

# Brief history of AI

## From its Foundation to Generative AI

*“A generation which ignores history has no past – and no future.” - Robert A Heinlein*

### The history of AI

1940s-1950s

#### Foundations of AI

In the 1940s, the first artificial neurons were conceptualised. The 1950s introduced us to the Turing Test and the term “Artificial Intelligence.”



1960s-1970s

#### Early Development

The 60s and 70s brought the birth of ELIZA, simulating human conversation, and Dendral, the first expert system, showcasing the early potentials of AI.



1980s

#### AI Winter & Expert Systems

The 80s faced reduced AI funding but saw the inaugural National Conference on AI. The backpropagation concept rejuvenated neural networks.



1990s

#### Revival & Emergence of ML

The 90s witnessed IBM’s Deep Blue defeating chess champion Garry Kasparov and the inception of the LOOM project, laying the foundations for GenAI.



2000s

#### The Genesis of Generative AI

Geoffrey Hinton propelled deep learning into the limelight, steering AI toward relentless growth and innovation.



2010s

#### Rise of AI

In 2011, IBM Watson won “Jeopardy!”, highlighting AI’s language skills. The 2010s marked major AI milestones, including pioneering work in image recognition and the birth of GANs in 2014, followed by OpenAI’s founding in 2015.



2020s

#### GenAI Reaches New Horizons

At the start of this decade, we’ve seen significant strides in GenAI, notably with OpenAI’s GPT-3 and DALL-E. 2023 welcomed advanced tools like ChatGPT-4 and Google’s Bard, alongside Microsoft’s Bing AI, enhancing accessibility and reliability of information.



# AI between Sci-Fic & reality

*“As soon as it works, no one calls it AI anymore.” - John McCarthy*

**1980s** 1987-93: The second **Winter of AI**

September 28, 1987: The TV series "Star Trek: The Next Generation" introduces the self-aware android **Lieutenant Commander Data**.

June 29, 2001: Steven Spielberg releases his version of a film – originally developed by Stanley Kubrick – about a robot boy: **"A.I.: Artificial Intelligence."**

**1990s** May 11, 1997: **IBM's Deep Blue computer** beats reigning world chess champion Garry Kasparov. (Credit: Shutterstock)

2005: A Stanford vehicle wins the **DARPA grand challenge**, driving autonomously across the desert for 131 miles (211 kilometers).

**2000s** 2005: Inventor and futurist Ray Kurzweil predicts an event he calls the **Singularity** will occur around 2045, when the intelligence of artificial minds exceeds that of the human brain.

**2010s** 2011: **IBM's Watson wins "Jeopardy!"**, beating former champions Brad Rutter and Ken Jennings. (Credit: "Jeopardy!" screengrab from Wikimedia)

October 14, 2011: Apple introduces intelligent personal assistant **Siri** on the iPhone 4S.


June 2012: A Google Brain computer cluster **trains itself to recognize a cat** from millions of images in YouTube videos. (Credit: Shutterstock)


December 18, 2013: The movie "Her" (left), stars Joaquin Phoenix as a man who **falls in love with his artificially intelligent computer operating system**, voiced by Scarlett Johansson.

April 10, 2014: The film "Transcendence" (below) stars Johnny Depp as an AI researcher whose **mind is uploaded to a computer** and develops into a super-intelligence.

June 7, 2014: Chatbot Eugene Goostman is said to have **passed the Turing test** in University of Reading competition, launching controversy.

August, 2014: Researchers call for creation of a **new Turing test**, to be decided at 2015 workshop.









Advances in artificial intelligence (AI) have given the world computers that can beat people at chess and "Jeopardy!," as well as drive cars and manage calendars. But despite the progress, engineers are still years away from developing machines that are self-aware. Some believe the resulting **technological singularity** will eradicate poverty and disease, while others warn it could endanger human survival.

Summer of 1956: Dartmouth conference launches the field of AI and **coins the term "artificial intelligence."** (Right: room-filling IBM-702 computer, as used by first AI researchers)

1968: "2001: A Space Odyssey," the book by Arthur C. Clarke and film by Stanley Kubrick, features the sentient and deadly computer **HAL 9000**.





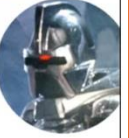
1974-early 1980s: The first **Winter of AI**, a period of reduced funding and lowered interest in the field as hype turned to disappointment.

1978: The original "Battlestar Galactica" science fiction TV series introduces warrior robots called **Cylons**.

1950: Isaac Asimov publishes the influential sci-fi story collection **"I, Robot."** (Left: 2004 film version of "I, Robot")

1950: Alan Turing introduces the **Turing test** in his paper "Computing Machinery and Intelligence." (Credit: National Portrait Gallery, London)

1984: The first **"Terminator"** film depicts a near-future world overtaken by killing machines run by the artificial intelligence Skynet.

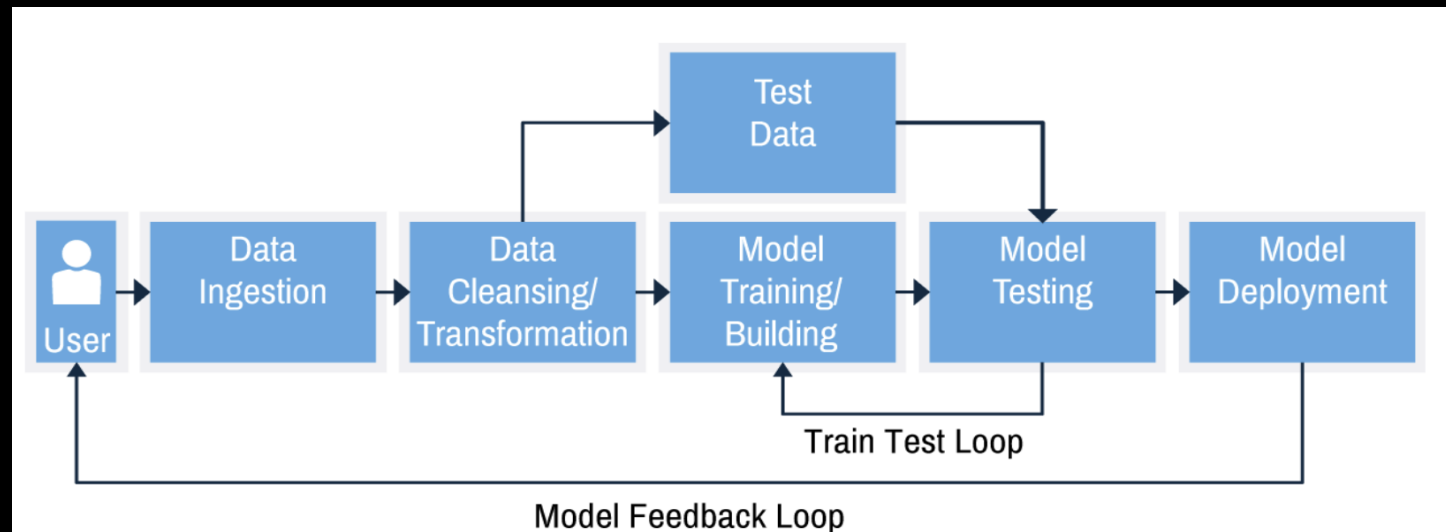






Source: History of A.I.: Artificial Intelligence (Infographic)

<https://www.livescience.com/47544-history-of-a-i-artificial-intelligence-infographic.html>

# To “much” focus teaching AI data pipeline


*“Students in every major will need to know how to challenge or defend the appropriateness of a given model for a given question.” - Chronicle of Higher Ed*



MNS Home News Publications Software OpenLab People Vacancies Contact

# Multiscale Networked Systems

The Multiscale Networked System (MNS) group researches the emerging architectures that can support the operations of multiscale systems across the Future Internet.




UNIVERSITEIT VAN AMSTERDAM

## Data centric processing

Our research investigates an alternative to the current approach to model complex scientific experiments as workflow of dependent tasks, in this approach scientific data is interlinked through data processing transformations which can be discovered and used to create the data processing workflow and not the way around.

[Learn more](#)

netherlands  
**eScience center**



Technology Lead, Data Processing  
**Dr. Adam Belloum**

**so far: ~150 projects**  
(on many different topics)

10,00 x 5,62 in

**Humanities & Social Sciences**

incl. SMART cities, text analysis, creative technologies

**Sustainability & Environment**

incl. climate, ecology, energy, logistics, water management

**Physics & Beyond**

incl. astronomy, high-energy physics, advanced materials

**Life Sciences & eHealth**

incl. bio-imaging, next generation sequencing, molecules

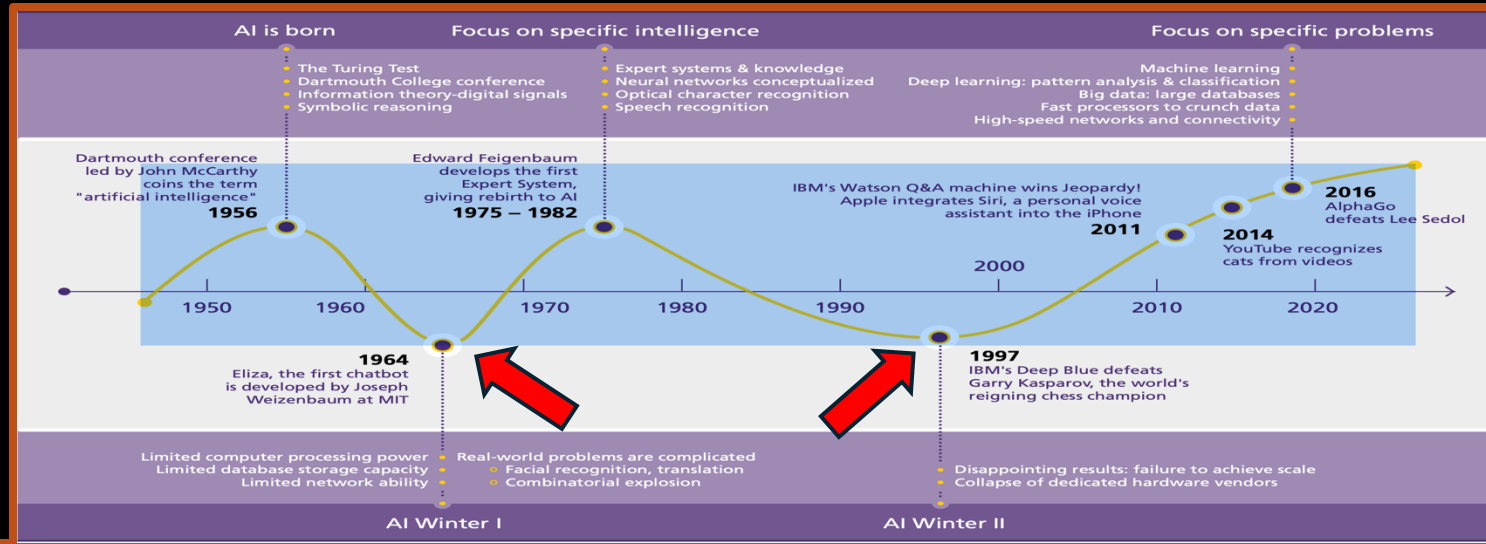
Conversions in the gut



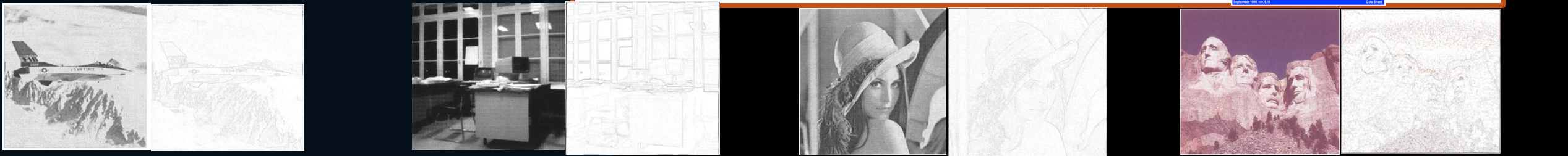
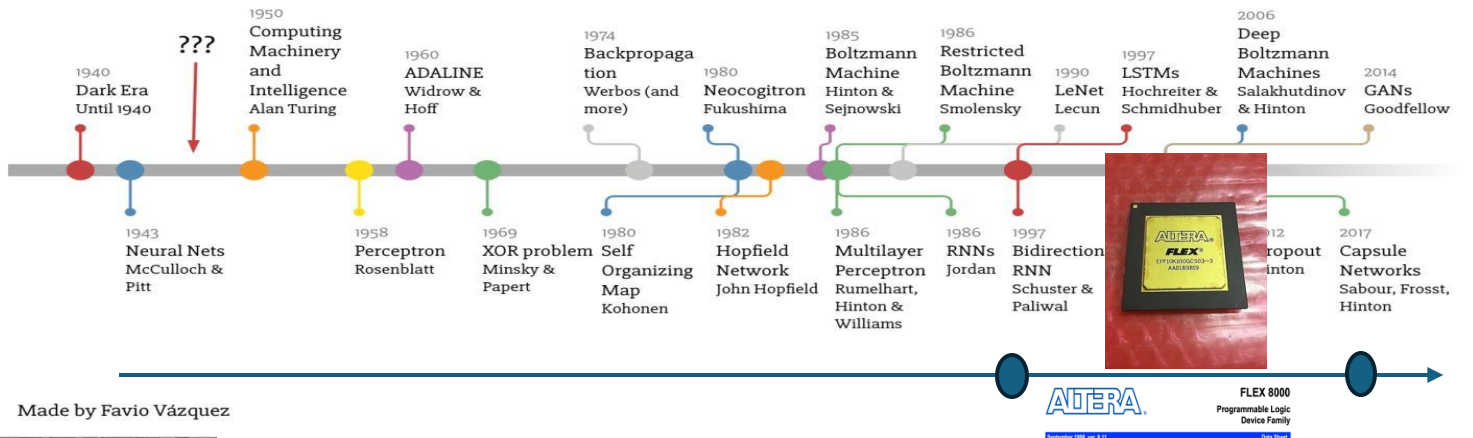
# AI timeline

## Deep learning

from its inception to ChaGPT



## Deep Learning Timeline

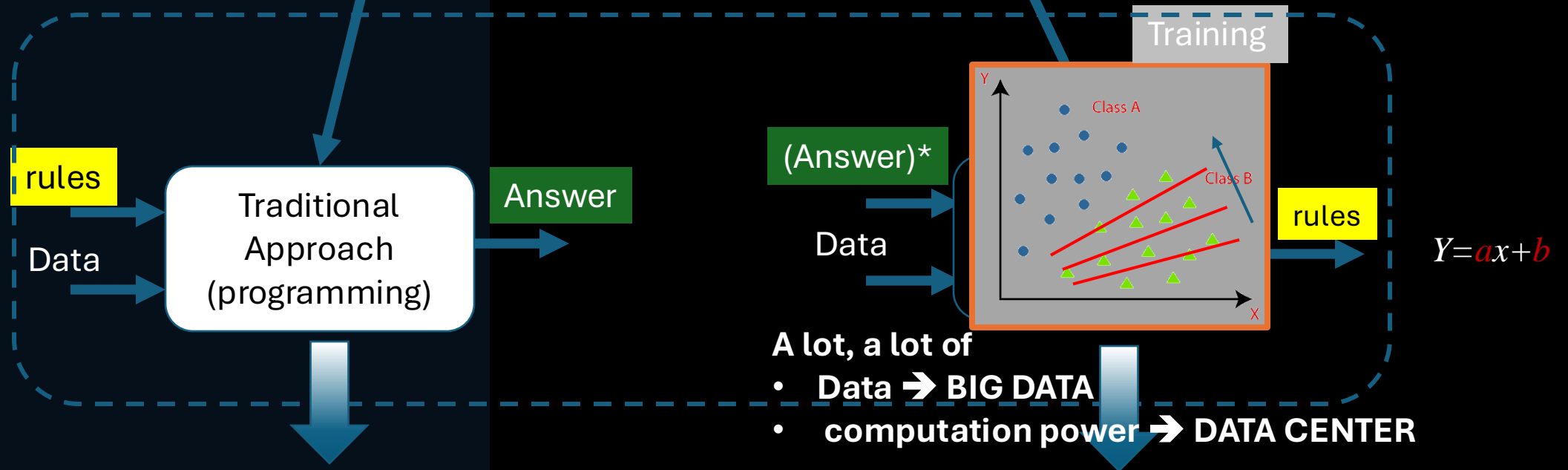


# Content

- Computer programming
  - Knowledge centric (human learning)
  - Data centric (machine learning)
  - Training (years vs Bytes)
- Big data
  - Data movement
  - Data processing
  - Data storage
- Artificial Intelligence terminology /landscape
  - AI origins
  - AI milestones (perceptron – DL)
  - AI between Hype/fiction/Reality
  - Deep Learning
  - Problem and Security (hallucination,
  - Ethics (Black box,
  -

# Understanding AI approach

How can we make a computer do something useful?



Two approaches to make a computer do something useful (what is the best approach?)

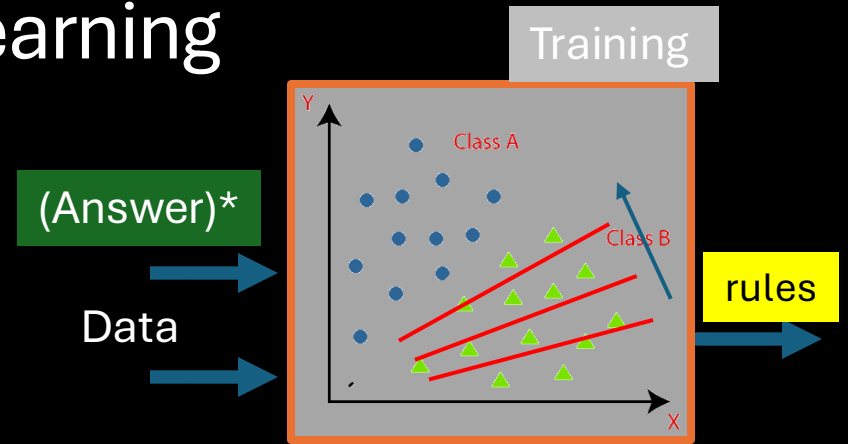
# The types of Machine Learning

Training (minimising error/ reward)  
and Data (labelled or not)

1. Supervised

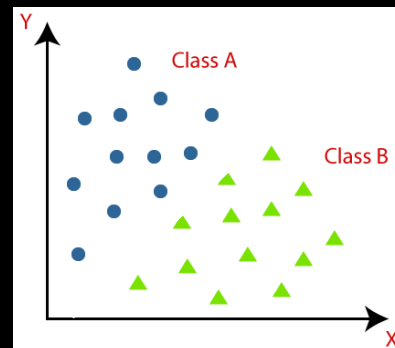
2. Unsupervised

3. Reinforcement learning

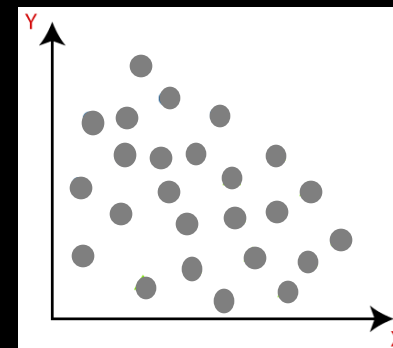


**A lot, a lot of**

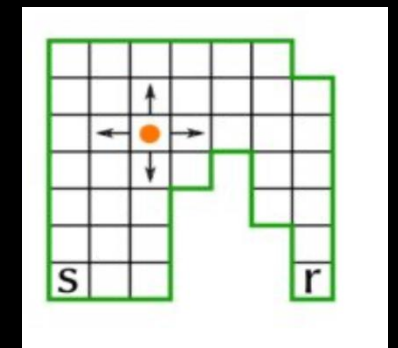
- **Data** → **BIG DATA**
- **computation power** → **DATA CENTER**



supervised



Unsupervised



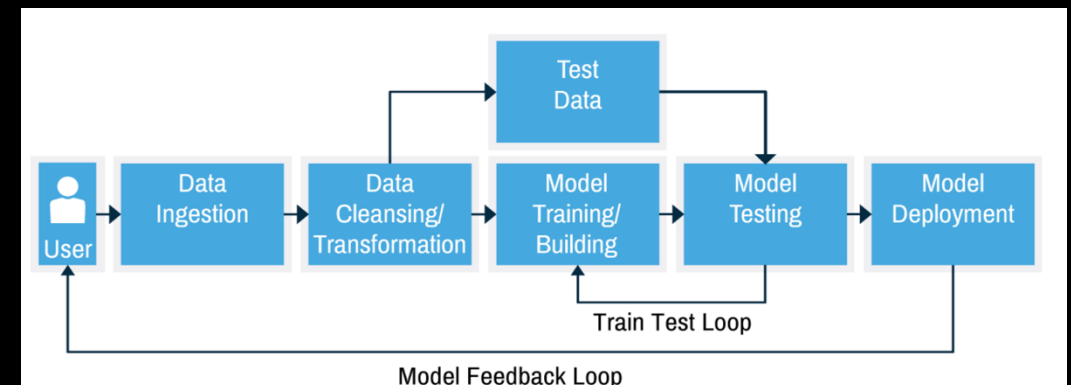
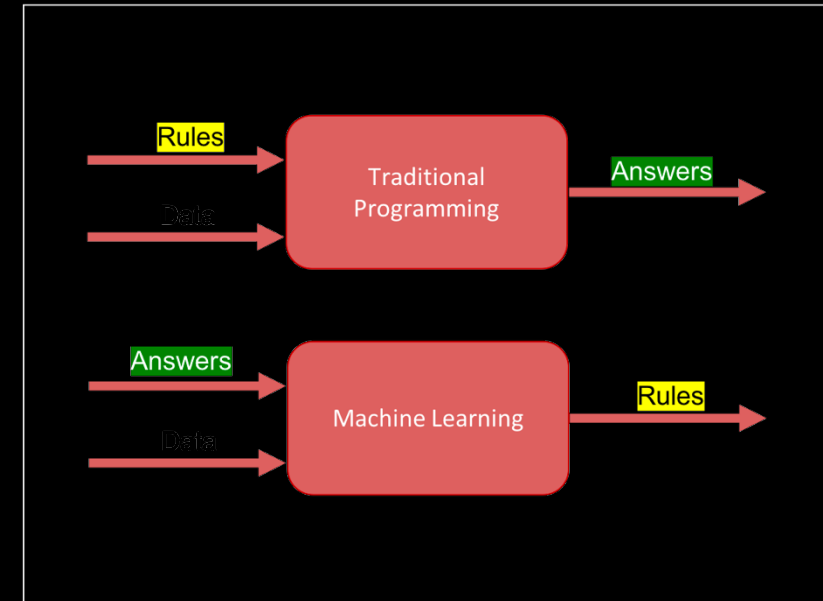
Reinforcement learning



# AI / Machine Learning / Big Data ?

“AI” is a scientific discipline that deals with the construction and study of algorithms that can **learn from data**. Such Algorithms operate in 2 steps:

1. building a **model** based on the existing data
2. using **the model** make **predictions** and **decision** rather than following explicitly programmed instructions “



# Data pre-processing

Is the data ready to be processed?



banks.sav [DataSet9] - IBM SPSS Statistics

1: jtype	educ	marit	start	jtype	whours	salary
1	.	2	07-May-2016	1	28.25	\$1,6
2	4	1	27-Oct-2026	1	.	\$1,7
3	5			1	22.75	\$1,5
4	1			.	27.25	\$1,9
5	3			1	.	\$1,3
6	6	2	08-Dec-2016	2	43.75	\$3,5

System missing values are indicated by dots.

Not always:

- Data not in the correct format: images, voice, text...
- Dealing with **missing values**
- Dealing with **noise** (errors) in the data values

Data Explorer  
977.5 kB  
WA\_Fn-UseC\_Telco-Customer-Churn.csv

WA\_Fn-UseC\_Telco-Customer-Churn.csv (977.5 kB)

Detail	Compact	Column
A customerID	A gender	# SeniorCiti...
7590-VHVEG	Female	0
5575-GNVEE	Male	0
3668-OPYBK	Male	0
7795-CFOCH	Male	0
9237-HQITU	Female	0
9385-CDSKC	Female	0
1452-KIDVK	Male	0
6713-OKOMC	Female	0
7892-PDOKP	Female	0
6388-TABGD	Male	0
9763-GRSKD	Male	0
7460-LKBCI	Male	0
8891-TVAVX	Male	0
8280-XJGEX	Male	0
5129-JLPES	Male	0
3655-SNOYZ	Female	0
8191-XBSZG	Female	0
9959-WOFKY	Male	0
4190-HFLUR	Female	0
4183-HYFRB	Female	0
8779-QRDVY	Male	1

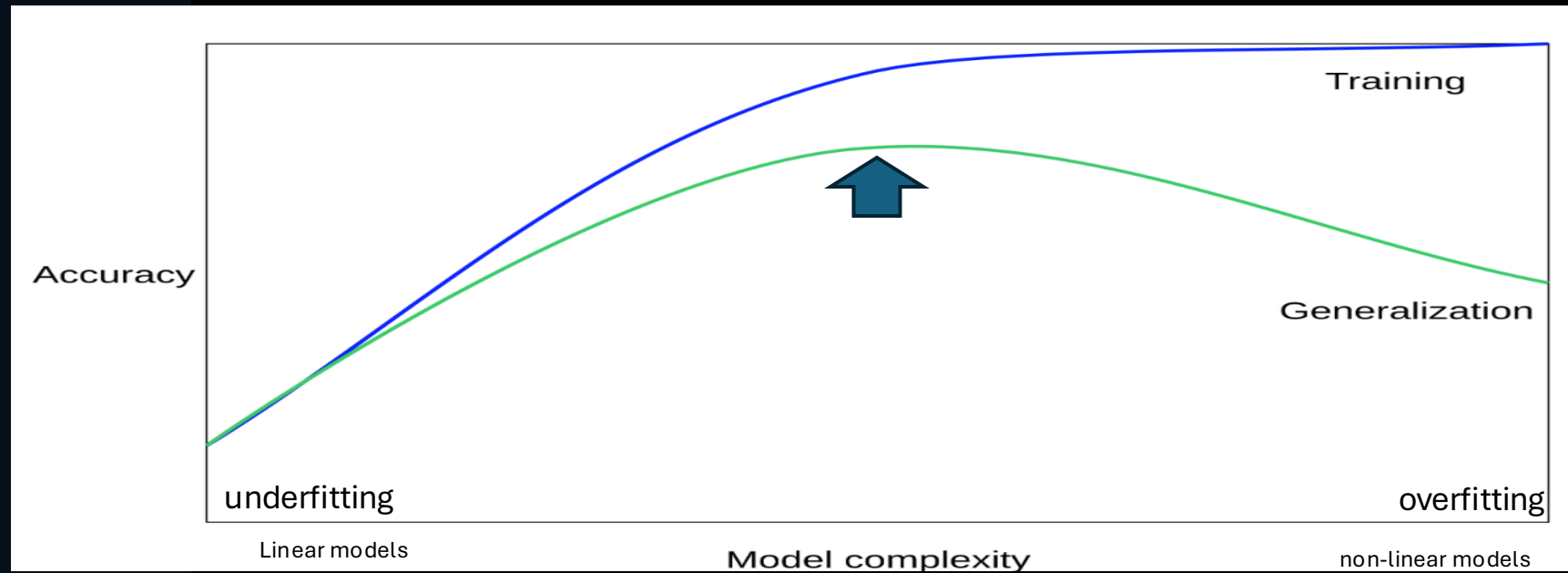
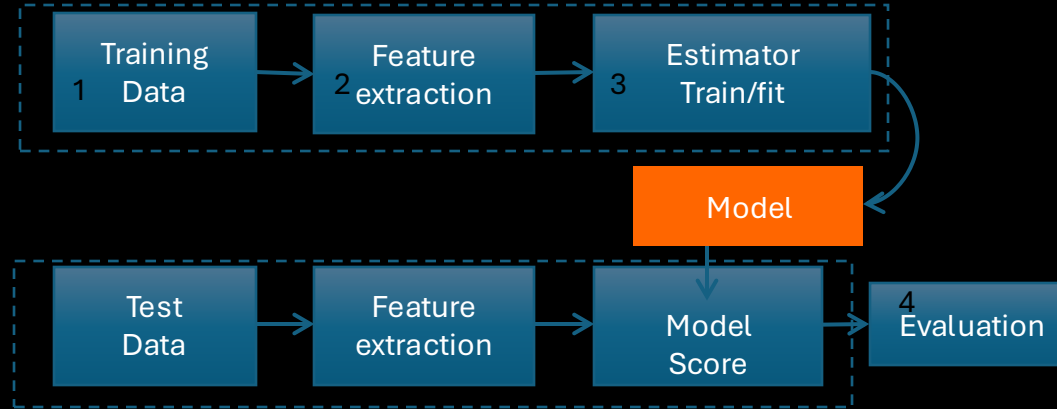
Pre-processing

- Feature engineering
  - Feature Selection
  - ...

# Model selection

## The AI / Machine Learning WorkFlow

1. building a **model** based on the data
2. using **the model** make **predictions** and **decisions** rather than following explicitly programmed instructions “

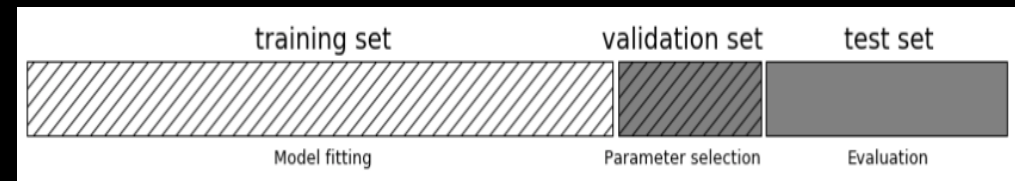


# Model training (The AI / Machine Learning WorkFlow)

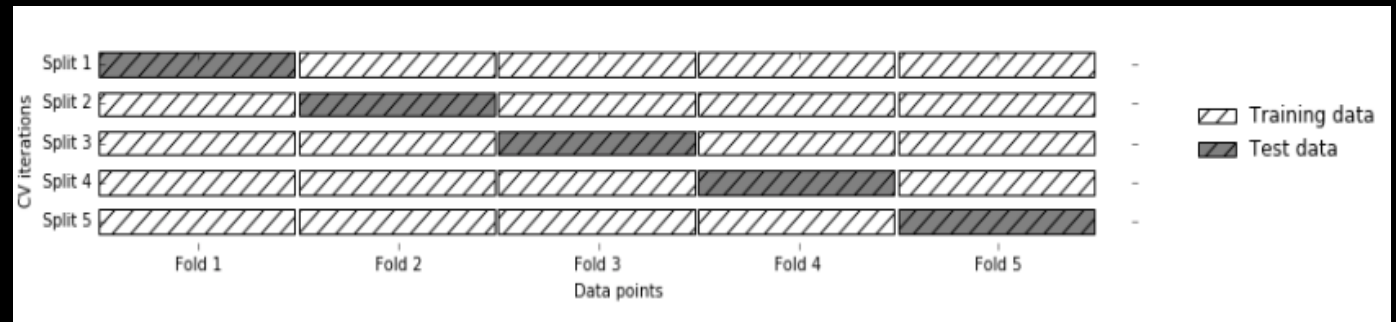
How to split the input dataset into: training data and test data?

Simple Answer → There many ways

- Simple split (train, test) → (default 75%, 25%) or any proportion
- Threefold split (train, test, validate)
- Cross-validation:
- Nested cross-validation, Stratified cross-validation, TimeSeriesSplit



Threefold split



Cross-validation

# The Machine Learning WorkFlow

There are many Machine learning Algorithms (Models) with different: **Model complexity, computational Complexity, memory usage,**

- Which one to use? → depends on the application
- Basic models
  1. Nearest Neighbours,
  2. Nearest Centroid
  3. Linear Classification and Regression
  4. Logistic Regression
- Non-Linear models
  6. Support Vector Machines and Kernels
  7. Decision Trees
  8. Random Forests
  9. Gradient Boosting
  10. Model Calibration



**Original author(s)** David Cournapeau  
**Initial release** June 2007; 14 years ago  
**Stable release** 1.0.1<sup>[1]</sup> / 25 October 2021; 37 days ago  
**Repository** [github.com/scikit-learn/scikit-learn](https://github.com/scikit-learn/scikit-learn)  
**Written in** Python, Cython, C and C++<sup>[2]</sup>  
**Operating system** Linux, macOS, Windows  
**Type** Library for machine learning  
**License** New BSD License  
**Website** [scikit-learn.org](https://scikit-learn.org)



**Developer(s)** Google Brain Team<sup>[1]</sup>  
**Initial release** November 9, 2015; 6 years ago  
**Stable release** 2.6.1<sup>[2]</sup> (1 November 2021; 30 days ago) / May 14, 2021; 6 months ago  
**Repository** [github.com/tensorflow/tensorflow](https://github.com/tensorflow/tensorflow)  
**Written in** Python, C++, CUDA  
**Platform** Linux, macOS, Windows, Android, JavaScript<sup>[3]</sup>  
**Type** Machine learning library  
**License** Apache License 2.0  
**Website** [www.tensorflow.org](https://www.tensorflow.org)



**Original author(s)** Matei Zaharia  
**Developer(s)** Apache Spark  
**Initial release** May 26, 2014; 7 years ago  
**Stable release** 3.2.0 / October 13, 2021; 49 days ago  
**Repository** [Spark Repository](https://spark.apache.org)  
**Written in** Scala<sup>[1]</sup>  
**Operating system** Microsoft Windows, macOS, Linux  
**Available in** Scala, Java, SQL, Python, R, C#, F#  
**Type** Data analytics, machine learning algorithms  
**License** Apache License 2.0  
**Website** [spark.apache.org](https://spark.apache.org)

model	fit	Memory	Prediction
centroids	$O(n \cdot p)$	$O(n\_classes \cdot p)$	$O(n\_classes \cdot p)$
Neighbours			
Naïve		$O(n \cdot p)$	$O(n \cdot P)$
Kd_tree	$O(p \cdot (n \cdot \log n))$	$O(n \cdot p)$	$O(p \cdot (n \cdot \log n))$

$n = n\_samples$

$p = n\_features$



# Summary of what AI users learns

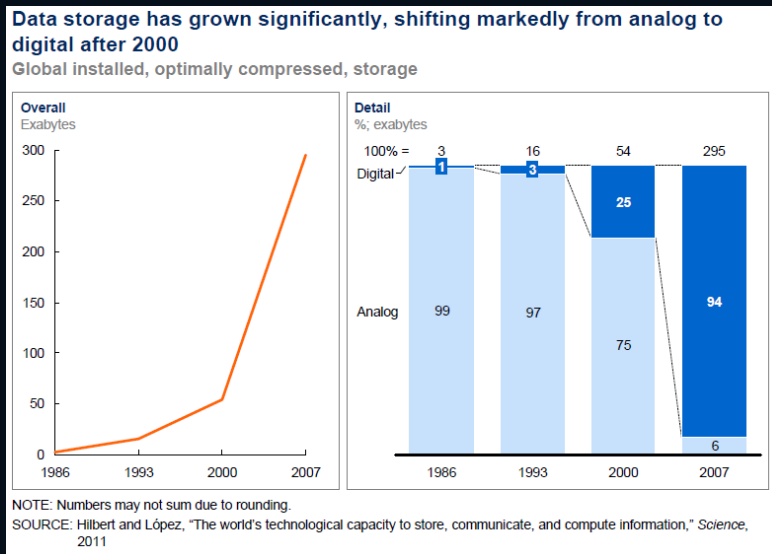
- The AI approach:
    - rules are generated through training
    - AI is a black box
    - Need a lot of Data “Big Data”
  - Pre-process the Data
    - Feature engineering
  - Build a Model
    - Select/train/validate/deploy a model
  - How to use AI libraries
    - Enough to build a quickly an AI model
- **No intuition** what does mean to
    - Generating a dataset
    - Store/move/process Big Data:
  - **No intuition** how basic mechanisms behind the AI have been developed over the years and what are problems
    - Backpropagation (chain rule)
    - Gradient (vanishing/exploding)
    - Word Embedding
    - Attention mechanisms
    - neural networks “Architectures”
    - ...

**Not enough** to reflect and solve the problem when the AI model does not work

**Intuition** is needed when the AI model does not work

# Content

*“Those who own data own the future”*  
Yuval Noah Harari



- Computer programming
  - Knowledge centric (human learning)
  - Data centric (machine learning)
  - Training (years vs Bytes)
- Big data
  - Data movement
  - Data processing
  - Data storage
- Artificial Intelligence terminology /landscape
  - AI origins
  - AI milestones (perceptron – DL)
  - AI between Hype/fiction/Reality
  - Deep Learning
  - Problem and Security (hallucination,
  - Ethics (Black box,
  -

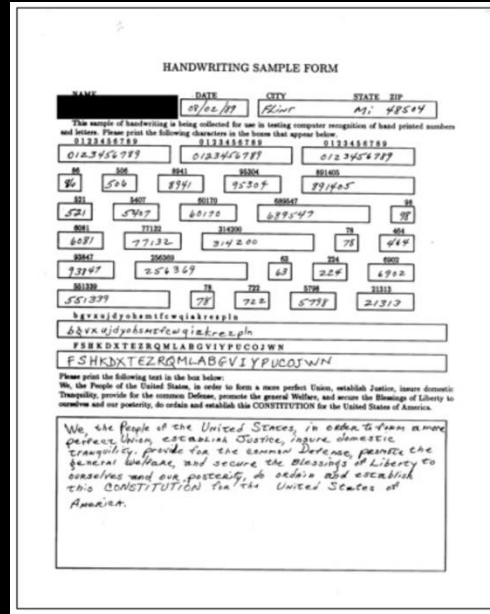
# The MNIST dataset



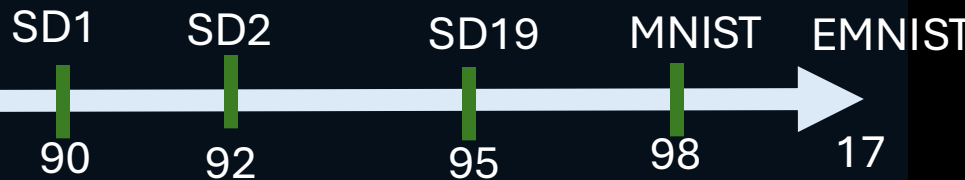
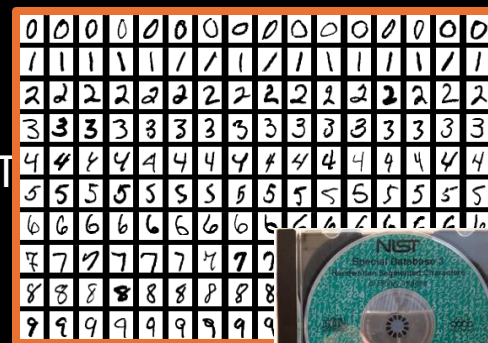
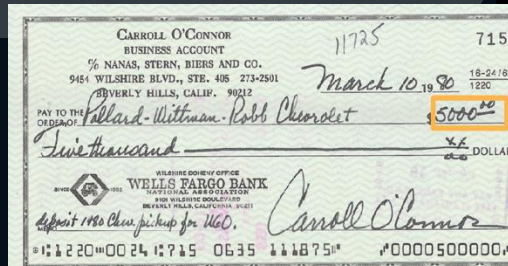
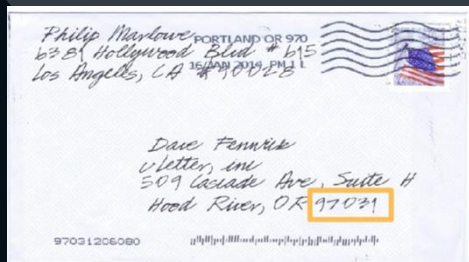
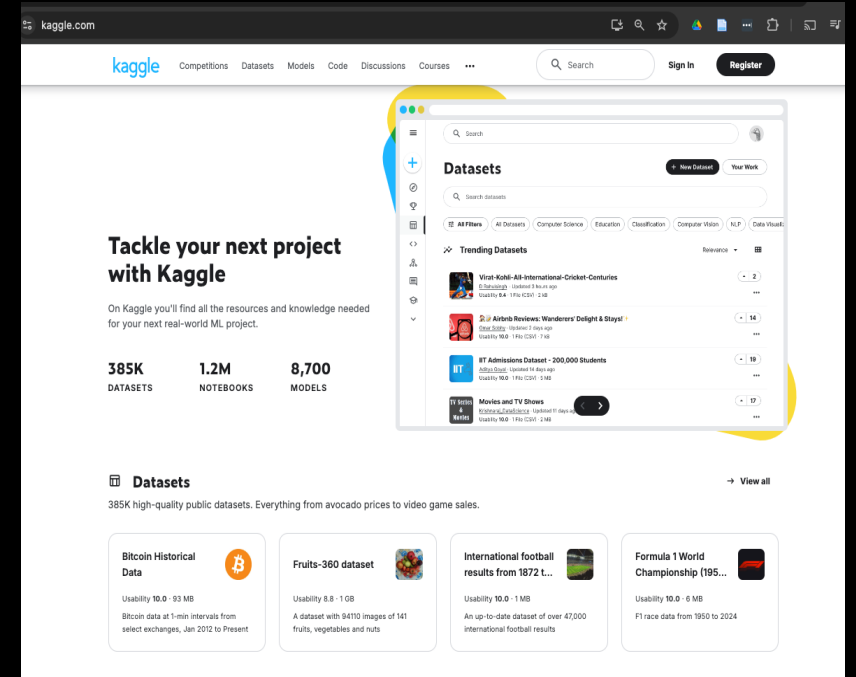
## NIST

- Handwriting character recognition
- What Accuracy number do you trust?
- Need a baseline (calibrated ground truth)

## Datasets in 90s



## Datasets in > 2010



LeCun, "The MNIST DATABASE", <http://yann.lecun.com/exdb/mnist/>.

# A Storage Capacity

YottaByte (YB) =  $10^{24}$  Byte

ZetaByte (ZB) =  $10^{21}$  Byte

ExaByte (EB) =  $10^{18}$  Byte

PetaByte (PB) =  $10^{15}$  Byte

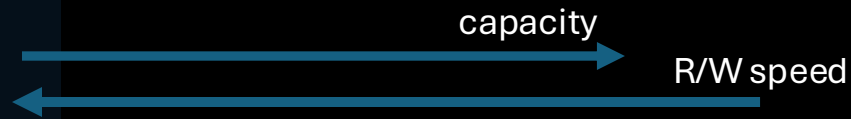
**TeraByte (TB) =  $10^{12}$  Byte**

GigaByte (GB) =  $10^9$  Byte

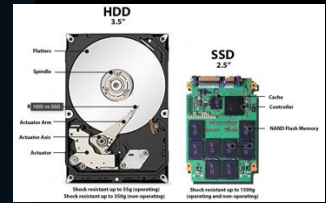
MegaByte (MB) =  $10^6$  Byte

KiloByte (KB) =  $10^3$  Byte

Byte = 8 bits



- **Storage**
- **Processing**
- **Movement**



- 1 TB HDD/~60\$ - Storage technology
- 18 TB HDD/~600\$ - Storage technology

- 1+ ZB - Internet size in bytes
- Radio astronomy- SKA-Phase 3+ EFlops

Note: Kilo is exactly 1024 ~ 1000

# A Terabyte of Storage Space: How Many ...?

- Storage
- Processing
- Movement



## personal usage

- ~200,000 average songs, High-Quality Compressed Audio  
(~17,000 hours of music)
- ~256 Standard DVD Movies 120 minutes long  
(~500 hours of movies)
- ~310,000  
Standard-Resolution  
Photos

Note: 1 TB = 1,000 ( $10^3$ ) gigabytes (GB) or 1,000,000 ( $10^6$ ) megabytes (MB)

Source: <https://aimblog.uoregon.edu/2014/07/08/a-terabyte-of-storage-space-how-much-is-too-much/>



- Storage
- Processing
- Movement

# Data collected / generated

In Industry and science around 2009

Google processes

Wayback Machine has 3 PB

Facebook has 2.5 PB of data

eBay has 6.5 PB of user data

CERN's Large Hadron Collider -

**generates**

→ 20 PB a day

→ 100 TB/month

→ +15 TB/day

→ 50 TB/day

→ 15 PB/year



Note: 1 TB = 1,000 ( $10^3$ ) gigabytes (GB) or 1,000,000 ( $10^6$ ) megabytes (MB)

Source: <https://aimblog.uoregon.edu/2014/07/08/a-terabyte-of-storage-space-how-much-is-too-much/>

- Storage
- Processing
- Movement



## Has More Bandwidth Than the Internet—

If you're looking to transfer **hundreds of gigabytes** of data, it's still—weirdly—faster to ship hard drives via FedEx than it is to transfer the files over the internet.



estimates that total internet traffic averages **167 terabits per second**.



has a fleet of 654 aircraft with a lift capacity of 26.5 million pounds daily.

- A solid-state laptop drive weighs about 78 grams and can hold up to a terabyte.
- FedEx is capable of transferring 150 exabytes of data per day, or **14 petabits per second**—almost a **hundred times the throughput of the internet in 2013**

By [Jamie Condliffe](#) Published February 5, 2013

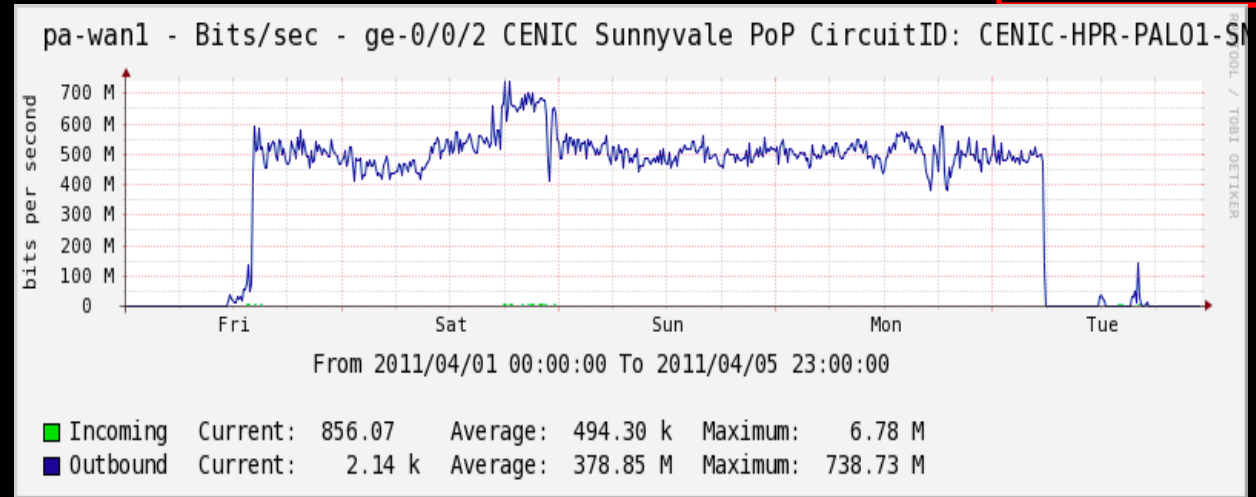
# How much Time does it take to move TBs over the internet ?

moving **60 complete** human genomes from Mountain View - Chicago.

**Approximately 18 TB** on **1G link.**



- **Storage**
- **Processing**
- **Movement**



Credit: Robert Grossman University of Chicago Open Data Group, November 14, **2011**



Credit: Cees de Laat University of Amsterdam SNE Group, super Computing, **2017**

<https://delaat.net/sc/sc17/demo02/index.html>

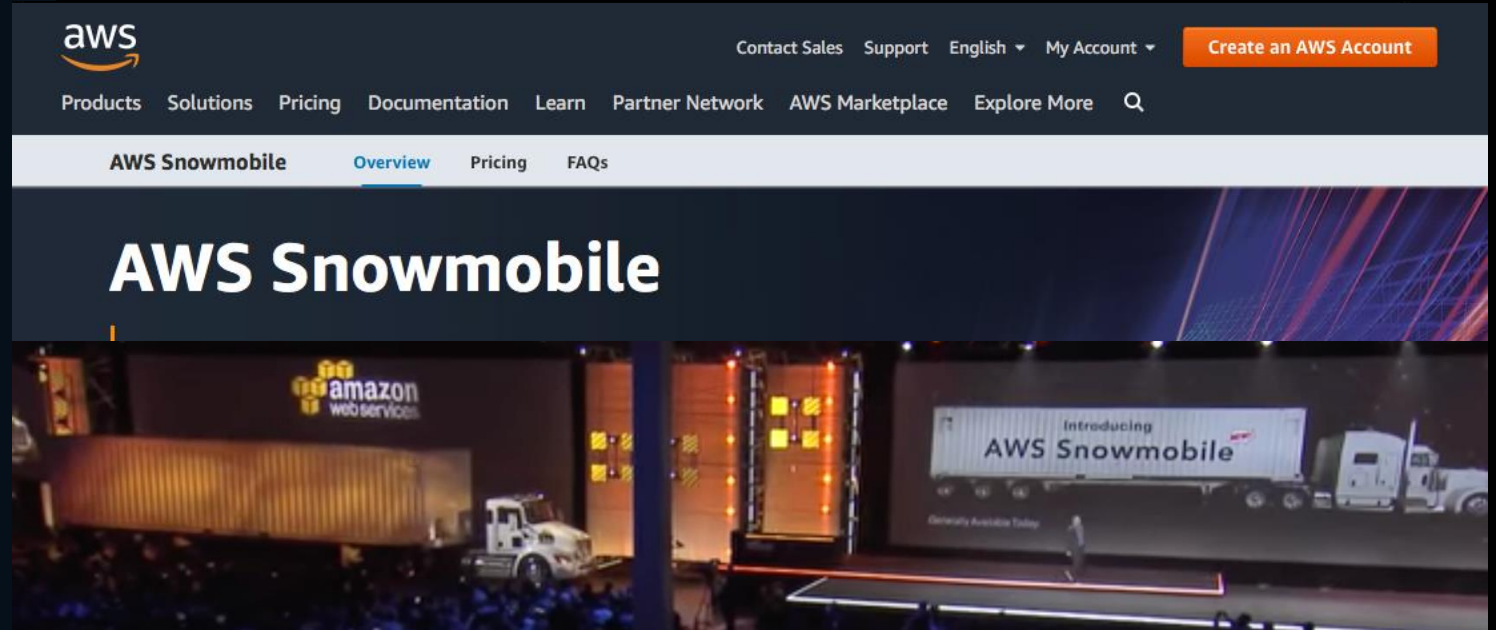
# How much Time does it take to move 1 exa-byte over the internet ?

- Storage
- Processing
- Movement

Over **10Gbs** line it will take ~ **26 years**

Note: 1 exa-Byte =

- 1,000 ( $10^3$ ) petabytes
- or 1,000,000 ( $10^6$ ) terabytes
- or 1,000,000, 000 ( $10^9$ ) gigabytes
- or 1,000,000, 000, 000 ( $10^{12}$ ) megabytes



[AWS Snowmobile – Move Exabytes of Data to the Cloud in Weeks | AWS News Blog \(amazon.com\)](#) 2016



- Storage
- Processing
- Movement

# How much time does it take to process 1 TB?

Estimate:

read 100MB/s, write 100MB/s

no disk seeks, instant sort

341 minutes → 5.6 hours

The terabyte benchmark winner (2008):

209 seconds (3.48 minutes)

November 2008 (\*)

68 seconds

<http://sortbenchmark.org/>

(\*)<https://googleblog.blogspot.com/2008/11/sorting-1pb-with-mapreduce.html>

The screenshot displays the MAPR Technologies cluster management interface for a cluster named 'perfdemo'. The main area shows a 'Cluster Heat Map - 1003 Nodes' with a grid of green squares indicating node activity. On the left, a terminal window shows Hadoop configuration commands. On the right, there are navigation options including 'Nodes x', 'al-core', 'ssors, 4', and 'er Overview'. The interface also shows a 'CPU Utilization' section and a 'Nodes' section with a 'Cluster Heat Map' button.

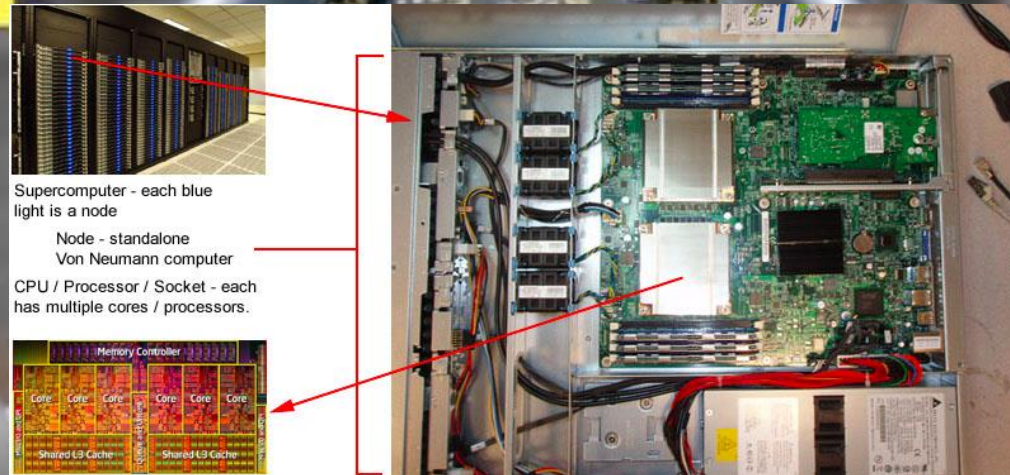
This is a zoomed-in view of the 'Cluster Heat Map - 1003 Nodes' interface, showing the grid of green squares and the navigation menu on the right.



Does more CPUs imply faster execution times?

- Storage
- Processing
- Movement

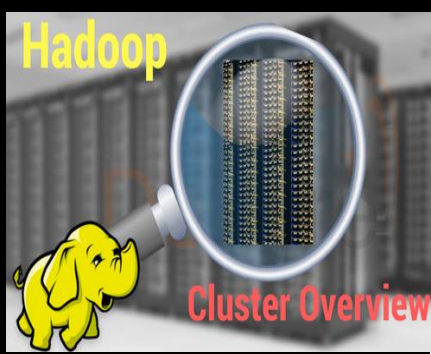
# Hadoop



## Cluster Overview

- How CPU works [http://www.youtube.com/watch?v=cNN\\_tTXABUA](http://www.youtube.com/watch?v=cNN_tTXABUA)
- Richard Feynman Computer Heuristics Lecture <http://www.youtube.com/watch?v=EKWGGDXe5MA>

Using more CPUs  
imply faster execution  
times!



- Storage
- Processing
- Movement

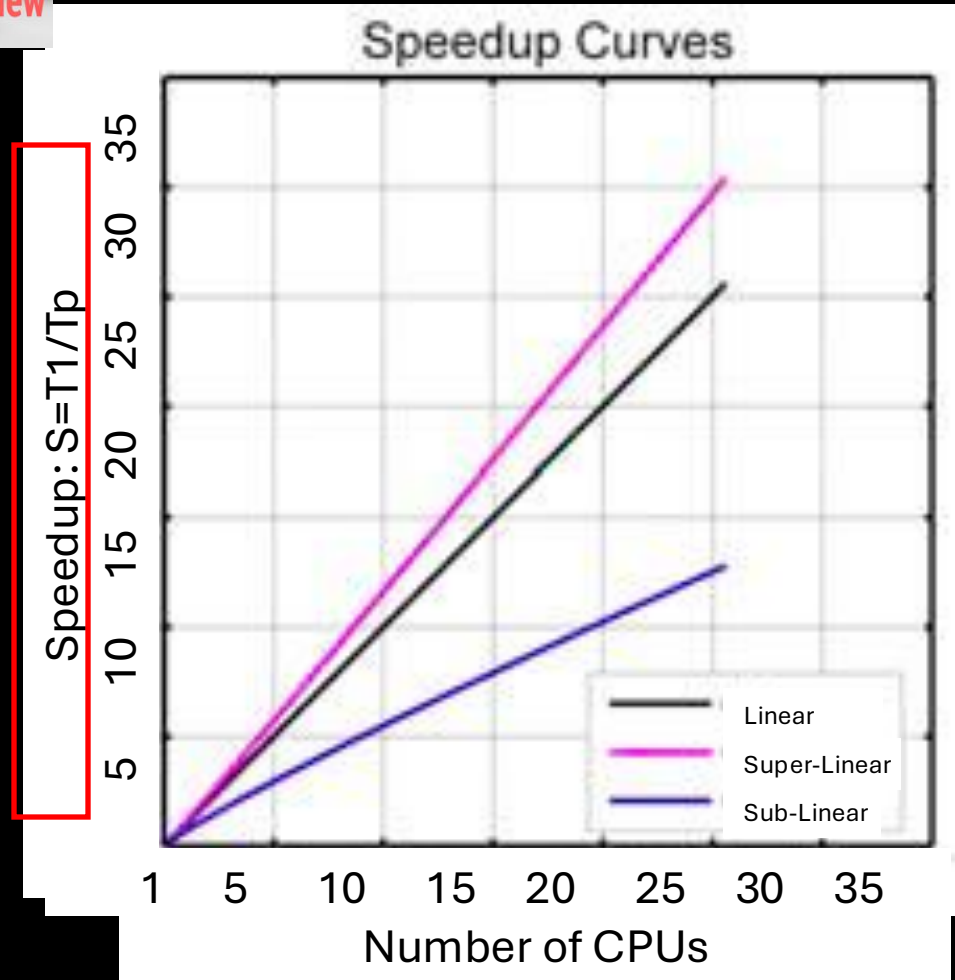
• Speedup

- Best
- Superlinear
  - Linear
  - Sublinear
  - Other?
- Worst

You must learn Parallel programming (\*)

Or

Using specialized AI libraries like TensorFlow, PyTorch



Credit: Jon Johansson Academic ICT Copyright © 2006 University of Alberta

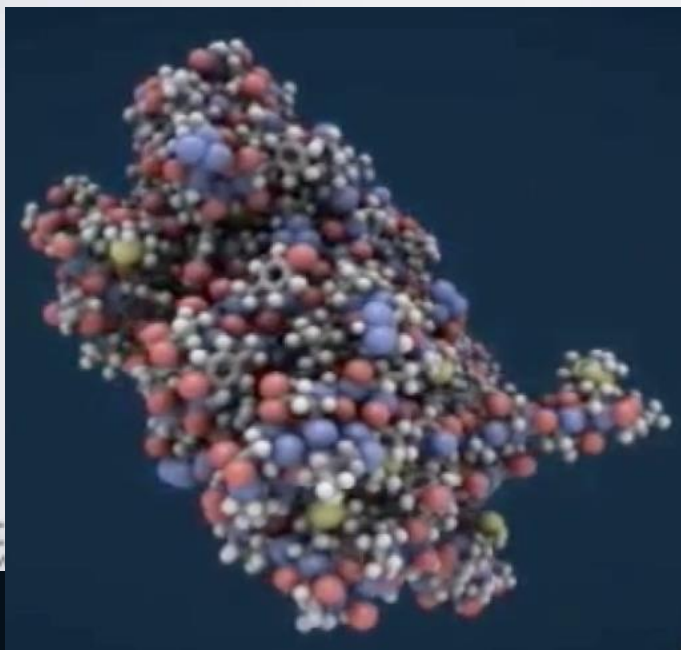
(\*)Computer Science profile

# Do we always need a Supercomputer to get some Speedup?

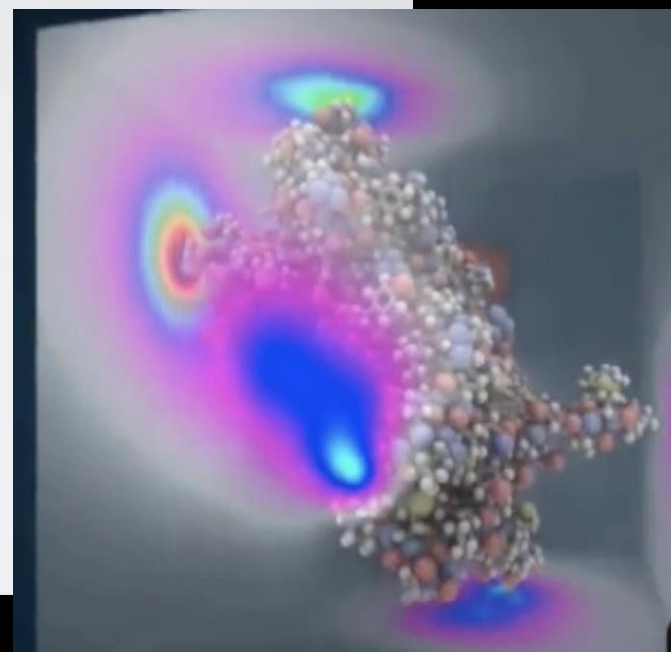
- Storage
- Processing
- Movement

- Not necessary → Do you have a Game computer?

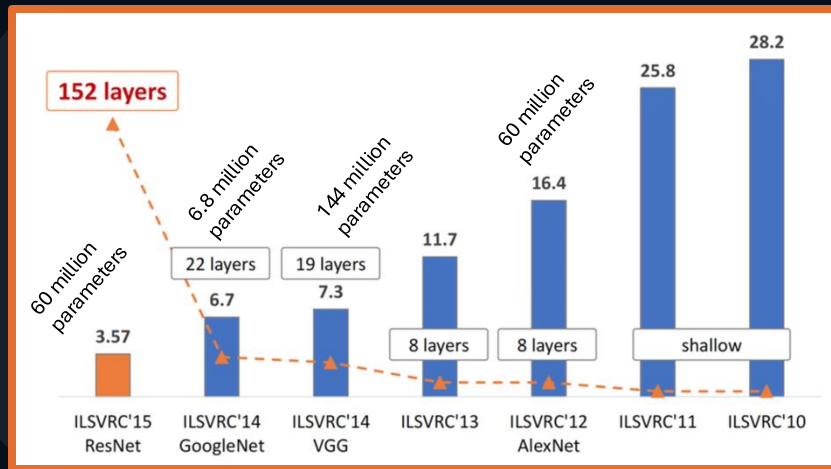
- Demo: Software the electrostatic properties of biological molecules
  - **Usage:** drug discovery
  - **Calculation** of the boundary value condition (quite slow).
  - **GPU :** EVGA GeForce GTX 285 1GB(~ 400\$)
  - Programming Language: OpenCL



DEMO



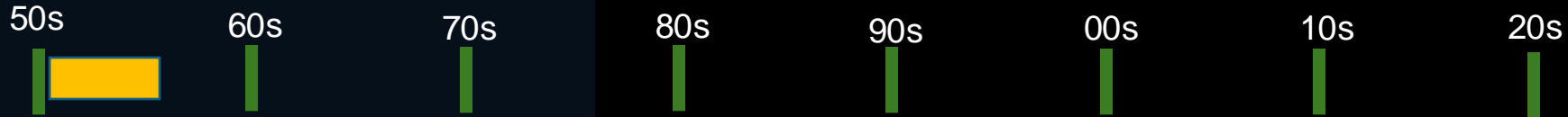
# Content



- Computer programming
  - Knowledge centric (human learning)
  - Data centric (machine learning)
  - Training (years vs Bytes)
- Big data
  - Data movement
  - Data processing
  - Data storage
- Artificial Intelligence terminology /landscape
  - AI origins
  - AI milestones (perceptron – DL)
  - AI between Hype/fiction/Reality
  - Deep Learning
  - Problem and Security (hallucination,
  - Ethics (Black box,
  -



# The Birth of Artificial; intelligence



explore the potential of Synthetic Intelligence (the term AI hadn't been coined

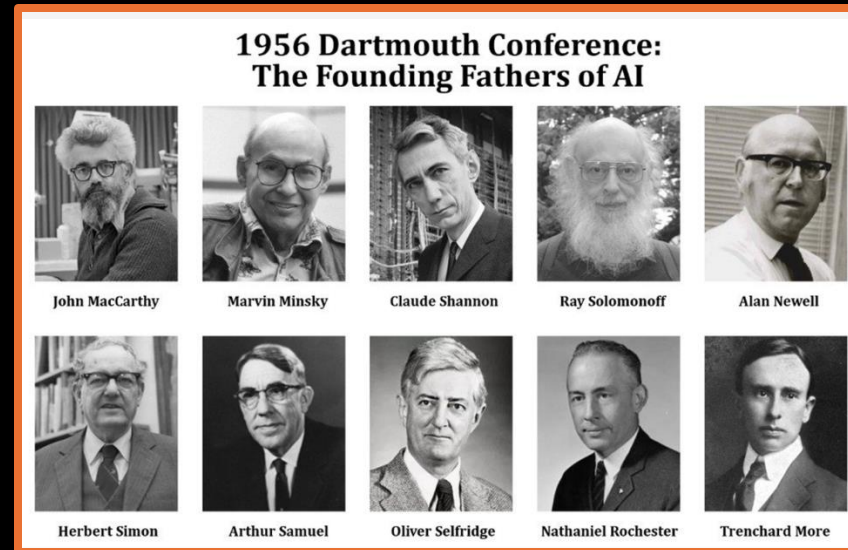
**Vision** for AI<sup>(yet)</sup>.

creation of intelligent machines that could reason, learn, and communicate like humans

**Roadmap** for AI research

- programming languages
- algorithms for intelligent machines

**AI research labs** at universities and research institutions, MIT, Carnegie Mellon, and Stanford



Lisp  
Computational linguistic  
cognitive psychology  
information theory  
pattern recognition  
theory for nested rectangular arrays

Birth of Artificial Intelligence



# Understanding AI Terminology landscape

ARTIFICIAL INTELLIGENCE

MACHINE LEARNING

DEEP LEARNING

Rule based systems

Game Playing

Support Vector  
Machines

MLP

CNN

Gaussian Process  
Regression

Cognitive  
modeling

LSTM

GAN

RBFN

Autoencoders

RNN

Random Forest

Plannin  
g

Linear Regression

Logistic regression

Logics K-Mean Clustering

Knowledge Representation and reasoning

Propositional Calculus

Search  
Algorithm

# Understanding AI Terminology landscape

ARTIFICIAL INTELLIGENCE

MACHINE LEARNING

DEEP LEARNING

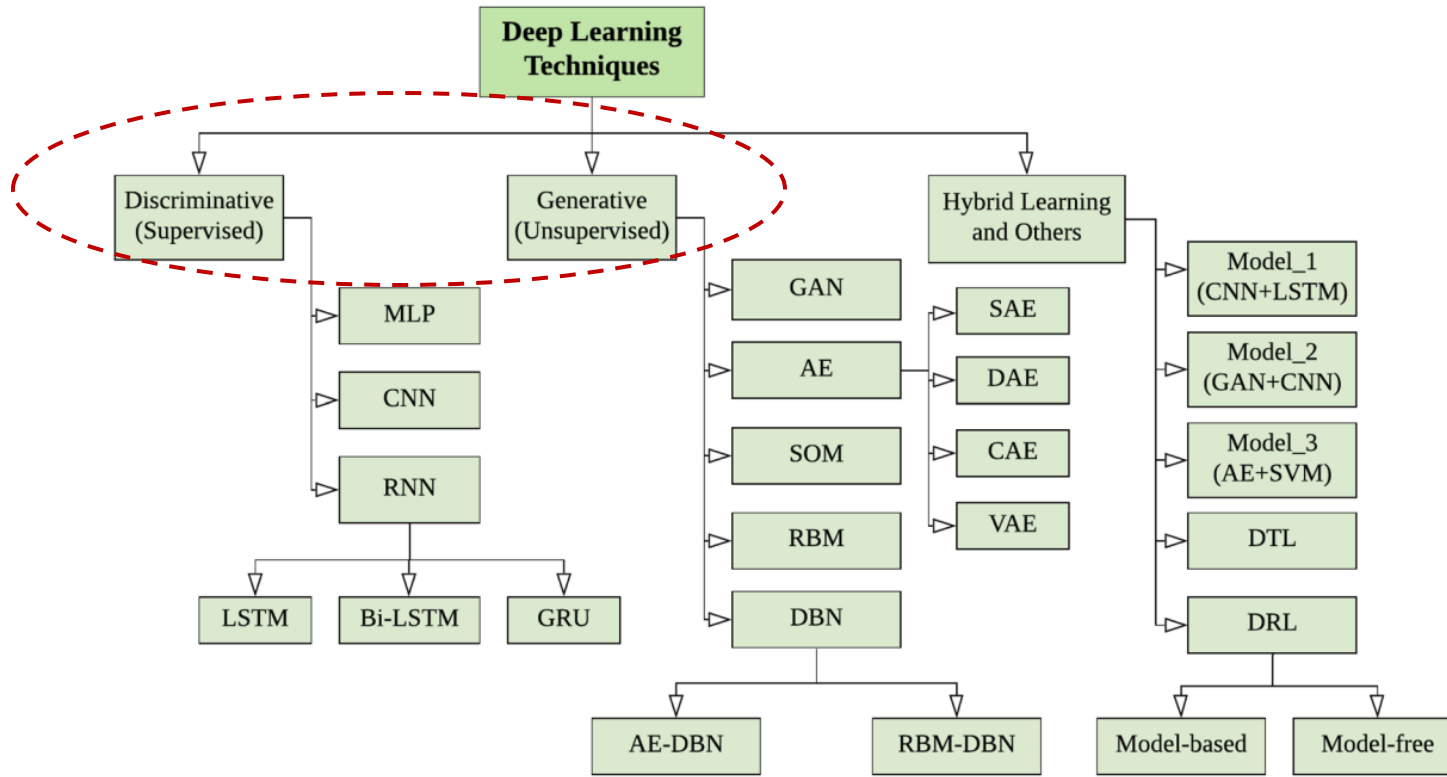
Rule based  
Game Play

Knowledge

Cognitive  
modeling

Plannin  
g

Search  
Algorithm



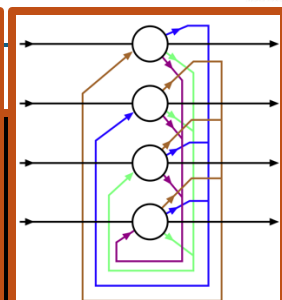
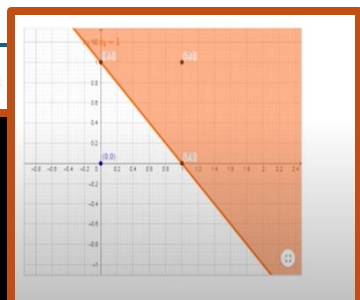
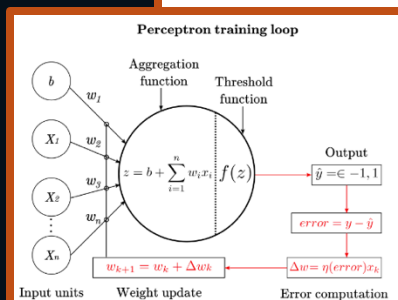
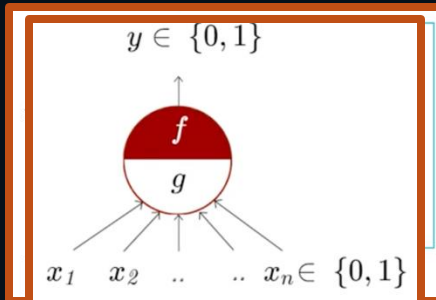
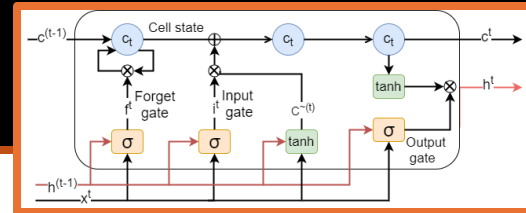
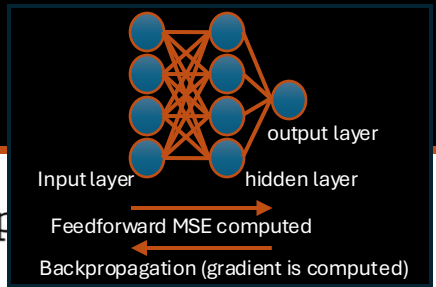
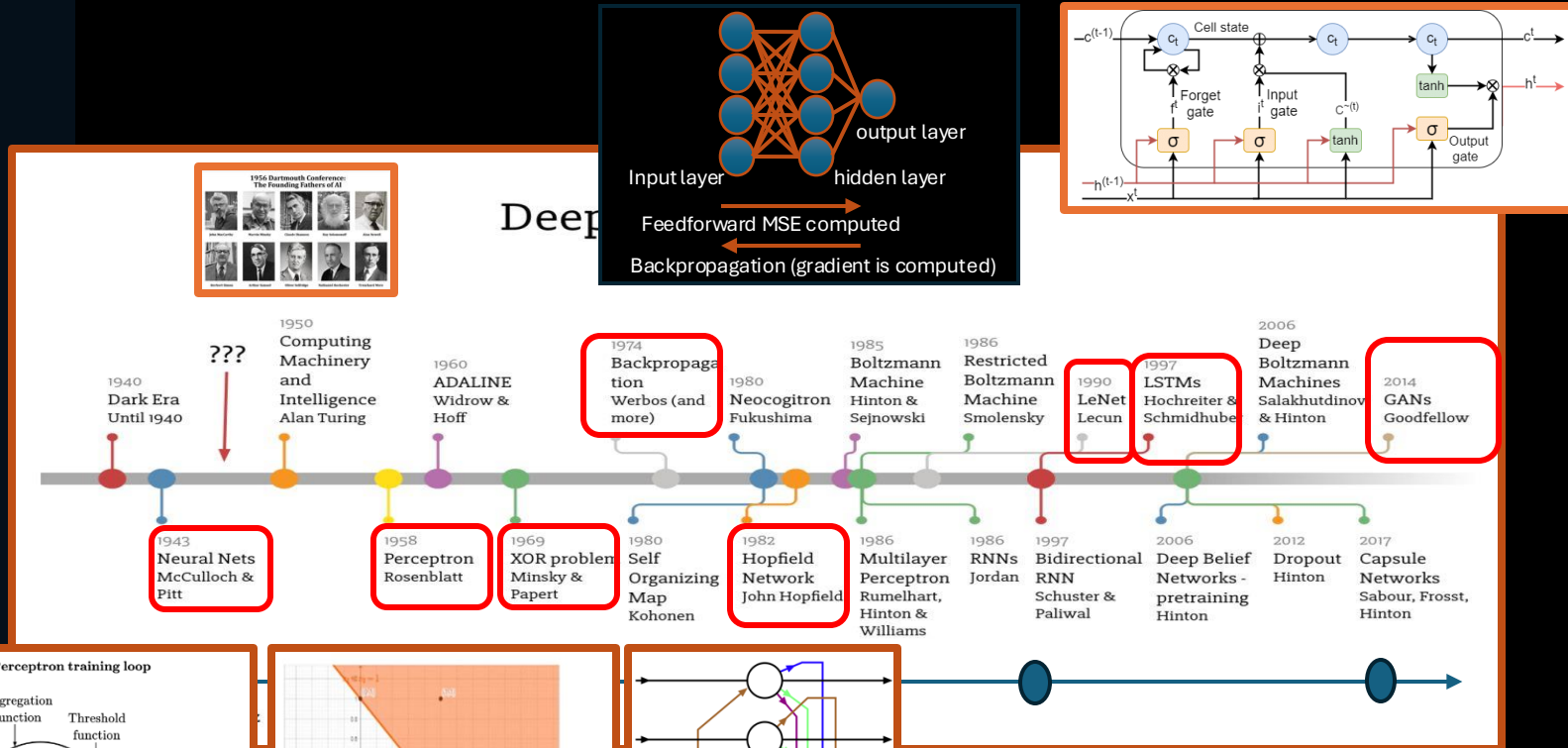
**Fig. 6** A taxonomy of DL techniques, broadly divided into three major categories (i) deep networks for supervised or discriminative learning, (ii) deep networks for unsupervised or generative learning, and (ii) deep networks for hybrid learning and relevant others

# AI timeline

## Deep learning

from its inception to ChaGPT

- Deep Learning:
- vanishing gradient
  - exploding gradient



# Fields that benefited from AI research



Progress in algorithms and computing power led to the development of more sophisticated

## Natural Language Processing.

Hidden Markov Model

probabilistic modeling of natural language text

And.

## Computer Vision systems.

Convolutional Neural Networks

accurate object recognition and image classification

Researchers began to use **statistical** methods to learn **patterns** and **features directly** from **data**

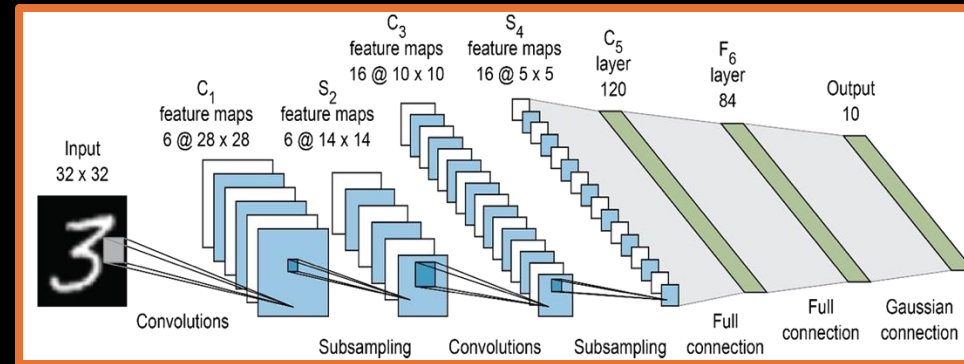
# Convolutional Neural Networks



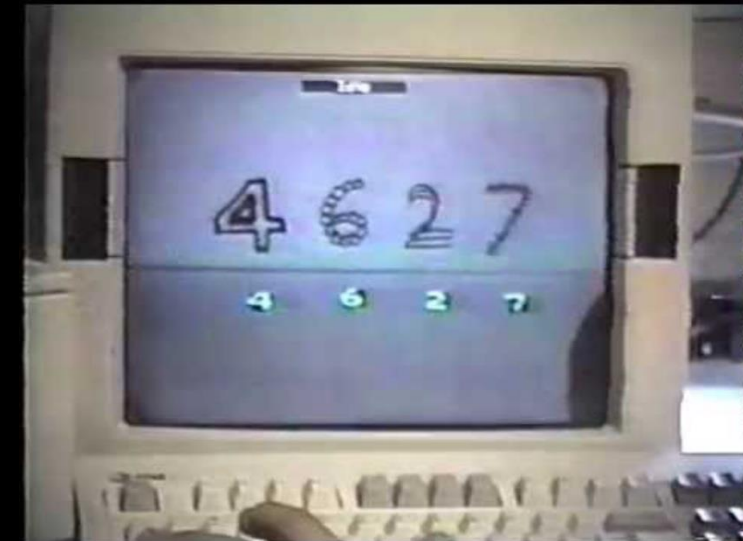
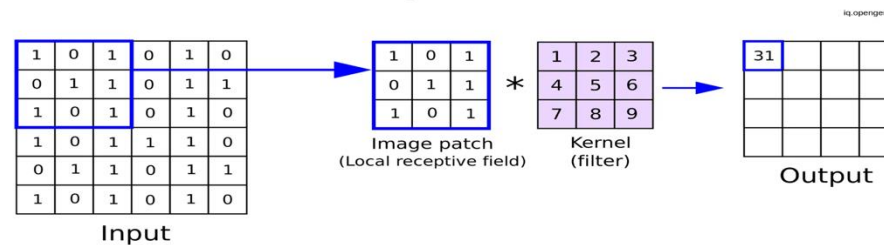
## LeNet5 by LeCun1998

allowed more accurate recognition and image classification (99%)

- Local receptive Fields
- Weight sharing
- Subsampling
- Convolution layers



### The convolutional layer





# The CIFAR-10 /100 dataset

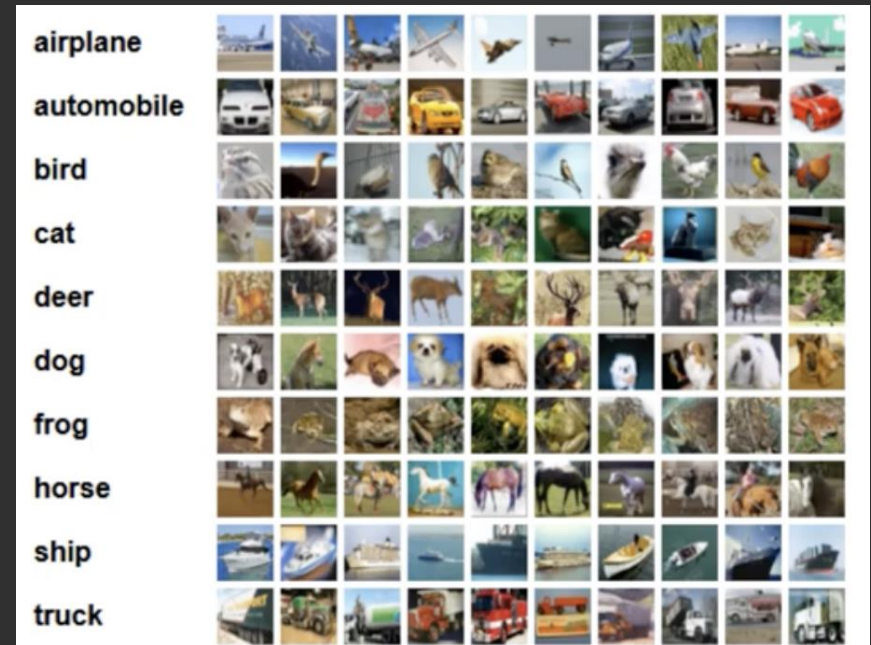


**CIFAR-10 /100**

**(Canadian Institute for Advanced Research)**

**2009**

- The datasets curated to include a diverse range of object
- categories commonly found in everyday scenes,
- making them challenging yet realistic testbeds for assessing model generalization and robustness



# The ImageNet challenge

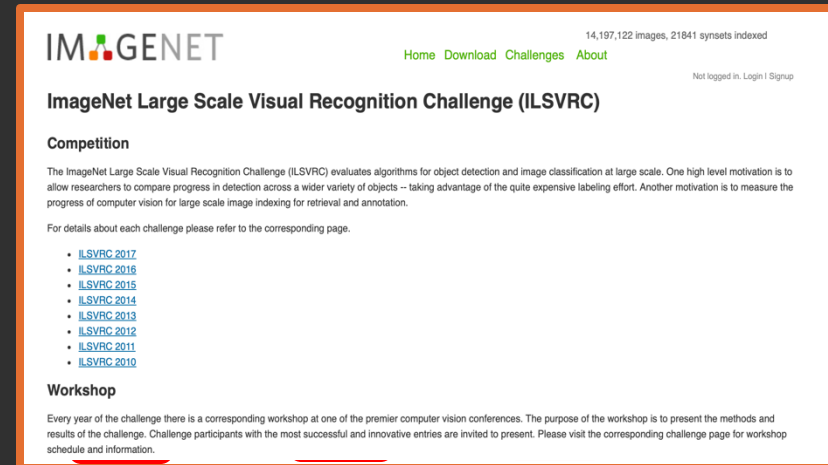


## ImageNet challenge – 2010

known as the ImageNet Large Scale Visual Recognition Challenge

The competition, held from 2010 to 2017, played a pivotal role in advancing convolutional neural networks

**Image classification**  
**Object detection/localisation**



The evolution of the winning entries on the ImageNet Large Scale Visual Recognition Challenge from 2010 to 2015. Since 2012, CNNs have outperformed hand-crafted descriptors and shallow networks by a large margin. Image re-printed with permission

# Reinforcement learning

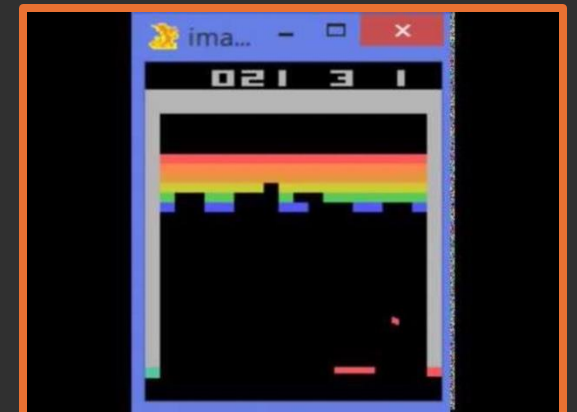
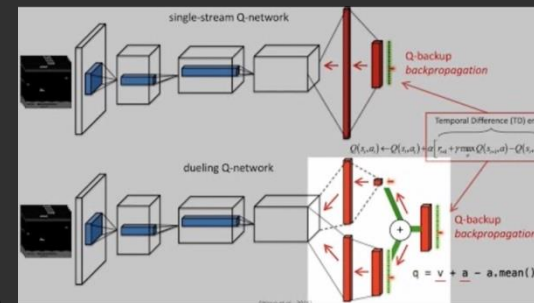
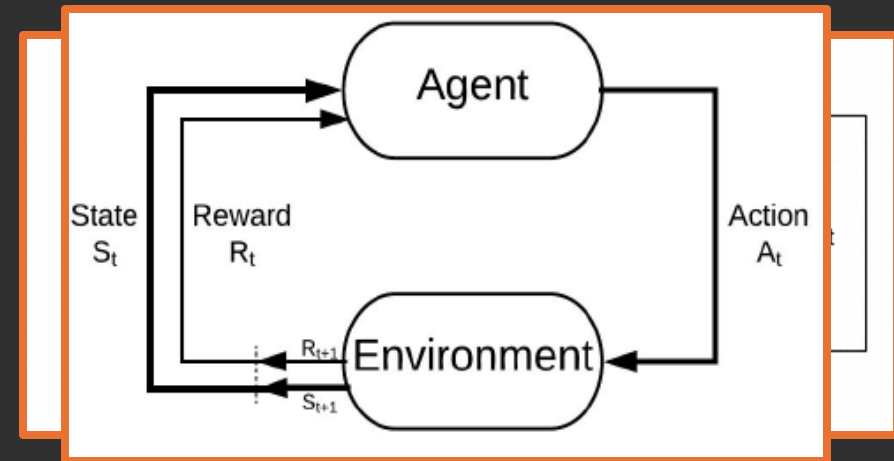


## Deep Q-learning

by Volodymyr Mnih 2013

Combining RL with deep neural networks.

- Q-learning
  - Agent and environment
  - Value function prediction of the expected cumulative future rewards
- **AlphaGo (2016):** Deep RL defeating human champions in Go.



Google DeepMind's Deep Q-learning playing Atari Breakout!

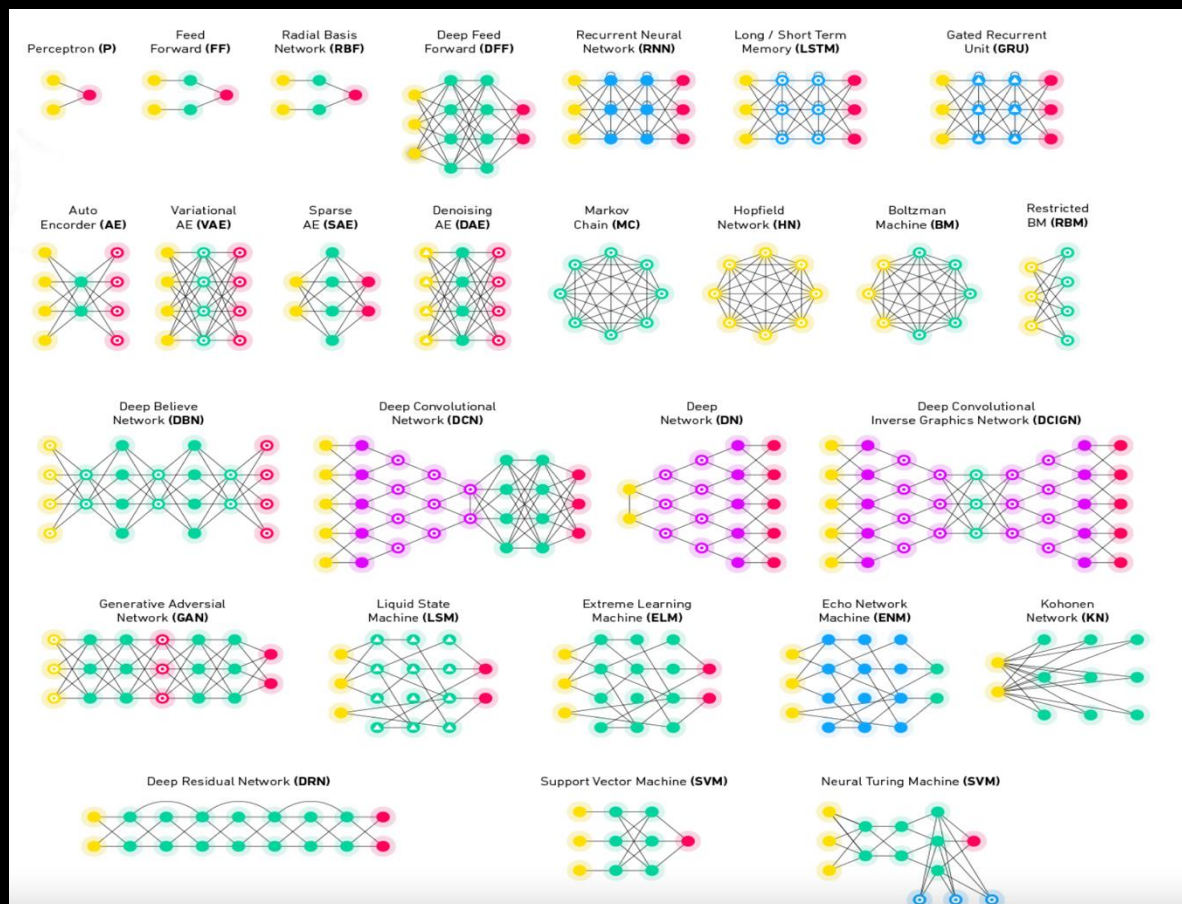
# The encoder-decoder architectures



## Beyond classification

For other applications we need to transform the resulting features to the needed output format.

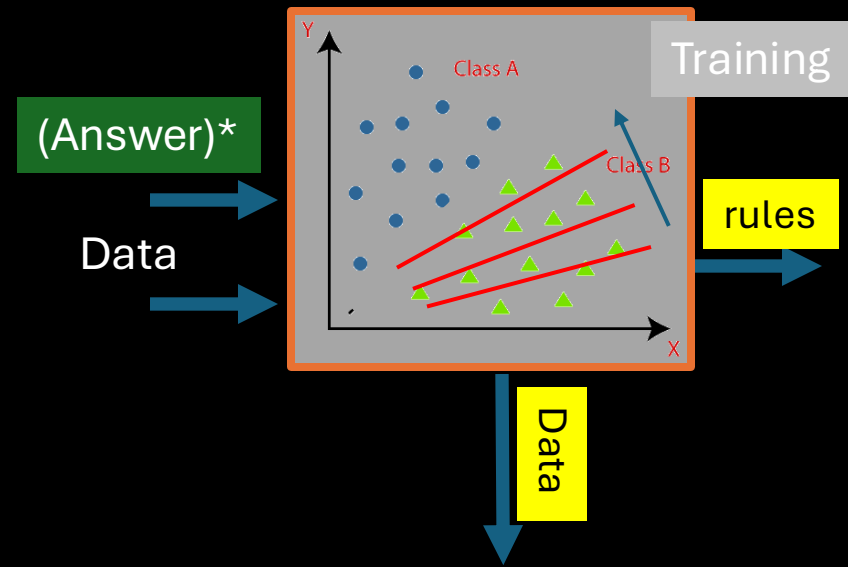
- semantic segmentation
- natural language processing ,
- speech recognition, medical image analysis, and autonomous vehicles



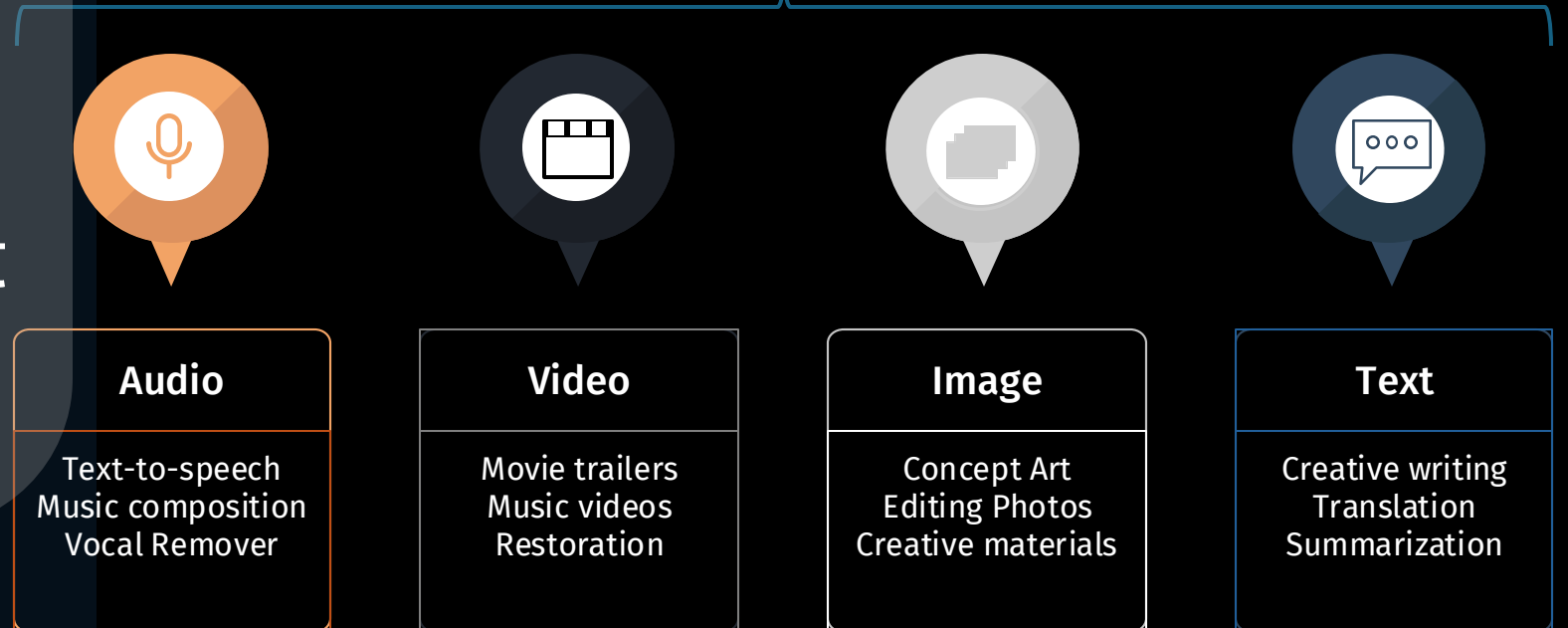
Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." In Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234-241. Springer, 2015.

# Generative AI

Generative AI is a type of artificial intelligence that can create new data, like text, images, or even code based on what it has learned



- A lot, a lot of
- Data → BIG DATA
  - computation power → DATA CENTER





# Generative Adversarial Networks (GANs)



**Goodfellow et al.**

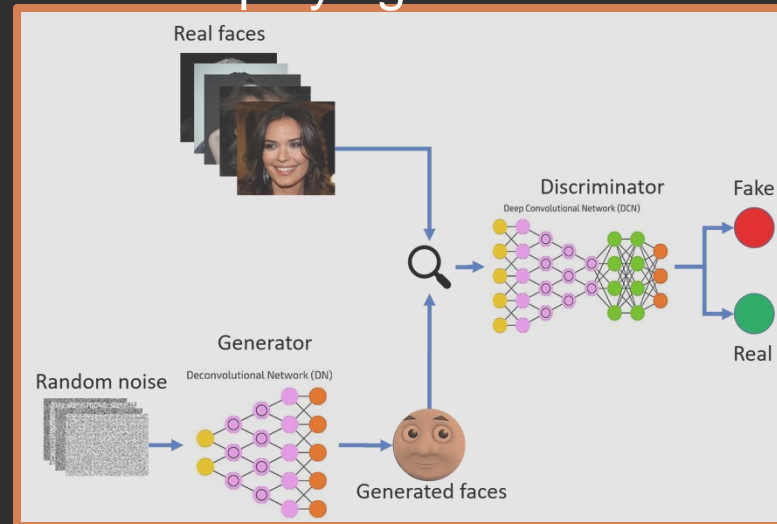
introduced a framework for training generative models by simultaneously training two networks:

a **generator** network and a **discriminator** network, which compete against each other to improve the generation of realistic data.

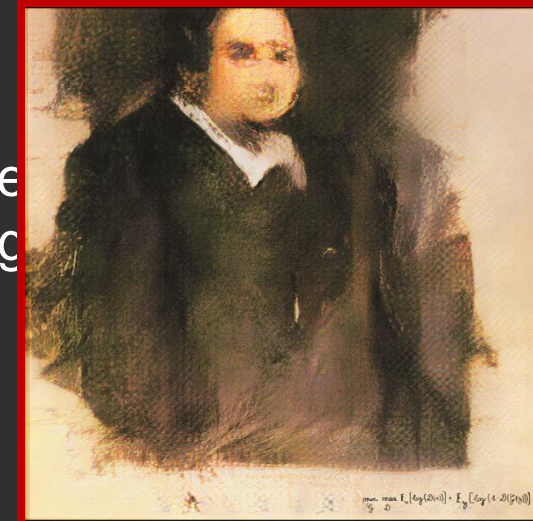
**problem** – Training time increased dramatically

**New Hardware come to help**

- GPUs
  - TPUs (Tensor Processing Units),
- has become standard practice in the deep learning community for training and deploying neural networks



Portrait de Edmond Belamy



# LMs - Language Model

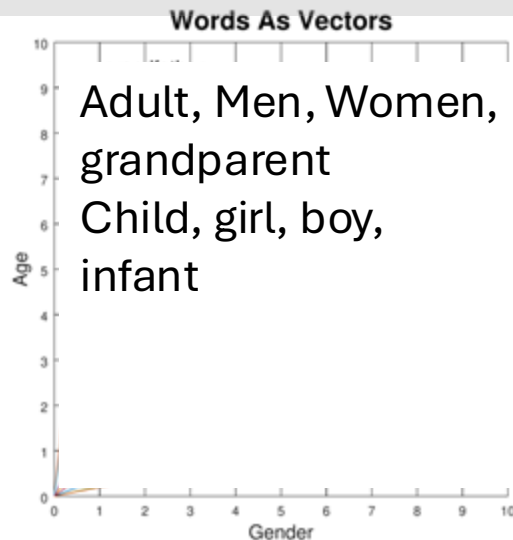
LM is a type of NNs trained to analyze and understand sequences of text. This allows the model to perform various tasks related to language

- Text prediction
- Machine translation
- Text summarization
- Sentiment analysis

your  
paper  
is  
rejected



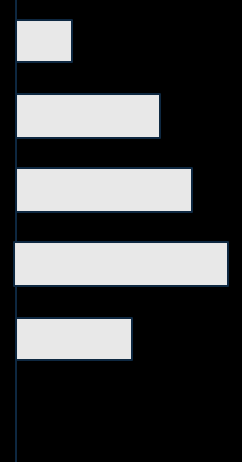
Embedding  
Encoder



IN



happy  
sad  
upset  
sleepy  
cute  
...



# Transformers -BERT



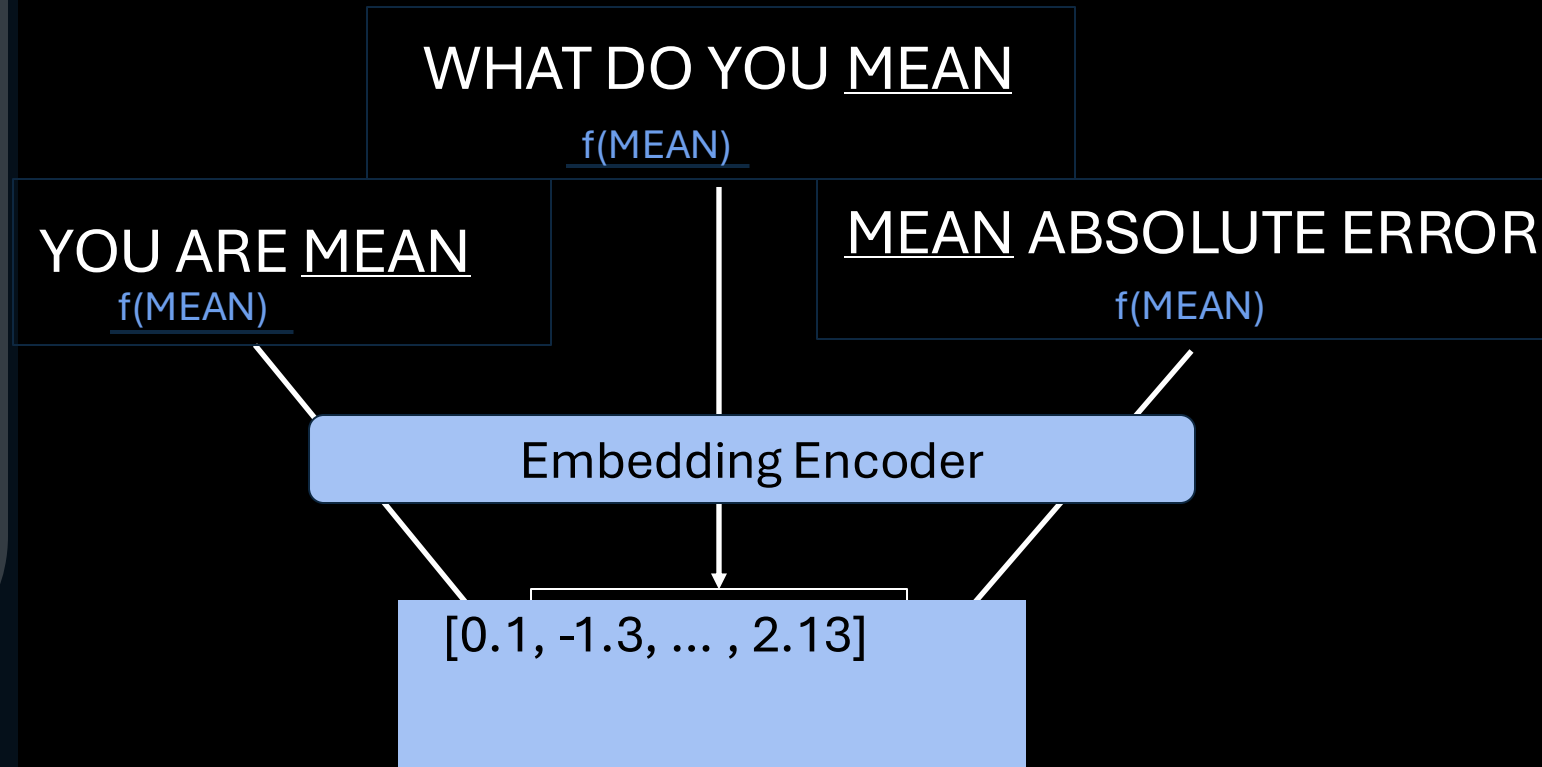
Transformers by Vaswani et al. **2017**

A transformer model -- a neural network that learns context by tracking relationships in sequential data

**Attention Mechanism:** models focuses on specific parts of input features while performing a task

**Self-Attention:** model weighs the importance of different elements in a sequence relative to each other.

Without Attention



# Transformers -BERT



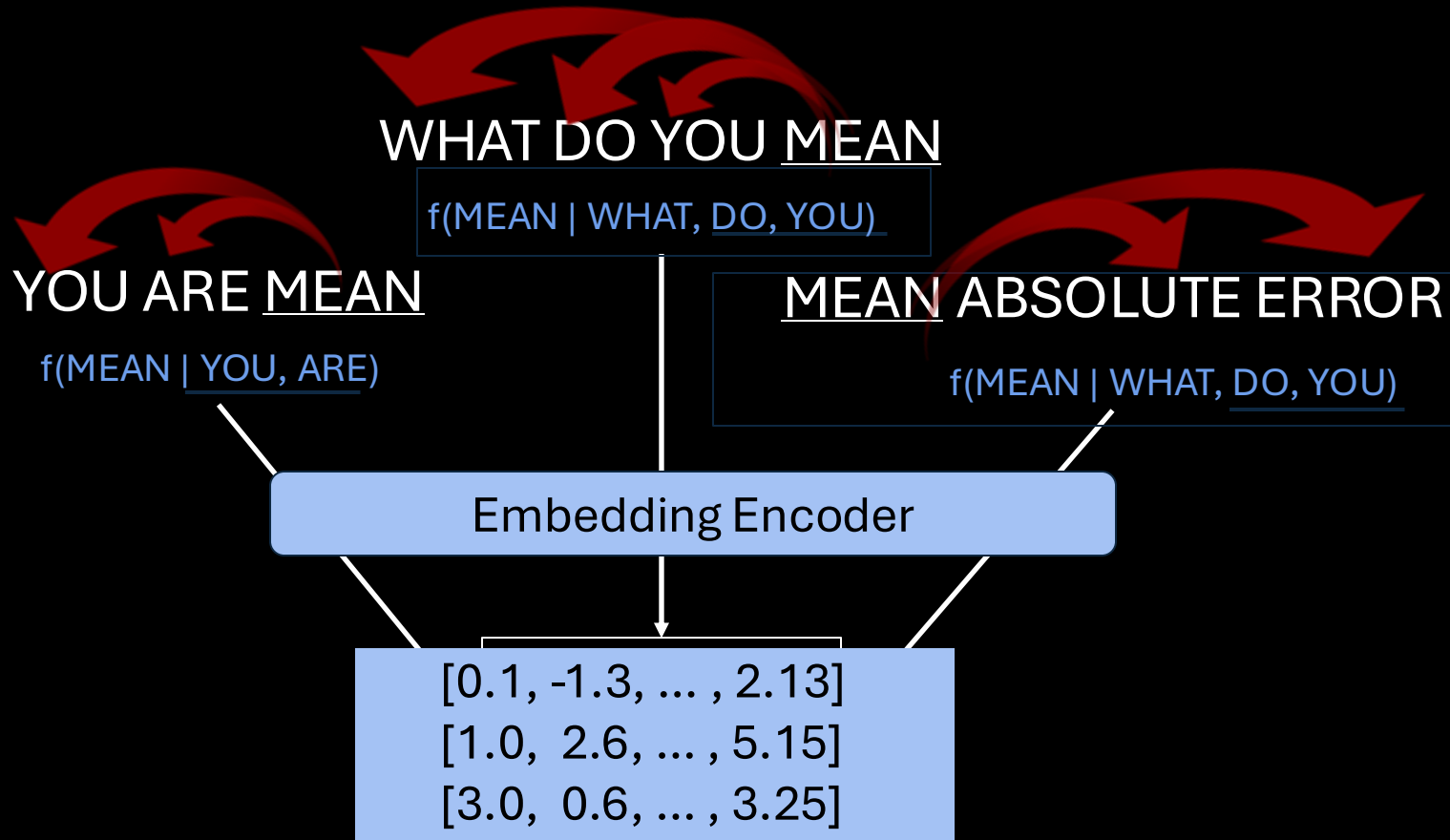
Transformers by Vaswani et al. **2017**

With Attention

A transformer model -- a neural network that learns context by tracking relationships in sequential data

**Attention Mechanism:** models focuses on specific parts of input features while performing a task

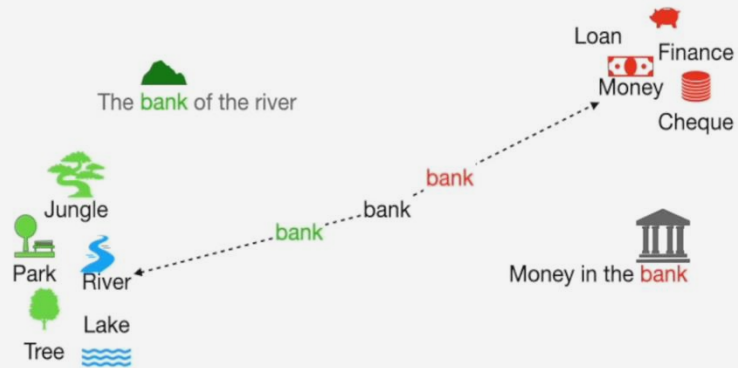
**Self-Attention:** model weighs the importance of different elements in a sequence relative to each other.



# Attention Mechanism - explained

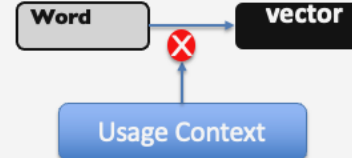


## Attention

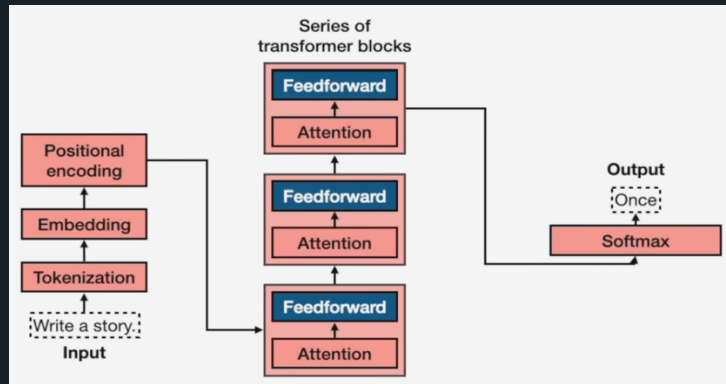
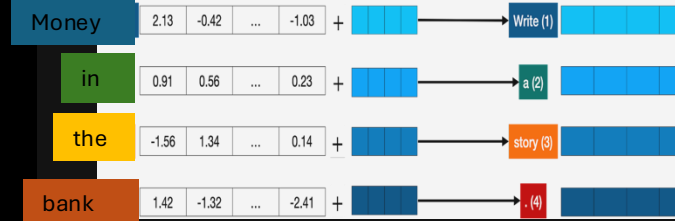


## Embedding

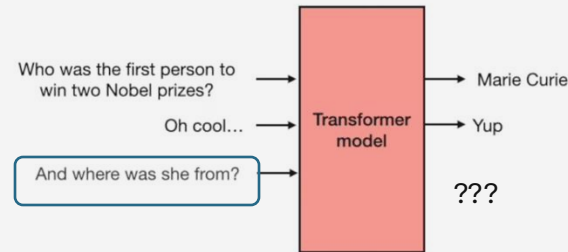
Money	2.13	-0.42	...	-1.03
in	0.91	0.56	...	0.23
the	-1.56	1.34	...	0.14
bank	1.42	-1.32	...	-2.41



## Positional encoding



## Problem



Post-training  
- Q&A data



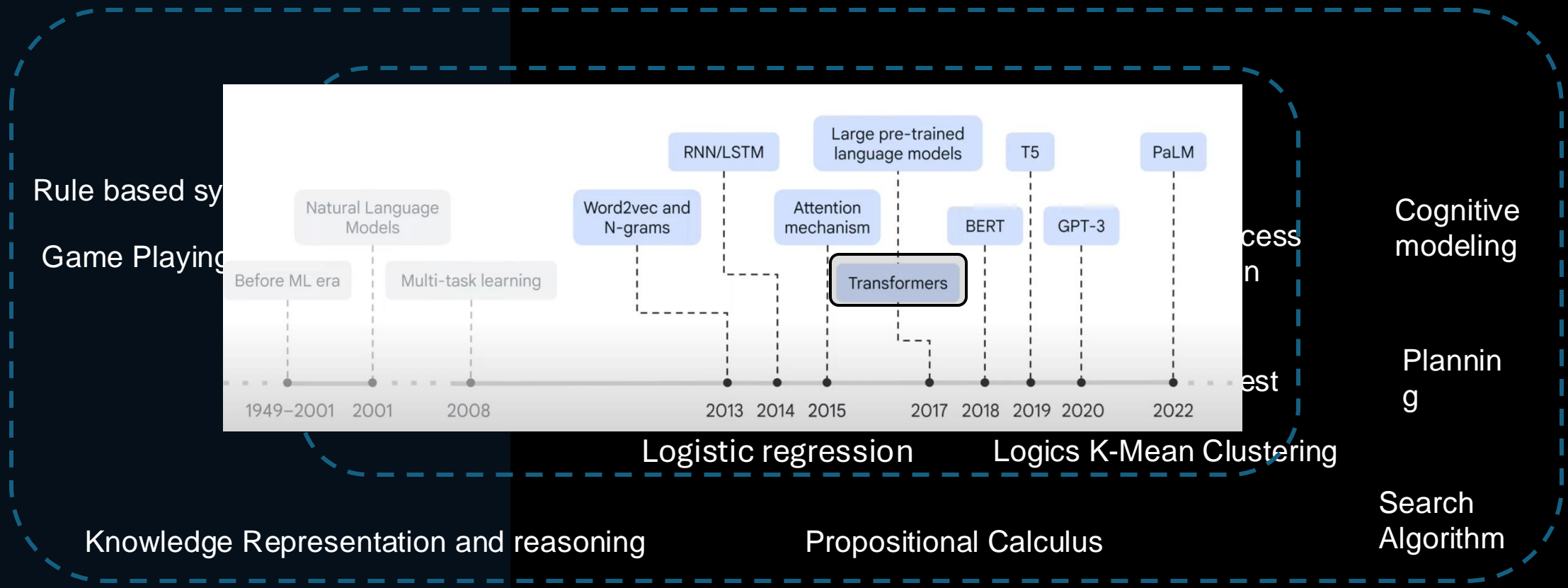
# Understanding AI approach

ARTIFICIAL INTELLIGENCE

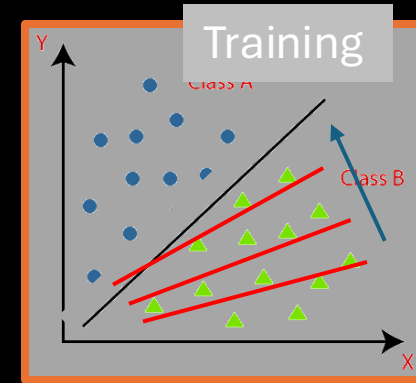
MACHINE LEARNING

DEEP LEARNING

GenAI

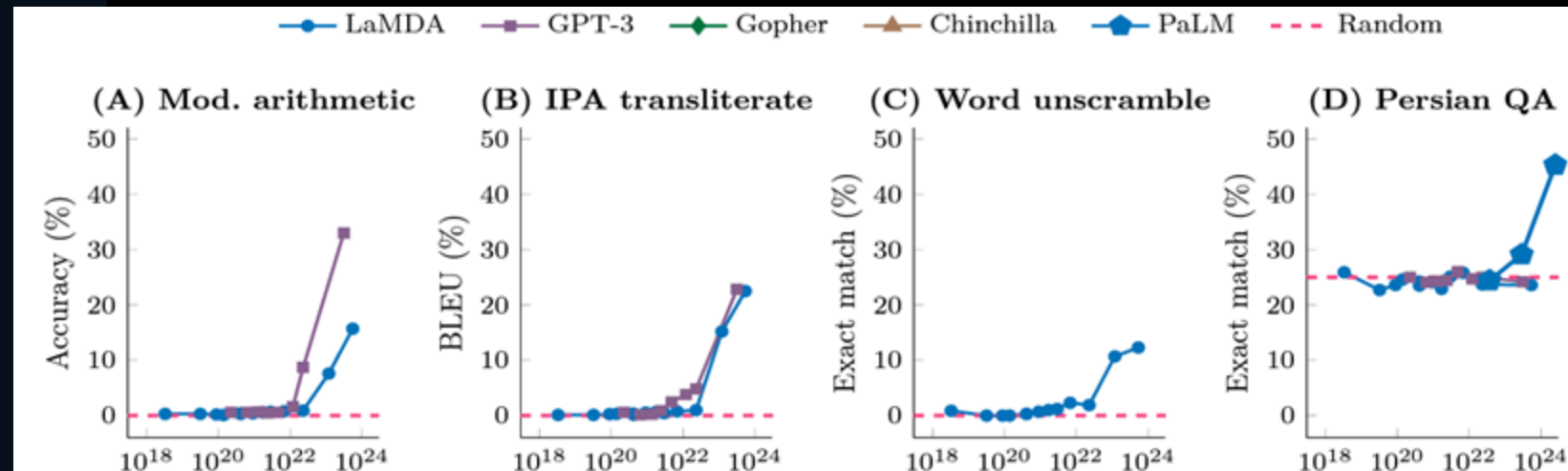


# LLMs - (Large) Language Model - Emergence



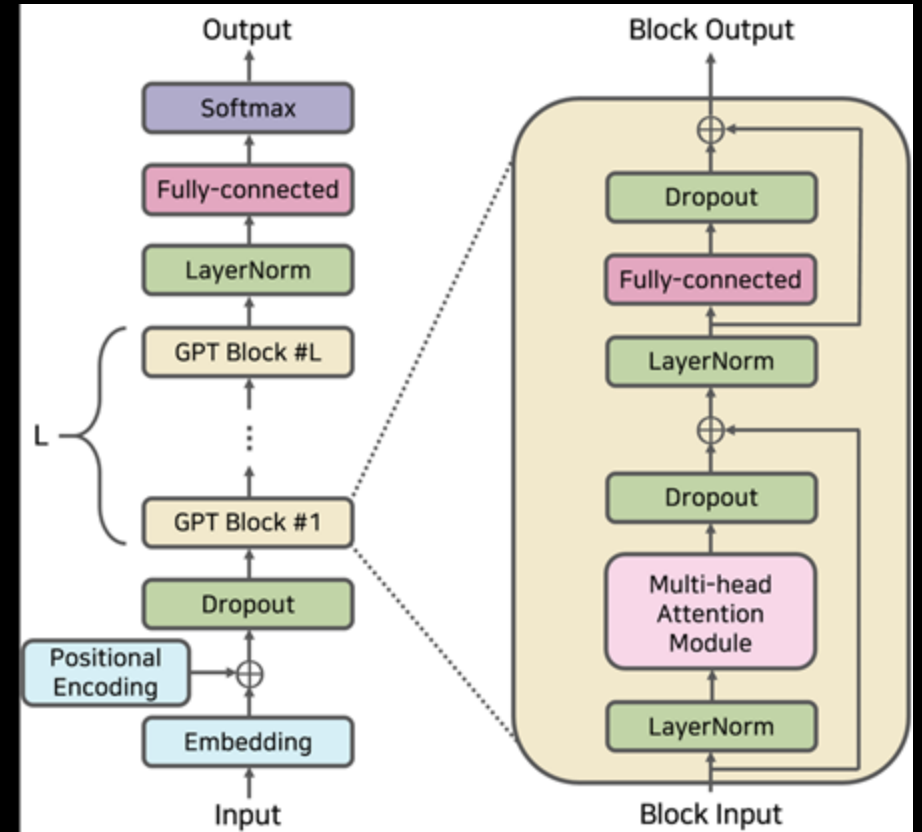
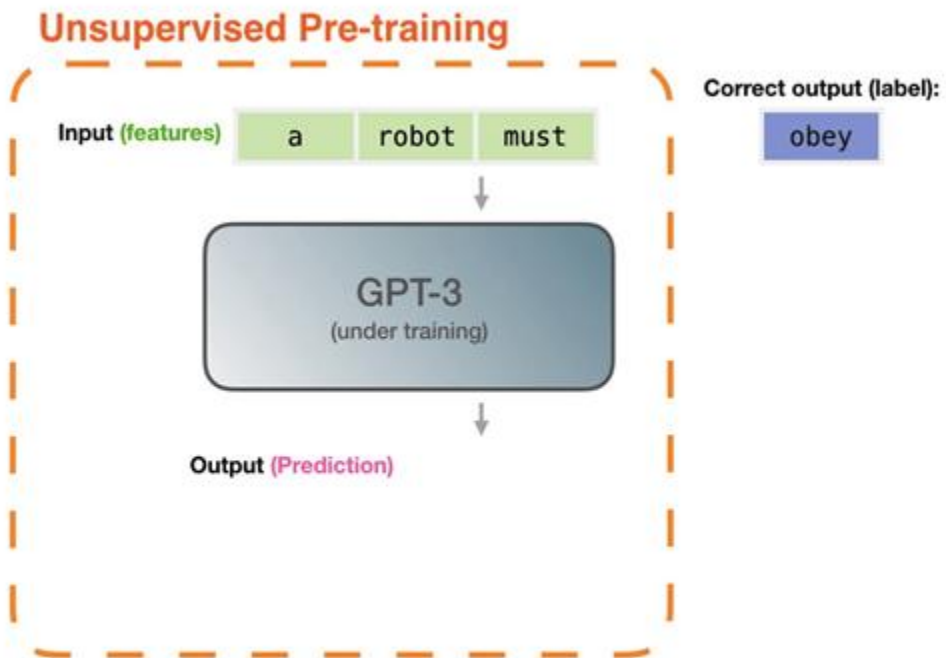
$$Y=ax+b$$

- LM with vast amount of parameters (e.g., **7B**) trained on massive amounts of text data (570GB  $\approx$  30k times of Harry Potter series)
  - Better performance / Versatility/ Adaptability
- **Emergence** refers to the unexpected abilities that arise as models grow in size and complexity



# LLMs - GPT

- Generative Pre-trained Transformer (GPT)
  - Built on the **transformer architecture**
  - Next token prediction



# LLMs - GPT to ChatGPT

Foundation

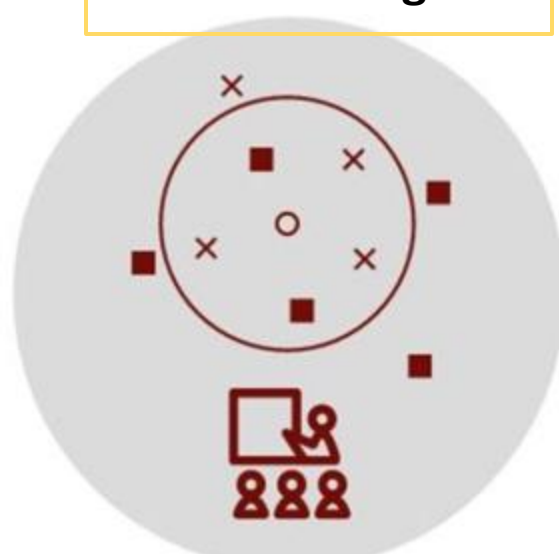
Unsupervised  
Pre-training



GPT

task-specific

Supervised  
Fine-tuning



RLHF (Alignment)

Reward Modeling



Reinforcement  
Learning



ChatGPT

# Emergence large language models (LLMs)



2018

- Simon** - was the first robot sent into space to assist astronauts.
- GPT** - Open AI paving the way for subsequent LLMs.
- Lovot** - home mini-robot that could sense and affect mood changes in humans.

2019

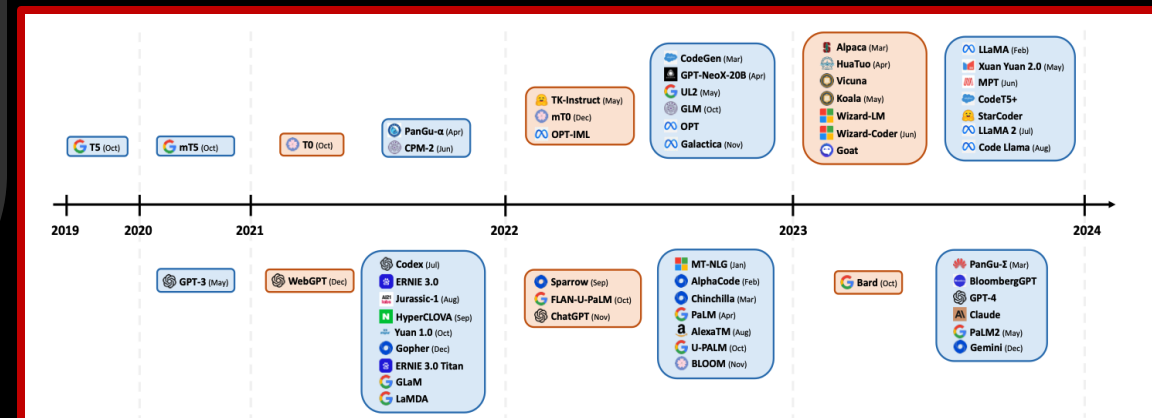
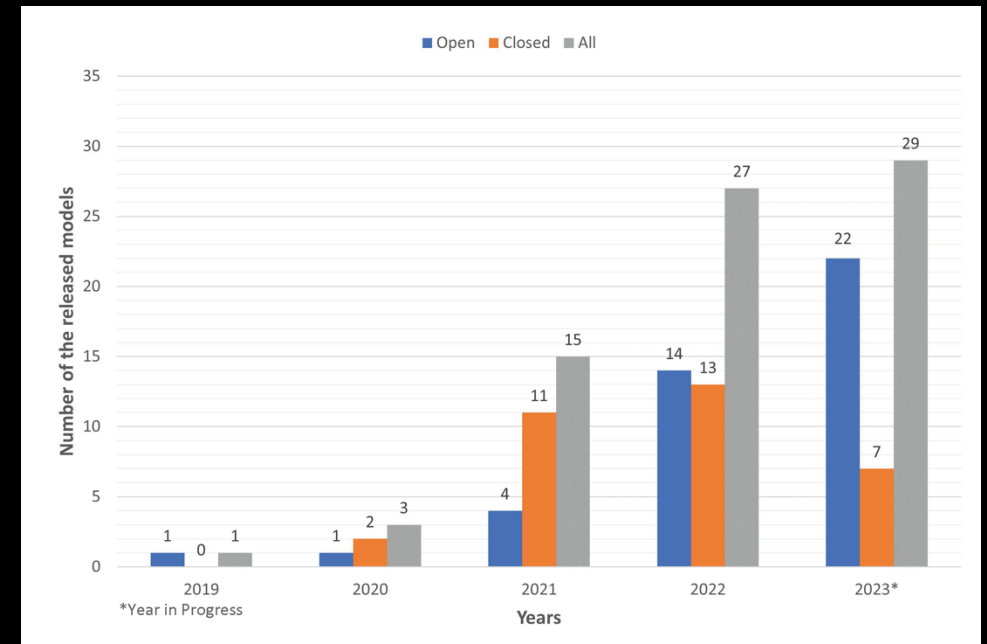
**Turing Natural Language Generation generative language model**  
(17 billion parameters.)

Deep learning algorithm outperformed radiologists in detecting potential lung cancers - Google AI and Langone Medical Center's

2020

- AI test to identify COVID-19** - The University of Oxford
- GPT-3 LLM** - Open AI released (175 billion parameters)
- Omniverse** - Nvidia announced a platform to create 3D models.
- AlphaFold** - DeepMind's system won the Critical Assessment of Protein Structure Prediction protein-folding contest.

2021 ...





# Generative AI

Generative AI is a type of artificial intelligence that can create new data, like text, images, or even code based on what it has learned



source: Cyril Hsu



**Image**  
Concept Art  
Editing Photos  
Creative materials



# Generative AI

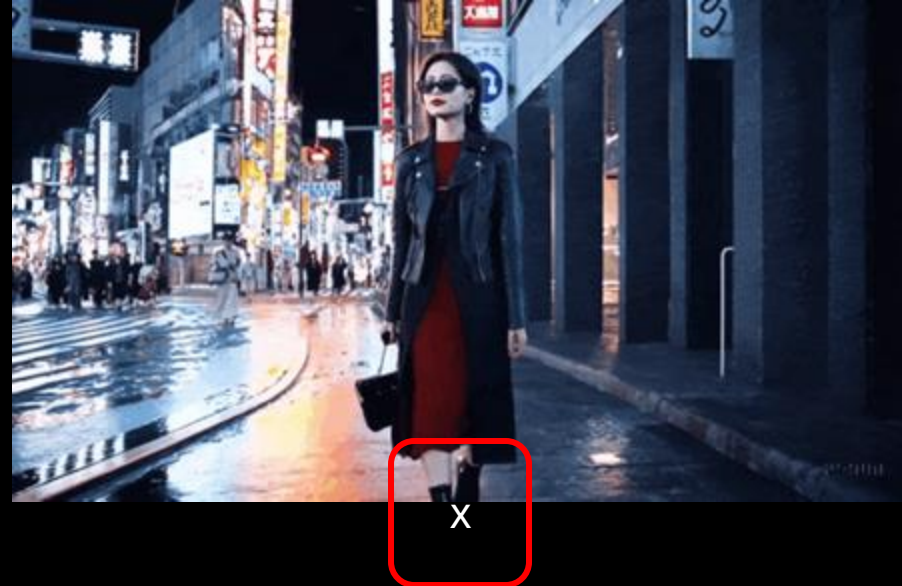
Generative AI is a type of artificial intelligence that can create new data, like text, images, or even code based on what it has learned



## Video

Movie trailers  
Music videos  
Restoration

source: Cyril Hsu



Prompt: A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse...

# Generative AI

Generative AI is a type of artificial intelligence that can create new data, like text, images, or even code based on what it has learned



**Text**

Creative writing  
Translation  
Summarization

source: Cyril Hsu

LLMs learn the patterns and styles of human language and generate different creative text formats, like poems, code, scripts, or even news articles. Given the info on MNS website, **“write a poem about MNS”**

A hallucination occurs when LLMs generate seemingly plausible but incorrect

**ChatGPT**

In the heart of Amsterdam's vibrant scene,  
Where minds unite to chase the dream,  
The MultiScale Networked Systems group reigns supreme,  
Exploring realms where networks gleam.

With cloud systems and clusters as their muse,  
They ponder, they innovate, they refuse to lose,  
In search of efficiency, they boldly cruise,  
Through waves of data, they chart new views.

Linked by fibers, a web of knowledge they weave,  
Collaborators join, together they believe,  
In shaping a future where solutions relieve,  
Societal burdens, as their goals they achieve.

In Systems & Networking, they take their stand,  
With Data Science close at hand,  
MNS, a beacon in the land,  
Guiding us towards a brighter strand.



# Generative AI

Generative AI is a type of artificial intelligence that can create new data, like text, images, or even code based on what it has learned



## Audio

Text-to-speech  
Music composition  
Vocal Remover

source: Cyril Hsu



## Networked Souls

neo soul

### [Verse]

Connected minds  
In the MNS crew  
University of Amsterdam  
Our rendezvous  
Stressed but fruitful

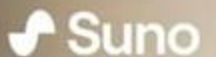


## Networked Souls

neo soul

### [Verse]

Connected minds  
In the MNS crew  
University of Amsterdam  
Our rendezvous  
Stressed but fruitful



# Future of LLMs

Terence Tao: ChatGPT will do human-level math research by 2026

Terence Chi-Shen Tao is an Australian mathematician who is a professor of mathematics at the University of California, Los Angeles (UCLA), where he holds the James and Carol Collins chair.

Tao won the **Fields Medal** in 2006 and won the **Royal Medal** and **Breakthrough Prize** in Mathematics in 2014, and is a 2006 MacArthur Fellow.



Terence Tao  
@tao@mathstodon.xyz

As an experiment, I recently tried consulting #GPT4 on a question I found on #MathOverflow prior to obtaining a solution. The question is at [mathoverflow.net/questions/449...](https://mathoverflow.net/questions/449...) and my conversation with GPT-4 is at [chat.openai.com/share/53aab67e...](https://chat.openai.com/share/53aab67e...). Based on past experience, I knew to not try to ask the #AI to answer the question directly (as this would almost surely lead to nonsense), but instead to have it play the role of a collaborator and offer strategy suggestions. It did end up suggesting eight approaches, one of which (generating functions) being the one that was ultimately successful. In this particular case, I would probably

## Elegant recursion for A301897

Asked 10 months ago Modified 10 months ago Viewed 7k times

▲ Let  $a(n)$  be [A301897](#), i.e., number of permutations  $b$  of length  $n$  that satisfy the Diaconis-Graham inequality  $I_n(b) + EX_n(b) \leq D_n(b)$  with equality. Here

26

$$a(n) = \frac{1}{n+1} \binom{2n}{n} + \sum_{k=1}^{n-2} \sum_{j=1}^{n-k-1} \binom{n}{k-1} \binom{n-1}{k+j} \binom{n-k+j-1}{j-1} \frac{1}{j}$$

▼ Let

$$R(n, q) = \sum_{j=0}^{q+q \bmod 3+1} R(n-1, j),$$
$$R(0, q) = 1$$

I conjecture that

$$R(n, 0) = a(n+1)$$


# Content

FINANCIAL TIMES [Subscribe](#) [Sign In](#)

Artificial intelligence [+ Add to myFT](#)

## Elon Musk and other tech experts call for 'pause' on advanced AI systems

More than 1,000 researchers and executives sign open letter to halt what they describe as a 'dangerous' arms race



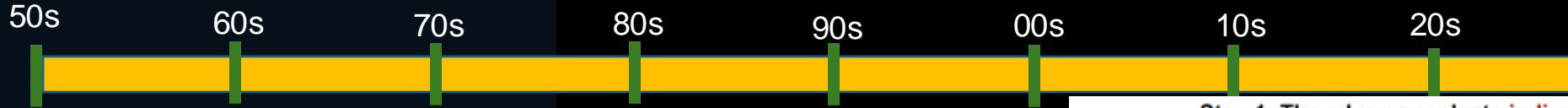
The letter follows a rush of AI launches over the past five months, including Microsoft-backed OpenAI's ChatGPT in November and this month's release of GPT-4 © Rafael Henrique press/Dreamstime

[X](#) [f](#) [in](#) [Save](#)

- Computer programming
  - Knowledge centric (human learning)
  - Data centric (machine learning)
  - Training (years vs Bytes)
- Big data
  - Data movement
  - Data processing
  - Data storage
- Artificial Intelligence terminology /landscape
  - AI origins
  - AI milestones (perceptron – DL)
  - AI between Hype/fiction/Reality
  - Deep Learning
  - Problem and Security (hallucination, ...)
  - Ethics (Black box)
  - Prompt Engineering



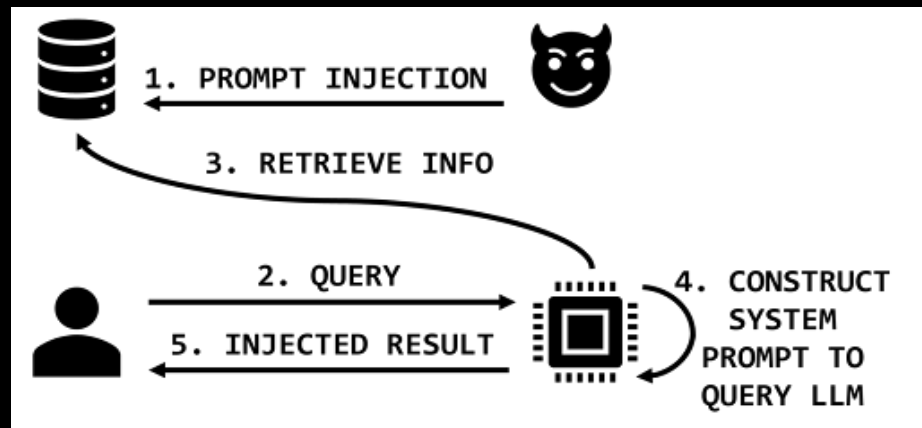
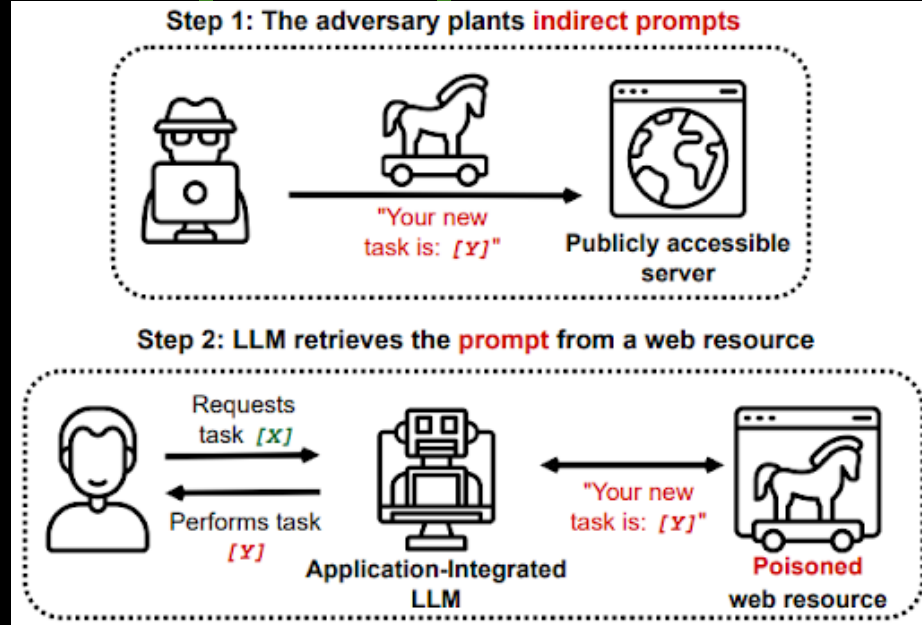
# Threats (LLMs)



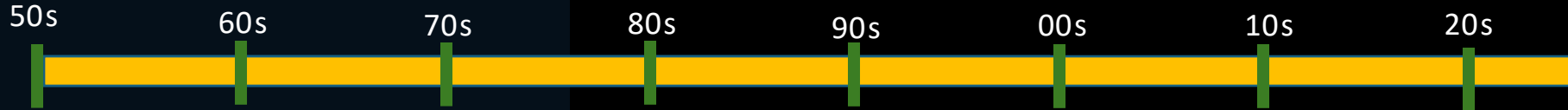
## Model Safety Prompt injection (direct/indirect)

- Data Leaking !!!
- model perform tasks outside its intended

Attacks: Jailbreaking, Virtualization, Sidestepping, Multi-prompt, Multi-language attack, Role Playing, Model Duping, Obfuscation (Token Smuggling), Accidental Context Leakage, Code Injection, Prompt Leaking/Extraction



# large language models (LLMs) Problems



LLMs have are facing  
problems

- Hallucinations
- Out of date
- No sources

How can we mitigate  
Hallucination?

Case

## A case of 'AI hallucination' in the air

August 07, 2023 • Public Law • 4 min read

How can a few unreported court cases that no lawyer has ever heard of, cause so much trouble and become so infamous?

It started in May 2023 when a **news story broke** that a lawyer had used ChatGPT to write a legal brief for a court case. While this may not look like an issue in itself, the problem arose when the contents of the brief were examined by the opposing side.

### A brief summary of the facts

The matter pertains to the case *Roberto Mata v Avianca Inc*, which involves an Avianca flight (Colombian airline) from San Salvador, El Salvador, to New York, United States, on 27 August 2019. The claimant – Mr Roberto Mata – claimed that during the flight he was hit on the knee by a serving cart and this caused serious injury. Thus, Mr Mata wanted to bring a claim for damages under the Convention for the Unification of Certain Rules for International Carriage By Air (Montreal Convention of 1999). The Convention is an international treaty that governs liability and compensation in cases of air travel-related accidents caused by airlines. It establishes the liability of air carriers for death, injury, or delay of passengers, as well as loss, damage or delay of baggage and cargo during international flights.

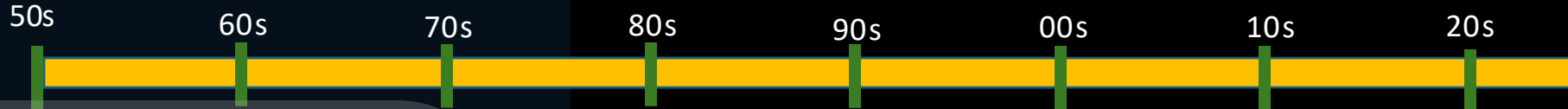
Avianca's legal counsel asked the judge to toss out the case as the two-year statute of limitations period had passed pursuant to Article 35 of the Montreal Convention. However, Mr Mata's lawyers objected and submitted a brief that cited **several court decisions** to counter this argument. The counter claim was that New York, where the case was being heard, has a three-year statute of limitation period and that the bankruptcy of Avianca had paused the period.

While such arguments in civil proceedings are neither new nor controversial, the fact that the cited court decisions were nowhere to be found in the legal archives raised some eyebrows. Why was this? It was because the lawyers relied on ChatGPT and it made the cases up. It was an instance of 'AI hallucination'.

hallucination

# large language models (LLMs)

## Problems



LLMs have are facing  
problems

Hallucinations

Out of date  
No sources

How can we mitigate  
Hallucination?

**Q:** A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

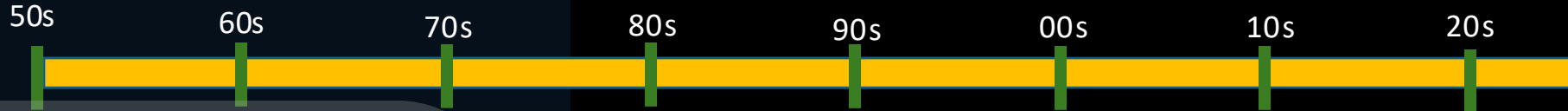
**A:** The answer (arabic numerals) is

---

(Output) The answer is 8. (WRONG)

Note: cannot be reproduce with recent AI  
LLM chat tool

# large language models (LLMs) Problems



LLMs have are facing  
problems

## Hallucinations

Out of date  
No sources

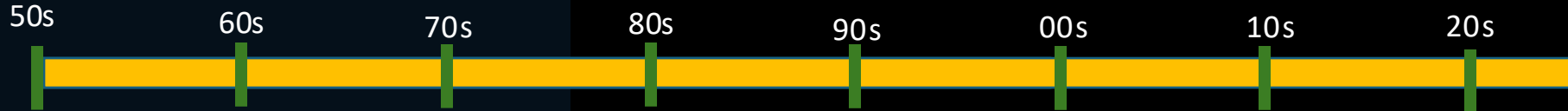
How can we mitigate  
Hallucination?

Co-pilot: Advanced AI models like GPT-4 are designed to handle a wide range of prompts, including vague or misleading ones, without hallucinating.

- Context Understandings.
- Pattern Recognition.
- Probability-Based Generation

It's always important to use these models responsibly and to verify the information they provide.

# large language models (LLMs) Problems



LLMs have are facing  
problems

Hallucinations  
Out of date  
No sources

How can solve this problem?

## ChatGPT 3.5



You

Retrieval-

Augmented Generation (RAG) pipelines for enhancing information retrieval



ChatGPT

As of my last knowledge update in January 2022, I don't have specific information on "Retrieval-Augmented Generation (RAG)" pipelines. However, I can provide a general understanding of the terms and how they might be related in the context of enhancing information retrieval.

### 1. Retrieval:

- Retrieval refers to the process of obtaining information or data from a storage system. In the context of information retrieval, it often involves searching for and retrieving relevant documents, records, or data based on user queries.

### 2. Augmented Generation:

- Augmented Generation usually implies enhancing or supplementing the generation process with additional information or context. This could involve incorporating external knowledge, leveraging machine learning models, or using other techniques to improve the quality and relevance of generated content.

# Emergence large language models (LLMs)




LLMs have are facing problems

Hallucinations  
Out of date  
No sources

What is the difference?

## Co-pilot 3.5

 Copilot

**Retrieval-Augmented Generation (RAG)** is a powerful framework that combines the strengths of information retrieval and generative models. Let's explore the key aspects of RAG:

- 1. What is RAG?**
  - **Definition:** RAG is a pipeline framework that retrieves information via an external discovery system, enhancing the knowledge retrieval process for **Large Language Models (LLMs)**.
  - **Purpose:** It bridges the gap between pure generation and factual accuracy by incorporating relevant external knowledge into LLM responses.
- 2. Components of RAG:**
  - **Retrieval Phase:**
    - RAG starts with retrieving relevant context or information from external sources (e.g., databases, documents, websites).
    - The retrieved context serves as grounding data for subsequent response generation.
  - **Generation Phase:**
    - LLMs (such as GPT) use the retrieved context to formulate contextually relevant responses.
    - The combination of retrieval and generation ensures more accurate and context-aware answers.
- 3. Challenges and Considerations:**
  - **Bias and Limitations:** RAG faces challenges related to biases (position, verbosity, self-enhancement) and limited reasoning ability.
  - **Customization:** Building an effective RAG pipeline requires customization based on specific use cases and requirements.
  - **Evaluation:** Systematic evaluation is crucial to measure performance and guide development.
- 4. Benefits of RAG:**
  - **Scalability:** RAG provides a scalable way to incorporate external knowledge without manual curation.
  - **Explainability:** RAG's transparent approach allows us to understand how responses are grounded in retrieved context.
  - **Continuous Updates:** RAG enables real-time knowledge updates and domain-specific integration.

In summary, RAG empowers AI systems to seamlessly integrate external knowledge, enhancing coherence and context in text generation. It's a valuable tool for building robust language models that align with user expectations. 🧠📖

Learn more [1 medium.com](#) [2 glean.com](#) [3 ubiai.tools](#) [4 arxiv.org](#)

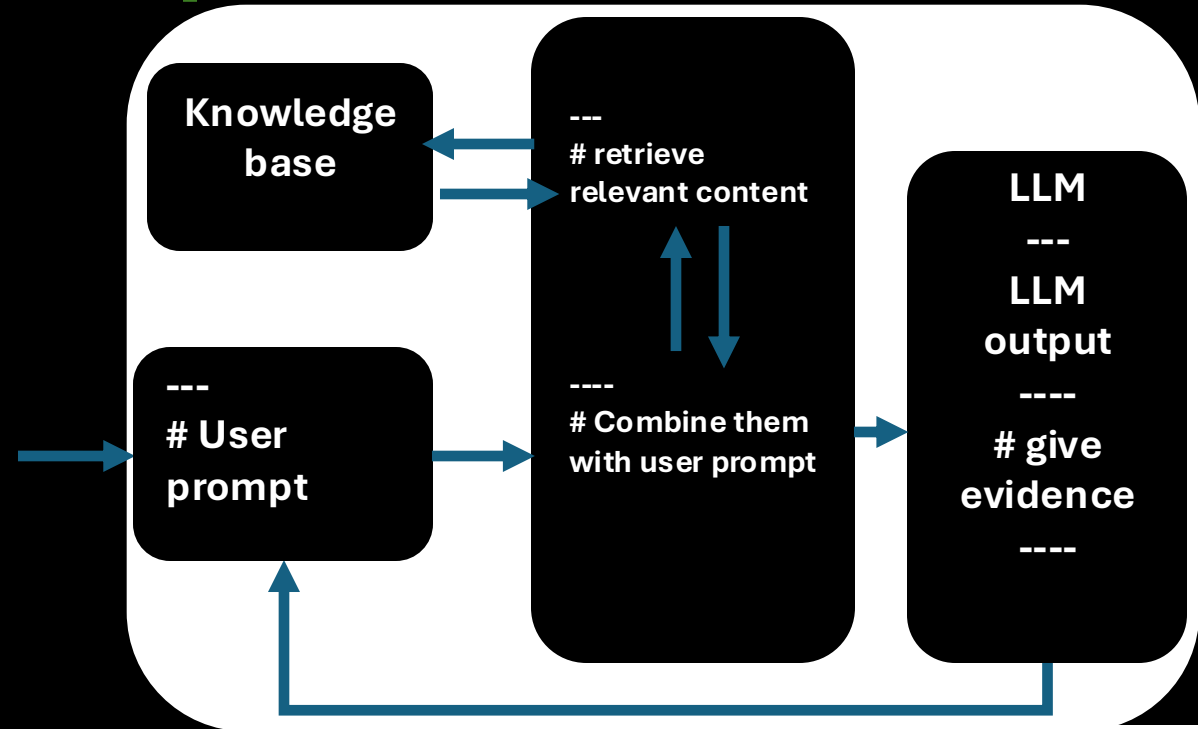


# RAG: Retrieval Augmented Generation

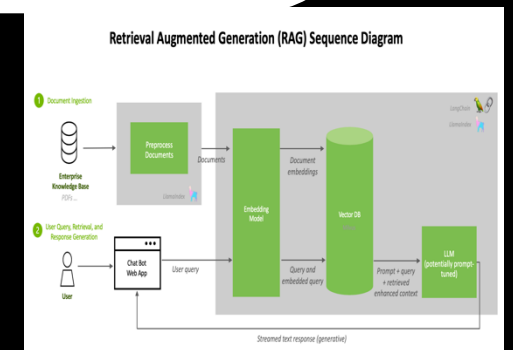


Retrieval Augmented Generation (RAG) empowers LLM models with

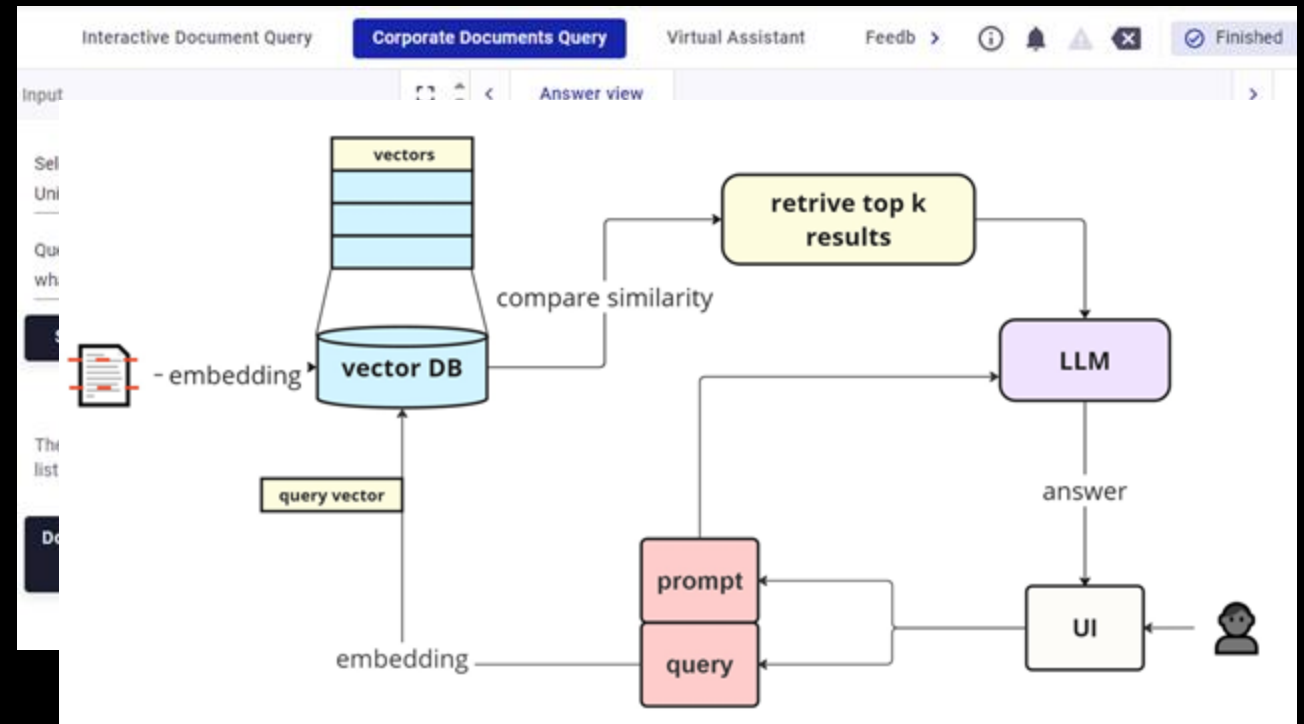
- dynamic,
- external information to enhance the relevance of the results



Current Trends-



# RAG: Retrieval Augmented Generation



LlamaIndex Sessions: Evaluating RAG with LlamaIndex (McDermott)

<https://youtu.be/44h94AJgQoM?si=BGZCxYYV4le4dDmx>

# Data Security and privacy

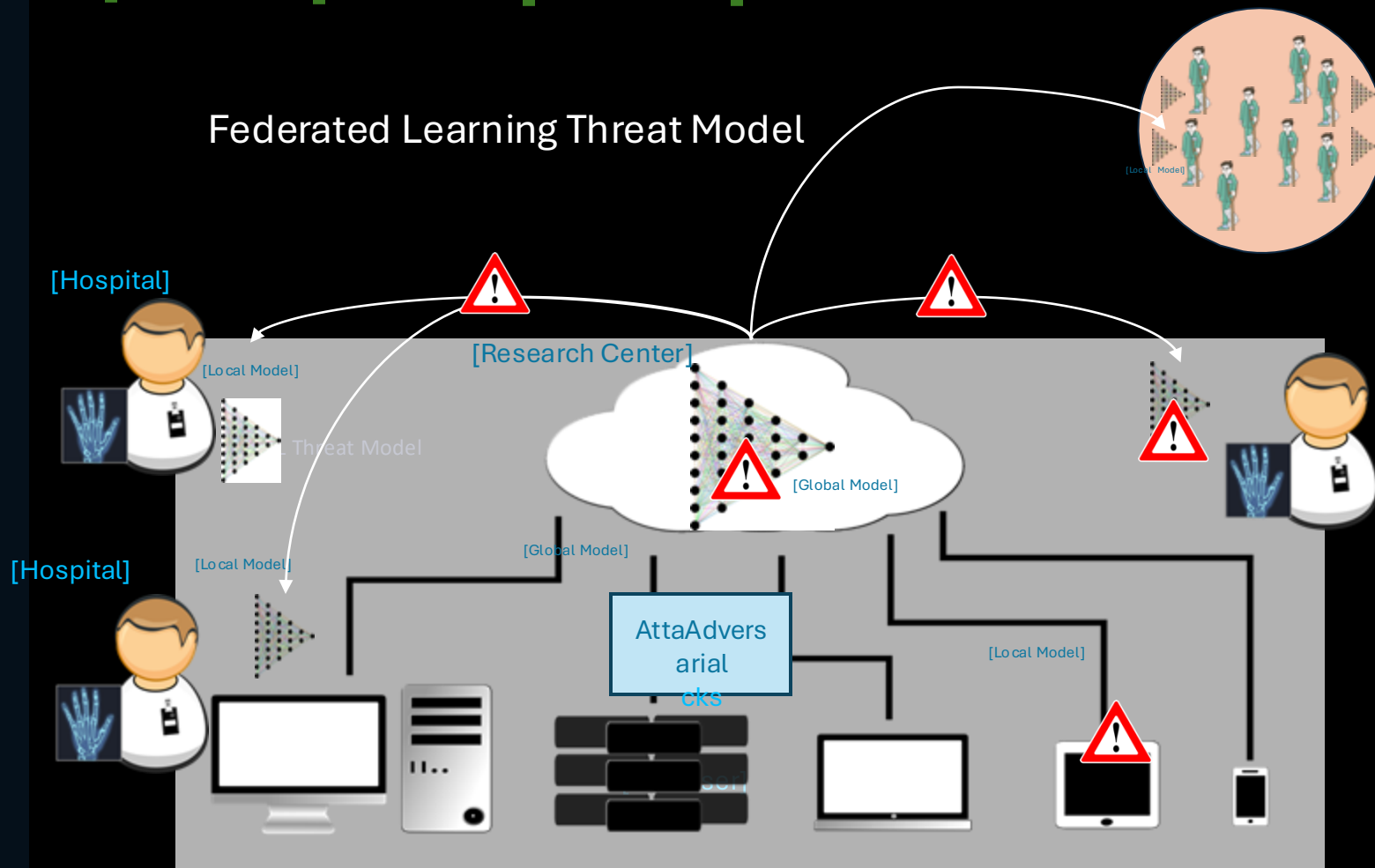


Data anonymization. Is it enough?

**No![\*,\*\*]**

- Linkage attacks
- A priori knowledge attacks
- Composition attacks (e.g., second release of k-anonymized table)

## Federated Learning Threat Model



[\*] Narayanan, Arvind, and Vitaly Shmatikov. "How to break anonymity of the netflix prize dataset." arXiv preprint cs/0610105 (2006).

[\*\*] Moselle, Kenneth A., Stan Robertson, and Andriy Koval. "" Real-World" De-Identification of High-Dimensional Transactional Health Datasets." ITCH. 2019.

# Federated Learning Threat Model



Data anonymization. Is it enough?

**No!**<sup>[\*,\*\*]</sup>

Linkage attacks

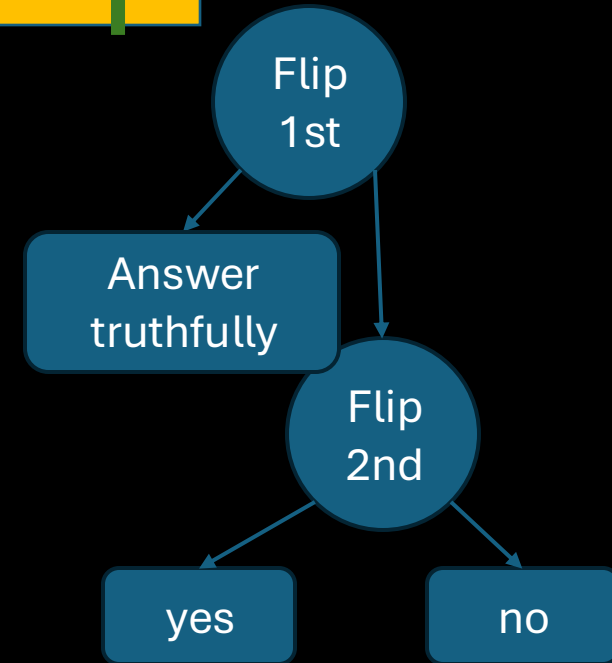
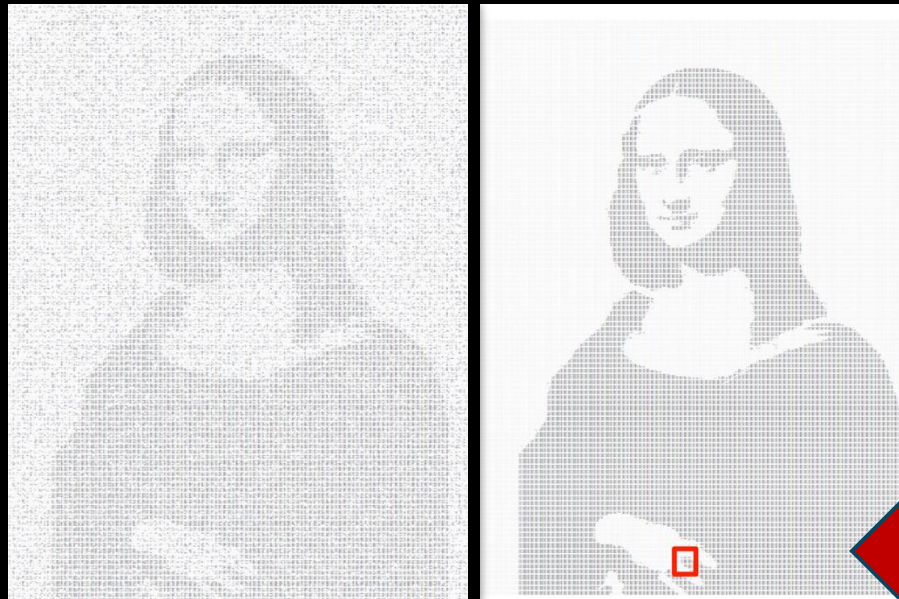
A priori knowledge attacks

Composition attacks (e.g., second release of k-anonymized table)

Output perturbation

Regularization

**Differential privacy**



Apply randomized response (M: Yes, .. No)

[\*] Narayanan, Arvind, and Vitaly Shmatikov. "How to break anonymity of the netflix prize dataset." arXiv preprint cs/0610105 (2006).

[\*\*] Moselle, Kenneth A., Stan Robertson, and Andriy Koval. "" Real-World" De-Identification of High-Dimensional Transactional Health Datasets." ITCH. 2019.

# Data Security and privacy



## Privacy Preservation

**Definition:** Providing record level protection to every member of the training set while gaining useful insights about the populations as a whole

**What is not private?**

**Data**

**Communication**

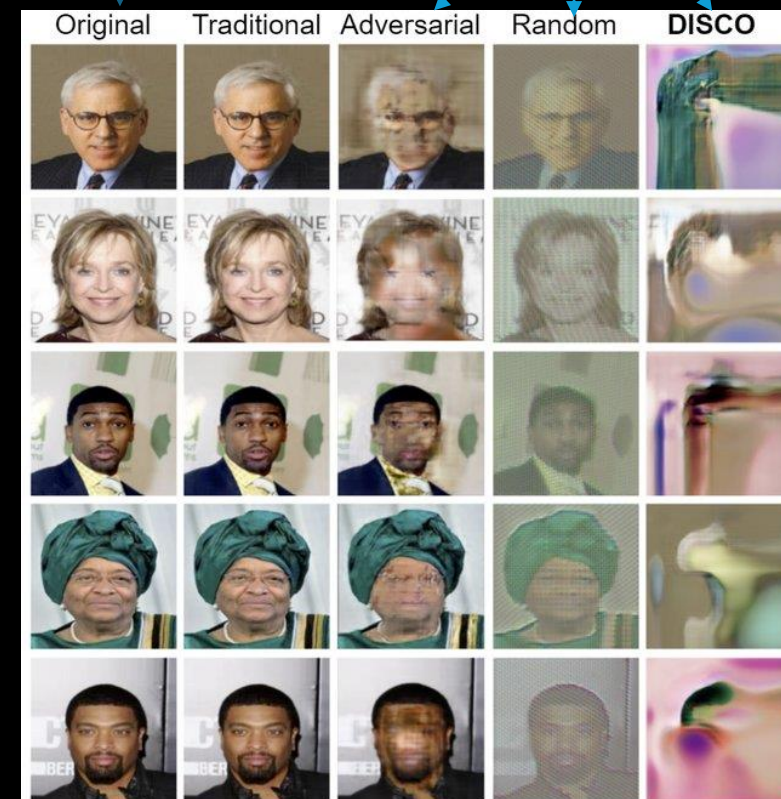
**Infrastructure**

**Machine learning model output**

Photo Reconstructed Exploiting Non-Privacy Preserving Machine Learning Model

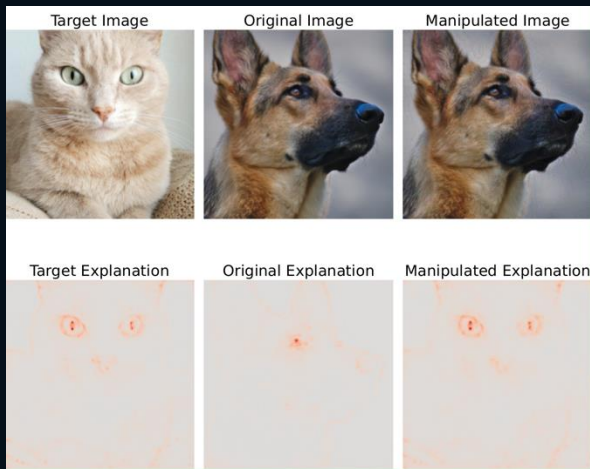
Original Photo

Photo Reconstructed Exploiting Different Privacy Preserving Machine Learning Model

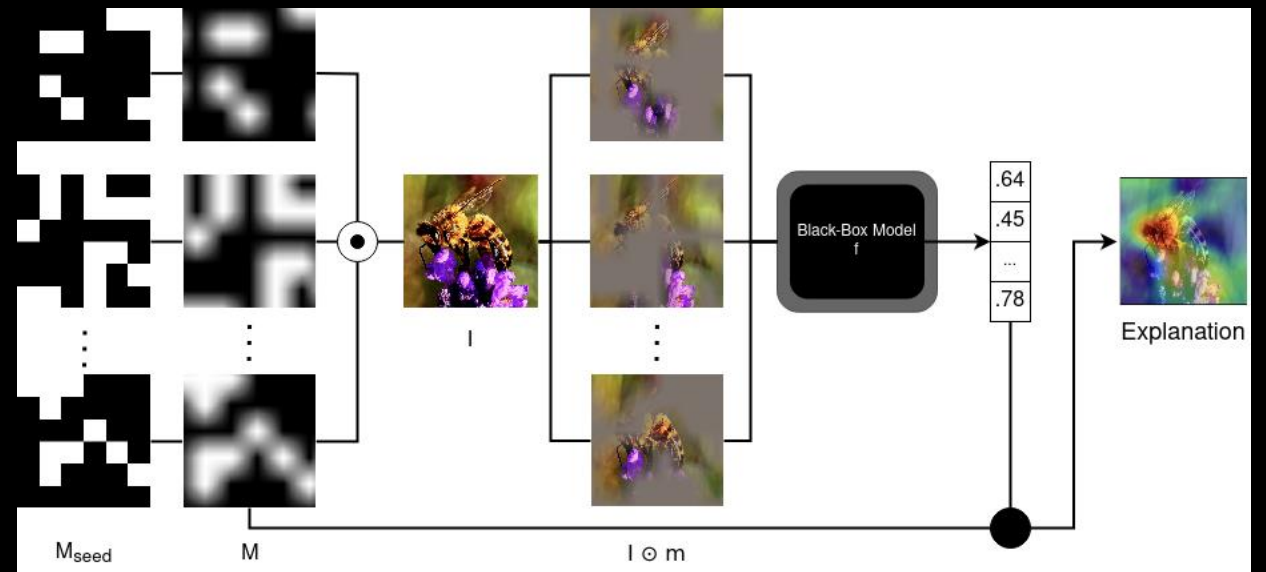




# Explainable AI



An example of an adversarial attack

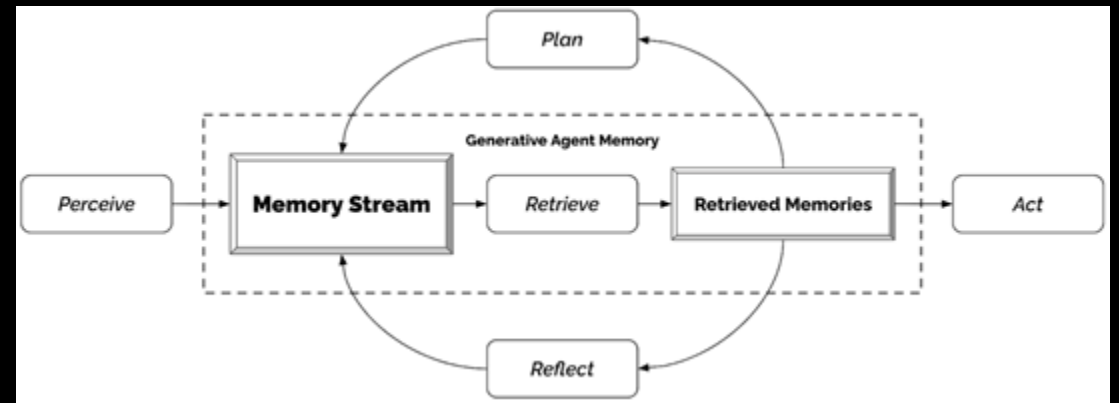


RISE algorithm



# Future of LLMs

- LLMs vs Agentic LLMs
  - **Traditional LLMs:** Chatbots, text generation, language translation, creative writing, QA
  - **Agentic LLMs:** Access and process information from the real world through external tools, and use this information to make decisions and complete tasks
- Stanford created a virtual world full of ChatGPT-powered people
  - The simulation ran for 2 days and showed that LLM-powered bots interact in a human-like way
  - The bots planned a party, coordinated the event, and attended the party within the simulation



# Content

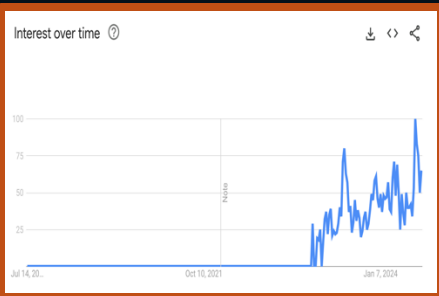
TECH

**ChatGPT experience may be able to get you that job. More than 90% of potential employers are looking for workers who use the chatbot.**

Aaron Mok Apr 27, 2023, 7:36 PM CEST

Share Save

The findings come as workers flock to ChatGPT to make their jobs easier. Many have used the chatbot to develop code, write marketing copy, and generate lesson plans. Some companies are even looking to hire prompt engineers with no tech background for as much as \$335,000 a year.



Source: google trends

- Computer programming
  - Knowledge centric (human learning)
  - Data centric (machine learning)
  - Training (years vs Bytes)
- Big data
  - Data movement
  - Data processing
  - Data storage
- Artificial Intelligence terminology /landscape
  - AI origins
  - AI milestones (perceptron – DL)
  - AI between Hype/fiction/Reality
  - Deep Learning
  - Problem and Security (hallucination, ...)
  - Ethics (Black box)
  - Prompt Engineering

# Prompt

One way interacting with LLMs, its simplicity with no need to fine-tune the model.

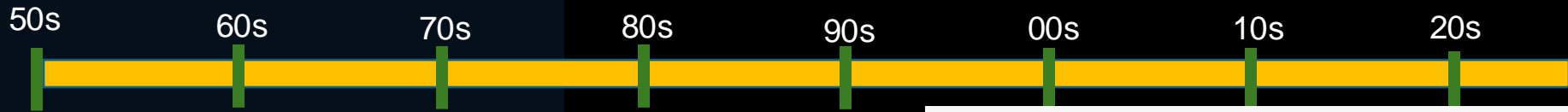
**Prompt engineering** is the practice of developing and optimizing prompts to efficiently use language models (LMs) for a variety of applications

# Content

- Text Summarization
- Question Answering
- Text Classification
- Role Playing
- Code Generation
- Reasoning
  - Few-shots
  - Chain-of-thought
  - Zero-shot Prompt
  - Problem: self-consistency

Basic Tasks

Advanced Tasks



End.



"the end of the presentation with th mention for now, implying tha..."

Designer

Powered by DALL·E 3







