ESCOLA SUPERIOR DE TECNOLOGIA IPCA

Artificial Intelligence Fundamentals

Fundamentals of Artificial Intelligence

MSc in Applied Artificial Intelligence, 2023-24

Contents

- Topics included
  - Artificial Intelligence (AI)
  - AI types and areas

- These slides were based essentially on the following bibliography:
  - Norvig, P, Russell, S. (2021). Artificial Intelligence: A Modern Approach, 4th Edition. Pearson, ISBN-13: 978-1292401133
  - Most of the slides with a left blue bar are authored by Prof. Joaquim Gonçalves

Artificial intelligence

Intelligence

- What the genius of Physics say?
  - “Intelligence is the **ability to adapt to change**.”, allegedly assigned to Stephen Hawking[1]
  - "The measure of intelligence is the **ability to change**", Albert Einstein

- **Intelligence** can be defined in terms of
  - (a) fidelity to **human performance** or
  - (b) **rationality**, i.e., "doing the right thing"

- Intelligence as ***rationality*** also varies ...
  - a property of **internal thought** processes and reasoning
  - demonstrated externally by the **intelligent behavior**

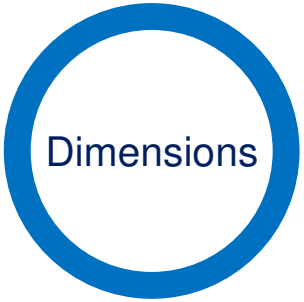
Retrieved on 2022.08.15 from
https://www.pinterest.pt/pin/155514993359432994

[1] https://www.washingtonpost.com/news/answer-sheet/wp/2018/03/29/stephen-hawking-famously-said-intelligence-is-the-ability-to-adapt-to-change-but-did-he-really-say-it/

Intelligence

- the faculty of knowing, understanding, **reasoning**, thinking and interpreting;
- ability to **choose** between several options;
- ability to **adapt** and improvise in a new environment;
- **generalize** knowledge (patterns) by applying it to unfamiliar settings.

Intelligence



Analytics

Solving theoretical problems, generally using abstraction

Creative

Introduction of new concepts and ideas or perspectives on a particular problem

Practice

Select the best solutions for everyday problems

Artificial

What is?

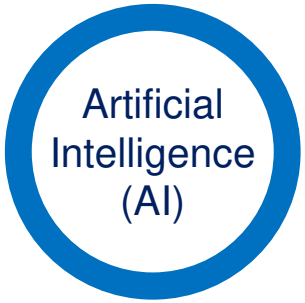
Tries to **appear to be** the original thing
it's not what it appears to be → *artificial flower*

Replaces the original thing
that's what it appears to be → *artificial light*

Tries to **appear to be** the original thing and **replaces** the original thing → *artificial tooth*

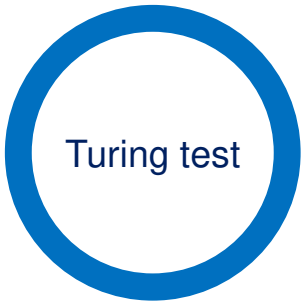
Intelligence *versus* Artificial Intelligence

- Ability to learn and understand how to deal with new situations
- Skilful use of knowledge



- Study of the ideas that enable computers to be intelligent
- Study of the computations that make perception, reasoning and action possible
- Create systems to perform tasks that require some features of human intelligence.

Intelligence *versus* Artificial Intelligence



The test is based on the impossibility of distinguishing two intelligent entities (human versus machine) in a given question

If it is not possible to distinguish the provenance of the answers, then the machine is endowed with intelligence.

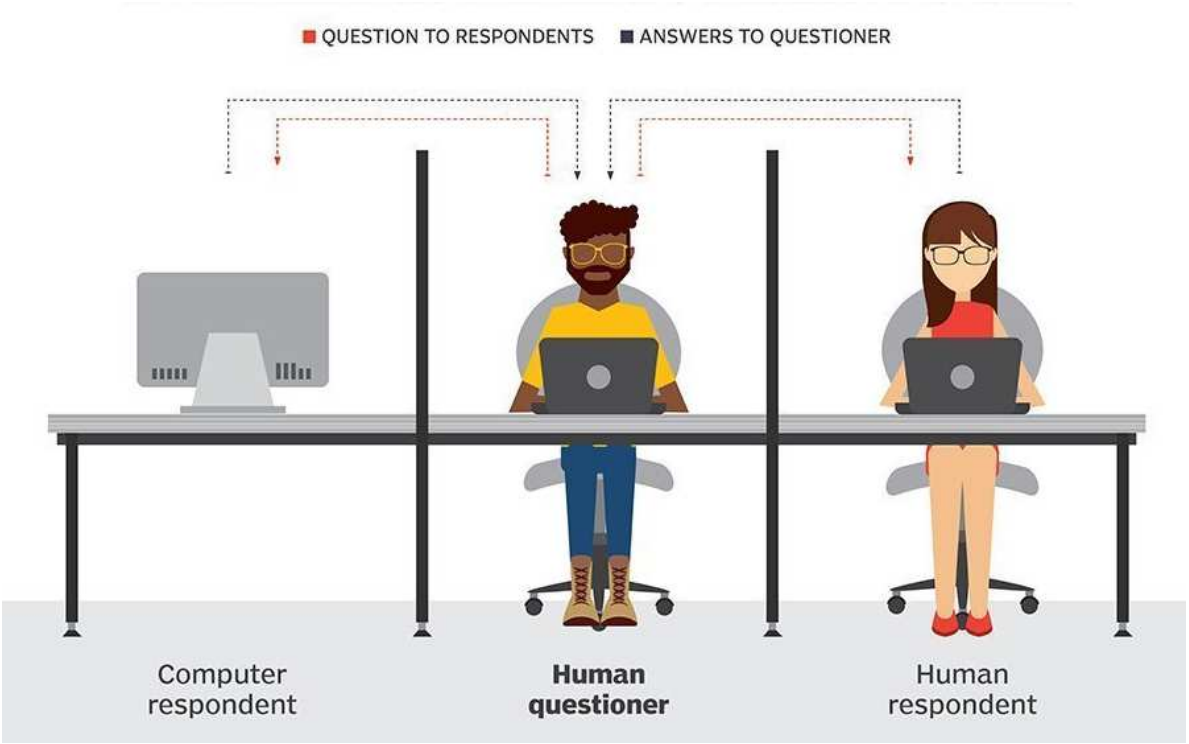
"If a machine behaves intelligently then it is intelligent"

Alan Turing

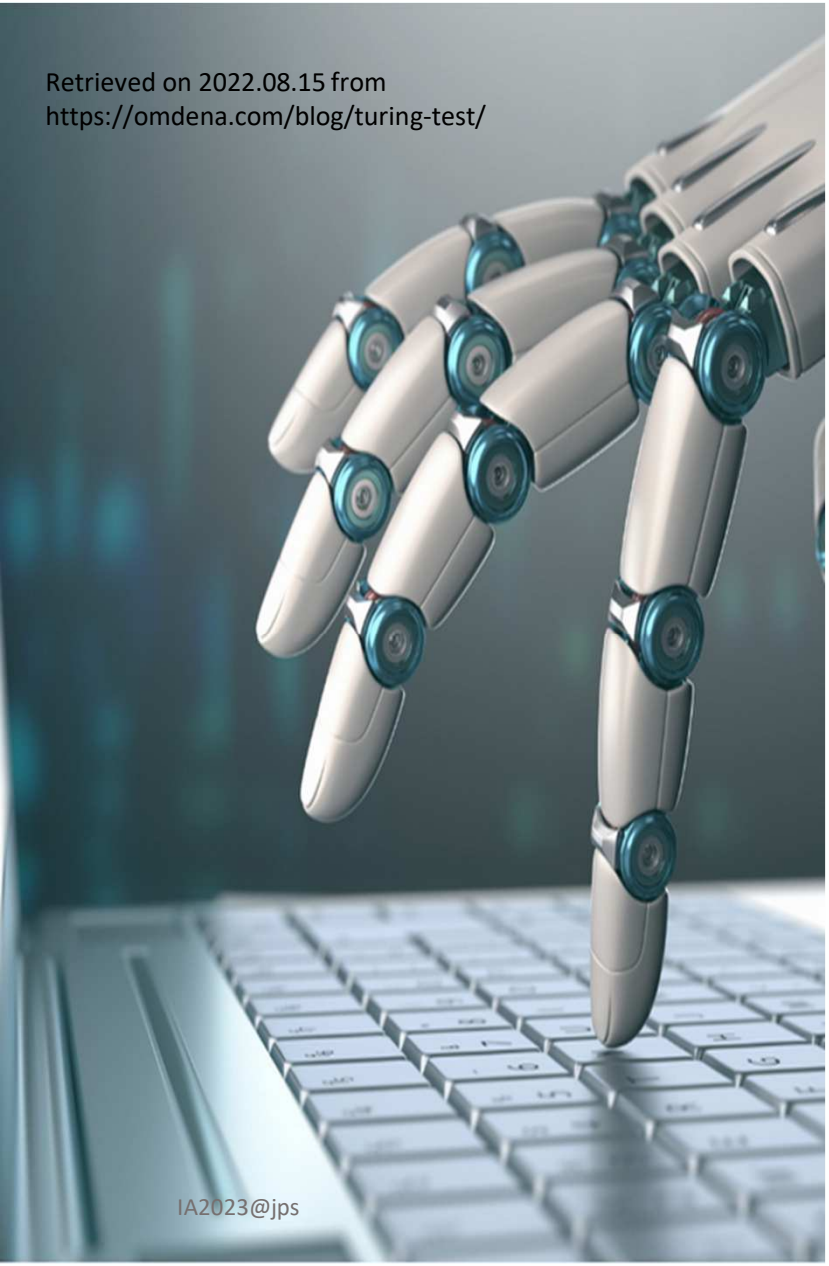
The Turing test approach

- The Turing test, proposed by Alan Turing (1950)
  - “Can a machine think?”
  - A computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or from a computer

- Recommended links
  - https://www.youtube.com/watch?v=4VROUIAF2Do&t=129s
  - https://plato.stanford.edu/entries/turing-test/



Retrieved on 2022.08.15 from https://www.techtarget.com/searchenterpriseai/definition/Turing-test


Retrieved on 2022.08.15 from
https://omdena.com/blog/turing-test/

The Turing test approach (2)

- In the previous scenario, the computer would need the following capabilities:
  - **natural language processing (1)** to communicate in a human language;
  - **knowledge representation (2)** to store what it knows or hears;
  - **automated reasoning (3)** to answer questions and to draw new conclusions;
  - **machine learning (4)** to adapt to new circumstances and to detect and extrapolate patterns.

- Later, other researchers have proposed a **total Turing test**, which requires interaction with objects and people in the real world that would require
  - **computer vision** and **speech recognition (5)** to perceive the world;
  - **robotics (6)** to manipulate objects and move about.

- These **six disciplines** compose most of AI application domains

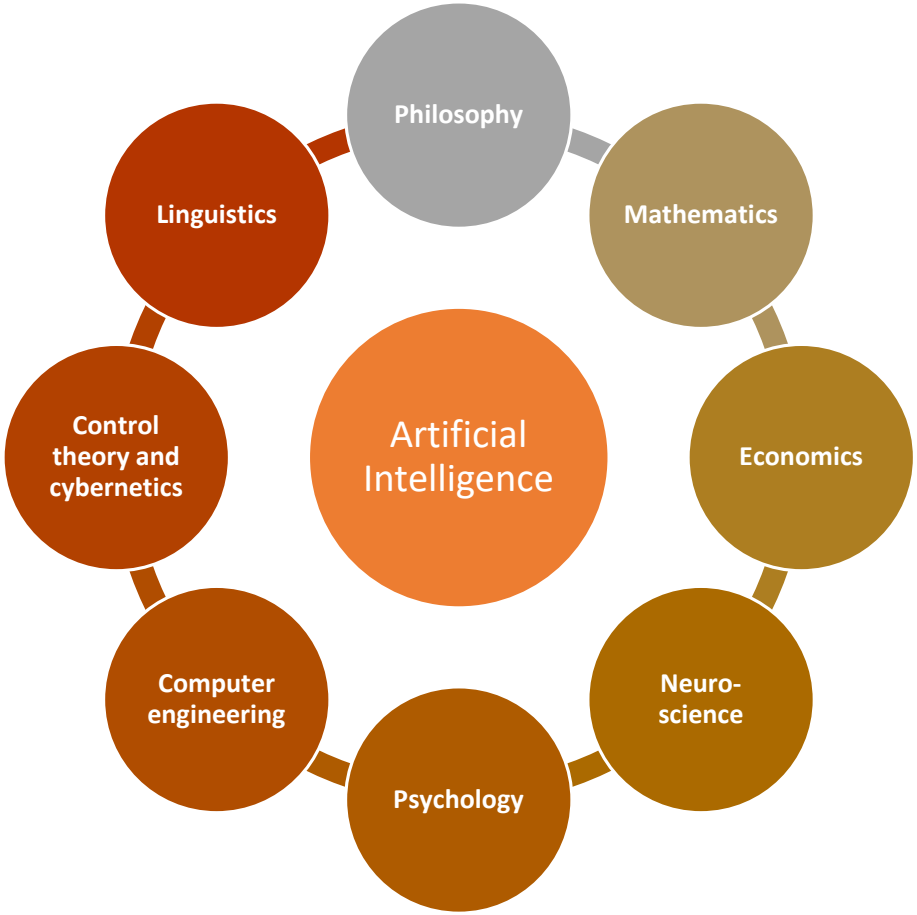
Value alignment problem

- The standard model assumes that we will supply a fully specified objective to the machine.
  - **Defined task**, such as chess or shortest-path computation, the task comes with an objective built in — so the standard model is applicable
  - However, in real world is **difficult to specify the objective completely and correctly**.
- For example, the objective of a **self-driving car** is to reach the destination safely
  - the strict goal of safety requires staying in the garage
  - there is a **tradeoff** between making progress towards the destination and incurring a risk of injury
  - it poses many questions about acceleration, steering, braking, etc. that are difficult to answer a priori.
- The values put into the machine must be aligned with those of the human

Retrieved on 2022.08.15 from https://www.pinterest.pt/pin/vintage-future--475270566904194319/

The Foundations of Artificial Intelligence

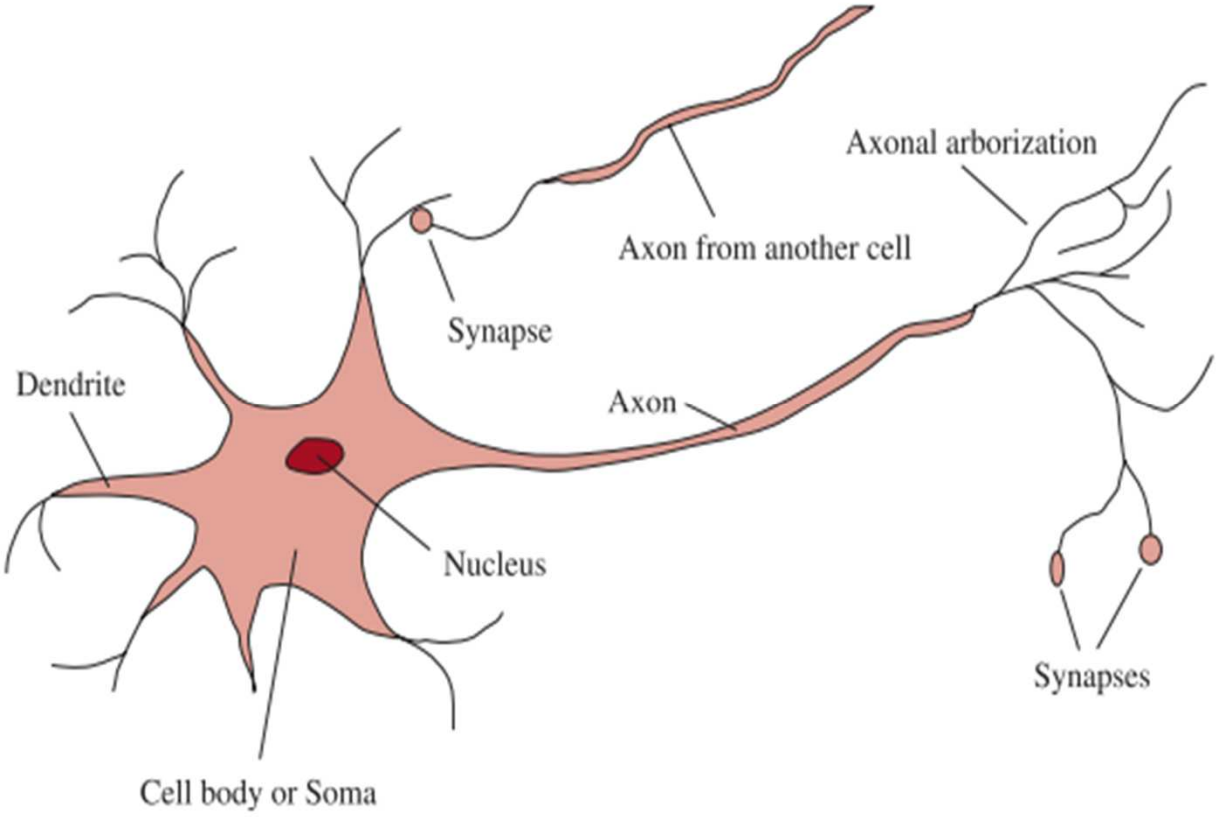
- **Philosophy**: empiricism, principle of induction, deontological ethics, ...
  - Can formal rules be used to draw valid conclusions?
  - How does knowledge lead to action?
- **Mathematics** and statistics: formal logic, theory of probability, algorithmics, ...
  - What are the formal rules to draw valid conclusions? What can be computed?
  - How do we reason with uncertain information?
- **Economics**: decision theory, game theory, Markov decision processes, ...
  - How should we make decisions in accordance with our preferences?
  - How should we do this when the payoff may be far in the future?



The Foundations of Artificial Intelligence (2)

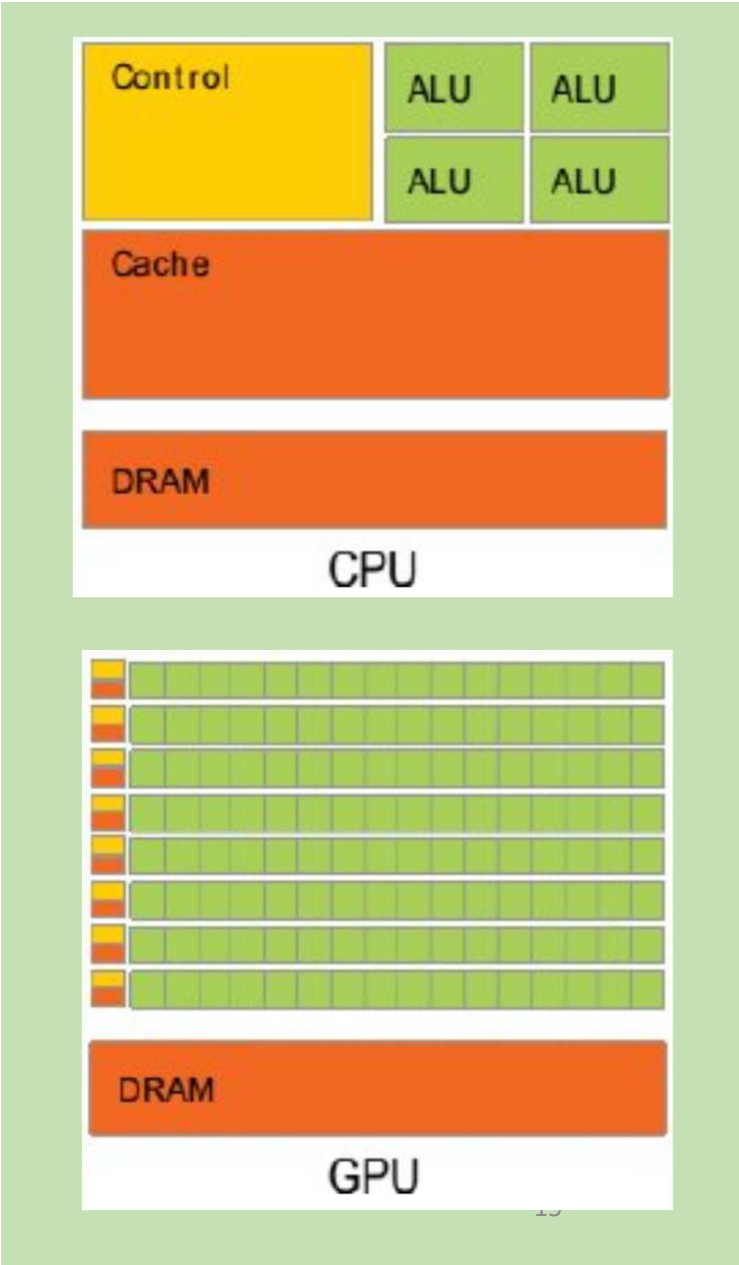
- **Neuroscience**: neuronal organization, optogenetics, brain–machine interfaces, ...
  - How do brains process information?

- Neurons
  - Each neuron has several dendrites, but just a single axon
  - The axon stretches out for a long distance
  - Typically, an axon is 1 cm long, but can reach up to 1 meter
  - A neuron makes connections with 10 to 100,000 other neurons at junctions called synapses



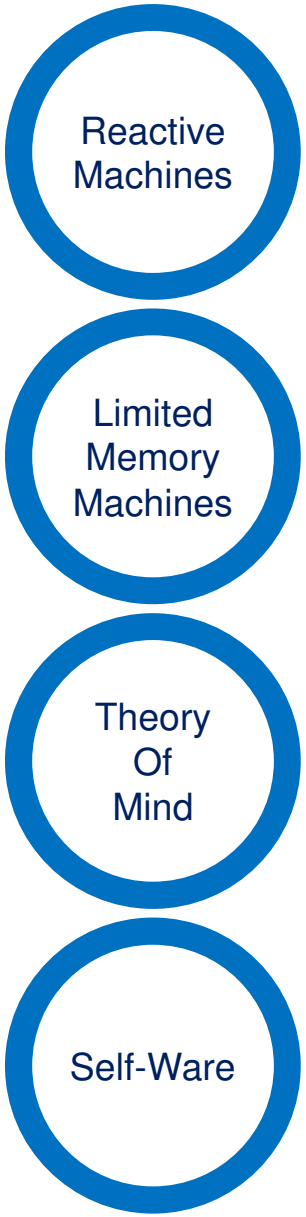
The Foundations of Artificial Intelligence (3)

- **Psychology**: cognitive psychology, cognitive science, human–computer interaction ...
  - How do humans and animals think and act?
  - How does knowledge lead to action?
- **Computer engineering**: graphics processing unit (GPU), tensor processing unit (TPU), wafer scale engine (WSE), ...
  - How can we run faster and efficient AI applications?
- **Control theory and cybernetics**: control theory, Cybernetics, homeostatic devices, cost function ...
  - How can artifacts operate under their own control?
- **Linguistics**: computational linguistics, knowledge representation
  - How does language relate to thought?



AI types and areas

AI Types



Systems that have **no memory-based functionality** and, therefore, no capacity to learn from past experiences or interactions

Systems have the same ability as reactive machines, but are also **capable of learning** from historical data

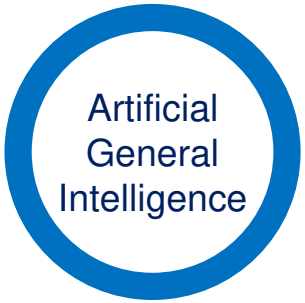
Systems will be able to **discern the needs, emotions, beliefs, and thought** processes of the humans it's dealing with, and be able to understand humans as individuals whose minds can be shaped by multiple factors

Systems develops **self-awareness**, with emotions, needs, beliefs, and potentially desires of its own, could develop ideas like self-preservation, which directly or indirectly conflict with the interests of humanity

AI Levels

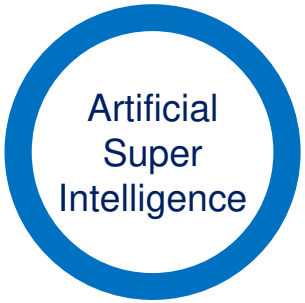
Narrow AI or Week AI

Systems that can only perform a specific task autonomously using human-like capabilities



Strong AI or Deep AI

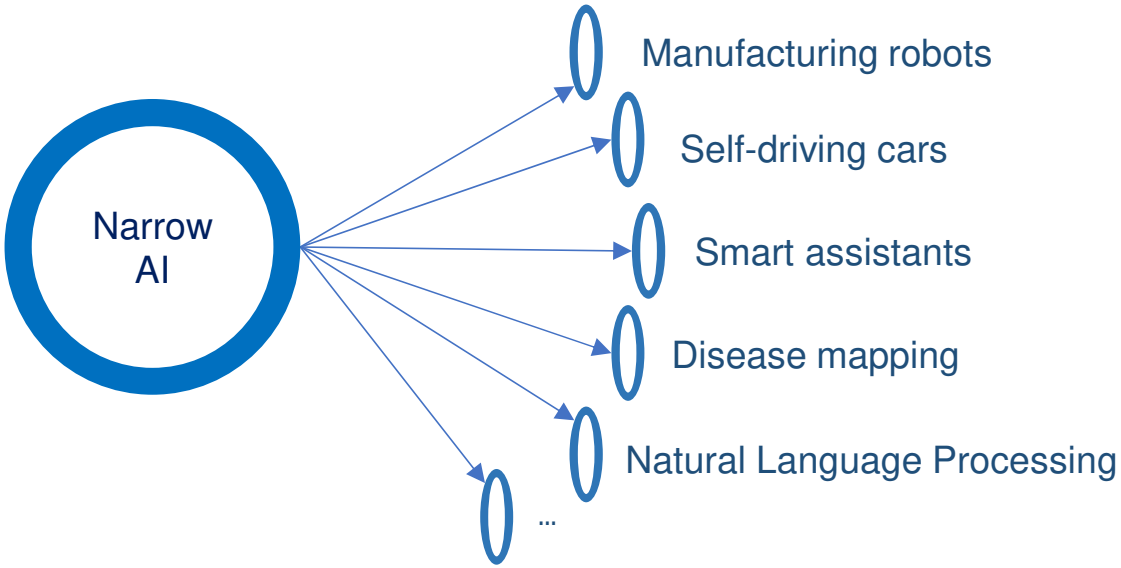
Systems that are just as capable as humans by replicating our **multi-functional capabilities**



Super AI

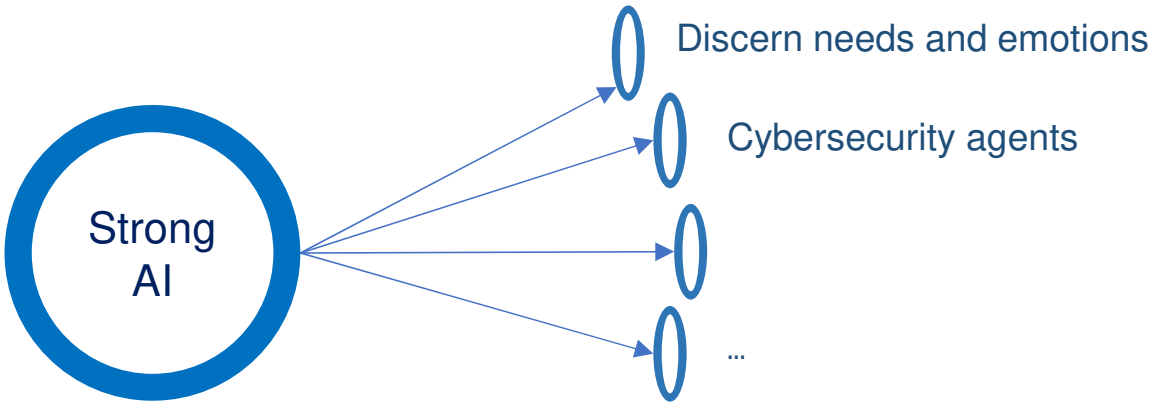
Hypothetical concept that would make the system the **most capable form of intelligence**

AI Levels



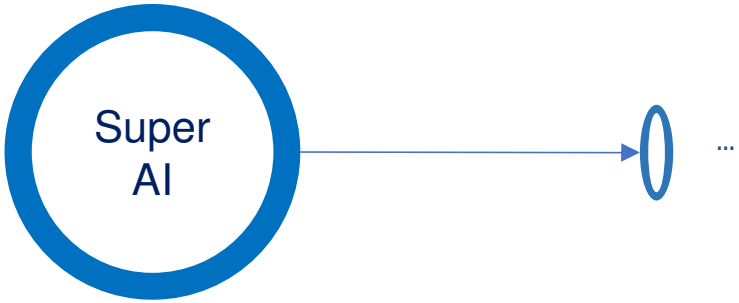
- **Narrow AI is goal-oriented**, designed to perform specific tasks
- It aims to simulate human behaviour, but not to replicate human intelligence.

AI Levels



- System that **mimics human intelligence and behaviour**
- The system have the ability to learn and apply your intelligence in solving any problem
- A system with GAI (General AI) will be indistinguishable from a human in understanding, thinking, behaviour and action

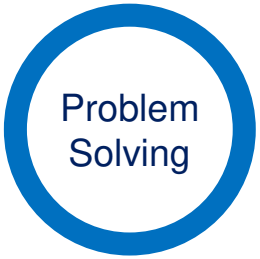
AI Levels



It is a hypothetical formulation in which systems are endowed with self-awareness and their intelligence **overcomes human intelligence**.

A system with this type of intelligence, which also has greater calculation and memory capacity, would make its decision-making and problem-solving capacity far superior to that of human beings.

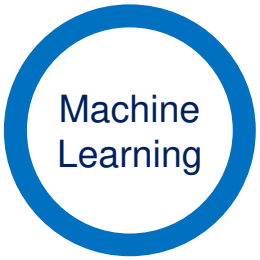
AI Areas



Mathematical formulation of a problem for the systematic construction of solutions

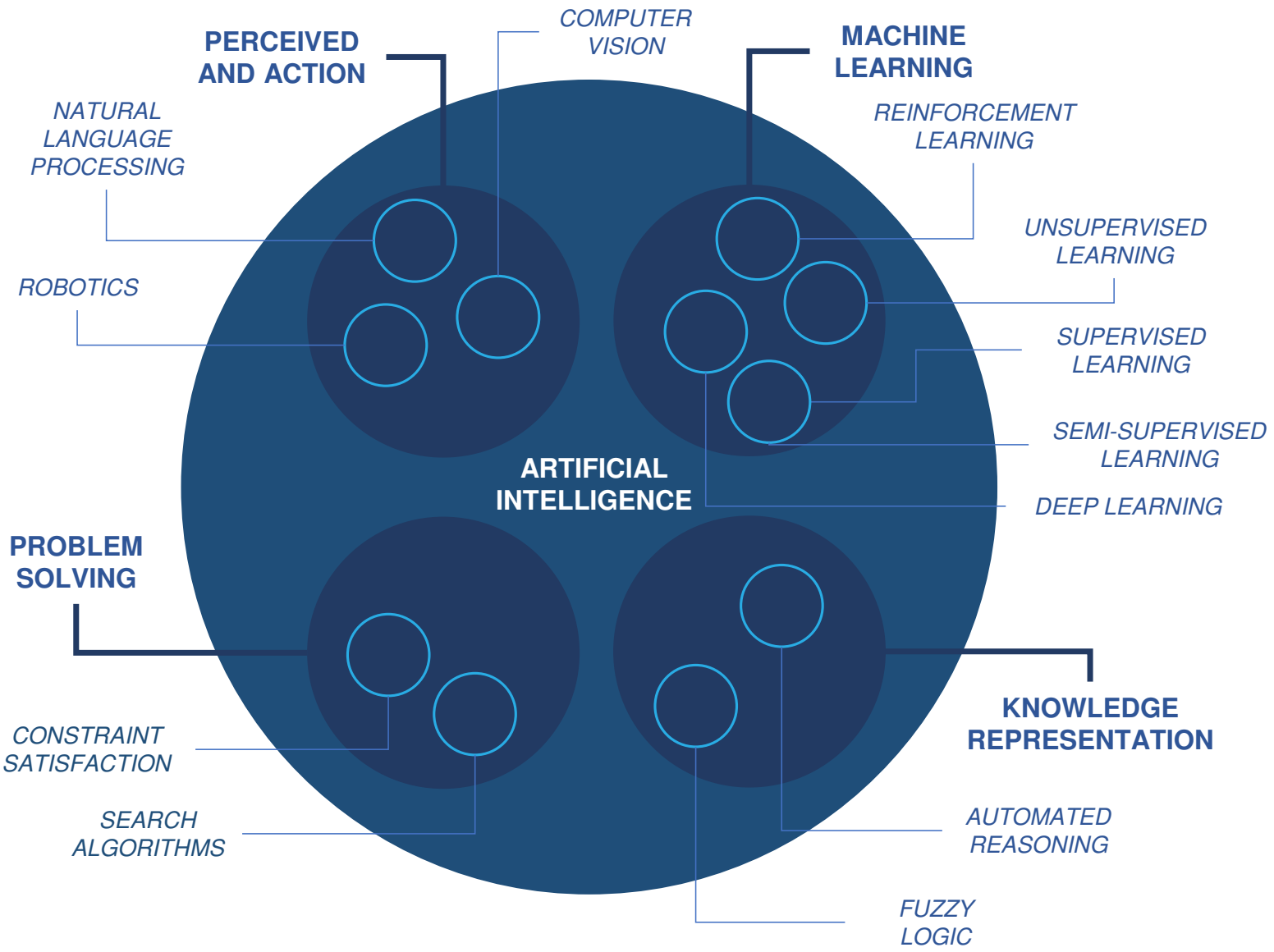
A way of expressing value judgments and intention of an intelligent agent, in a formal and transparent symbolic notation, suitable for automated reasoning

Process of capturing information from the environment and defining appropriate actions to respond to a problem

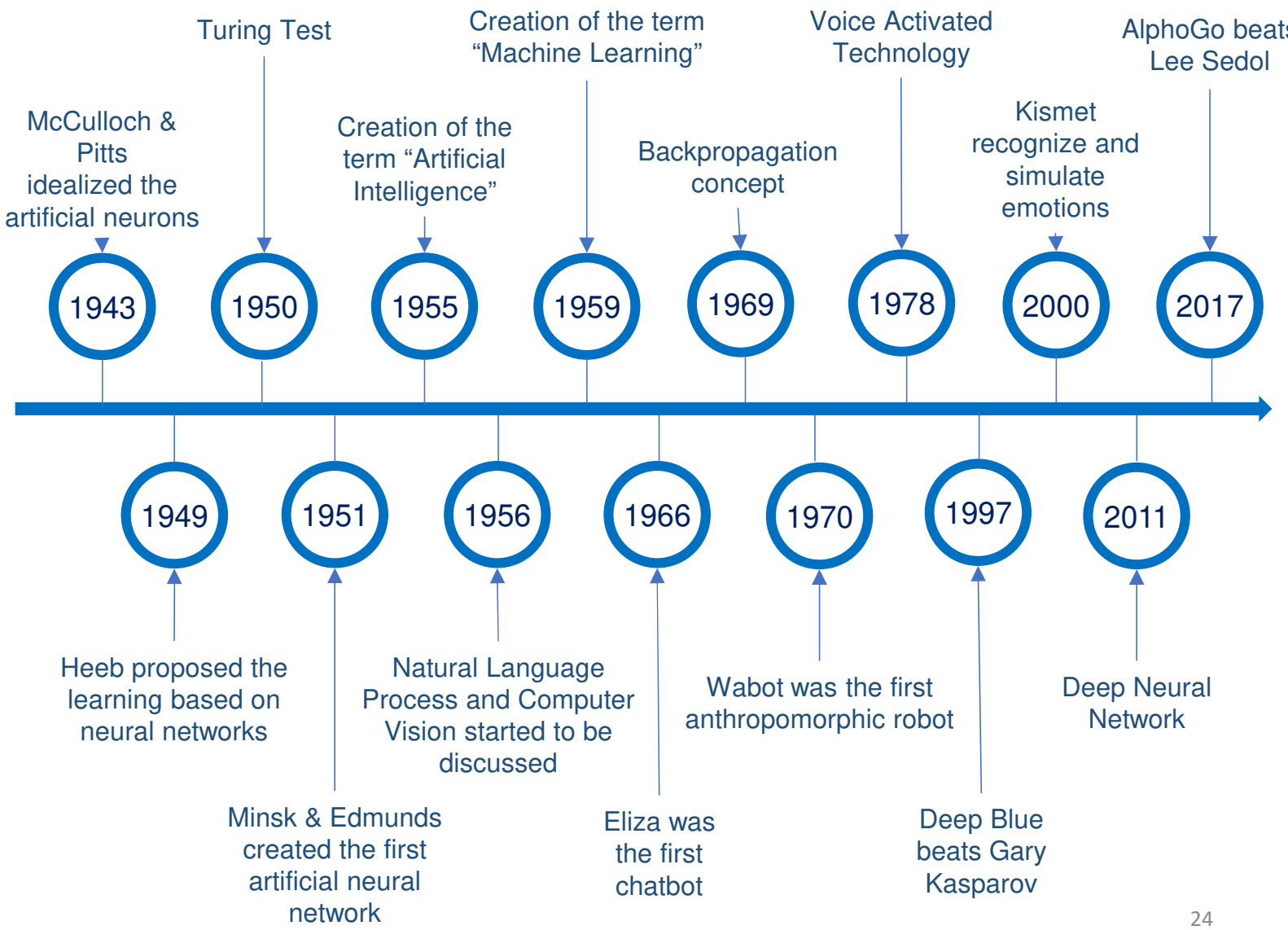


Building models that learn logic from data

AI Areas


ARTIFICIAL INTELLIGENCE
PERCEIVED AND ACTION
COMPUTER VISION
NATURAL LANGUAGE PROCESSING
ROBOTICS
MACHINE LEARNING
REINFORCEMENT LEARNING
UNSUPERVISED LEARNING
SUPERVISED LEARNING
SEMI-SUPERVISED LEARNING
DEEP LEARNING
PROBLEM SOLVING
CONSTRAINT SATISFACTION
SEARCH ALGORITHMS
KNOWLEDGE REPRESENTATION
AUTOMATED REASONING
FUZZY LOGIC

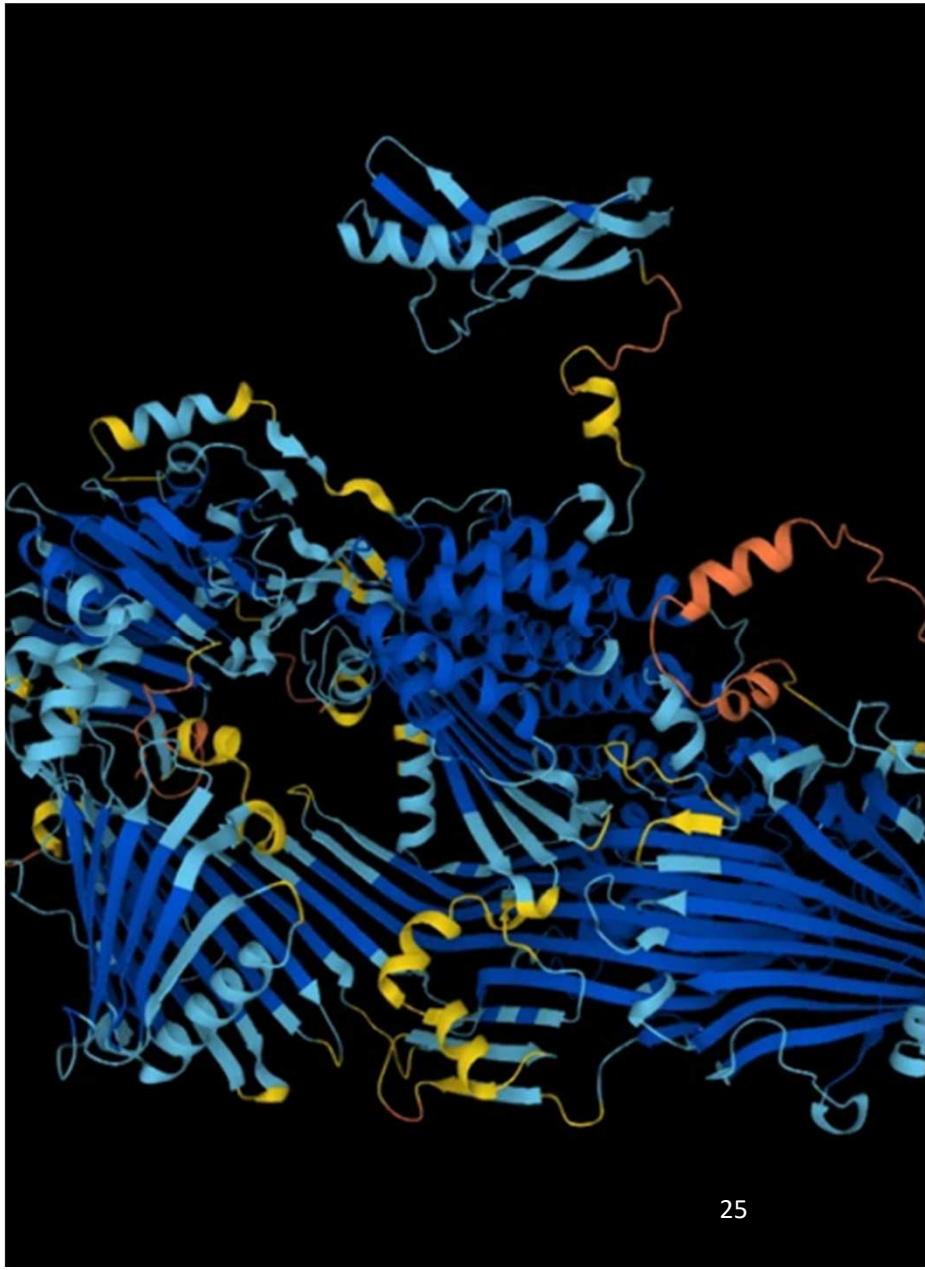
AI Milestones



Risks and benefits of AI

- **Benefits** of AI are well known:
  - AI is used by many services and applications we use daily.
  - The potential for AI and robotics to free humanity from menial repetitive work and to dramatically increase the production of goods and services could presage an era of peace and plenty.
  - The capacity to **accelerate scientific research** could result in cures for disease and solutions for climate change and resource shortages.

- As Demis Hassabis, CEO of Google DeepMind, has suggested "**First solve AI, then use AI to solve everything else**."
- "The entire protein universe': **AI predicts shape of nearly every known protein**", Nature journal, 28th July 2022



Retrieved on 2022.09.22 from https://www.nature.com/articles/d41586-022-02083-2

Risks of misuse of AI

- **Lethal autonomous weapons**: These are defined by the United Nations as weapons that can locate, select, and eliminate human targets without human intervention (drones).
- **Surveillance and persuasion**: speech recognition, computer vision, and natural language understanding can be used in a scalable fashion to perform mass surveillance of individuals.
- **Biased decision making**: Careless or deliberate misuse of machine learning algorithms for tasks such as loan applications can result in decisions that are biased by race, gender, or other protected categories
- **Impact on employment**: Concerns about machines eliminating jobs.
- **Safety-critical applications**: Concerns about AI used in safety-critical applications such as driving cars and managing the water supplies.
- **Cybersecurity**: AI techniques are useful in defending against cyberattack, but they will also contribute to the potency, survivability, and proliferation capability of malware.

Retrieved on 2022.08.15 from
https://www.thequint.com/news/world/2015-isnt-quite-different-from-orwells-1984

Some facts about AI

From https://www.valuer.ai/blog/75-facts-about-artificial-intelligence:

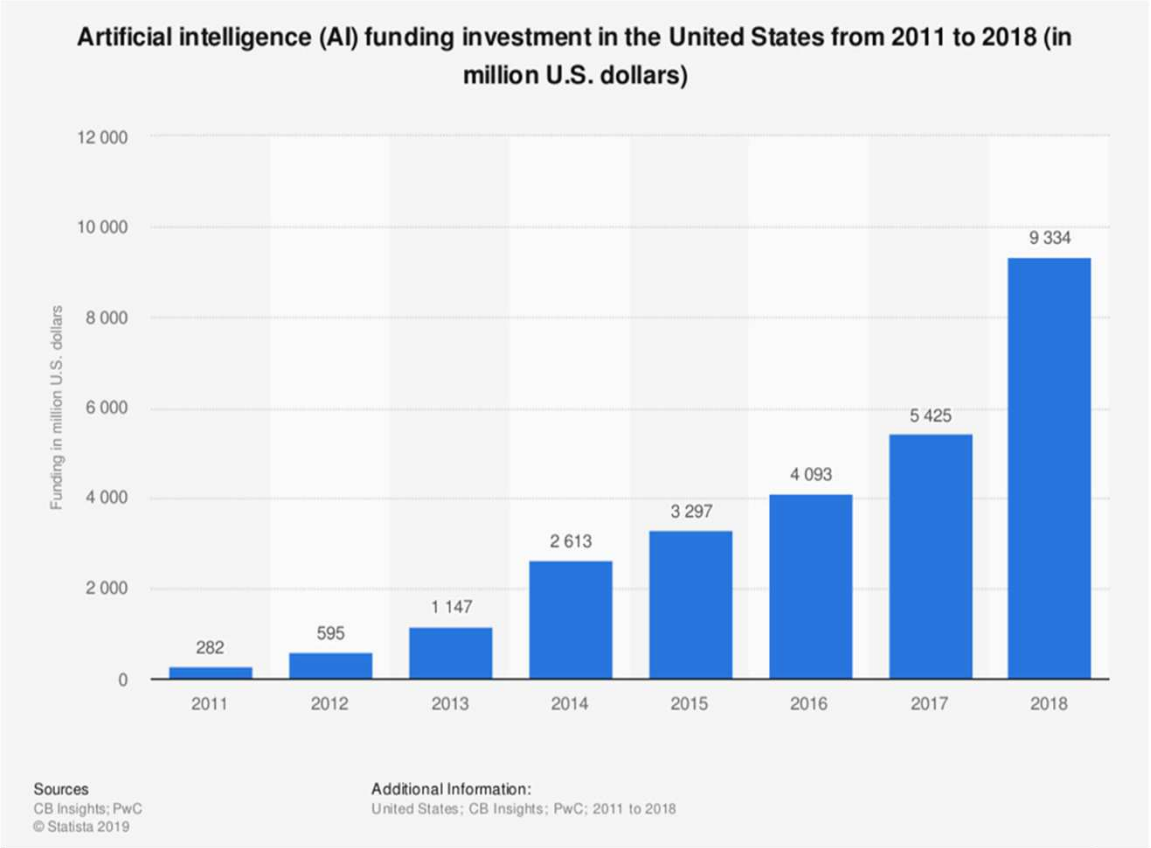
- Google CEO claimed that **AI would be more transformative to humanity than electricity and fire**.
- A McKinsey survey from 2021 found that 56% of companies have adopted AI in at least one function.

From https://dataprot.net/statistics/ai-statistics/:

- The rise of AI will eliminate 85 million jobs and create 97 million new ones by 2025.

But (https://dataprot.net/statistics/ai-statistics/),

- Only 19% of people have faith that AI will improve productivity in their workplace
- 49% of consumers say they don't think automation in healthcare is safe.



Retrieved on 2022.08.15 from https://dataprot.net/statistics/ai-statistics/

Thank you!