



# Artificial Intelligence (ACCA7026)

## Introduction to Artificial Intelligence

Dr Seena Joseph

# Warm-up: what AI have you already used today?



## Chat assistants

ChatGPT, Gemini, Claude, Copilot



## Voice assistants

Siri, Alexa, Google Assistant



## Recommendations

Netflix, Spotify, TikTok, shopping



## Maps & travel times

Route choice, traffic, ride-hailing ETAs



## Face & photo features

Face unlock, photo search, filters



## Writing & code help

Autocomplete, grammar, summarising

Hands up for each one. Then: which did you use without realising it was AI?

# Artificial Intelligence

## Artificial Intelligence [AI] – Definition

- *A system's ability to correctly interpret external data, to learn from such data and to use those learnings to achieve specific goals and tasks through flexible adaptation [1]*
- *Science and Engineering of making intelligent machines [2]*
- *Any computer that passes Turing test [3]*
  - *Turing Test ?*
- *Theory and development of computer systems capable of performing tasks that historically required human intelligence, such as recognizing speech, making decisions, and identifying patterns*
- *Programs that can sense, reason, act and adapt*

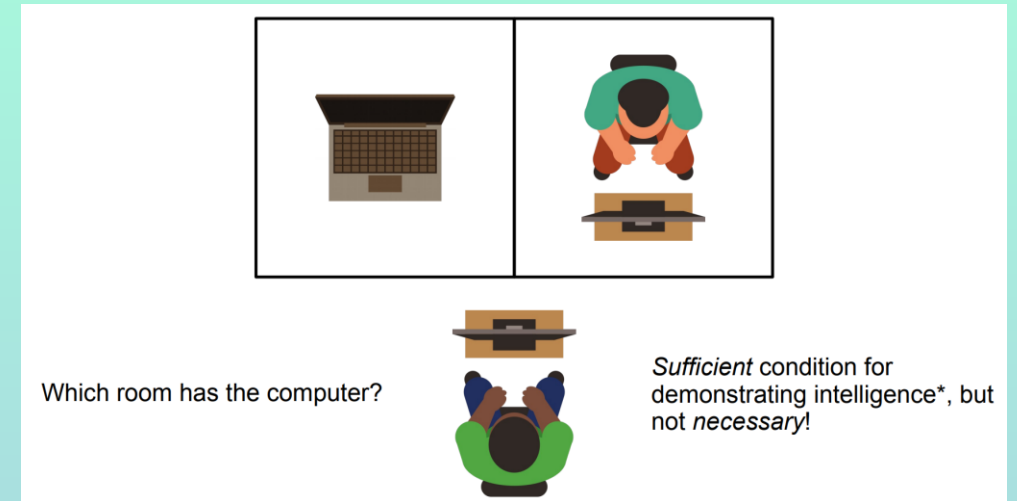
# Artificial Intelligence

## Artificial Intelligence [AI] – Definition

- AI – study of how to make computers perform intelligent tasks that in the past could only be performed by humans [4].
- AI - a multidisciplinary field with the capability of integrating cognition, machine learning, emotion recognition , human computer interaction, data storage and decision making.

# Artificial Intelligence

Turing Test: The Turing test checks whether a machine can chat by text so much like a person that someone talking to both cannot reliably tell which is the machine.



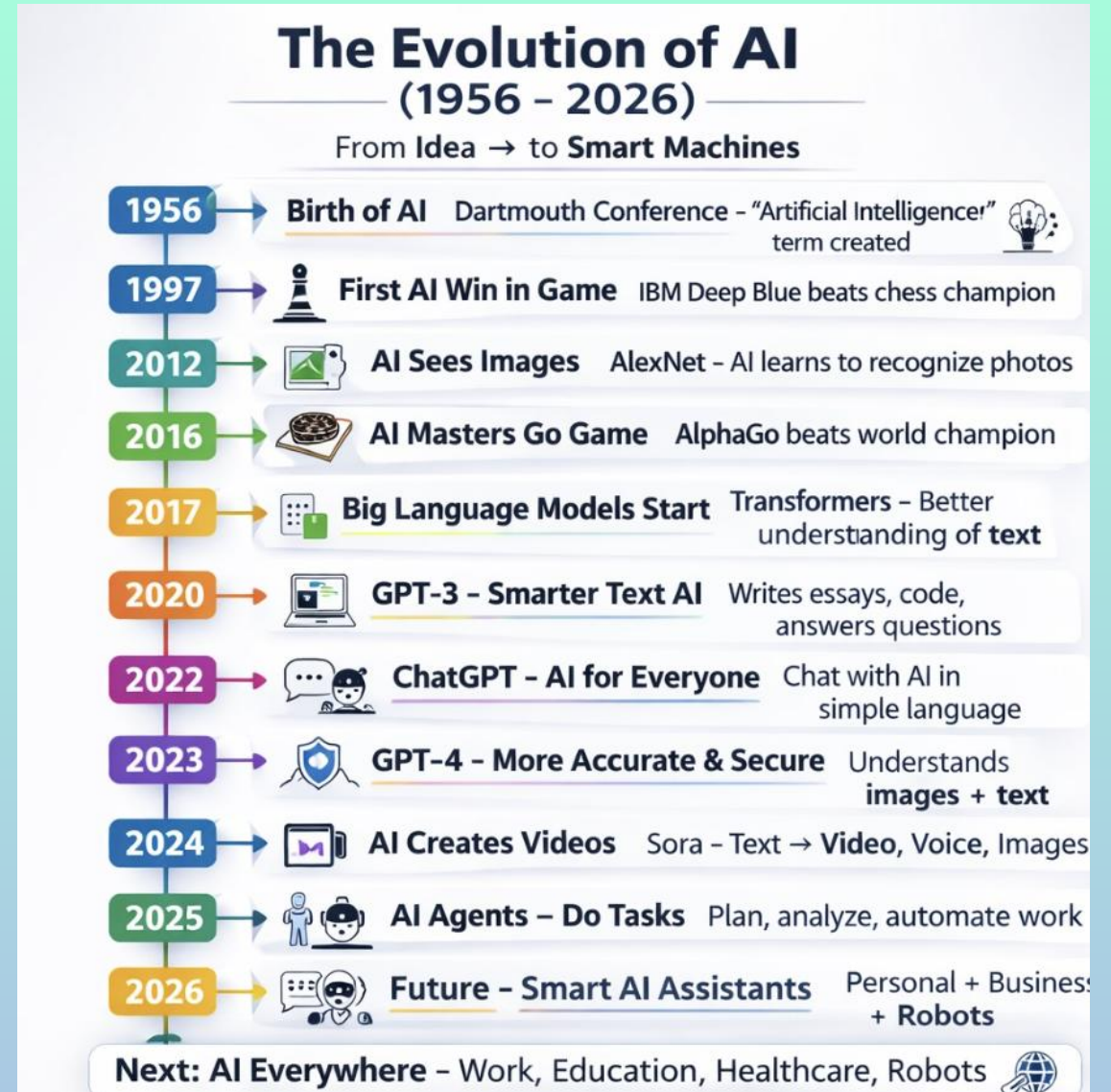
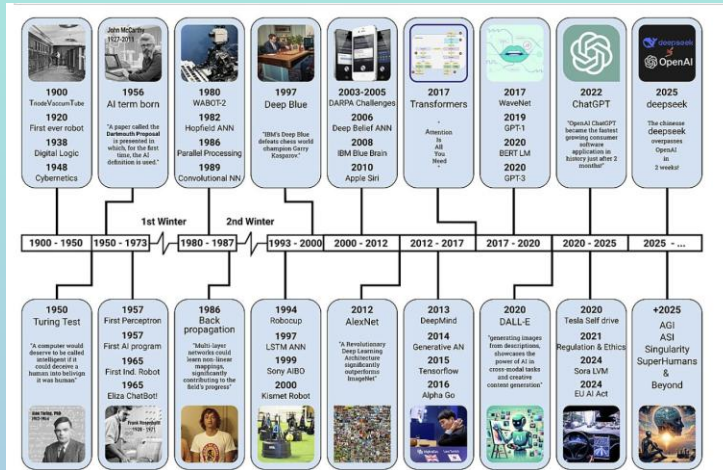
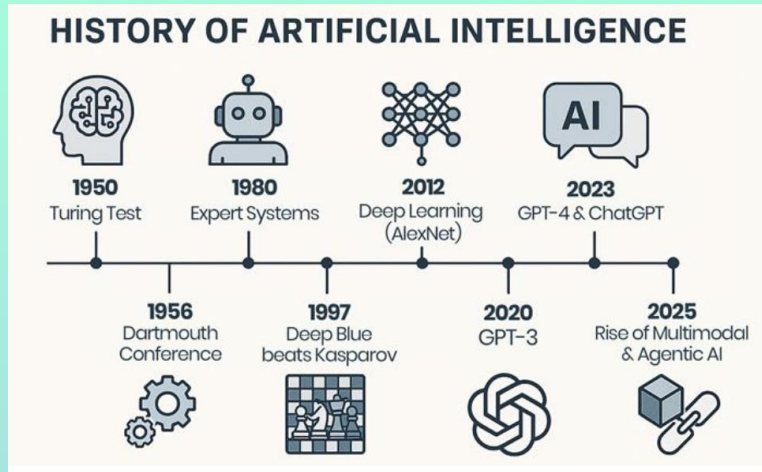
73%

of the time, GPT-4.5 (with a persona prompt) was picked as the human in a 2025 three-party Turing test – more often than the real humans. [5]

## DISCUSS

**Pairs, 2 minutes:** Does passing the Turing test mean a machine is intelligent? What is the test actually measuring?

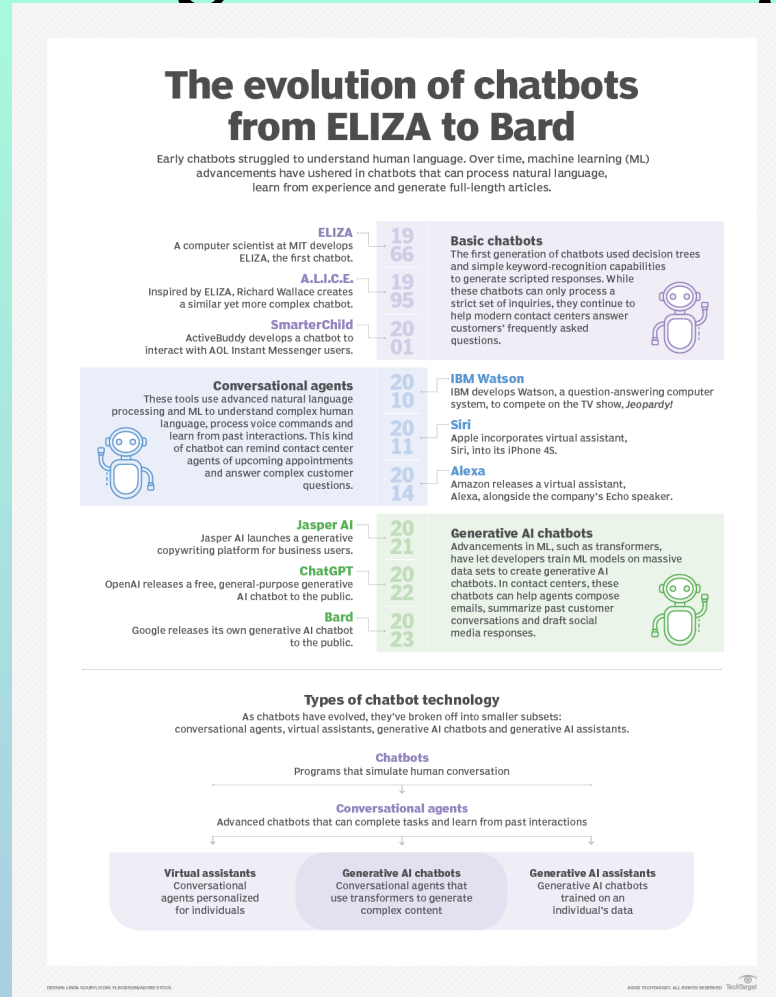
# Artificial Intelligence Development Timeline



- What is the position of AI in 2026 and beyond ?

# Artificial Intelligence Development Timeline

- Evolution of chatbots



# The map: seven eras and two winters



1943–1956

## Birth of AI

Neurons as logic, Turing's question, Dartmouth names the field.



1956–1974

## Golden years

Logic Theorist, ELIZA, perceptrons and bold promises.



1974–1980

## First AI winter

Lighthill report, combinatorial explosion, funding cuts.



1980–1987

## Expert systems

Hand-written rules from experts: MYCIN, XCON, big business.



1987–1993

## Second AI winter

Specialist hardware crashes; rule bases prove brittle.



1993–2016

## Machines start winning

Deep Blue, Watson, AlexNet, AlphaGo: data plus GPUs.



2017–2022

## Transformers & LLMs

Attention, BERT, GPT-3, ChatGPT.



2023 – now

## Agentic AI

LLMs plus memory, tools and multi-agent teams.



You decide

## What's next?

AGI, another winter, or something we have not imagined?

## ACTIVITY

# Put the milestones in order

In pairs: write the letters from earliest to latest. 3 minutes.

**A**

ChatGPT is released to the public

**B**

Deep Blue beats the world chess champion

**C**

“Attention Is All You Need” introduces the Transformer

**D**

ELIZA, the first famous chatbot, is written

**E**

AlexNet wins the ImageNet challenge

**F**

The Dartmouth workshop names the field

**G**

AlphaGo beats Lee Sedol at Go

### BONUS

Both AI winters fall between which two events?

Your order: \_ \_ \_ \_ \_

# The order, and why it happened when it did

1956 F Dartmouth workshop

1966 D ELIZA

Both AI winters fall in this gap (1974–1980, 1987–1993)

1997 B Deep Blue

2012 E AlexNet

2016 G AlphaGo

2017 C The Transformer

2022 A ChatGPT

## WHY NOW?

Old ideas finally had the fuel they needed.



### Data

The web, digitised text, labelled images



### Compute

GPUs and cloud scale



### Algorithms

Deep learning and the Transformer

*The cost: energy, hardware, concentration of power.*

# Artificial Intelligence Development Timeline

 [The Ultimate Evolution of Artificial Intelligence Timeline: From Turing to Agentic AI \(1950–2025\) - Ai Digital Friend](#)

- What is the position of AI in 2026 and beyond ?

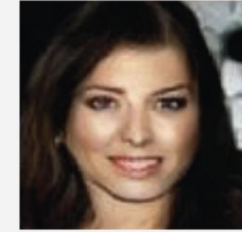
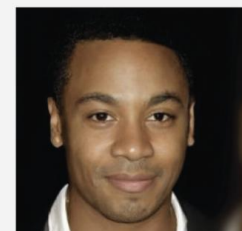
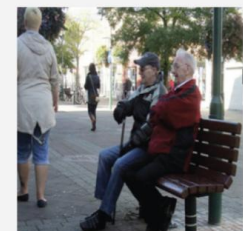
# Artificial Intelligence Development Timeline

- The rise of AI over the years

Source: [2]

Timeline of images generated by artificial intelligence  
These people don't exist. All images were generated by artificial intelligence.

Our World in Data

|                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                 |                                                                                                                                                                                                                                                                                                                                                                                               |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>2014</p>  <p>Goodfellow et al. (2014) - Generative Adversarial Networks</p>                                        | <p>2015</p>  <p>Radford, Metz, and Chintala (2015) - Unsupervised Representation Learning with Deep Convolutional GANs</p>                                                                   | <p>2016</p>  <p>Liu and Tuzel (2016) - Coupled GANs</p>                                                                                                                                                                                                                                                    |
| <p>2017</p>  <p>Karras et al. (2017) - Progressive Growing of GANs for Improved Quality, Stability, and Variation</p> | <p>2018</p>  <p>Karras, Laine, and Aila (2018) - A Style-Based Generator Architecture for Generative Adversarial Networks</p>                                                                | <p>2019</p>  <p>Karras et al. (2019) - Analyzing and Improving the Image Quality of StyleGAN</p>                                                                                                                                                                                                           |
| <p>2020</p>  <p>Ho, Jain, &amp; Abbeel (2020) - Denoising Diffusion Probabilistic Models</p>                        | <p>2021</p> <p>Image generated with the prompt: "a couple of people are sitting on a wood bench"</p>  <p>Ramesh et al. (2021) - Zero-Shot Text-to-Image Generation (OpenAI's DALL-E 1)</p> | <p>2022</p> <p>Image generated with the prompt: "A Pomeranian is sitting on the King's throne wearing a crown. Two tiger soldiers are standing next to the throne."</p>  <p>Saharia et al. (2022) - Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding (Google's Imagen)</p> |

# Finish the Timeline: AI Images 2023 → 2026

The chart stops at **2022**. For each year, find:

- 1 A new image or video generation model was released that year
- 2 Who made it + a credit line (paper or official announcement)
- 3 One thing it could do that earlier models couldn't

**2023**

Model: \_\_\_\_\_

By: \_\_\_\_\_

New: \_\_\_\_\_

**2024**

Model: \_\_\_\_\_

By: \_\_\_\_\_

New: \_\_\_\_\_

**2025**

Model: \_\_\_\_\_

By: \_\_\_\_\_

New: \_\_\_\_\_

**2026**

Model: \_\_\_\_\_

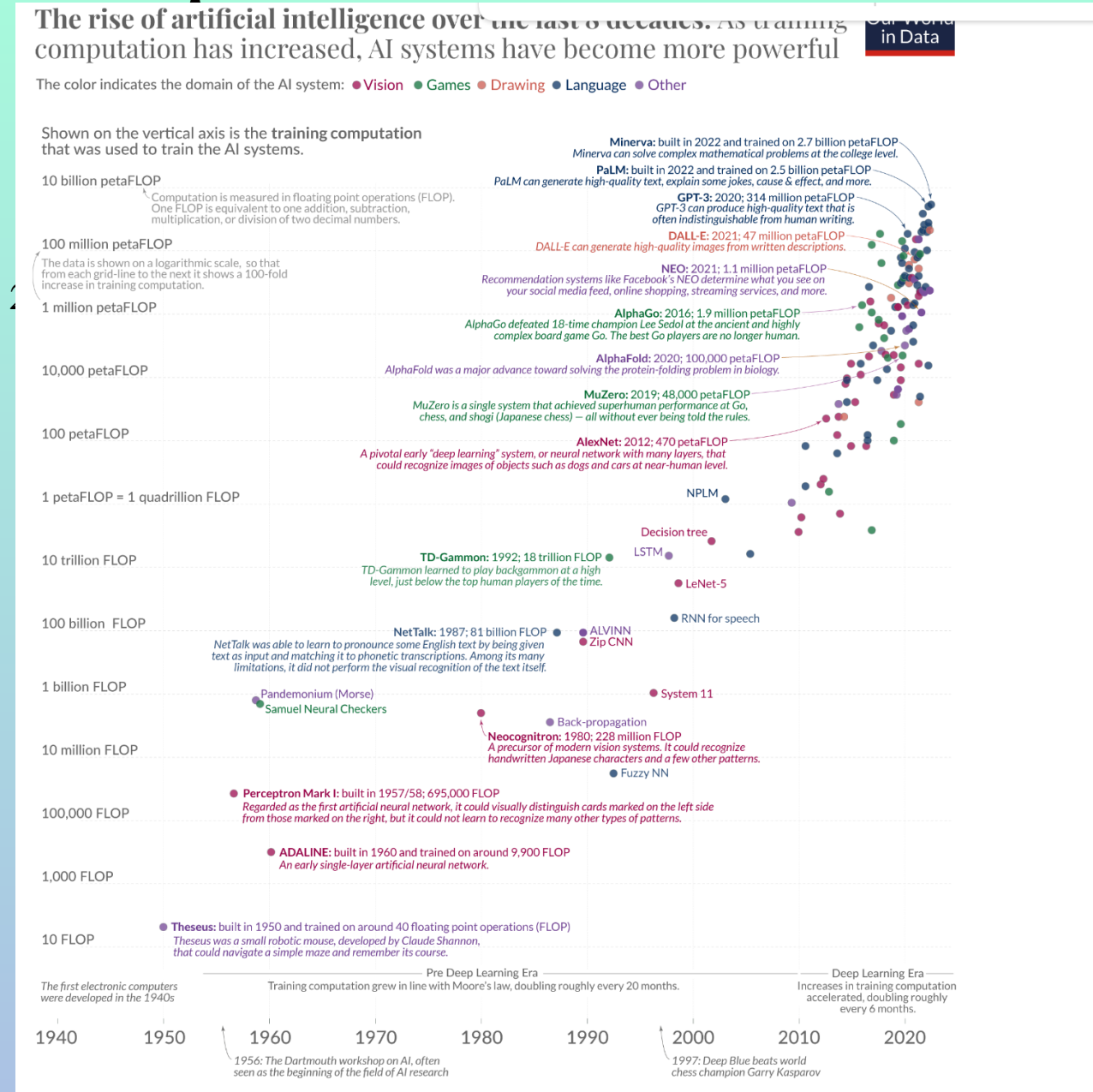
By: \_\_\_\_\_

New: \_\_\_\_\_

 Bonus: find an AI image that fooled people. What gave it away? · Cite real sources – don't just ask a chatbot!

# Artificial Intelligence Development Timeline

- The rise of AI over the years
- Computing capacity – by counting FLOPS
- Floating point operations per second
- A FLOP is a floating-point operation; e.g.  $3.1416 + 2$



Source: [2]

# The Three Ingredients of Modern AI



## Data

- The fuel: text, images, sensor, game telemetry
- More and richer data → better accuracy
- But: quality, bias and consent matter



## Algorithms

- The recipe: models that learn patterns
- Decision trees, Naïve Bayes, k-means, neural nets, Transformers
- Most of this module!



## Compute

- The engine: GPUs, TPUs, NPUs in phones
- Massively parallel matrix maths (FLOPs)
- Frontier training compute has grown ~4–5× per year

**Data** × **Algorithms** × **Compute** = **Modern AI**

# The Three Ingredients of Modern AI

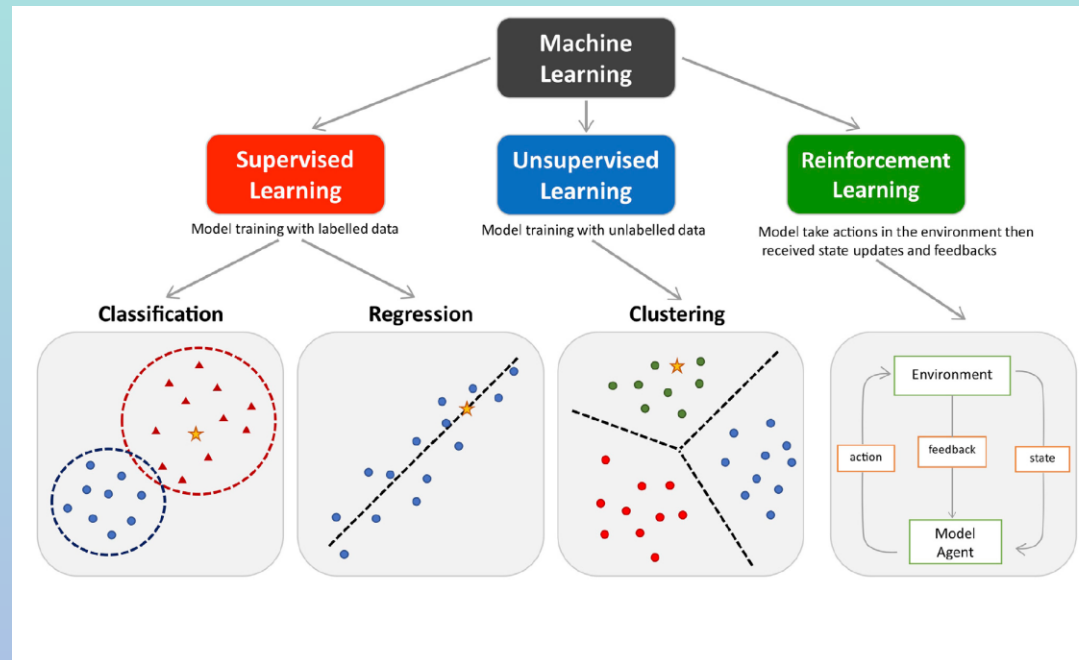
- Data
  - Prerequisite of AI
  - A core factor that promotes AI to improve recognition rate and accuracy
  - Increased growth in data generation
  - More comprehensive data – high dimensional data
- Algorithm
  - Mathematical models designed to mimic human intelligence by learning from data, recognize patterns, decision making and prediction.
    - Speech recognition, semantic analysis, recommendation system etc.

# Drivers and technologies of AI

- Machine learning (ML)
  - Developing algorithms that enables learning, prediction or decision making based on the given data, without being explicitly programmed.
  - Types of problem solved using ML – Clustering, Classification, Prediction and Dimensionality reduction
  - Four categories of ML algorithm – Supervised Learning, unsupervised learning, and reinforcement learning.
    - **Supervised learning**
      - Labelled data-> Model training-> Prediction ( Classification and Regression)
      - Classification – prediction of output of discrete values
      - Regression – prediction of output of continuous values

# Drivers and technologies of AI

- Machine learning (ML)
  - Unsupervised learning
    - Clustering – data is grouped into different classes without any tags or labels
  - Reinforcement learning
    - Obtain rewards by interacting with environment, assessing the action quality at reward levels, and then training the model.



# Transformers and large language models (2017–2022)

## Four generations of language models

1990s

### Statistical

n-gram models predict the next word from a short window.

2003

### Neural

Neural nets learn word embeddings (Bengio et al.).

2018

### Pre-trained

ELMo and BERT: pre-train on text, then fine-tune for a task.

2020

### Large

GPT-3: 175B parameters, few-shot learning from the prompt.

**2017** The Transformer: “Attention Is All You Need”

**Nov 2022** ChatGPT brings LLMs to the general public



**Your coursework lives here.** Fine-tune BERT-family models (DistilBERT, BERTweet) versus zero-shot and few-shot prompting of an LLM, and decide which is the better tool for the job.

# Agentic AI arrives (2023 – now)

Early 2023



## AutoGPT and BabyAGI

An LLM plus tool calls plus task-chaining loops: the first “generative agents”.

2023–2024



## Agent frameworks

MetaGPT, AutoGen, CrewAI, LangGraph: teams of role-based agents sharing memory.

Dec 2025



## Standards emerge

Model Context Protocol joins the Linux Foundation’s Agentic AI Foundation.

A CAUTION FROM THE ANALYSTS

# 40%+

of agentic AI projects are predicted to be cancelled by the end of 2027.

Reasons cited: escalating costs, unclear business value, inadequate risk controls.

Gartner press release, 25 June 2025 [10]

# Weak, Strong and Super AI

ANI

## Weak / Narrow AI

*All deployed AI today*

**Does one job, or a narrow family of jobs, very well.**

- Trained for specific tasks and data
- Often superhuman inside its niche
- No general understanding; brittle outside its training

*Chess engines, spam filters, translation, voice assistants, recommender systems*

AGI

## Strong / General AI

*Not yet achieved*

**Human-level competence across any intellectual task.**

- Learns new skills and transfers knowledge
- Reasons, plans and uses common sense
- Some definitions add genuine understanding

*Fictional so far: think of a film android*

ASI

## Super AI

*Hypothetical*

**Surpasses the best humans in every domain.**

- Speculative: no such system exists
- Central to AI-safety debates
- Raises questions of control and alignment

*Discussed in AI-safety research and policy*

**Where the terms come from:** Searle [11] - weak AI treats the computer as a tool for studying the mind; strong AI claims an appropriately programmed computer really is a mind.

# How do we decide? Three lenses



## Autonomy

How much does it decide and act without a human?



WEAK

STRONG

Acts when asked; goals set by people

Sets its own goals; acts independently



## Reasoning

Can it plan, infer and handle uncertainty across domains?



WEAK

STRONG

Pattern matching in one domain

General reasoning and common sense



## Adaptability

Can it take on new tasks without redesign or retraining?



WEAK

STRONG

Fails outside its training

Learns new skills like a person

### WORKED EXAMPLE · A ROBOT VACUUM

**Autonomy:** Medium

Runs alone, but the owner sets the goal

**Reasoning:** Low

Maps rooms and plans routes only

**Adaptability:** Low

Cannot take on a new kind of task

Verdict

**Weak AI**

# What would Strong AI need?

## Capabilities most definitions expect



**Reason and plan** under uncertainty



**Represent knowledge** including common sense



**Learn quickly** new skills from little experience



**Communicate** in natural language



**Transfer** knowledge across domains



**Understand?** consciousness is contested

## Searle's Chinese Room (1980)



### Fluent replies, yet no understanding inside.

Syntax is not semantics: symbol manipulation alone does not amount to understanding, so passing the Turing Test does not prove a mind.

**Open problems:** no agreed definition or test of consciousness; alignment and control.

# Weak, strong, or somewhere in between?

In groups: place each system on the line and justify it with at least two lenses. 4 minutes.

CLEARLY WEAK

STRONG (AGI)



## Chess engine

Beats every human; plays only chess



## Autonomous car

Drives itself in traffic and weather



## Robot vacuum

Maps your home and plans routes



## AI research agent

Searches, runs code, drafts a report



## LLM chat assistant

Writes, codes, translates and explains



## A fictional android

Like Data in Star Trek

# From rules to reactive text to agents

**Agentic does not mean strong.** Every rung is still narrow AI. The ladder measures autonomy and structure, not general intelligence.



1980s →

## Classical agents

Symbolic if-then rules and formal logic. Fixed goals, state-based memory. Cannot learn.

*STRIPS · Prolog · JADE · Jason*



2018 →

## Generative AI

Generates text, images or code from a prompt. Reactive, single-step, no persistent memory or tools.

*GPT-4 · Claude · Gemini · LLaMA*



2023 →

## AI agents

One LLM-driven agent uses predefined tools for a specific task. Memory is limited to a session.

*Copilot · plugins · voice assistants*



2023 →

## Agentic AI

Teams of specialised agents with persistent memory, dynamic tool use and multi-step planning.

*AutoGen · LangGraph · CrewAI · MetaGPT*

# What makes AI “agentic”? Four pillars

## AGENTIC AI

Systems that pursue goals with a degree of autonomy: they break tasks down, reason over many steps, remember, use tools and collaborate, often as a team of specialised agents.



### Multi-step reasoning

Break a goal into sub-problems, weigh options and revise the plan.

*Chain-of-Thought · Tree-of-Thought*



### Persistent memory

Remember across steps and sessions, not just inside one prompt.

*Context · retrieval (RAG) · scratchpad*



### Tool use

Act on the world: search, run code, call APIs, use software.

*Search · code · APIs · Agentic RAG*



### Active interaction

Work with users, their environment and other agents.

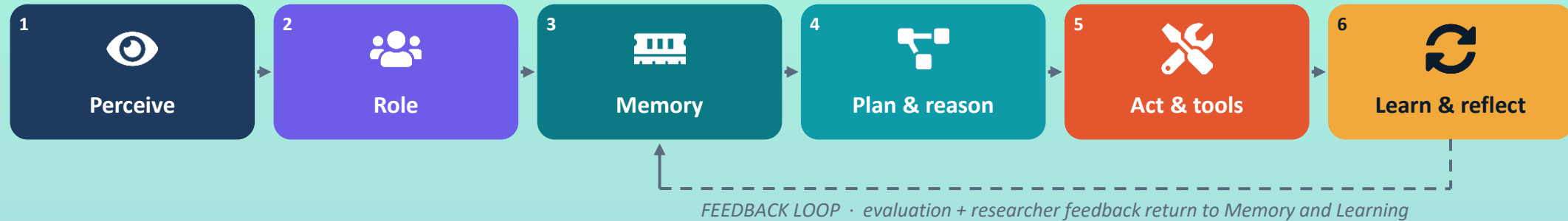
*Task division · shared results*

***The language model is one component of an agentic system, not the whole system.***

# Four kinds of system, side by side

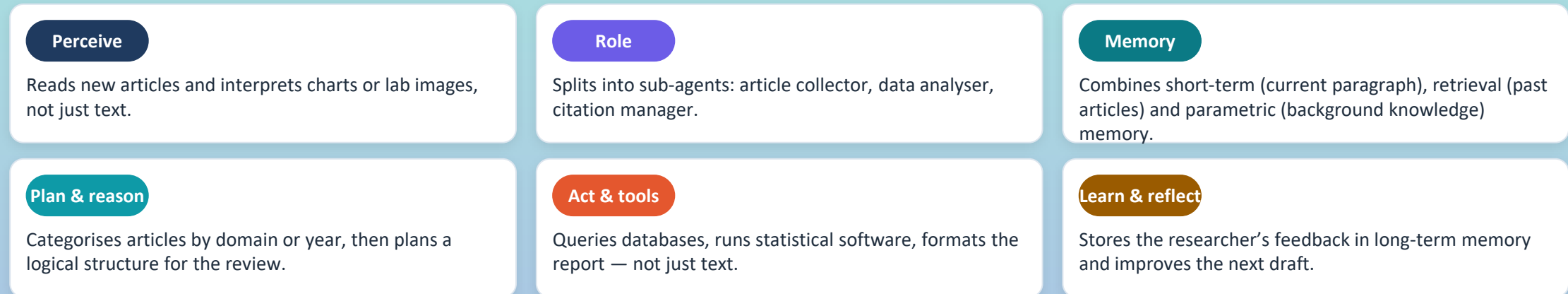
|                        | Classical rule-based agent                     | Generative AI                             | AI agent (single)                     | Agentic AI                                                |
|------------------------|------------------------------------------------|-------------------------------------------|---------------------------------------|-----------------------------------------------------------|
| <b>Reasoning</b>       | Symbolic if–then rules (e.g. BDI)              | Statistical pattern generation; reactive  | Limited, task-oriented                | Multi-step (Chain/Tree-of-Thought) with reflection        |
| <b>Memory</b>          | Working memory plus a fixed knowledge base     | Only the context window                   | Session-bound; nothing carried over   | Persistent, multi-layer (episodic, retrieval, scratchpad) |
| <b>Tools</b>           | Predefined and limited                         | None directly; needs external integration | Specific, predefined (APIs, search)   | Dynamic selection: search, code, Agentic RAG              |
| <b>Planning</b>        | Predefined goals; search-based planning        | Single-step response; no goals            | Limited planning for one task         | Task decomposition, adaptive planning, self-regulation    |
| <b>Main limitation</b> | Brittle in unforeseen situations; cannot learn | Hallucination; cannot act independently   | No collaboration or self-improvement  | Coordination complexity, cost, safety and reliability     |
| <b>Examples</b>        | STRIPS, Prolog, JADE, Jason                    | GPT-4, Claude, Gemini, LLaMA              | Copilot, ChatGPT plugins, Siri, Alexa | AutoGen, LangGraph, CrewAI, MetaGPT                       |

# Anatomy of an agent: six components



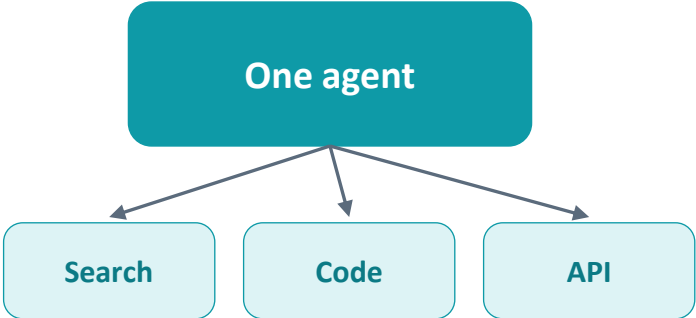
Not the same as the “four pillars” slide — that was why Agentic AI is different; this is what it is built from (Peykani et al., section 3).

## WORKED EXAMPLE · A LITERATURE-REVIEW ASSISTANT (adapted from the paper’s own example)

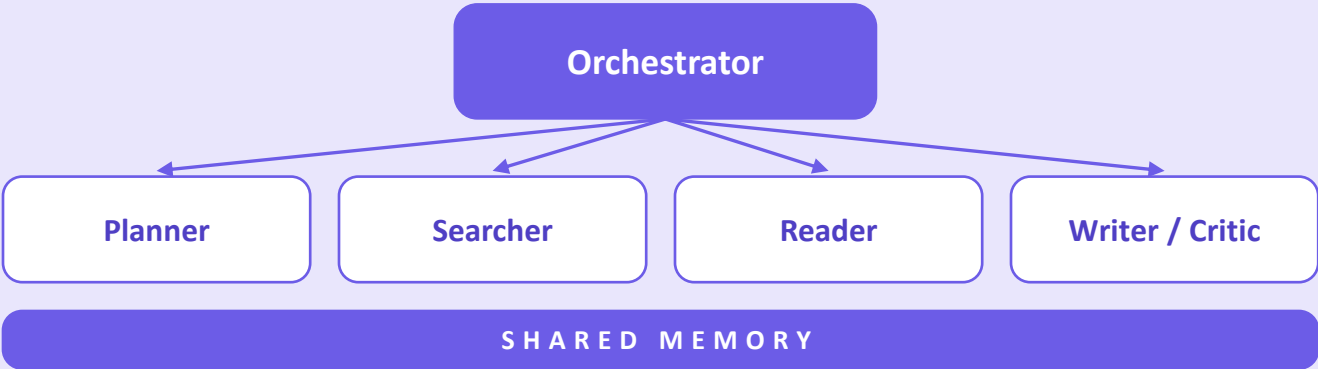


# One agent, or a team?

## Single agent



## Multi-agent team (a research assistant)



|                   | Single agent                   | Multi-agent                                   |
|-------------------|--------------------------------|-----------------------------------------------|
| Scalability       | Limited: harder as tasks grow  | High: add agents to add capability            |
| Specialisation    | One generalist does everything | Distinct roles and expertise                  |
| Fault tolerance   | Single point of failure        | Redundancy and distributed design             |
| Cost & complexity | Simple and cheap to build      | More design effort, compute and risk of loops |

# Design an agent team

In groups: pick one scenario and complete the design sheet. 5 minutes.

## A Airline disruption desk

A storm cancels 60 flights. Rebook passengers, keep them informed and apply compensation rules.

## B Smart-campus energy manager

Cut energy use across buildings without disturbing lectures or lab equipment.

## C Security operations triage

Thousands of alerts a day: investigate, prioritise and contain real incidents.

## D Software release manager

Review changes, run tests and deploy to production safely.

### YOUR DESIGN SHEET



#### Goal

What does success look like?



#### Agents and roles

Who does what?



#### Tools

What can each agent use?



#### Memory

What is remembered and shared?



#### Human checkpoint

When must a person approve?

# Who is accountable when an agent gets it wrong?



An airline's autonomous rebooking agent reads a customer email with hidden instructions (a prompt injection). Following them, it books 40 non-refundable hotel rooms.

**Vote: who bears the main responsibility?**

- ☐ The developer who built the agent
- ☐ The airline that deployed it
- ☐ The provider of the language model
- ☐ The customer who sent the email
- ☐ Nobody: the agent did it

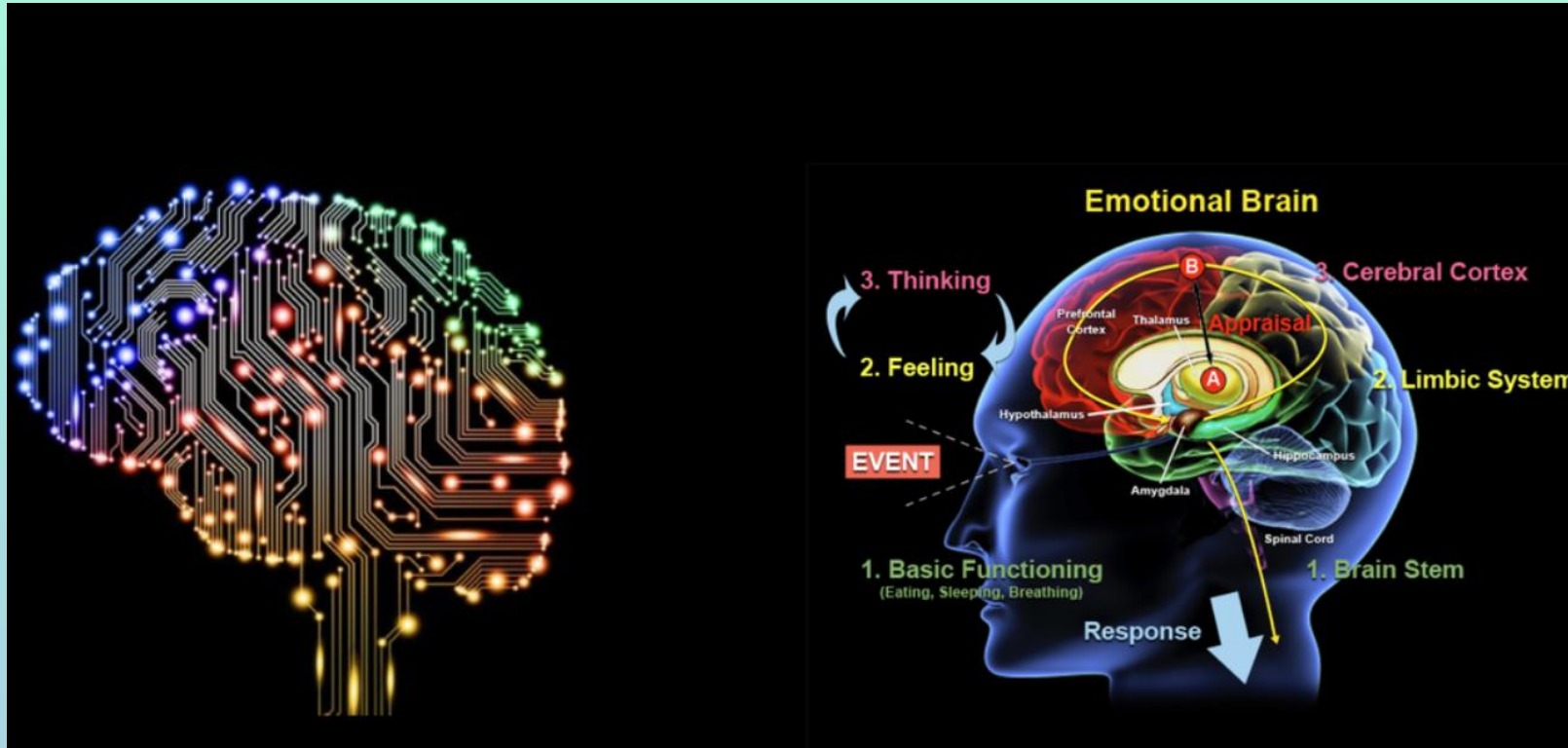
**What controls would have stopped it?**

- A spending limit per action
- Human approval above a threshold
- Treat email content as untrusted data
- Logs to trace the decision chain

# Artificial Intelligence V Human Intelligence

| Key insights       | A Intelligence                                                                                                                                                                                                                         | Human Intelligence                                                                                                                                                            |
|--------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Functionality      | Specific functionalities with high efficiency.<br>Depends on the input (data) and programming from human                                                                                                                               | Broader and adapt to any task; simple, to complex and abstract                                                                                                                |
| Learning ability   | Depends on ML algorithms ; learning is limited to data and training;<br>Hard facts and data                                                                                                                                            | Stems from various sources such as education, training, and life experiences.<br>Data, ethics, emotions and judgement                                                         |
| Scope              | Constantly expanding, finance and healthcare to education; the processes and abilities depends on human decision.                                                                                                                      | It is virtually endless. Not limited by predetermined parameters, allowing space for innovation.                                                                              |
| Flexibility        | Limited based on the training and data; ability to think outside the box or apply knowledge to unfamiliar situation is limited                                                                                                         | Highly flexible; absorb new information, learn and make decisions;                                                                                                            |
| Accessibility      | AI is increasingly accessible due to the availability of technology across industries. There is democratized access to its capabilities. However, its accessibility is still limited to people who have niche expertise in technology. | Human intelligence is extremely accessible on the individual and organizational level. It is not dependent on technology or expertise in data but is readily available within |
| Pace of Operations | One of the biggest advantages of AI is its ability to process and complete tasks at a pace exceeding human capabilities.                                                                                                               | Compared to AI, human intelligence is not as fast-paced.                                                                                                                      |

# Artificial Intelligence V Human Intelligence



## FOUR THINGS TO TAKE AWAY

1

### Definitions matter

AI can be defined by behaviour or by goals. Modern AI favours the rational agent, not imitation.

2

### History is cyclical

Breakthroughs, hype and winters. Data, compute and better algorithms explain today's boom.

3

### Judge with three lenses

Autonomy, reasoning, adaptability. Every deployed AI system today is still weak AI.

4

### Agentic ≠ strong

Agentic AI = language model + memory + tools + planning, often as a team.  
More autonomy, new risks.

**Question to carry away:** where would you never let an agent act on its own?

# Hands on AI Tool

## Class Activity

- [Image Model - Teachable Machines](#)
- To Explore
- [The brief history of artificial intelligence: the world has changed fast — what might be next? | Our World in Data](#)
- [www.ibm.com/think/topics/artificial-general-intelligence](http://www.ibm.com/think/topics/artificial-general-intelligence)
- [www.ibm.com/think/topics/agent-ai](http://www.ibm.com/think/topics/agent-ai)
- [www.ibm.com/think/topics/ai-agents](http://www.ibm.com/think/topics/ai-agents)
- [Building Effective AI Agents](#)
- [The 2026 AI Index Report | Stanford HAI](#)
- [www.elementsofai.com](http://www.elementsofai.com)

# Reference

- [1] A. Kaplan and M. Haenlein. A. Kaplan and M. Haenlein. Siri, siri, in my hand: Who's the fairest in the land? on the interpretations, illustrations, and implications of artificial intelligence. Siri, siri, in my hand: Who's the fairest in the land? on the interpretations, illustrations, and implications of artificial intelligence. *Bus. Horiz. Bus. Horiz.* 6262(1), pp. 15–25. pp. 15–25. 2019
- [2] *The brief history of artificial intelligence: the world has changed fast — what might be next?*. Available: [The brief history of artificial intelligence: the world has changed fast — what might be next? | Our World in Data](#)
- [3] Alan Turing, "A Turing Computing Machinery and and Intelligence ," *A quarterly review of psychology and philosophy*.
- [4] B. Huang, Y. Huan, L. Xu, L. Zheng, Z. Zou, *Automated trading systems statistical and machine learning methods and hardware implementation: a survey*, *Enterprise Inf. Syst.* 13 (1) (2019) 132–144.
- [5]. C. R. Jones and B. K. Bergen, "Large language models pass the Turing test," arXiv:2503.23674, 2025
- [6] J. Peng et al, "Machine Learning Techniques for Personalised Medicine Approaches in Immune-Mediated Chronic Inflammatory Diseases: Applications and Challenges," *Front. Pharmacol.*, vol. 12, 2021. . DOI: 10.3389/fphar.2021.720694
- [7] W. X. Zhao et al., "A Survey of Large Language Models," *Frontiers of Computer Science*, vol. 20, no. 12, 2026, doi: 10.1007/s11704-026-60308-3
- [8] M. A. K. Raiaan et al., "A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges," *IEEE Access*, vol. 12, pp. 26839–26874, 2024, doi: 10.1109/access.2024.3365742
- [9] P. Peykani, S. Ghanidel, I. Javadi-Sisi, V. Snasel, and S. Mirjalili, "A Holistic Review of Agentic AI Frameworks, Applications, and Research Trajectories," *Archives of Computational Methods in Engineering*, 2026, doi: 10.1007/s11831-026-10675-8
- [10] Gartner, "Gartner predicts over 40% of agentic AI projects will be canceled by end of 2027," press release, Jun. 25, 2025.
- [11] J. R. Searle, "Minds, brains, and programs," *Behavioral and Brain Sciences*, vol. 3, no. 3, pp. 417–424, 1980.