

CSE422: Artificial intelligence

Introduction

Fundamentals of AI

Asif Shahriar
Lecturer, CSE, BRACU

WHAT is AI?

- The science and engineering of making **intelligent machines**, especially intelligent computer programs. – *John McCarthy, 1956*
 - Chess-playing AI
- The science of making machines do things that would require **intelligence if done by men**. – *Marvin Minsky, founder of MIT AI Lab*
 - Robot vacuum cleaner
- The study of agents that receive percepts from the environment and perform actions. – *Stuart Russell and Peter Norvig*
 - A self-driving car

Main goals of AI

- **Represent** and **store** knowledge
- **Retrieve** and **reason** about knowledge
- **Behave** intelligently in complex environments
- Develop interesting and useful **applications**
- **Interact** with people, agents, and the environment

AI: Real World Impact

- **Public imagination**
 - AI artist

AI: Real World Impact

- **Public imagination**

- AI artist
- *Prompt: “Draw an image of a forest, where a tiger, a deer and a cow are drinking water from the lake.”*

[This image is generated by GPT-4o.]



AI: Real World Impact

- **Public imagination**

- AI artist
- *Prompt: “Draw a Cyberpunk-style abstract picture about the sufferings of the people of Gaza.”*

[This image is generated by GPT-4o.]



AI: Real World Impact

- **Public imagination**

- AI artist
- AI story-writer
- *Prompt: “Write a joke about an early space program that had used monkeys as astronauts for many missions, and then sent the first human astronaut into space along with a monkey. Add a slightly dark tone to the story.”*

[This story is generated by GPT-4o.]

In the early days of the space program, they sent monkeys on mission after mission. Dozens of successful flights. Then finally, they launched their first human... along with a monkey.

Mid-flight, Mission Control radioed in:

“Monkey, engage thrusters.”

The monkey did.

“Monkey, initiate course correction.”

The monkey complied.

Then they called the human.

“Human, are you there?”

“Roger that,” he replied. “What’s my job?”

Mission Control paused, then said:

“Your job is to make sure the monkey survives. He’s the one we trained.”

A long silence.

Then a quiet reply:

“...What happens if I fail?”

“Let’s just say we’ve only packed one return capsule.”

AI: Real World Impact

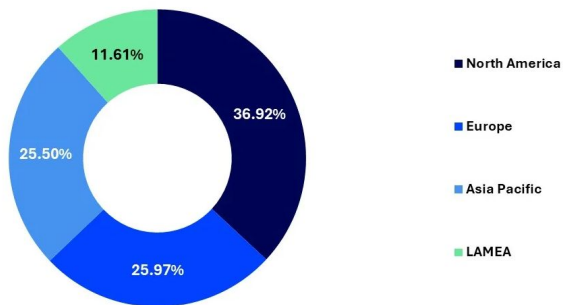
- **Public imagination**

- AI artist
- AI story-writer
- A serious debate
 - *Who owns AI-generated content?*
 - *Intellectual theft during training?*
 - *Truly creative, or just a remix?*
 - *Devaluing human art and storytelling?*

AI: Real World Impact

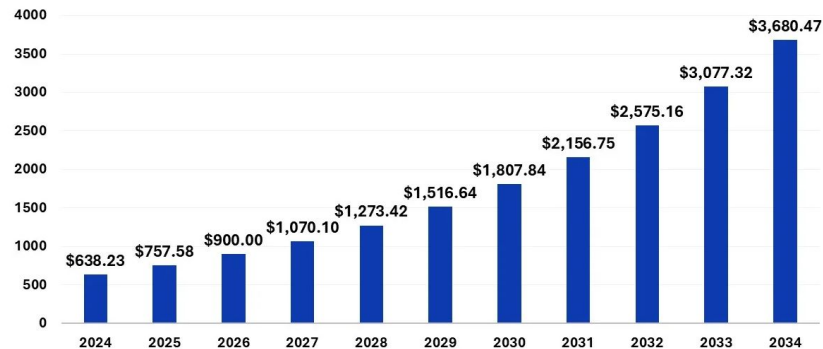
- Public imagination
- Economy

Artificial Intelligence (AI) Market Share, By Region, 2024(%)



Source: <https://www.precedenceresearch.com/artificial-intelligence-market>

Artificial Intelligence (AI) Market Size 2024 to 2034 (USD Billion)



Source: <https://www.precedenceresearch.com/artificial-intelligence-market>

AI: Real World Impact

- Public imagination
- Economy
- Politics



AI: Real World Impact

- **Public imagination**
- **Economy**
- **Politics**
- **Law**
 - Lots of lawsuits against AI companies involving copyright claims, training data, and human creativity devaluation
 - Who's responsible when AI causes harm?
 - Laws like the EU AI Act are attempting to classify and regulate risk levels in AI systems

US appeals court rejects copyrights for AI-generated art lacking 'human' creator

Getty Images and Stability AI face off in British copyright trial that will test AI industry

Reddit sues Anthropic for allegedly not paying for training data

AI: Real World Impact

- **Public imagination**
- **Economy**
- **Politics**
- **Law**
- **Labor**
 - Effects of automation
 - Demanding new skills in the face of AI
 - Wage polarization

AI: Real World Impact

- Public imagination
- Economy
- Politics
- Law
- Labor
- Scientific advancements

AlphaFold 3 predicts the structure and interactions of all of life's molecules

May 08, 2024
5 min read

Introducing AlphaFold 3, a new AI model developed by Google DeepMind and Isomorphic Labs. By accurately predicting the structure of proteins, DNA, RNA, ligands and more, and how they interact, we hope it will transform our understanding of the biological world and drug discovery.

DeepMind scientists say they trained an A.I. to control a nuclear fusion reactor

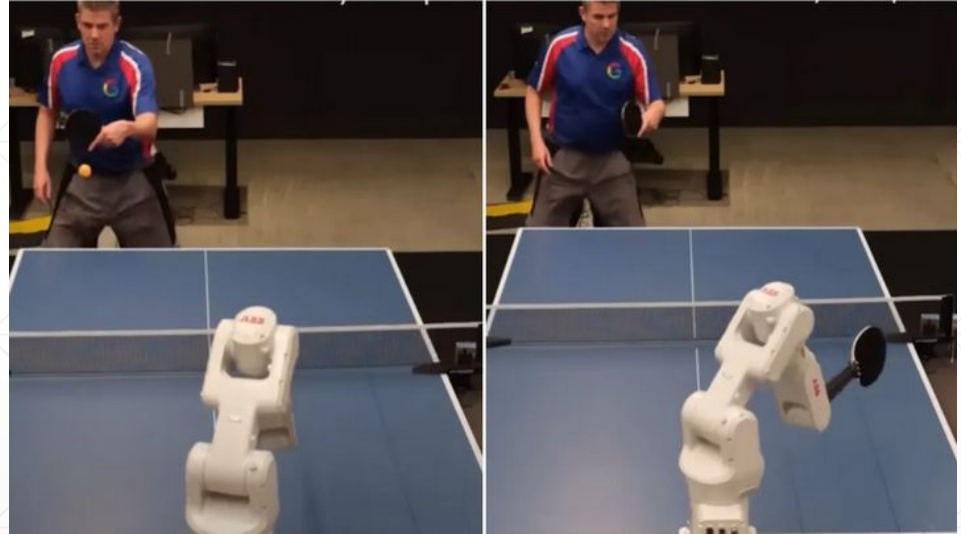
What can AI do?

- Win against humans in chess? ✓






What can AI do?





- Win against humans in chess? ✓
- Play a decent game of table tennis? ●








What can AI do?

- Win against humans in chess? 
- Play a decent game of table tennis? 
- Unload dishwashers at home? 







What can AI do?

- Win against humans in chess? 
- Play a decent game of table tennis? 
- Unload dishwashers at home? 
- Drive safely along the highway? 








What can AI do?

- Win against humans in chess? 
- Play a decent game of table tennis? 
- Unload dishwashers at home? 
- Drive safely along the highway? 
- Drive safely along the streets of San Francisco, or Dhaka? 

What can AI do?

- Win against humans in chess? 
- Play a decent game of table tennis? 
- Unload dishwashers at home? 
- Drive safely along the highway? 
- Drive safely along the streets of San Francisco, or Dhaka? 
- Shop from the internet? 

What can AI do?

- Win against humans in chess? 
- Play a decent game of table tennis? 
- Unload dishwashers at home? 
- Drive safely along the highway? 
- Drive safely along the streets of San Francisco, or Dhaka? 
- Shop from the internet? 
- Shop from supershops like Shopno? 

What can AI do?

- Win against humans in chess? ✓
- Play a decent game of table tennis? ●
- Unload dishwashers at home? ●
- Drive safely along the highway? ✓
- Drive safely along the streets of San Francisco, or Dhaka? ✗
- Shop from the internet? ✓
- Shop from supershops like Shopno? ✗
- Discover or prove new mathematical theorems? ●

What can AI do?

- Win against humans in chess? ✓
- Play a decent game of table tennis? ●
- Unload dishwashers at home? ●
- Drive safely along the highway? ✓
- Drive safely along the streets of San Francisco, or Dhaka? ✗
- Shop from the internet? ✓
- Shop from supershops like Shopno? ✗
- Discover or prove new mathematical theorems? ●
- Perform a surgical operation? ✗

What can AI do?

- Win against humans in chess? ✓
- Play a decent game of table tennis? ●
- Unload dishwashers at home? ●
- Drive safely along the highway? ✓
- Drive safely along the streets of San Francisco, or Dhaka? ✗
- Shop from the internet? ✓
- Shop from supershops like Shopno? ✗
- Discover or prove new mathematical theorems? ●
- Perform a surgical operation? ✗
- Identify a disease from symptoms? ●

What can AI do?

- Win against humans in chess? ✓
- Play a decent game of table tennis? ●
- Unload dishwashers at home? ●
- Drive safely along the highway? ✓
- Drive safely along the streets of San Francisco, or Dhaka? ✗
- Shop from the internet? ✓
- Shop from supershops like Shopno? ✗
- Discover or prove new mathematical theorems? ●
- Perform a surgical operation? ✗
- Identify a disease from symptoms? ●
- Win an art competition? ✓



What can AI do?

- Win against humans in chess? ✓
- Play a decent game of table tennis? ●
- Unload dishwashers at home? ●
- Drive safely along the highway? ✓
- Drive safely along the streets of San Francisco, or Dhaka? ✗
- Shop from the internet? ✓
- Shop from supershops like Shopno? ✗
- Discover or prove new mathematical theorems? ●
- Perform a surgical operation? ✗
- Identify a disease from symptoms? ●
- Win an art competition? ✓
- Write a funny story? ✓

What can AI do?

- Win against humans in chess? ✓
- Play a decent game of table tennis? ●
- Unload dishwashers at home? ●
- Drive safely along the highway? ✓
- Drive safely along the streets of San Francisco, or Dhaka? ✗
- Shop from the internet? ✓
- Shop from supershops like Shopno? ✗
- Discover or prove new mathematical theorems? ●
- Perform a surgical operation? ✗
- Identify a disease from symptoms? ●
- Win an art competition? ✓
- Write a funny story? ✓
- Construct a building? ✗

A (short) History of AI

- **1940—1950: Early days: neural and computer science meet**
 - 1943: McCulloch & Pitts: Boolean circuit model of brain
 - 1950: Turing's "Computing Machinery and Intelligence"
- **1950—70: Excitement! Logic-driven**
 - 1950s: Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
 - 1956: Dartmouth meeting: "Artificial Intelligence" adopted
 - 1957: Newell, Shaw, Simon GPS, General Problem Solver
 - 1965: Robinson's complete algorithm for logical reasoning
- **1970—90: Knowledge-based approaches**
 - 1969—79: Early development of knowledge-based systems
 - 1980—88: Expert systems industry booms
 - 1988—93: Expert systems industry busts: "AI Winter"

A (short) History of AI

- **1990—2010: Statistical approaches**

- Resurgence of probability, focus on uncertainty
- 1992: TD-Gammon attains human-level performance
- 1996: Kasparov defeats Deep Blue at chess
- 1997: Deep Blue defeats Kasparov at chess
- 2002: Embodied AI; Roomba vacuum invented
- 2004: First DARPA Grand Challenge robot race
- 2007: Checkers is solved!

- **2010—2017: Big Data, GPUs, Deep Learning**

- 2010: Google self-driving cars reach their 1000th mile
- 2011: Apple releases SIRI
- 2012: AlexNet (neural net) wins ImageNet (image recognition) competition
- 2015: DeepMind achieves human-level control in Atari games
- 2016: DeepMind's AlphaGo beat Go Master, Lee Sedol
- 2016: Google Translate migrates to neural networks

A (short) History of AI

- **2017—: Scaling up, Large Language Models**
 - 2017: Google invents Transformer architecture “*Attention is All You Need*”
 - 2017: DeepStack and Libratus defeat humans at poker
 - 2018-2020: AlphaFold predicts protein structure from amino acids
 - 2021-2022: Modern text-to-image generation
 - 2022: OpenAI release ChatGPT

What are we using AI for Today?

- **Healthcare:** medical imaging analysis, disease prediction, personalized medicine, drug discovery
- **Business & finance:** fraud detection, algorithmic trading, customer service
- **Retail & E-commerce:** recommendation system, inventory management, AI shopping assistants
- **Transportation:** self-driving cars, traffic management, route planning, autonomous delivery
- **Education:** personalized learning platforms, automated grading, virtual teaching assistants
- **Entertainment:** AI-generated music and art, content development, AI opponents in games
- **Cyber security:** threat detection, intrusion prevention, phishing detection
- **Industry:** Quality control, predictive maintenance, automation
- **Environment & agriculture:** precision agriculture, wildlife conservation, climate change modelling
- **Scientific research:** Used in almost every branch of science
- **Legal systems:** Legal document review, predictive case outcomes
- **And everywhere else:** Think of any field, AI is probably being used there in some form or capacity

Some Things AI CANNOT DO (YET)

- Physical tasks in unstructured environments
- General reasoning and common sense
 - Understand abstract humor, irony, or sarcasm across cultures
 - Solve problems requiring deep common-sense knowledge
- Truly creative and unaided scientific discovery
 - AI excels in **pattern recognition** and **remixing**, but struggles with paradigm-shifting **originality** and **insight**
- Ethics, emotion, and empathy
 - AI lacks consciousness, emotional states, and lived experience — it may **simulate** emotion, but cannot **feel**
- Intent and accountability
 - AI cannot define its own goals — current AIs only optimize human-defined objectives
 - Cannot be legally or morally accountable

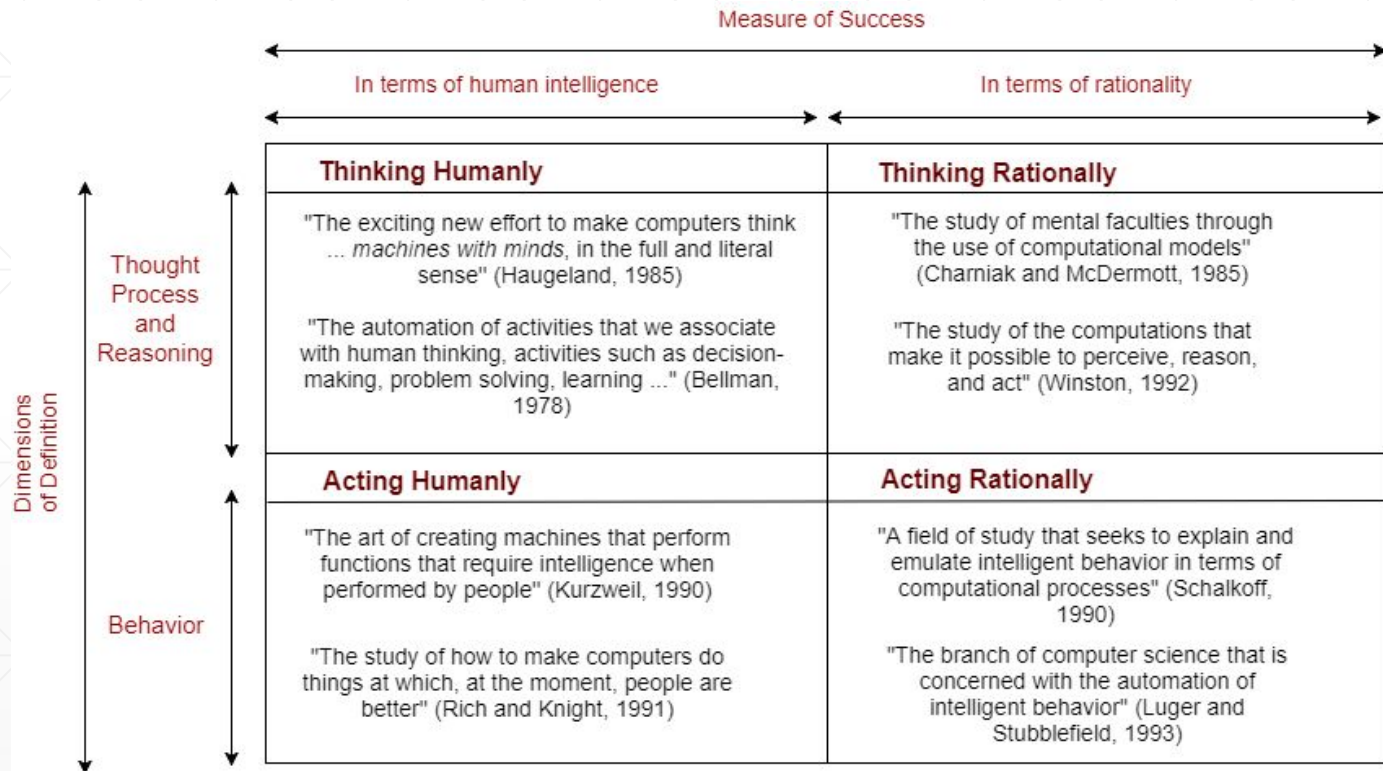
What is EASY and what is HARD?

- It has been easier to mechanize many of the high-level tasks we usually associate with “intelligence” in people
 - e.g., symbolic integration, proving theorems, playing chess, medical diagnosis
- However, it has been very hard to mechanize tasks that lots of animals can do
 - walking around without running into things
 - catching prey and avoiding predators
 - interpreting complex sensory information (e.g., visual, aural, ...)
 - modeling the internal states of other animals from their behavior
 - working as a team (e.g., with pack animals)
- *Is there a fundamental difference between the two categories?*

Human-brains vs Rationality

- Brains (human minds) are very good at making rational decisions, but not perfect
- Rational: *maximally achieving pre-defined goals*
- Humans are **often irrational** — we make emotional, biased, or inconsistent decisions, not focused on achieving optimal results
- “Brains are to intelligence as wings are to flight”
- Unlike software, human-brain is hard to reverse engineer
- Nonetheless, human-brain is still a great inspiration for AI
- We cannot yet build AI systems on the same scale as human-brain
- An AI acting rationally may not align with human values unless those values are encoded — a core challenge in AI safety

What makes a Successful AI?



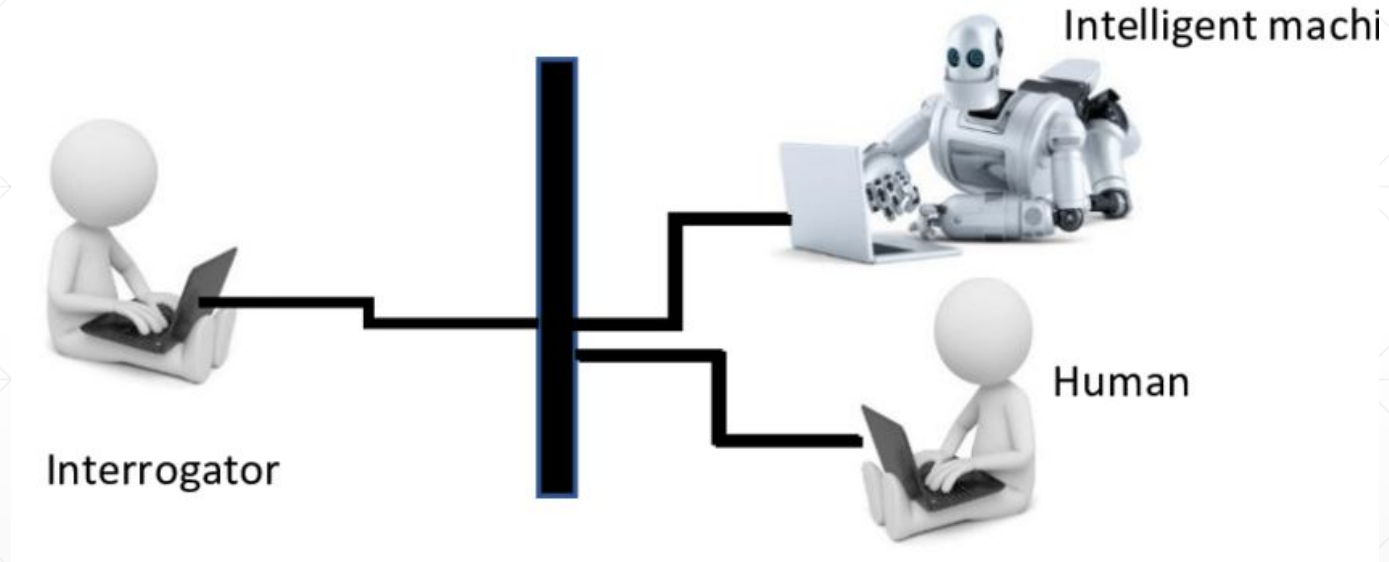
Thinking like Human

- A **cognitive** approach [Cognitive means being or involving conscious intellectual activity (such as thinking, reasoning, or remembering)]
- **Goal:** Model how humans actually *reason* — with flaws, intuitions, biases, etc
 - Not just to produce human-like behavior, but to produce a sequence of steps of the reasoning process that would be similar to what a human would do to solve the same problem
- **Example:**
 - Cognitive architectures like ACT-R or SOAR, which try to simulate human thought processes, including human-like errors or biases
 - ChatGPT mimicking human conversation styles, often using analogies and informal tone
- **Measure of success:** Does the AI reason in the same way a human would?
- **Useful for:** Cognitive science, psychology, human-AI interaction research

Acting like Human

- A **behaviourist** approach
- **Goal:** Make the AI behave indistinguishably from a human
 - Not interested in how you get results, just the similarity to what human results are
- **Example: The Turing Test** — if a machine can hold a conversation so natural that a human can't tell it's a machine
- **Measure of success:** Can the AI appear human in action?
- **Useful for:** Chatbots, virtual assistants, humanoid robots

TURING TEST



Thinking Rationally

- **Goal:** Use formal logic, rules, and reasoning **(that can be rendered in algorithms)** to reach correct conclusions
 - Emphasis on systems that are probably correct, and guarantee finding an optimal solution
- **Example:**
 - Expert systems (e.g., logic-based medical diagnosis engines)
 - Chain-of-thought reasoning in LLMs
- **Measure of success:** Does the AI follow valid reasoning paths to the right answer?
- **Useful for:** Domains requiring logical inference — math, theorem proving, knowledge systems

Acting Rationally

- **Goal:** Given goals and knowledge, take actions that maximize expected outcomes
 - The emphasis is to choose actions that optimize goals in uncertain environments
 - Can make decisions that appear non-human or counterintuitive if they maximize winning chances
- **Example:** Intelligent agents, like self-driving cars or game-playing AIs (e.g., Stockfish for chess), that make optimal decisions
- **Measure of success:** Does the AI choose the best possible action under uncertainty?
- **Useful for:** Real-world decision-making, robotics, autonomous systems
- **Heuristic:** A rule of thumb, strategy, trick, simplification, or any other kind of device which drastically limits search for solutions in large problem spaces
 - Heuristics **do not guarantee** optimal solutions (they do not guarantee any solution at all); but they can offer solutions which are **good enough most of the time**

Strong AI

- **Strong AI:** (Hypothetical) artificial intelligence with general cognitive abilities that match or surpass human intelligence
 - Capable of learning, reasoning, and self-awareness (like a human)
 - Potential to exhibit intent, creativity, and emotional understanding
 - The goal of OpenAI is to create **Artificial General Intelligence (AGI)** – which is still a theoretical concept (some conjecture that OpenAI has already developed AGI but is holding off for commercial reasons)
 - Think of Samantha from the 2013 movie ***HER***

Weak AI

- **Weak AI:** Software or system designed and trained for a specific, narrow task
 - Does not possess (require) consciousness, self-awareness, or general reasoning ability
 - Does not truly *understand*, uses **pattern recognition** and **statistical learning**
 - Example:
 - Virtual assistants: Siri, Alexa, Google assistant
 - Large language models: GPT-4o, Gemini, Claude, Deepseek
 - Recommendation systems: Netflix, Amazon, Spotify recommendations
 - Specific applications: Gmail spam filter, Facial recognition systems, Autonomous vehicles, Stockfish (chess engine)

“The key question is not whether machines can think, but whether people can still think.” – Stuart Russell