytes
PB	petabyte	$1,000^5$ bytes	1,000,000,000,000,000 bytes
EB	exabyte	$1,000^6$ bytes	1,000,000,000,000,000,000 bytes
ZB	zettabyte	$1,000^7$ bytes	1,000,000,000,000,000,000,000 bytes
YB	yottabyte	$1,000^8$ bytes	1,000,000,000,000,000,000,000,000 bytes

*A lowercase "b" is used as an abbreviation for bits, while an uppercase "B" represents bytes.

65bn

messages sent over WhatsApp and two billion minutes of voice and video calls made

Facebook

ACCUMULATED DIGITAL UNIVERSE OF DATA

Searches made a day — 5bn

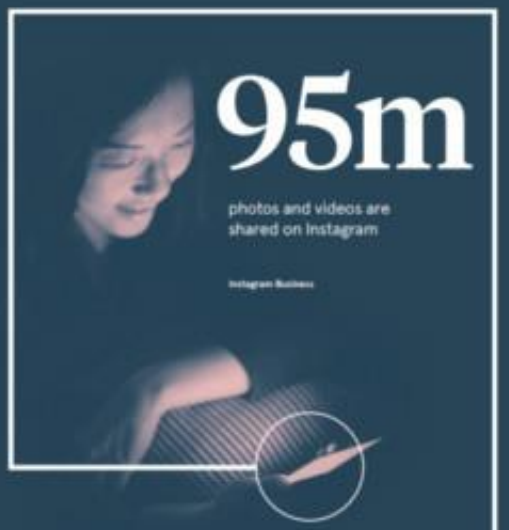
Searches made a day from Google — 3.5bn

Smart Insights

463EB

of data will be created every day by 2025

IDC



95m

photos and videos are shared on Instagram

Instagram Business

28PB

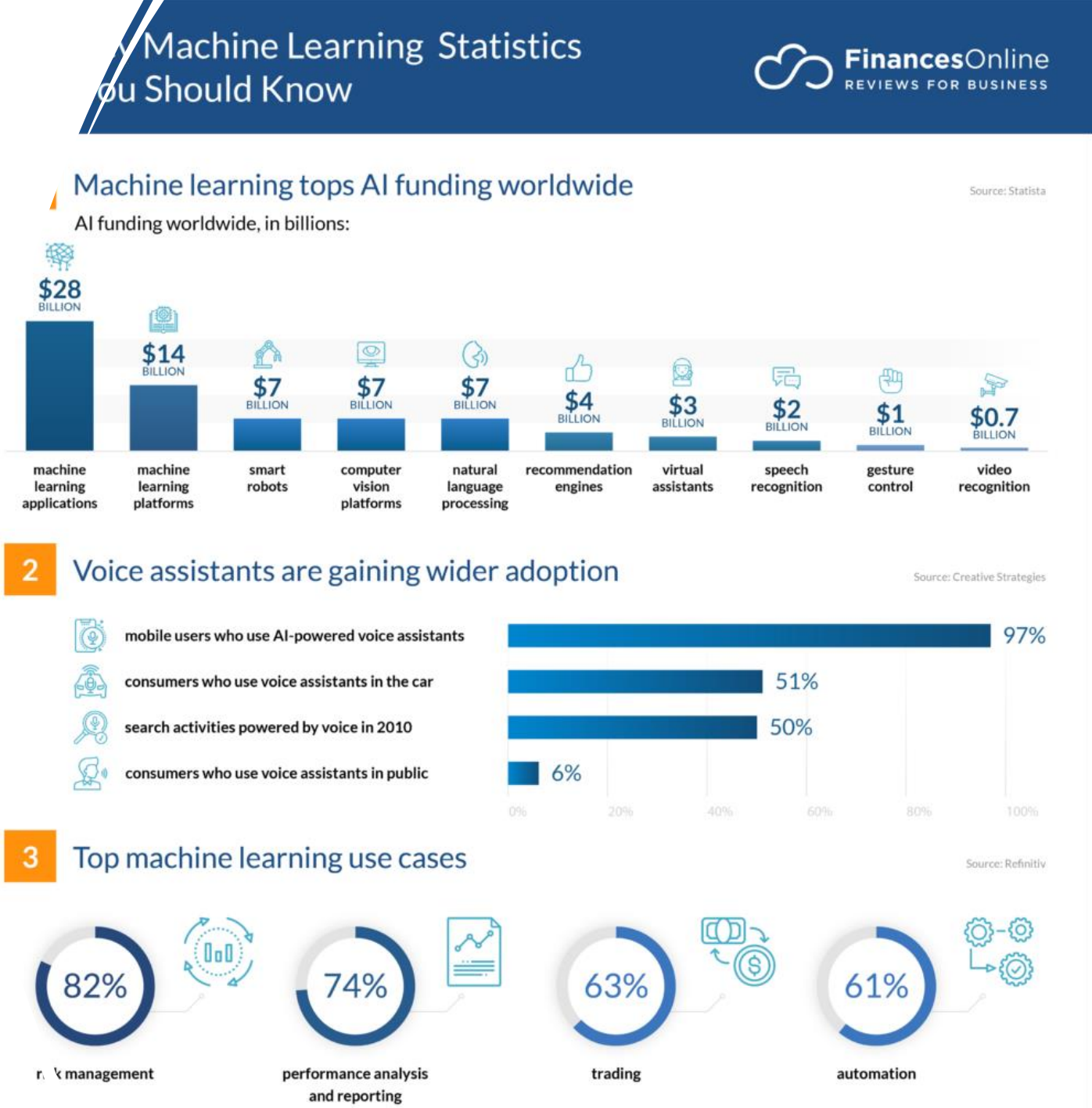
to be generated from wearable devices by 2020

Statista

From the Past to the future

- 2008-2013 Years of theoretical studies and hardware production

- 2016->…… **Time to bring out the application**
  - “Google and Movidius who have teamed up to increase adoption (of deep learning technology) within mobile devices.”
  - Google changed the «Page Rank» algorithm with «Rank Brain» Deep learning based
  - Facebook «face recognition» is deep learning based
  - Google and Apple cars use DL to drive autonomous vehicles
  - Toyota is spending $1 billion on AI in Silicon Valley for autonomous cars
  - GPT4 is the current standard de-facto for text analysis



Gartner's Top Strategic Technology Trends for 2023

Optimize	Scale	Pioneer
Digital Immune System	Industry Cloud Platforms	Superapps
Applied Observability	Platform Engineering	Adaptive AI
AI TRiSM	Wireless-Value Realization	Metaverse
	Sustainable Technology	

Gartner

2023 Gartner Emerging Technologies and Trends Impact Radar

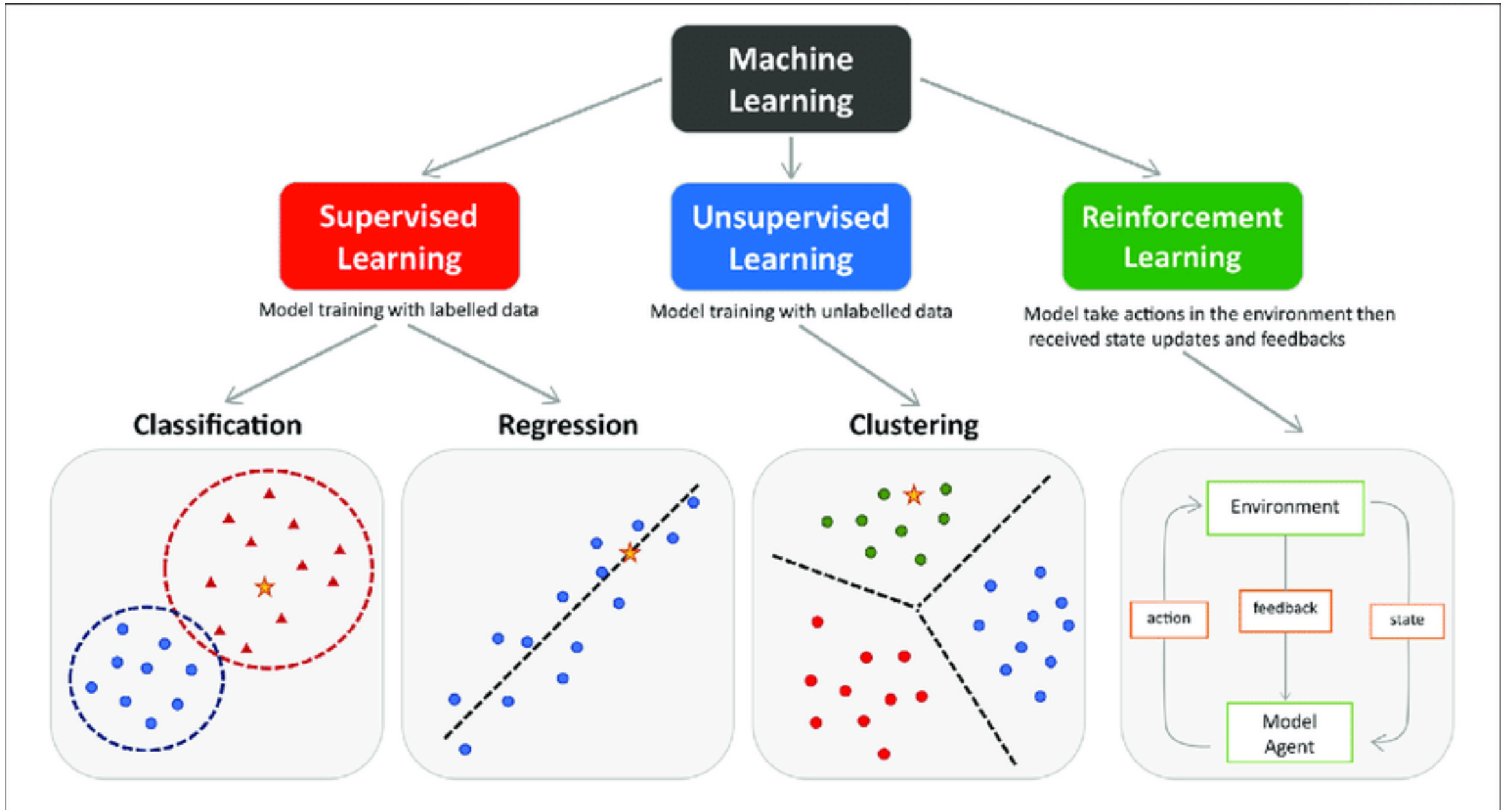
gartner.com

Note: Range measures number of years it will take the technology/trend to cross over from early adopter to early majority adoption. Mass indicates how substantial the impact of the technology or trend will be on existing products and markets.

Source: Gartner

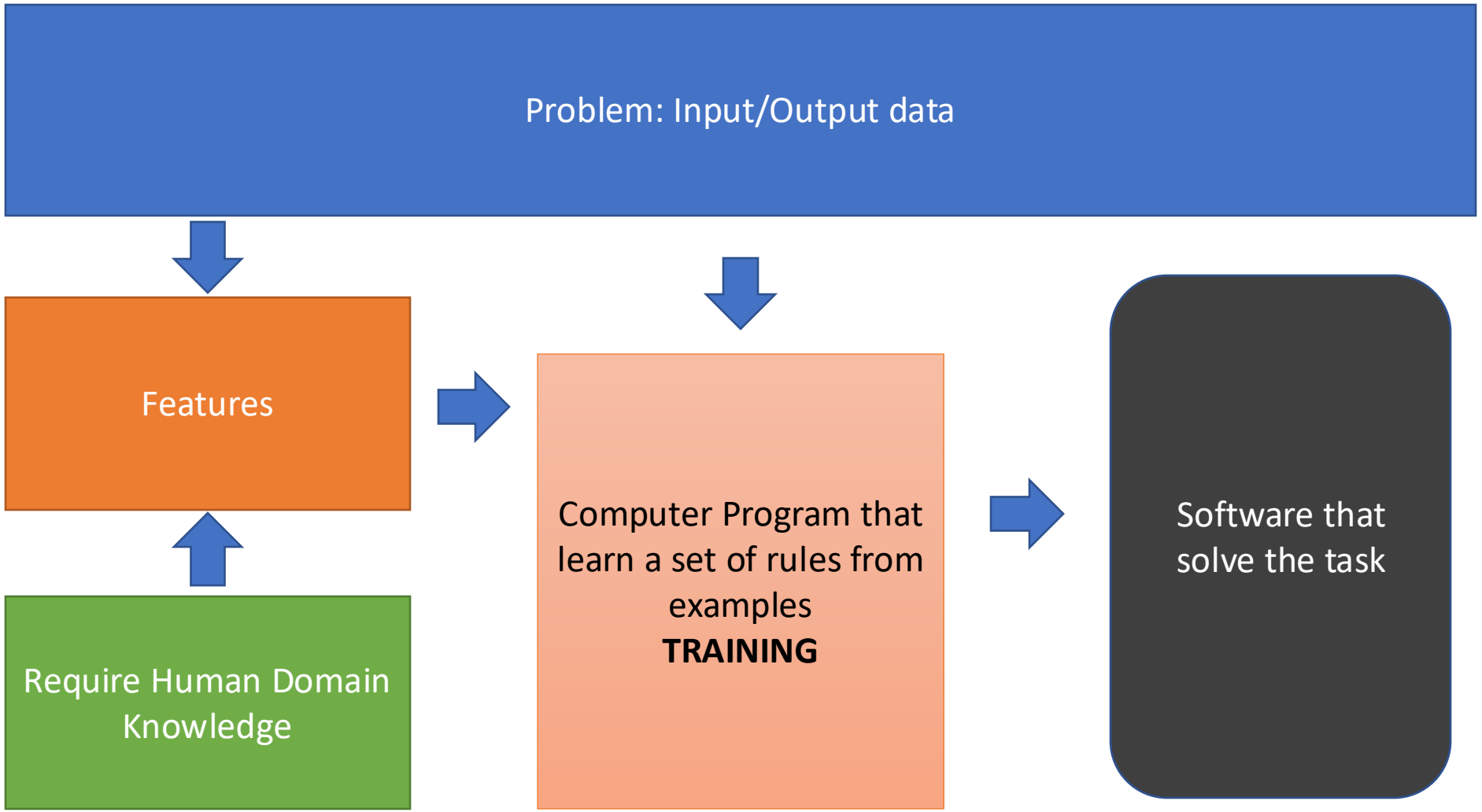
Gartner

What is learning and a learning machine

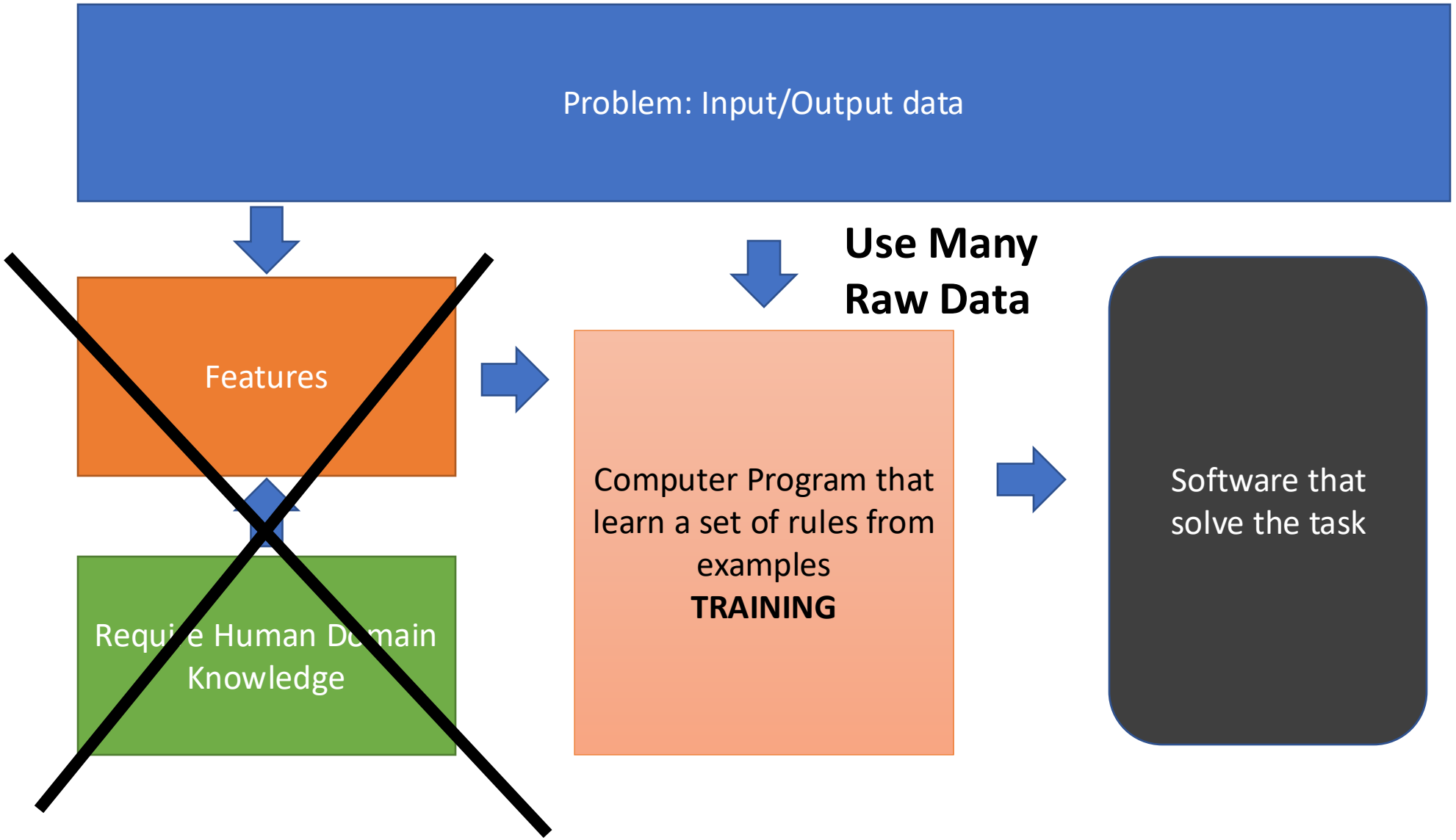

Machine Learning
Supervised Learning
Unsupervised Learning
Reinforcement Learning
Model training with labelled data
Model training with unlabelled data
Model take actions in the environment then received state updates and feedbacks
Classification
Regression
Clustering
Environment
action
feedback
state
Model Agent

General structure of a learning model

Learning Pipeline



Innovation of Deep Learning



Deep Networks For:

Numerical Data -> Deep Neural Network

Applications: Production management, Prediction, Controls and Robotics

Multimedia Data-> Convolutional Network/ViT

Applications: Image and Video classification, Face recognition, Licence Plate Detection, OCRs..

Time series/Text -> Recurrent Neural Network/Transformer

Applications: Financial Analysis, Audio and Speech analysis, Text analysis and traslation, Forecasting

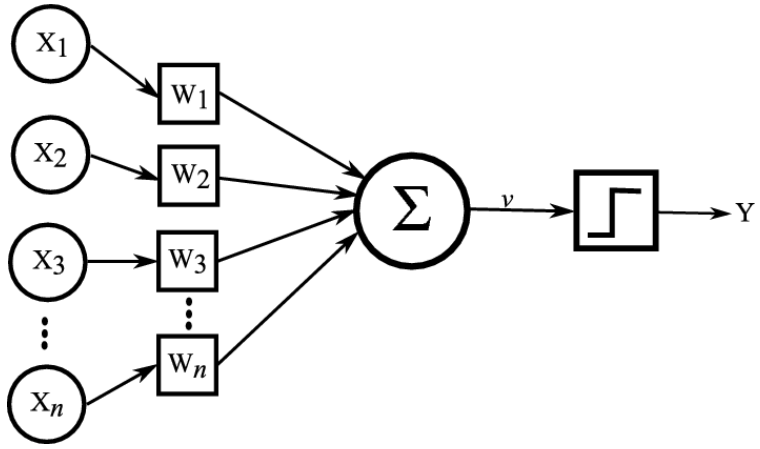
Artificial Neuron

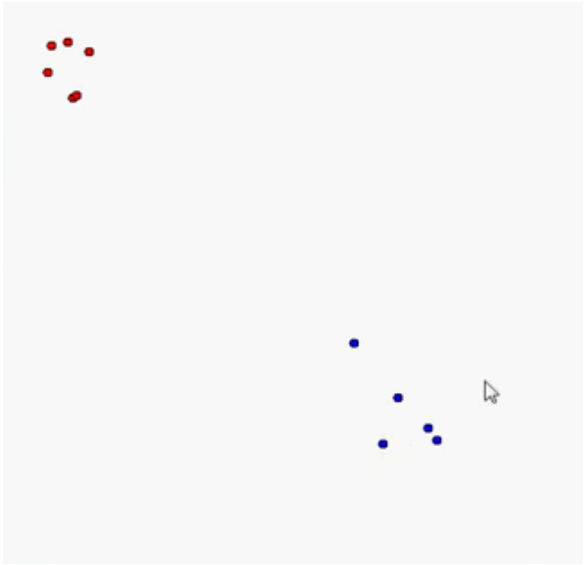
- An **Artificial Neuron** is a simple mathematical function that receives one or more inputs and yields a **real-valued** output.

- Its output is produced by
  - **Taking a weighted sum** of the input data according to a set of **parameters** called **weights**
  - Processing the result of the previous operation through a **non linear activation function**

$$\hat{y} = \sum_i w_i x_i$$

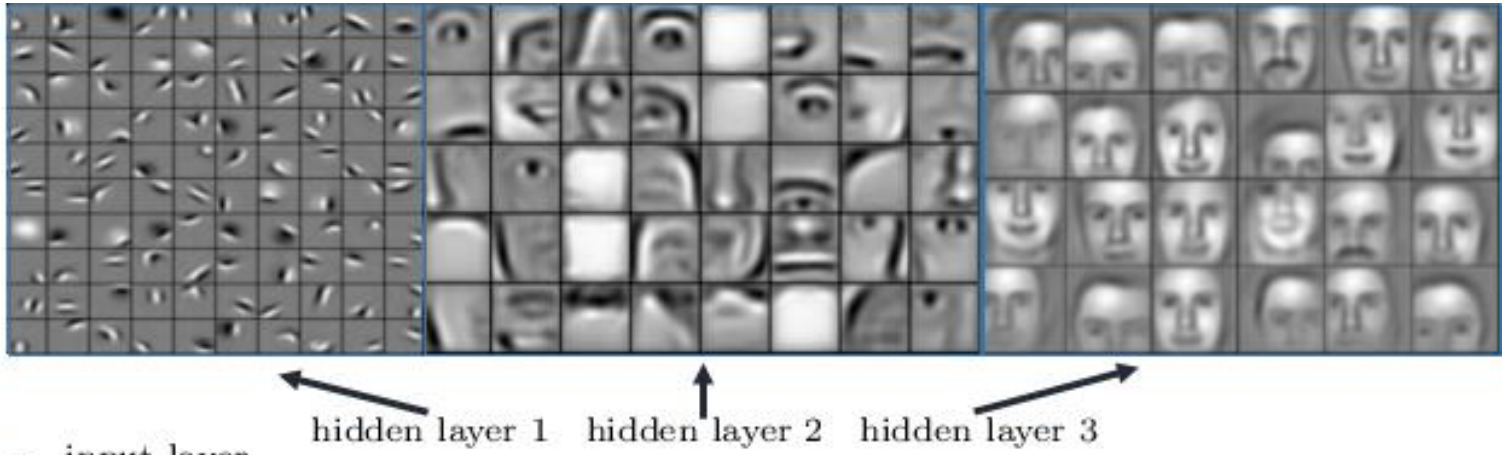
Perceptron (1958)

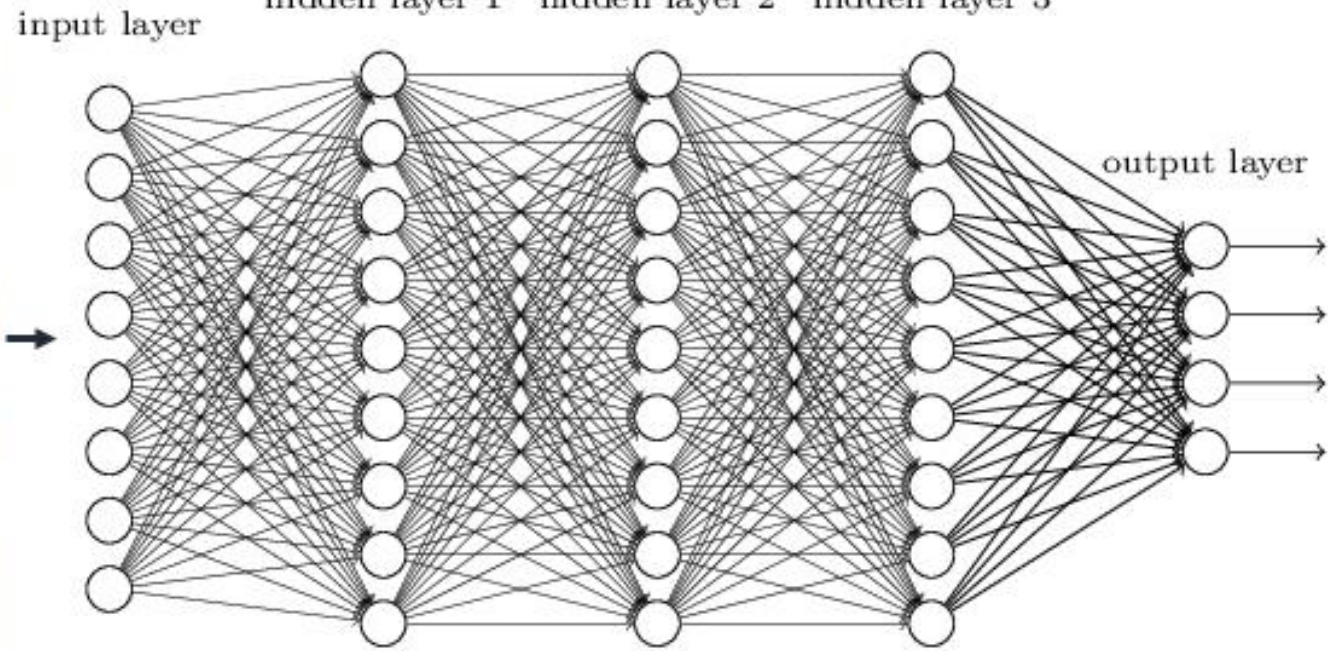




Going Deep

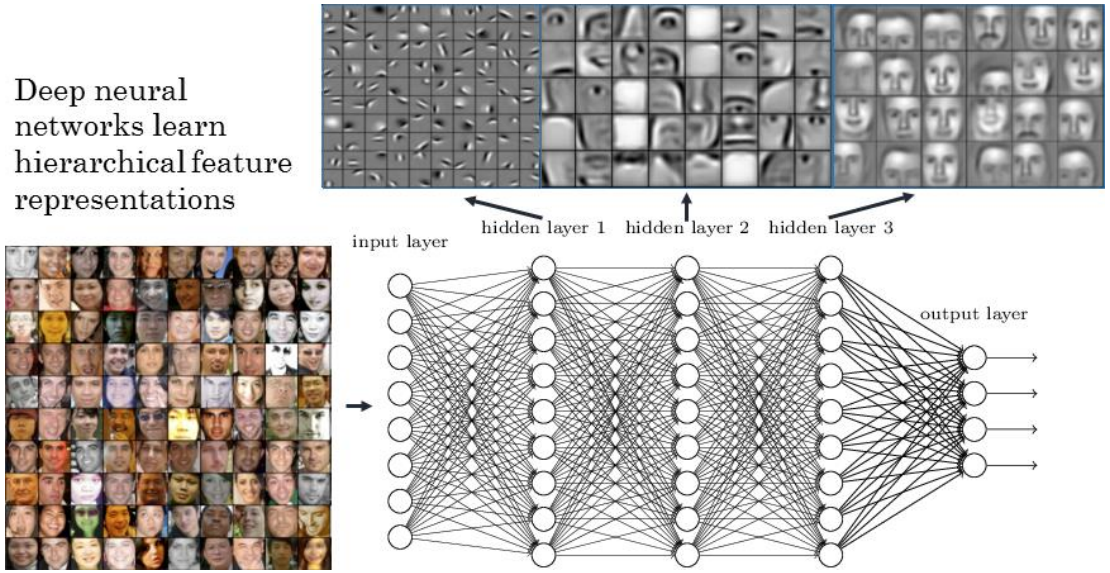
Deep neural networks learn hierarchical feature representations



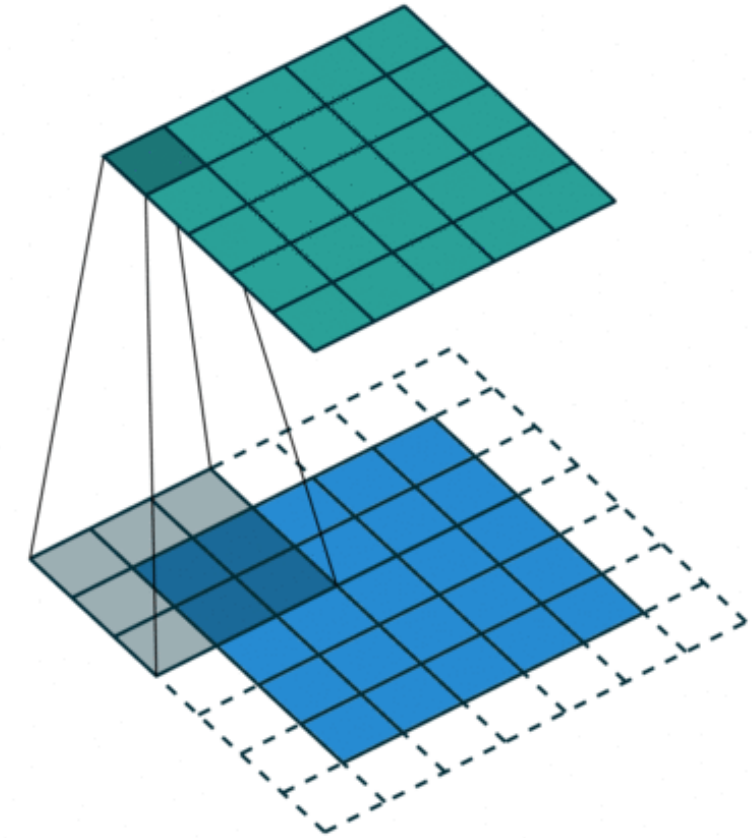


Convolution

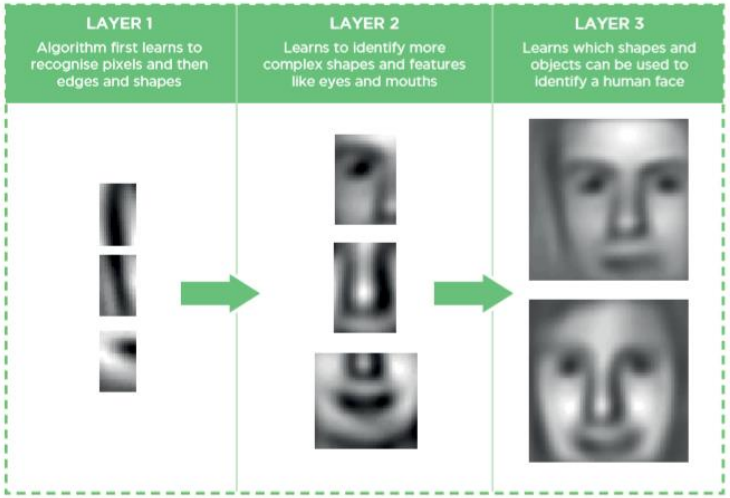
- Custom neurons have been proposed to deal with different data.
- Convolution based unit for images
  - Convolutional neural networks
  - Based on study 1960's about visual cortex
    - Different cells with different receptive fields



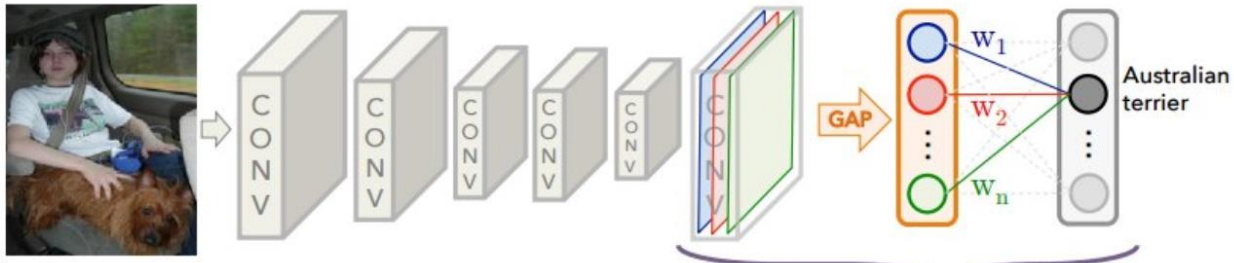
CNN (LeCunn 1989)

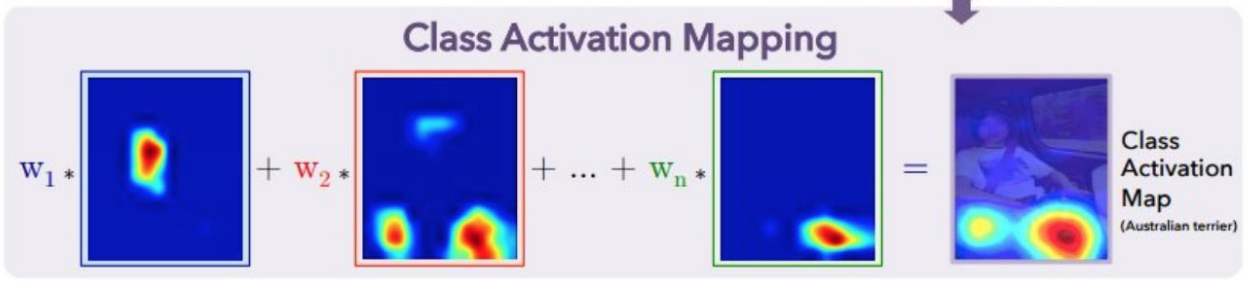


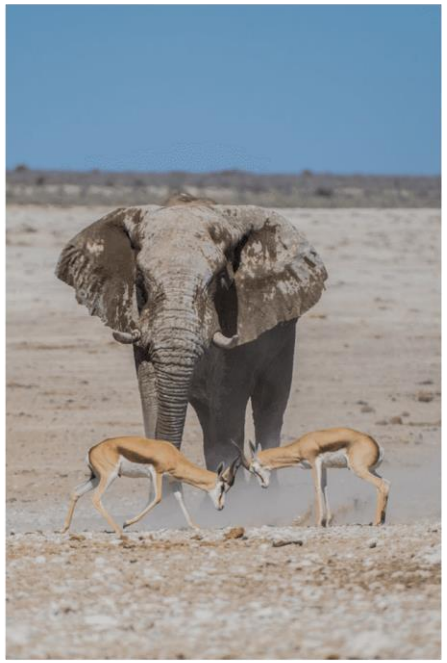
How the network learn the world

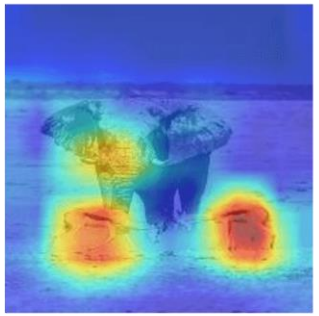


- Inner Representation is compositional
- Semantic representation must be inspected with additional rechniques e.g. Class Activation Maps

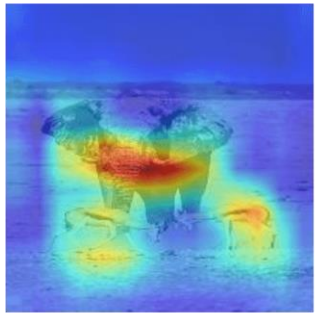




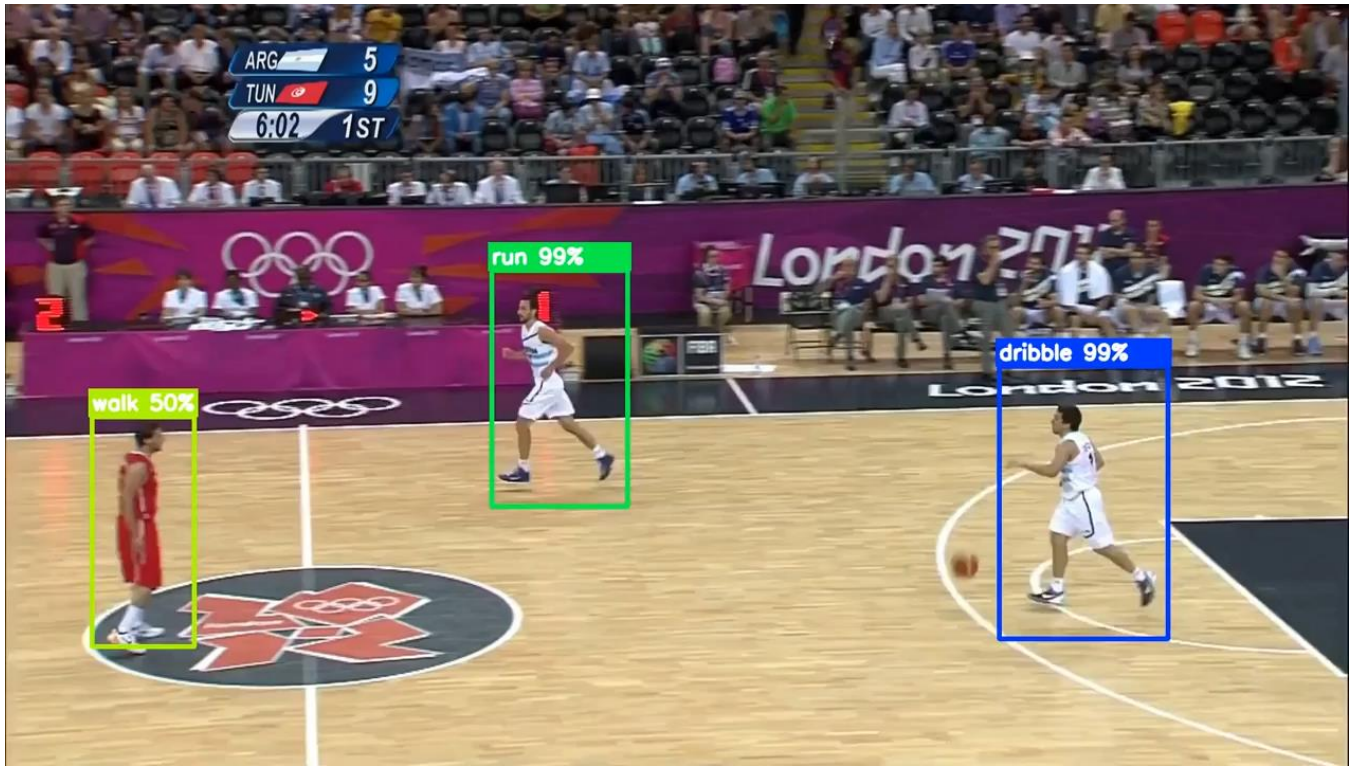


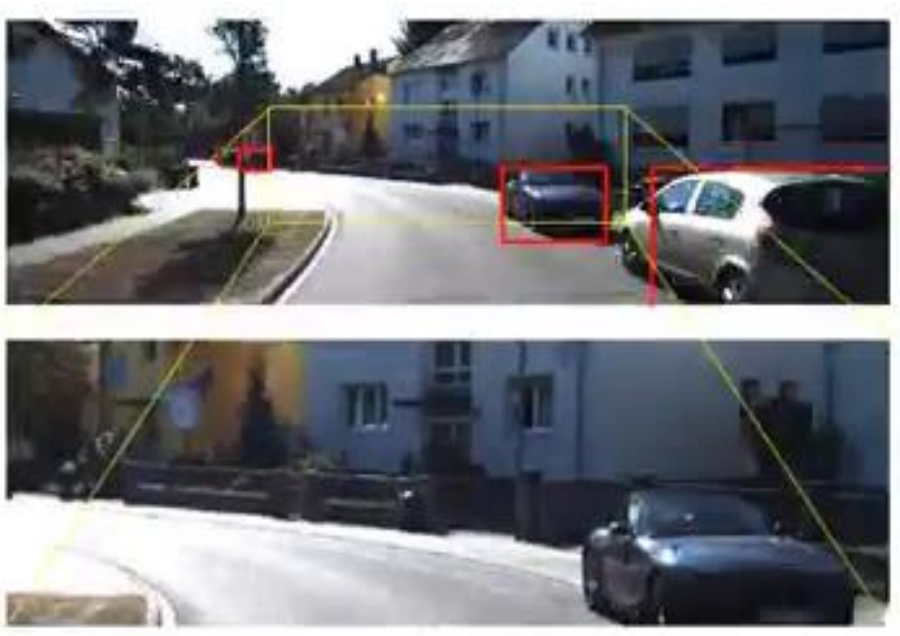


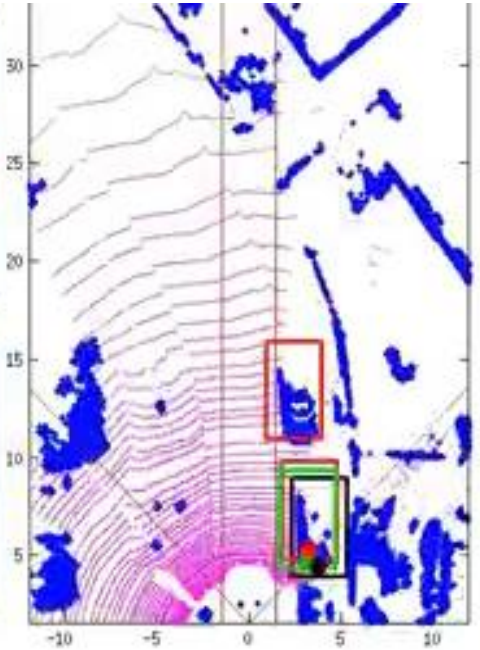
gazelle



elephant

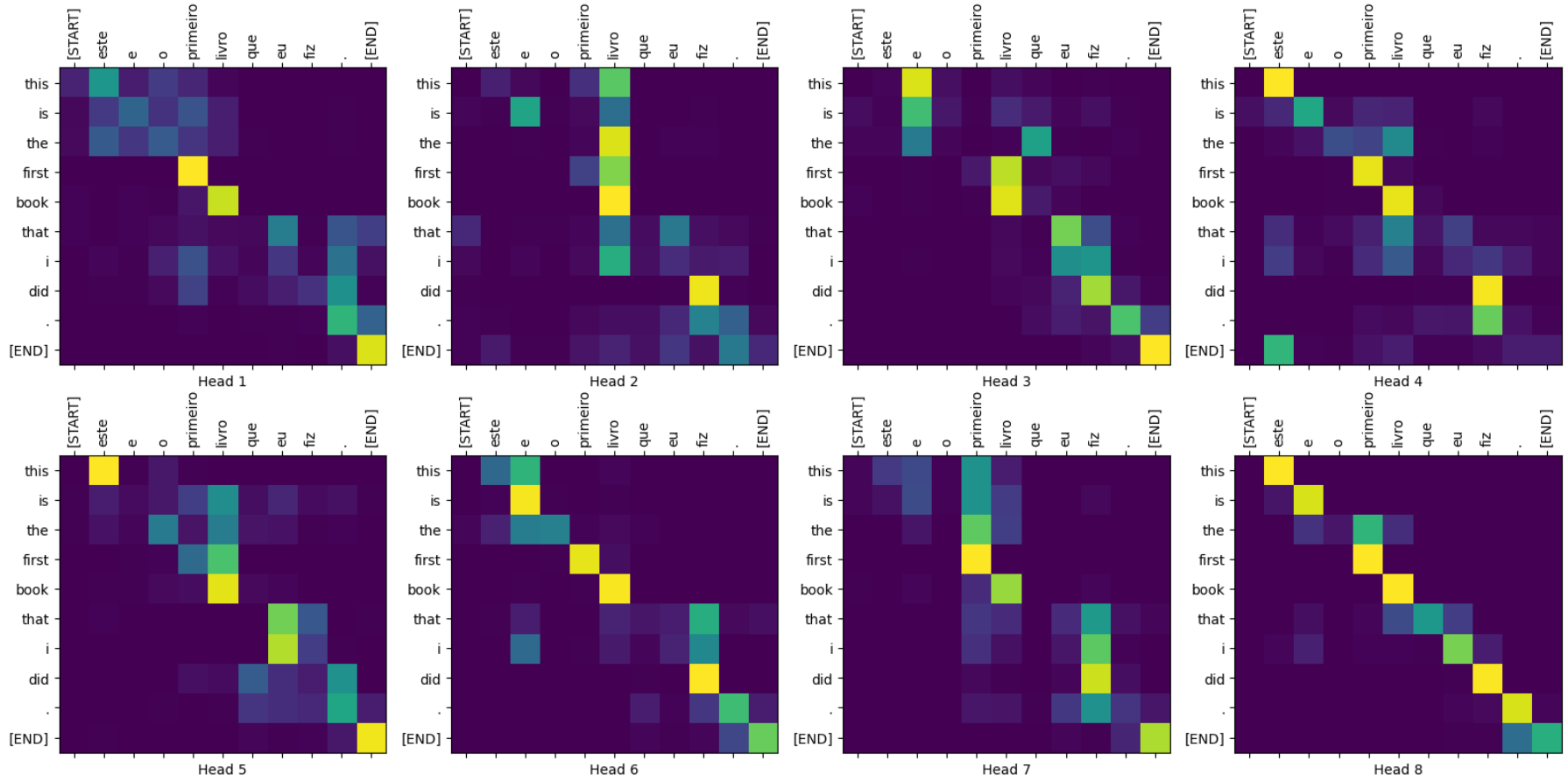

ARG 5
TUN 9
6:02 1ST
walk 50%
run 99%
dribble 99%
London 2012




30
25
20
15
10
5
-10
-5
0
5
10

Attention (Trasformer)

- Mos
  Lea
- Idea
  -
- Kille
  -
  -
  -



Vaswani 2017)

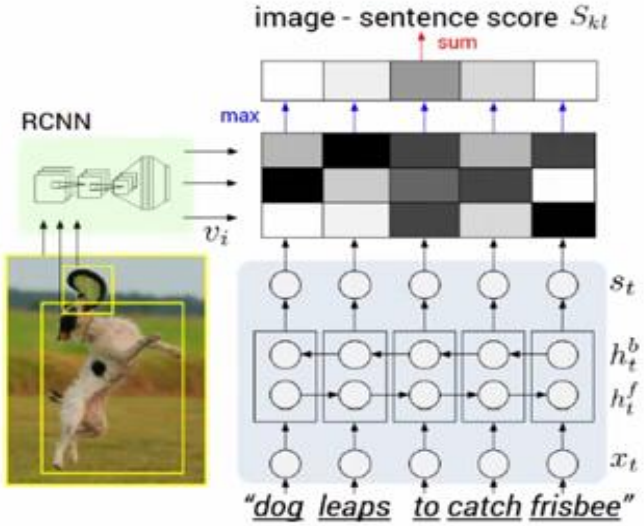
Source: Andrej Karpathy

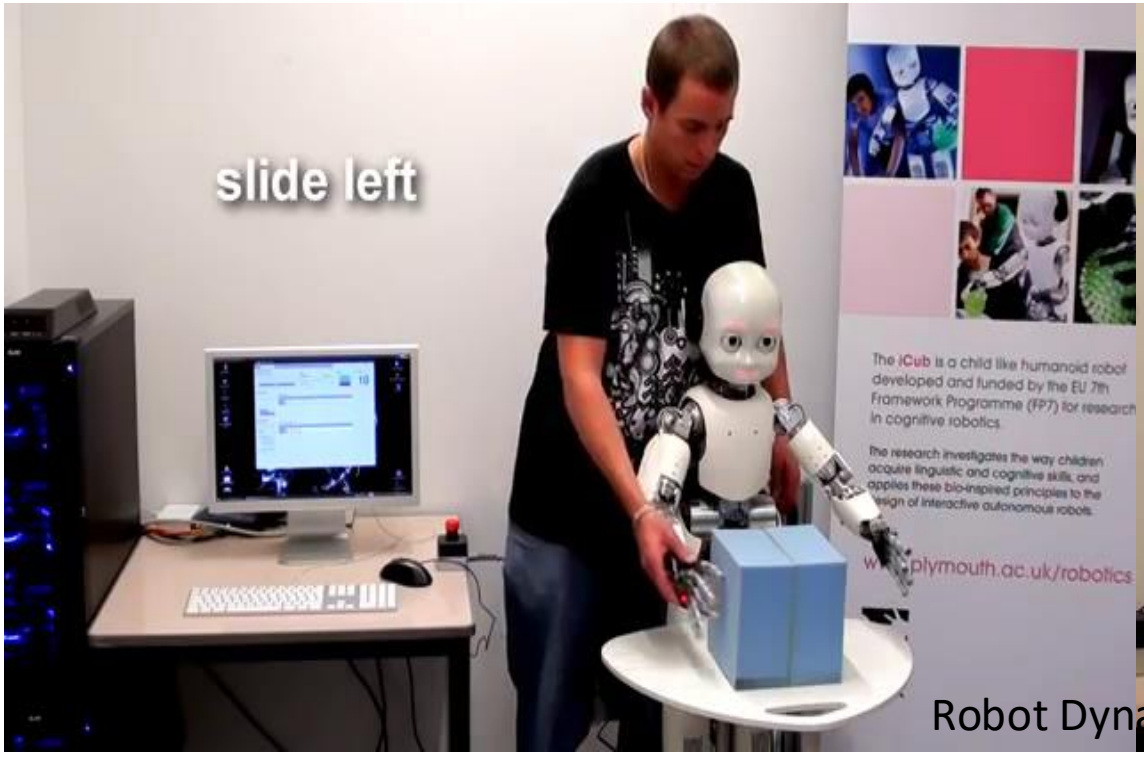
Text and Music Writing

```
tyntd-iafhatawiaoihrdemot  lytdws  e ,tfti, astai f ogoh eoase rranbyne 'nhthnee e
plia tklrgd t o idoe ns,smtt   h ne etie h,hregtrs nigtike,aoaenns lng
```

```
"Tmont thithey" fomesscerliund
Keushey. Thom here
sheulke, anmerenith ol sivh I lalterthend Bleipile shuwy fil on aseterlome
coaniogennc Phe lism thond hon at. MeiDimorotion in ther thize."
```

Source: [Karpathy and Fei-Fei 2015]

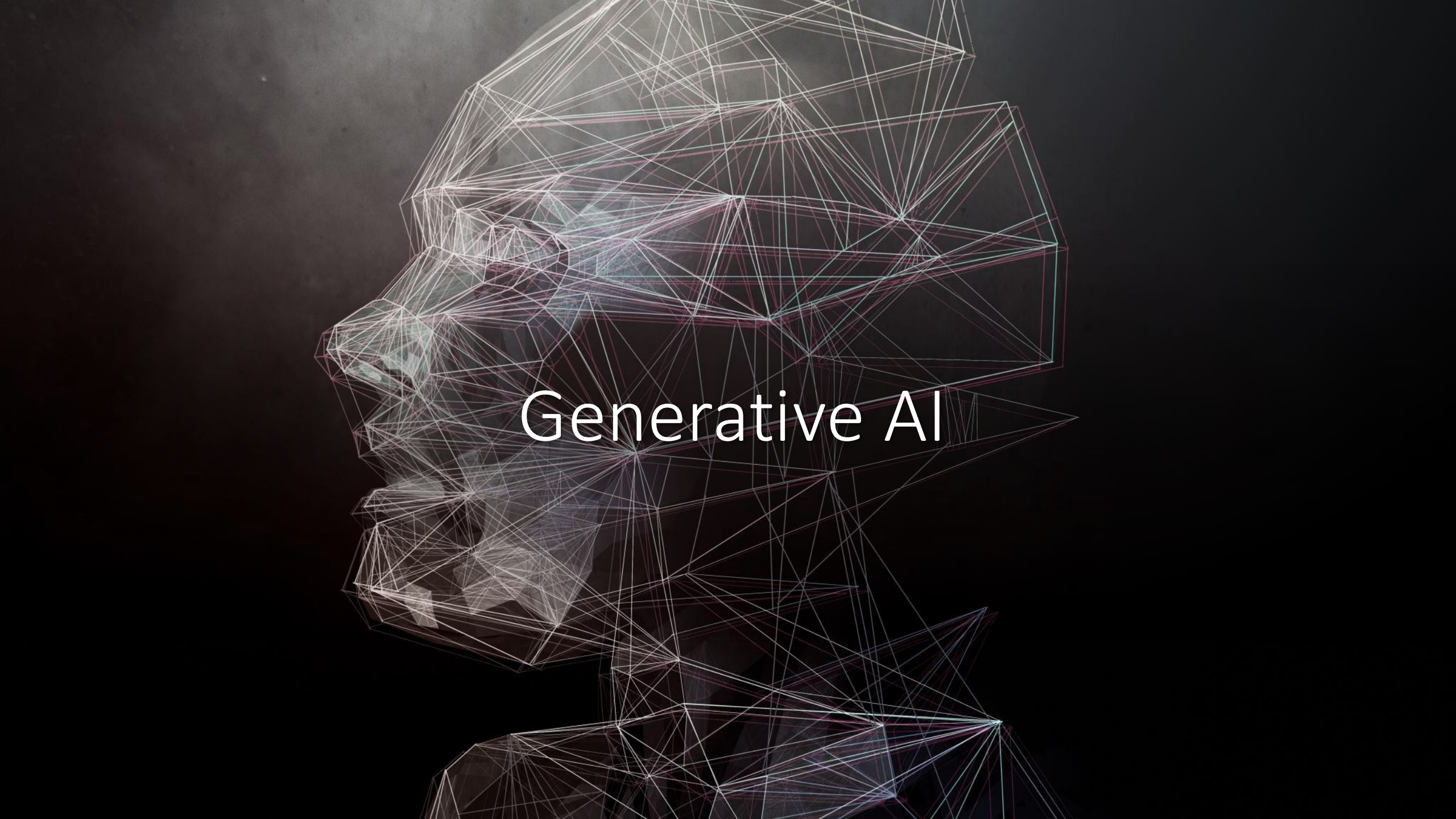




Robot Dynamics Control

New trajectories 2023 - 2030

- Generative Models
- Continual inremental e lifelong learning
- AI regulation and Trusthworthy AI


Generative AI

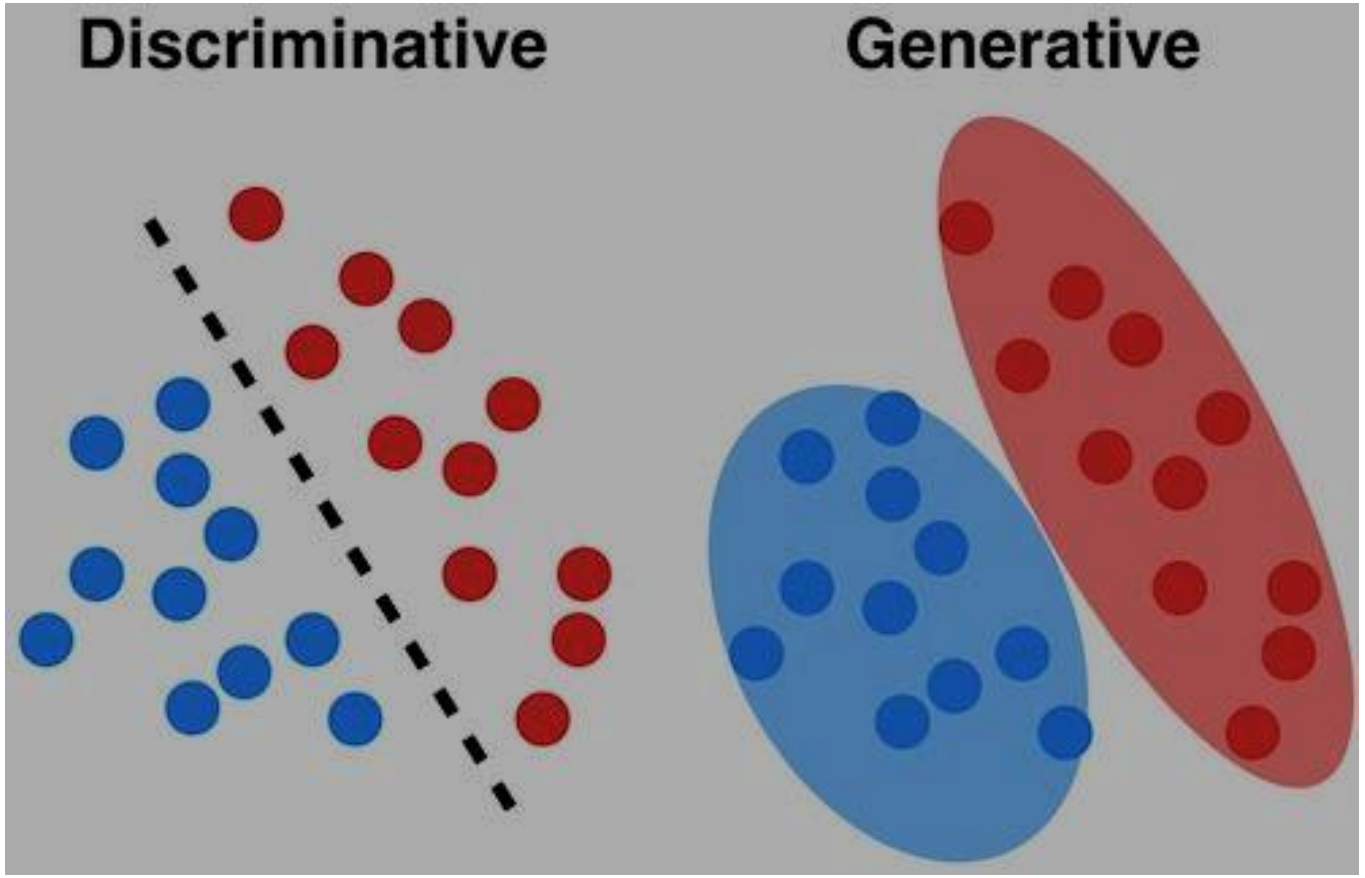
Dear Hiring Manager,

I am excited to apply for the position of Retail Sales Associate at All Saints. With one year of experience in the retail industry, I am confident in my ability to provide exceptional customer service, handle cash transactions accurately, and manage my time effectively.

During my time at my previous employer, I was responsible for handling cash transactions for an average of 50 customers per day with a 100% accuracy rate. I am confident in my ability to handle financial responsibilities and am excited to bring this experience to the team at All Saints.

I am passionate about the retail industry and enjoy creating a positive shopping experience for customers. I pride myself on being able to build a rapport with customers and anticipate their needs. In my previous role, I increased sales by 15% by providing exceptional customer service and building lasting relationships with customers.

I am excited to contribute to the All Saints team by utilizing my key skills in customer service, cash handling, and time management. I am confident that my passion for the



- Discriminative model:
  Learn by comparing differences and distilling Knowledge
- Generative Model: Learn by creating a sophisticated parametric model of the world from which we can sample

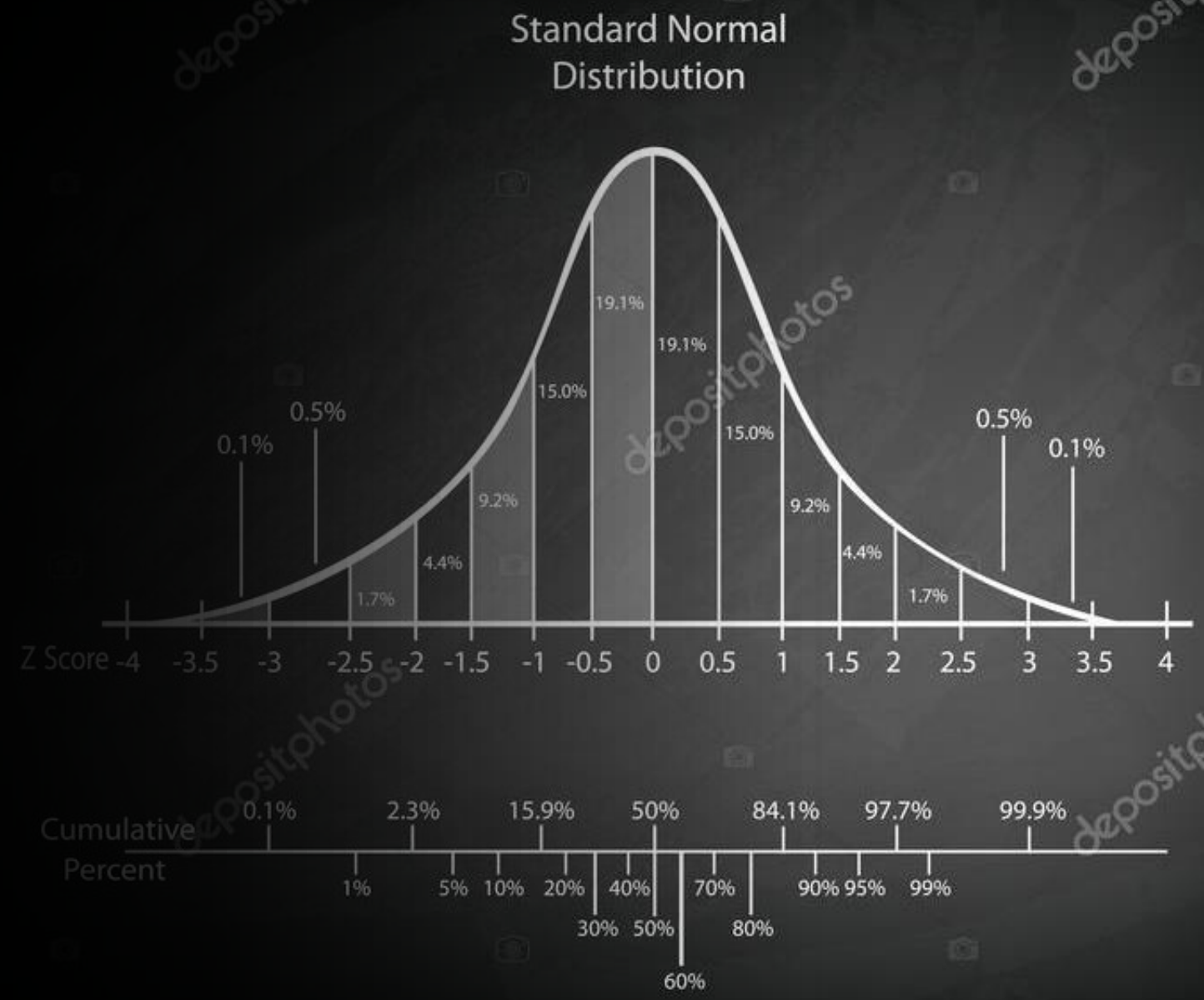
$$P(Y|X)=P(X|Y)P(Y)$$

The most famous Generative AI Model

$$N(x|\mu,\sigma)$$

$$\mu = \frac{\sum_{i=1}^{N} X\_i}{N}$$

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$



GPT3 (Text)

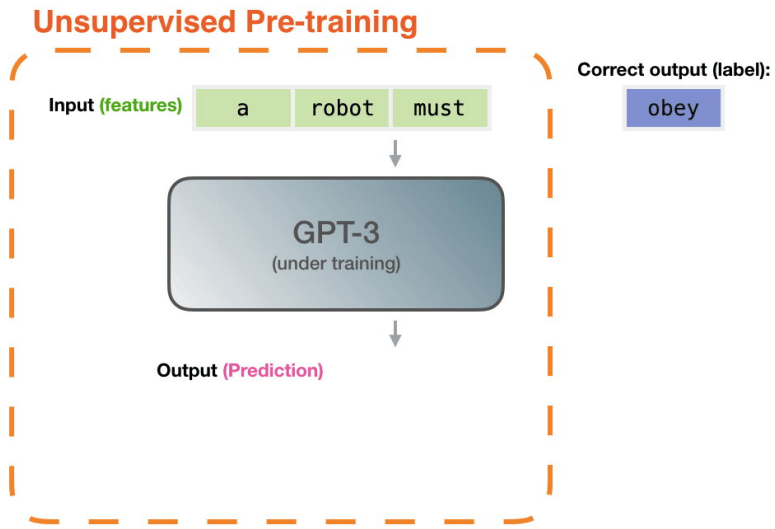
- Huge Trasformer decoder
  - Process one word at a time and predict the next (fully autoregressive)

175 Billions parameters

Text: Second Law of Robotics: A robot must obey the orders given it by human beings

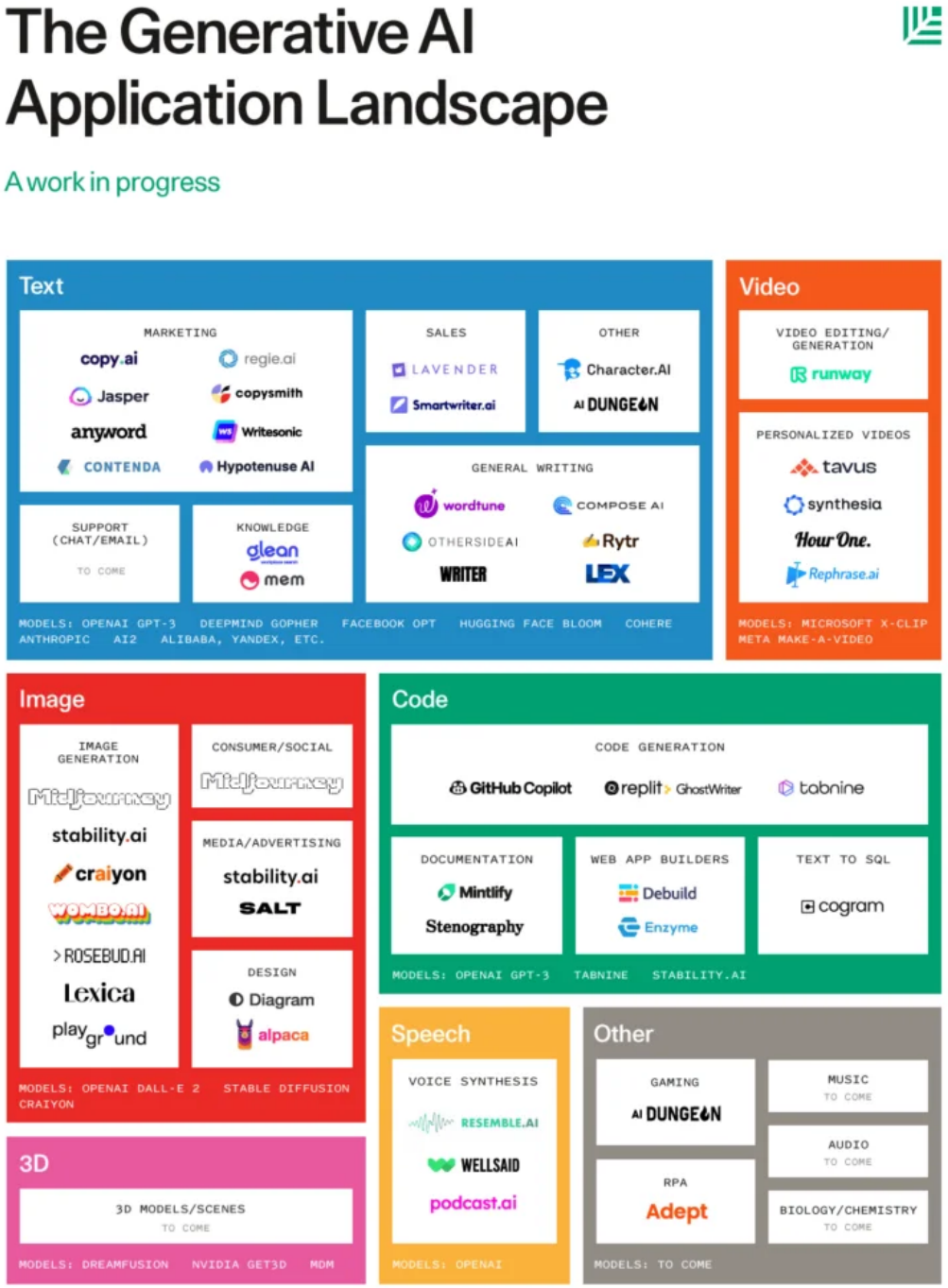
Generated training examples

Example #	Input (features)	Correct output (labels)
1	Second law of robotics :	a
2	Second law of robotics : a	robot
3	Second law of robotics : a robot	must
...		



- >500 companies based on generative AI applications (worldwide in the last 6 months)
- Millions of customized use of generative models (model structures resembles transformer encoder-decoder structure)

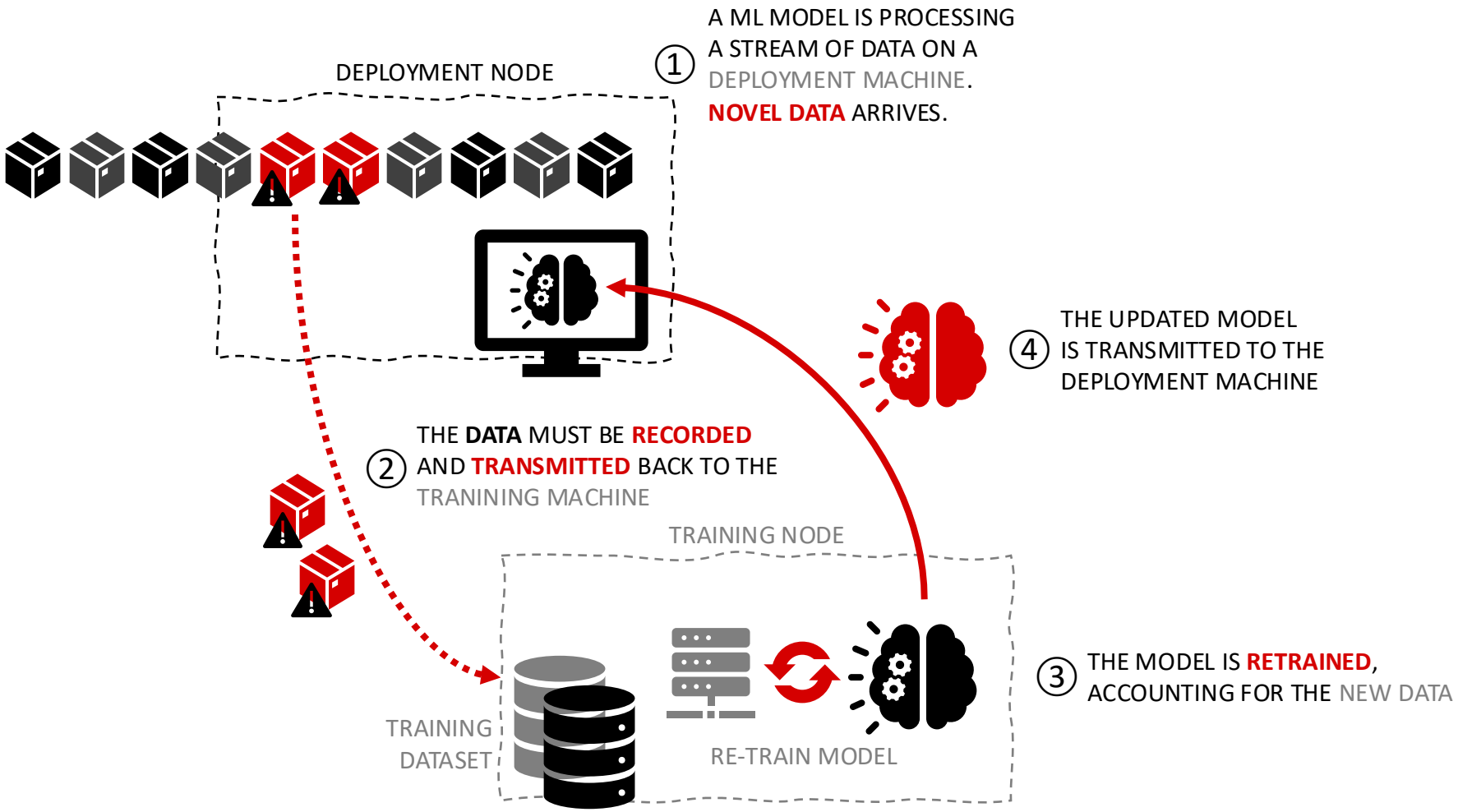
- Fostering the scale up of deep learning models (GPT4 1 trillion parameters GPT 2 1 billion 1 year ago)



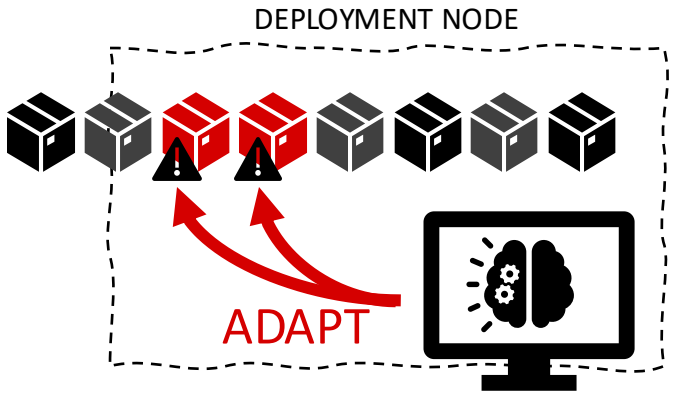

Continual and lifelong AI learning

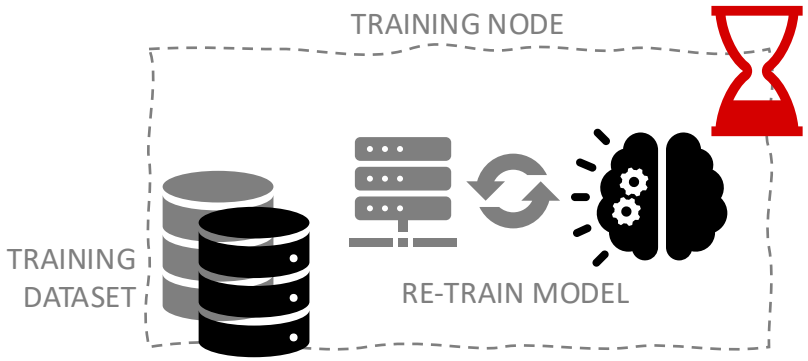
A Typical MLOPS setup

Without **CL**, when the domain data changes, we need to **retrain** and **deploy** the ML model.



CL Scenario I: Drift Prevention





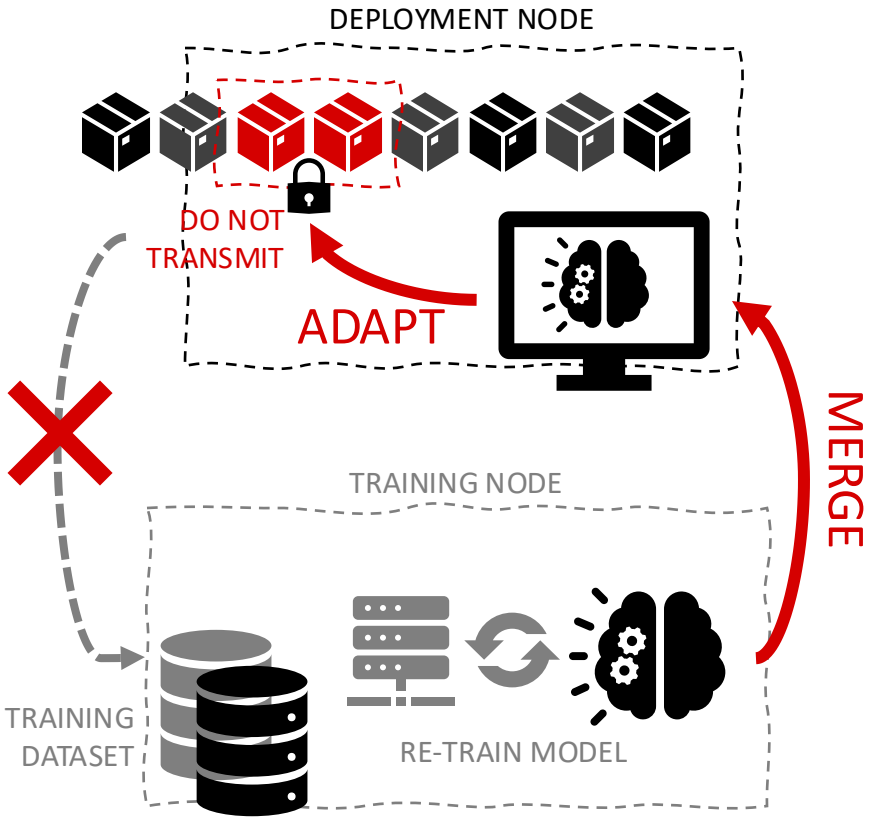
While the model is being re-trained, the deployment model is **incapable of handling new data**.

Its response on new data will either be **delayed** or **unreliable**.

CL methods allow for a **plastic behavior**, making the performance degrade gracefully until the update.

After re-training, the re-trained model **replaces** the continual learner on the deployment node.

CL Scenario II: Decoupled ADAPTATION



Some novel stream data-points might not be transmitted due to **security constraints**.

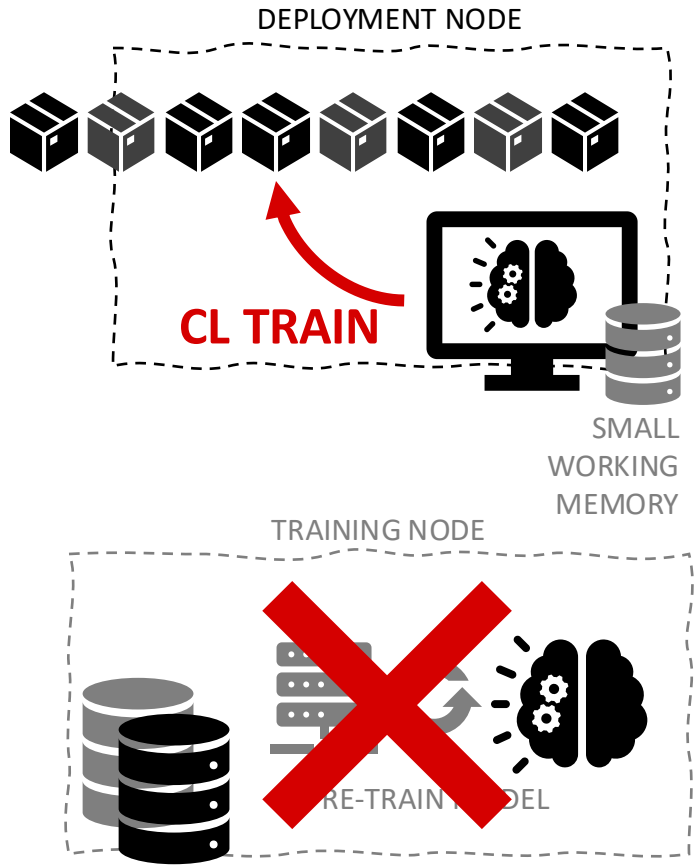
CL methods might still learn from them **without them leaving** the deployment node.

When the re-trained model (unaware of secure data) is available, the CL learner cannot be replaced.

Our recently proposed combined CL-transfer learning[3] approach allows for knowledge merging.

3. Boschini, Matteo, Lorenzo Bonicelli, Angelo Porrello, Giovanni Bellitto, Matteo Pennisi, Simone Palazzo, Concetto Spampinato, and Simone Calderara. Transfer without Forgetting. ECCV 2022.

CL SCENARIO III: Autonomous On-The-Edge Learning



Ideally, a deployed model with a strong CL algorithm does not need a separate re-training phase.

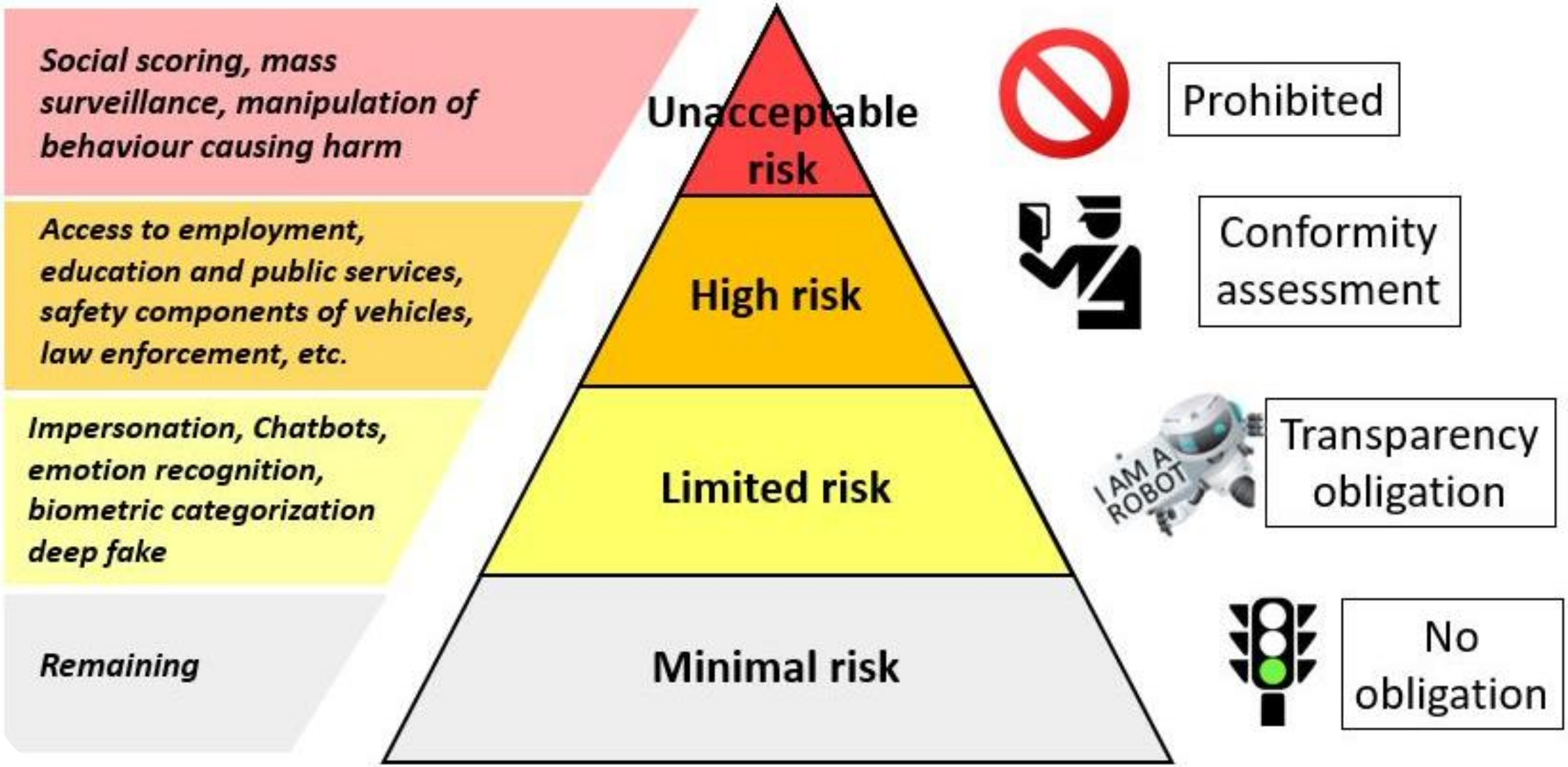
Instead of relying on a comprehensive database and training from scratch, the model might learns incrementally from live data and a small working memory.

We have a strong expertise in training memory-based accurate and unbiased continual learning models[4,5].

4. Buzzega, Pietro, Matteo Boschini, Angelo Porrello, and Simone Calderara. Rethinking experience replay: a bag of tricks for continual learning . ICPR 2020.
5. Buzzega, Pietro, Matteo Boschini, Angelo Porrello, Davide Abati, and Simone Calderara. Dark experience for general continual learning. NeurIPS 2020

Trustworthy AI

EU Artificial Intelligence Act: Risk levels

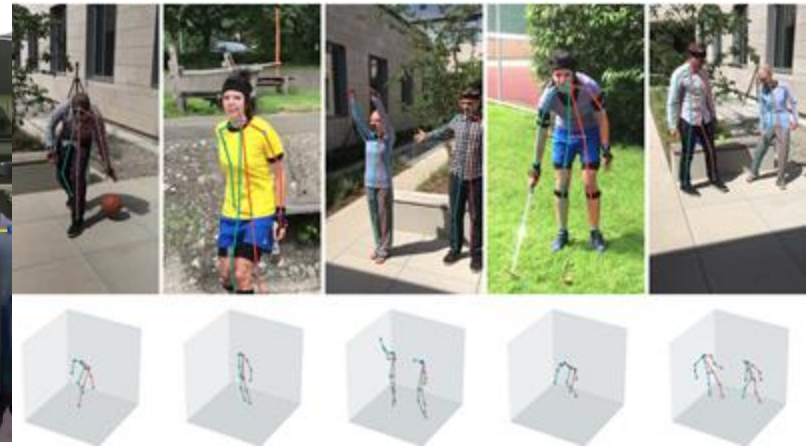


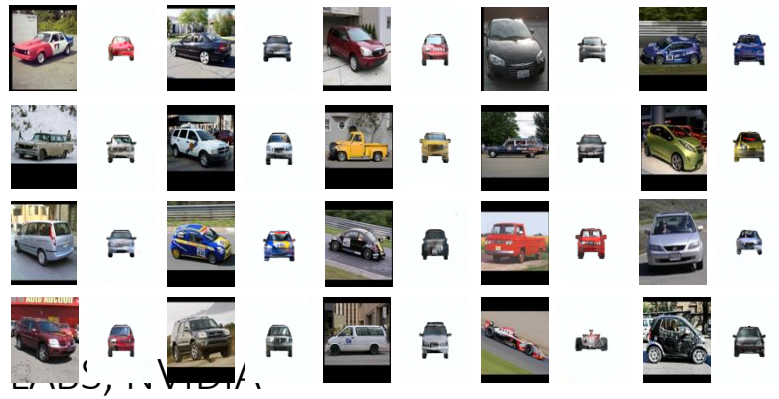
Applications made in ER

https://aimagelab.ing.unimore.it/

ML and CV for People Analysis in Smart Urban Environment

- **Study Deep Learning techniques** for:
- People tracking
- People detection 2D and 3D
- Human Behavior understanding
- Anomaly Detection
- Vehicles-human interaction
- Geometric view synthesis
- Conferences and Journals:
- CVPR, ICCV, TPAMI, TIP, TMM





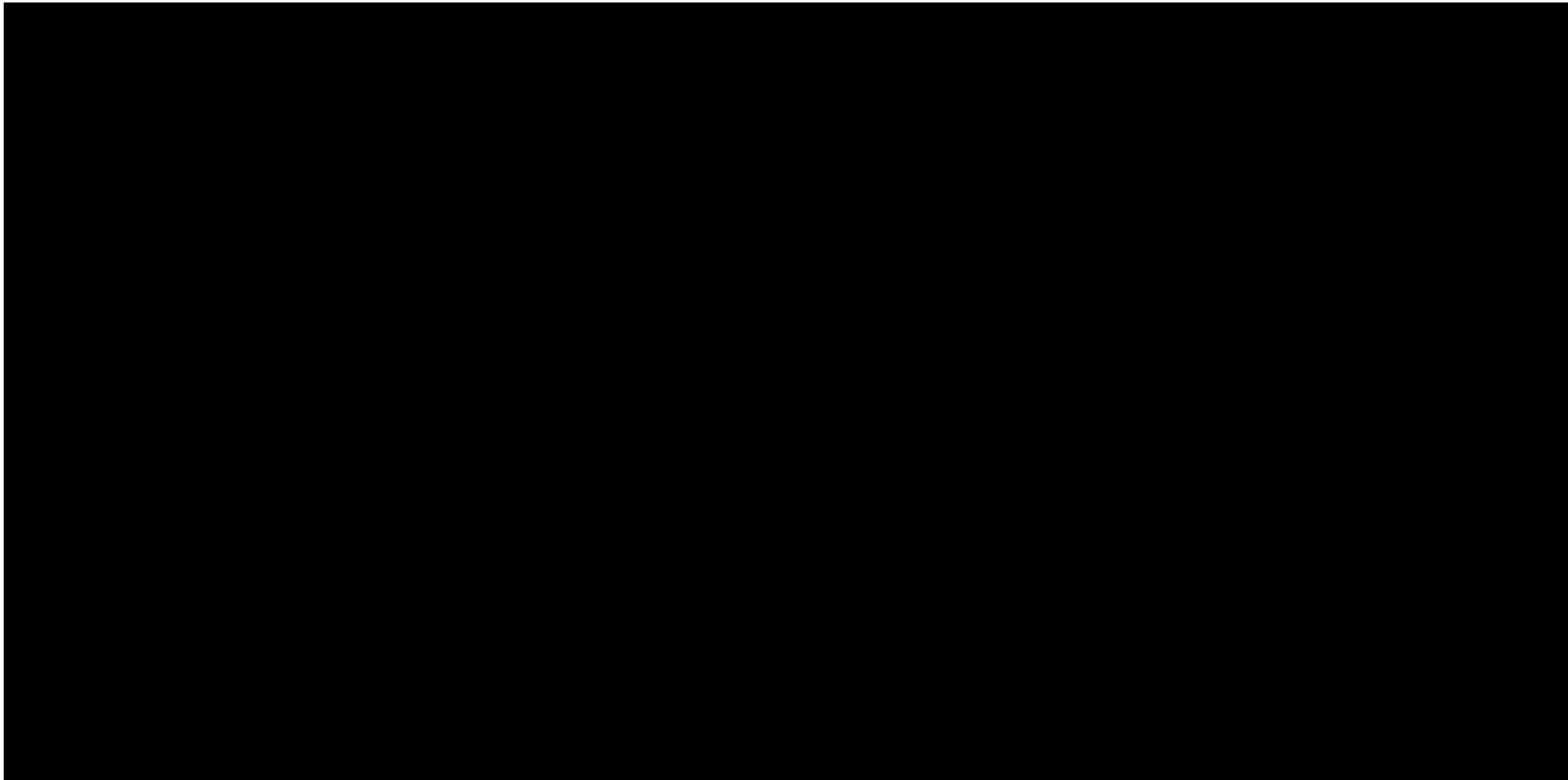
- Projects and Collaborations:
- PRIN COSMOS and PREVUE, EU PRYSTINE, E [illegible] LABS, NVIDIA

- AImageLab Group: Rita Cucchiara, Roberto Vezzani, Simone Calderara
- http://aimagelab.ing.unimore.it

AImage Lab
University of Modena and Reggio Emilia

AI academy

Car perception and analysis



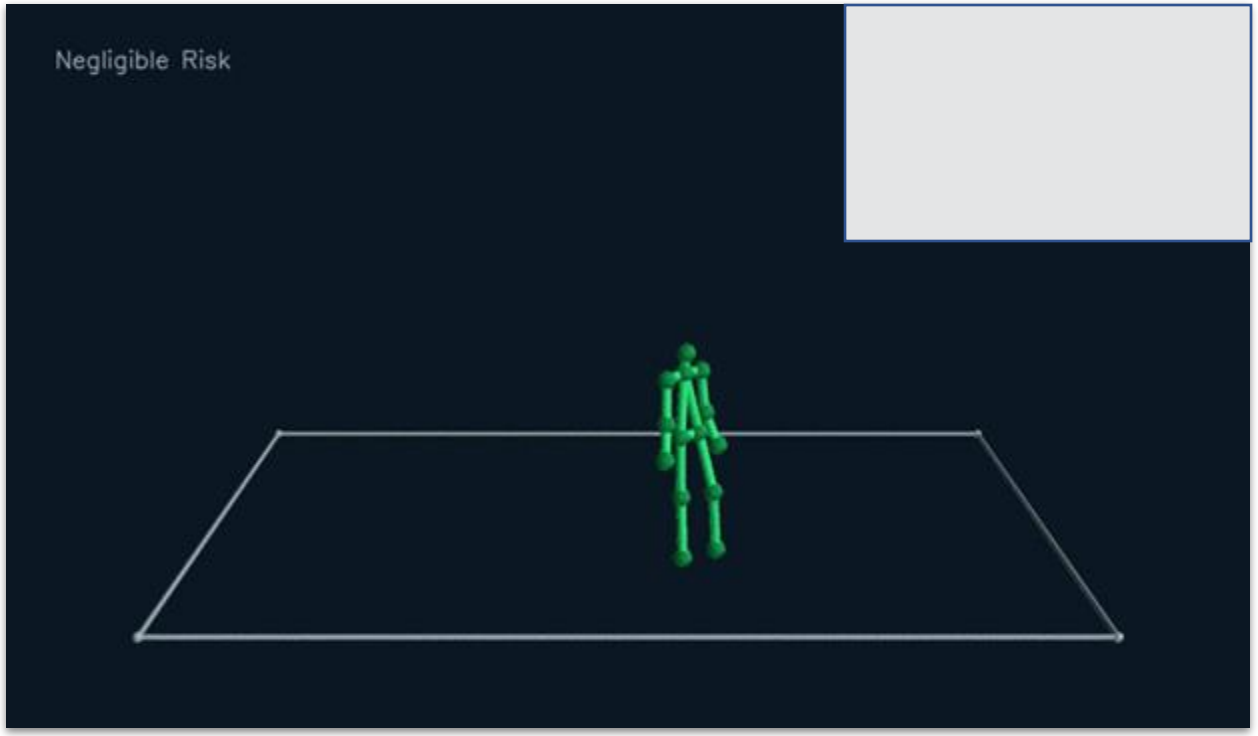
AI academy

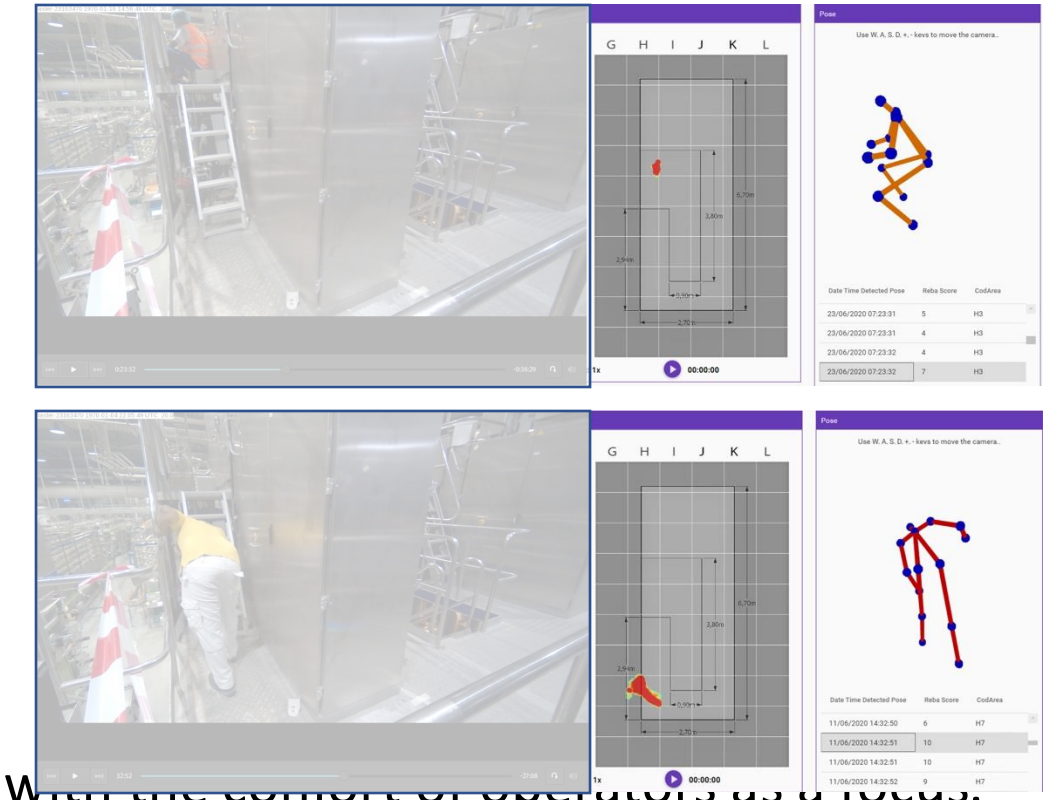
EyeCandy Results

Learning human in industrial environment

Can virtual data generalize to different environments?

Can we understand how people work and interact with machines?



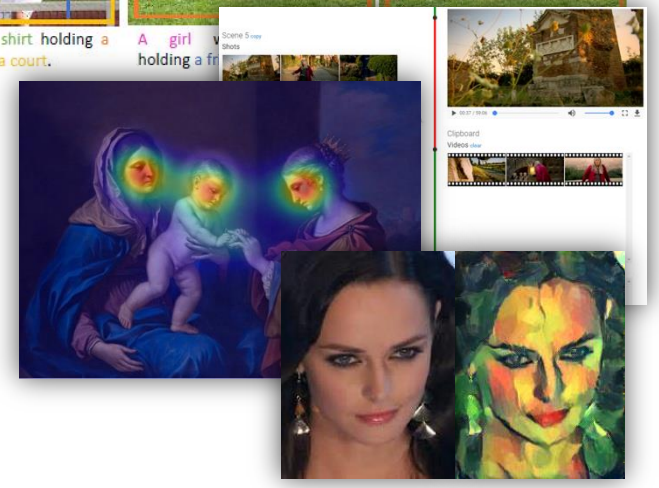


Machines of the future can reconfigure themselves with the comfort of operators as a focus.

Activity in conjunction with Tetra Pak (Automation & Digital) (Modena Lab) Emilia Romagna

Embodied AI and Digital Humanities

- **Embodied AI**: Integration between Vision, Language and Action
  - Automatic description of Images and Video
  - Natural Language and multi-modal retrieval
  - Vision and Language Navigation
  - Navigation of embodied agents in unseen environments
- Applications in Cultural Heritage and Digital Humanities



- Conferences and Journals:
- CVPR, ICCV, TPAMI, TIP, TMM

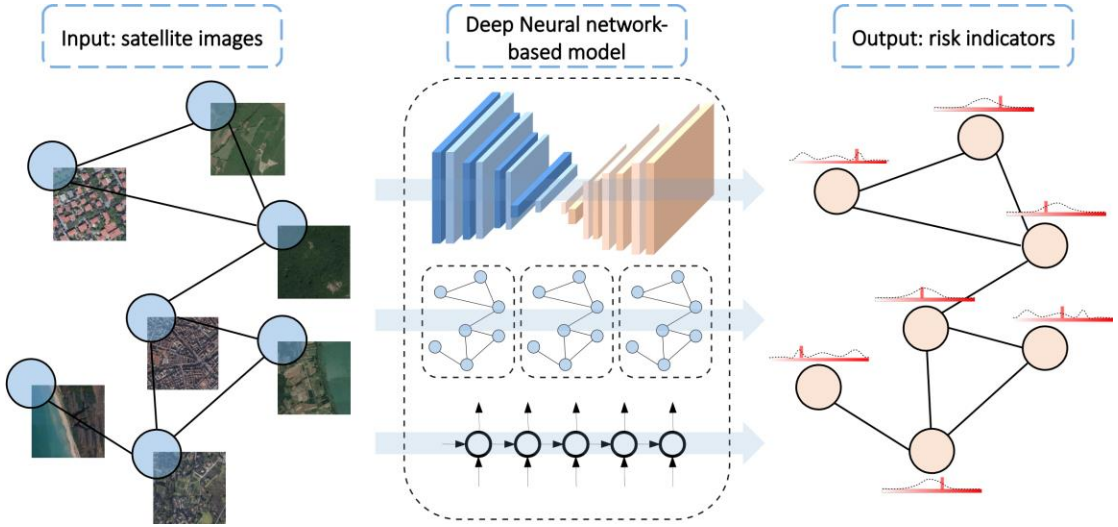
- Projects and Collaborations:
  - IDEHA, CULTMEDIA, AI4CH, AI4DH
  - Facebook AI Research, NVIDIA, University of Haifa (Israel)

- AImageLab Group: Rita Cucchiara, Lorenzo Baraldi, Marcella Cornia
- http://aimagelab.ing.unimore.it

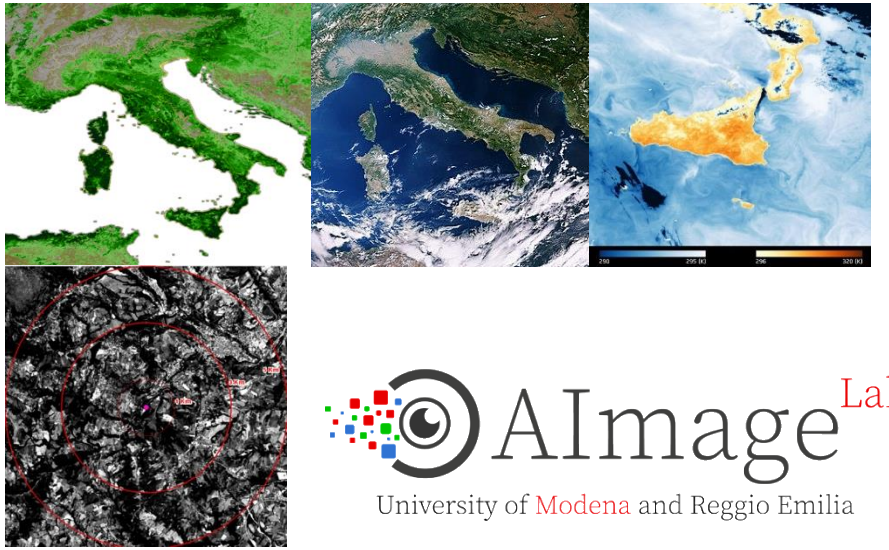
AImage Lab
University of Modena and Reggio Emilia

Machine learning for Earth Observation

- Deep Learning and Graph based analysis for:
- Satellite Images self-supervised feature extraction
- Inference of physical phenomena from EO
- Epidemic and vectors analysis using temporal EO



- **Projects and Collaborations:**
- AI4VECT Italian Ministry of Health,
  AIDEO European Spatial Agency



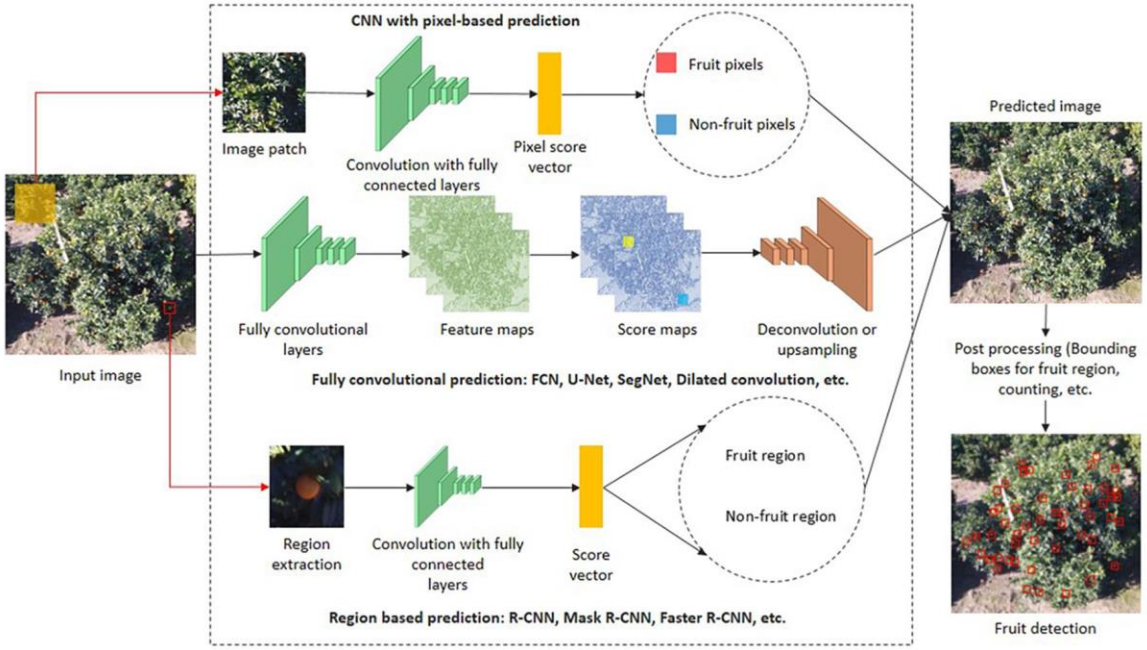
AImageLab Group: Simone Calderara, Angelo Porrello
http://aimagelab.ing.unimore.it

AImage Lab
University of Modena and Reggio Emilia

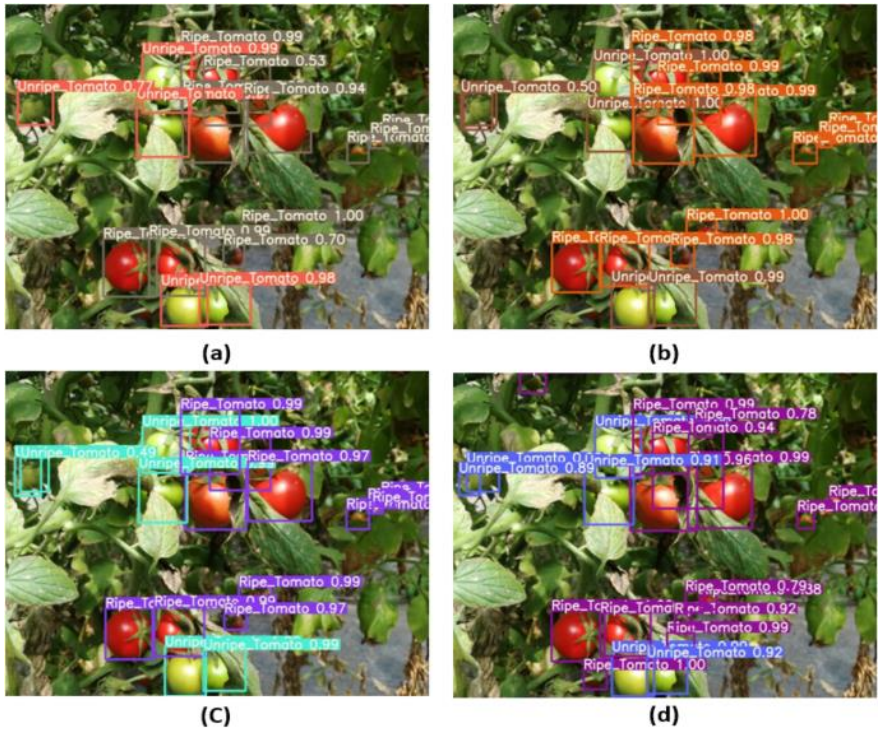
AI 4 Harvesting

- Detect and locate fruits or vegetable for harvesting
  - Exploit recent convolutional network for detection (e.g. YOLO, SSD, RCNN)
  - Exploit precise neural network for image segmentation (e.g. Fully Convolutional architecture RCNN, Diffusion Models, zero shot grounding DINO)

Maheswari, Prabhakar, et al. "Intelligent fruit yield estimation for orchards using deep learning based semantic segmentation techniques—a review." *Frontiers in Plant Science* 12 (2021

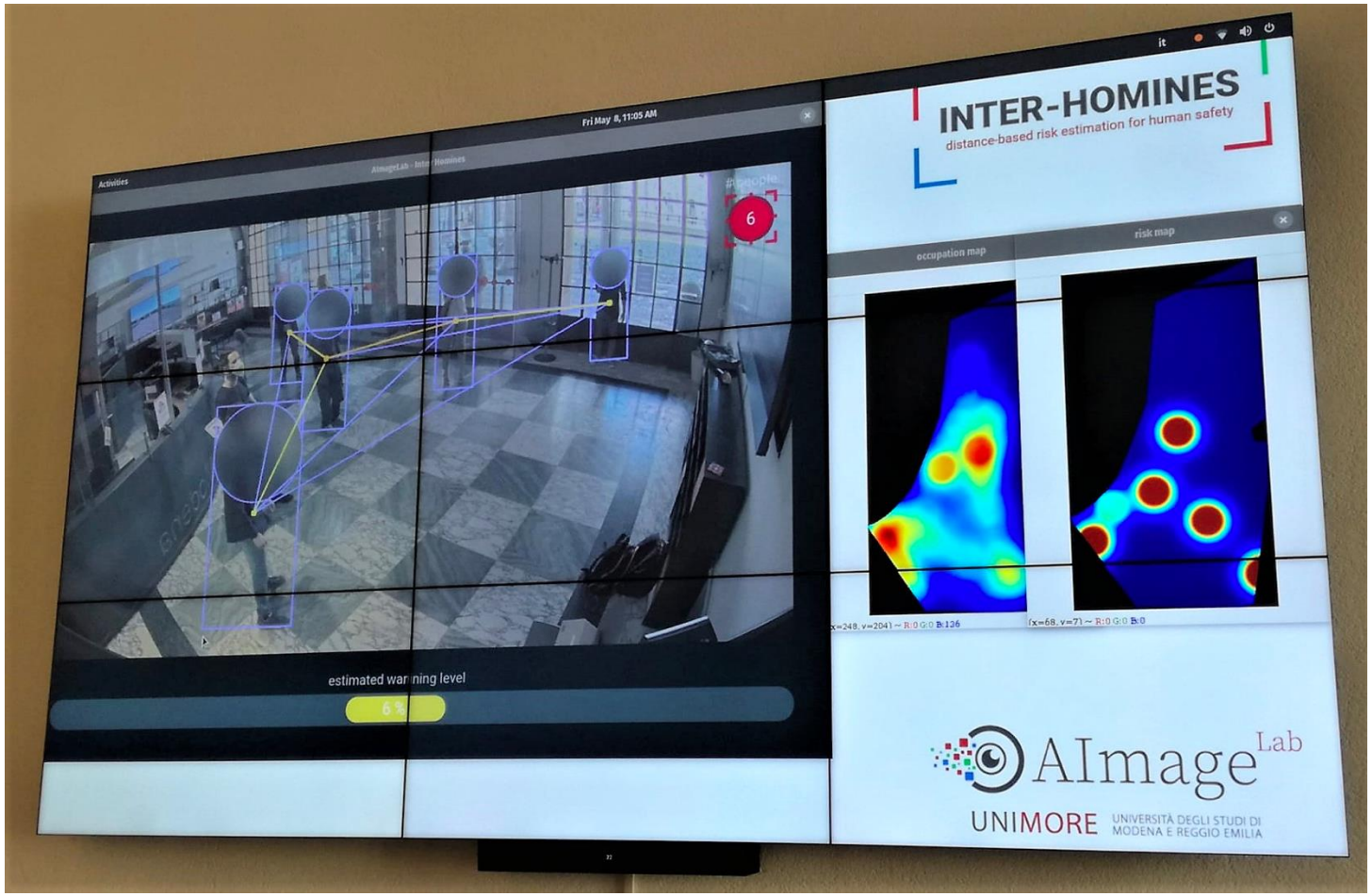


Lawal, M.O. Tomato detection based on modified YOLOv3 framework. *Sci Rep* **11,**

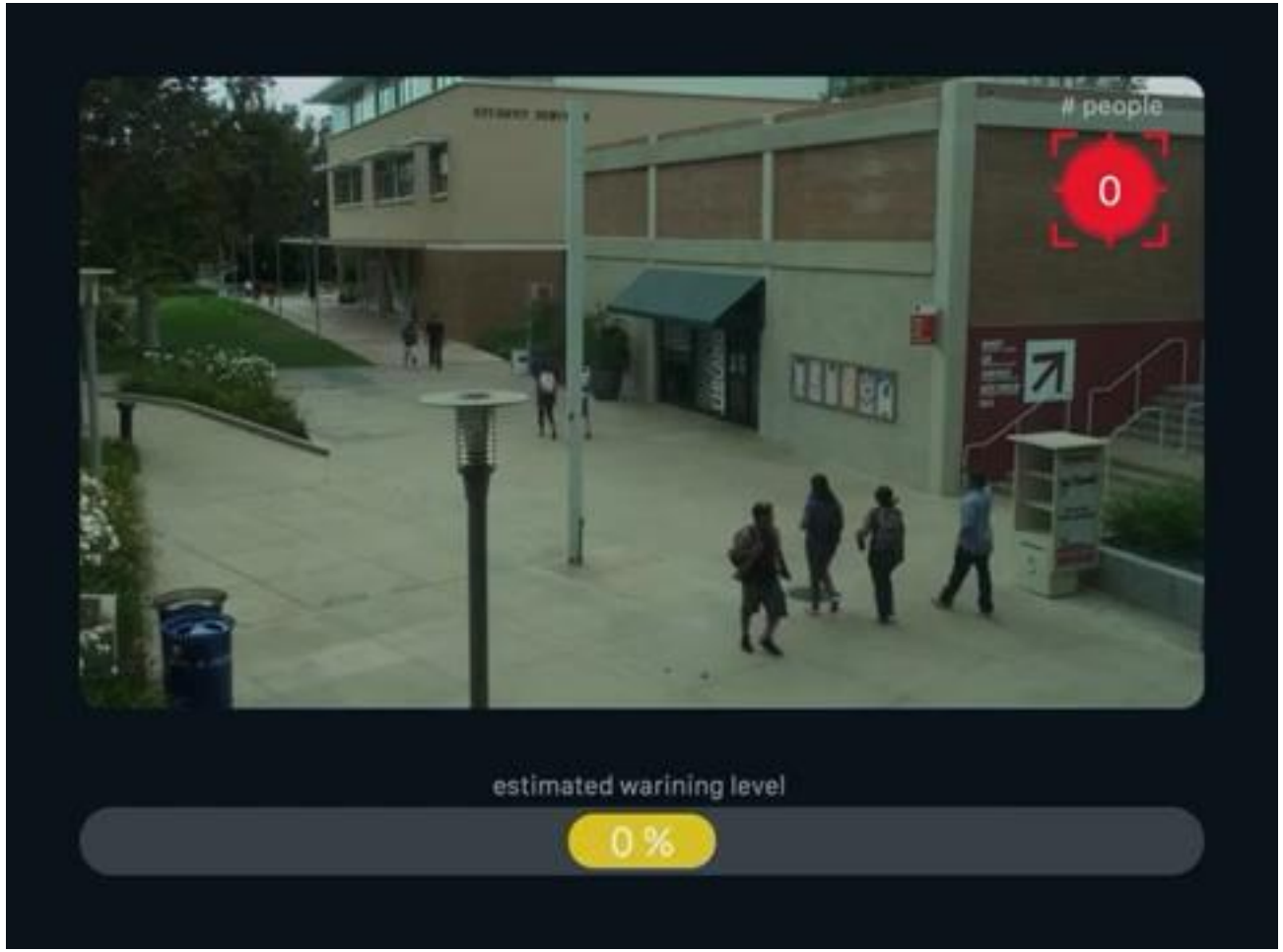


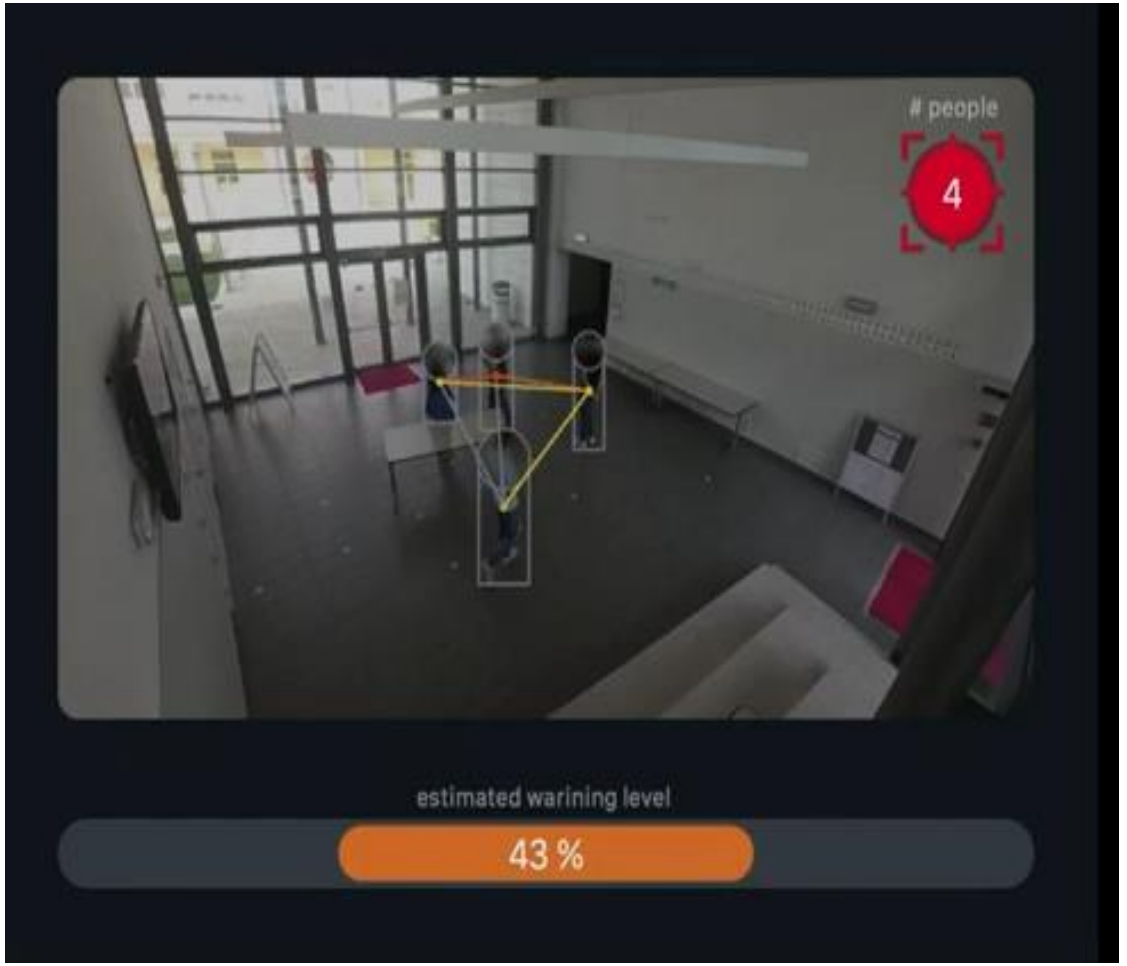
Inter-Homines - Distance-based risk estimation for human safety

- Using cameras and AI to detect interpersonal distances
- Assess the risk of an area
- Sophisticated behavior analysis models for system robustness and risk evaluation

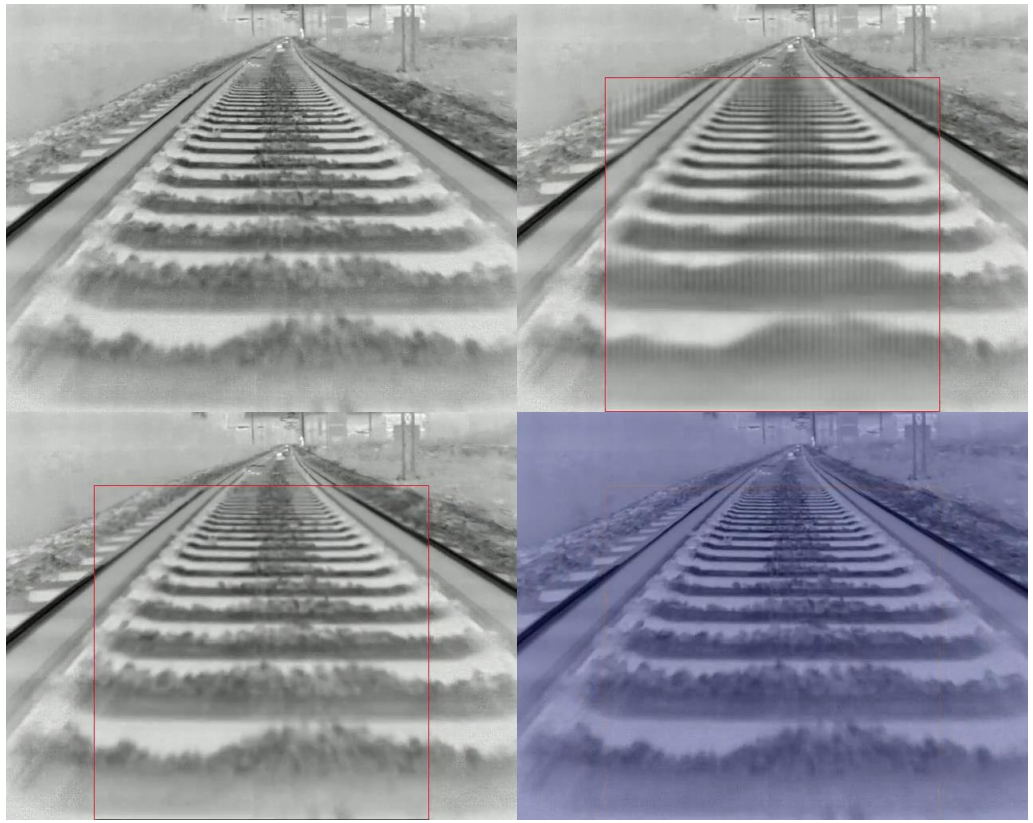


For more info https://aimagelab.ing.unimore.it/imagelab/project.asp?idprogetto=82
Prof. Rita Cucchiara rita.cucchiara@unimore.it Director of the Project
Dr. Matteo Fabbri matteo.fabbri@unimore.it


people
0
estimated warining level
0 %


people
4
estimated warining level
43 %

Anomaly Detection on Railways



RFI - “Computer Vision for Real-Time Obstacle and Anomaly Detection on Railway” (2019-2021)

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

RFI
RETE FERROVIARIA ITALIANA
GRUPPO FERROVIE DELLO STATO ITALIANE

Thank you for your attention

- **More on us**

- **Research:**
- AImageLab Research Group: http://aimagelab.unimore.it
- **Ellis Unit UNIMORE: https://ellis.eu/units/modena-unimore**

- **Tech Transfer and Life Long Learning:**
- AIAcademy UNIMORE: http://aiacademy.unimore.it/

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

AI Academy

AImage Lab