

# TDDE56: What is Artificial Intelligence?

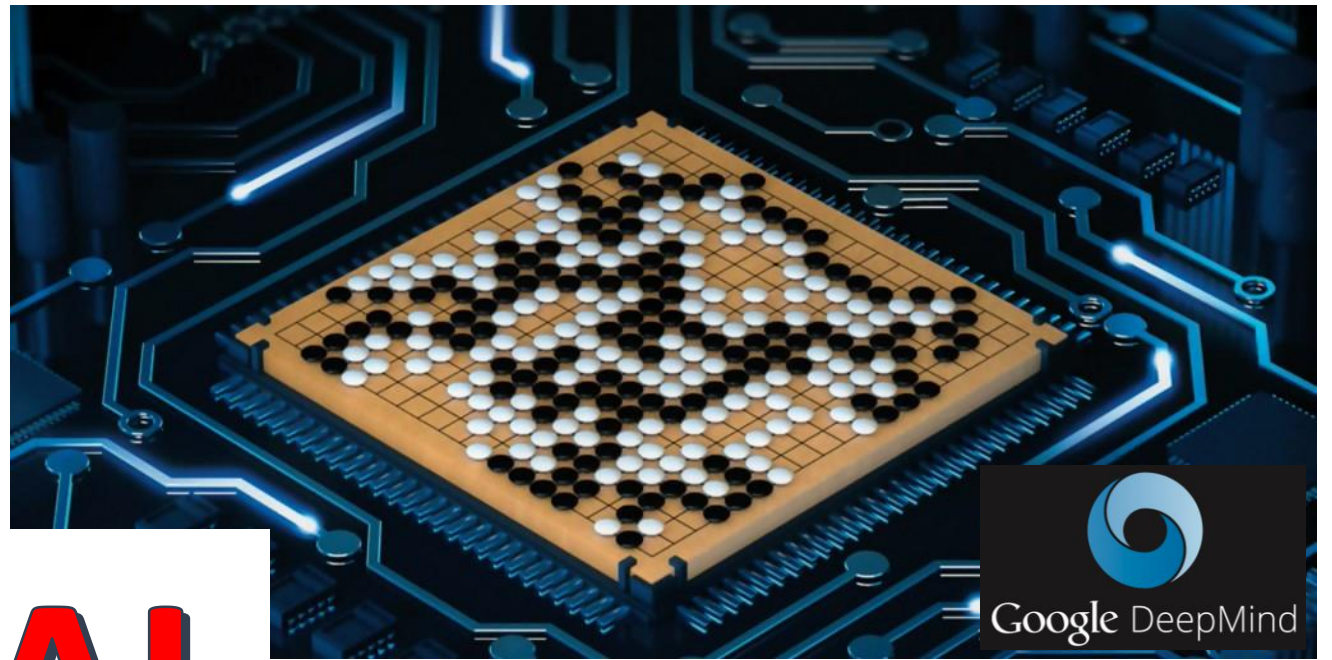
Fredrik Heintz

Dept. of Computer Science, Linköping University

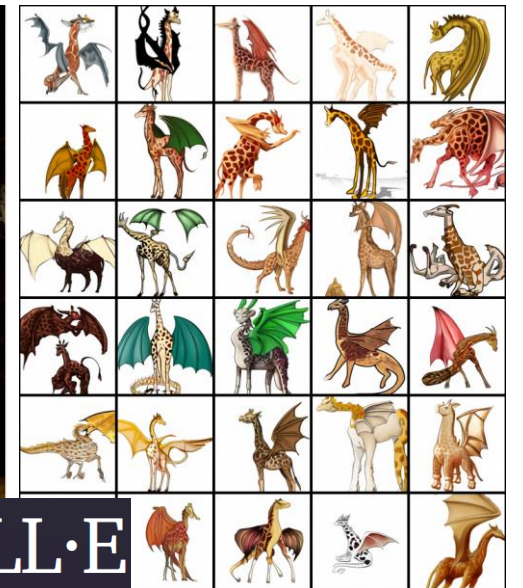
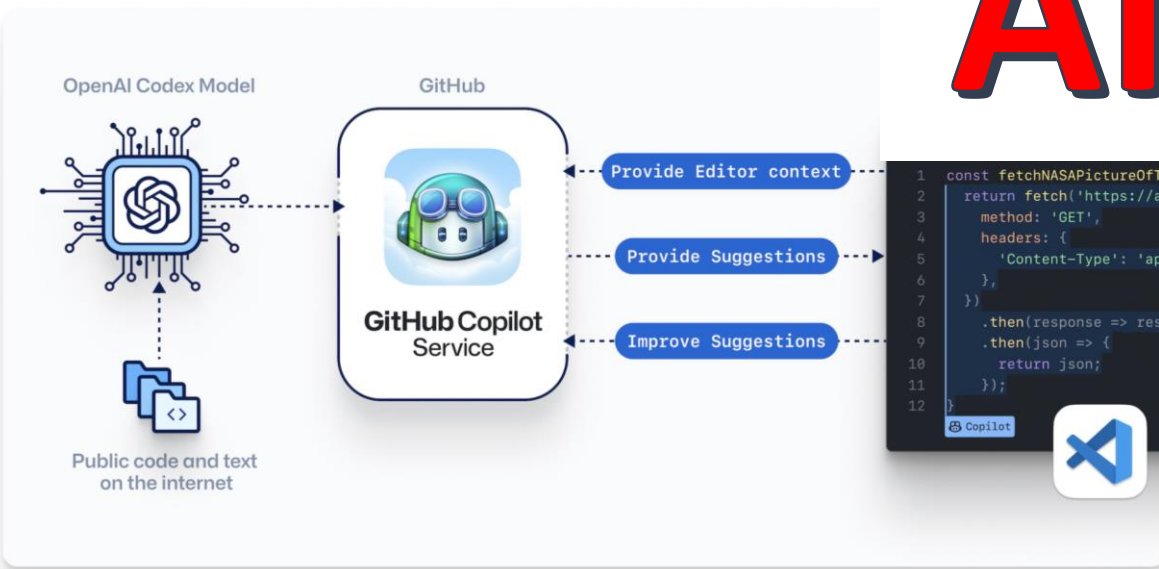
[fredrik.heintz@liu.se](mailto:fredrik.heintz@liu.se)

@FredrikHeintz



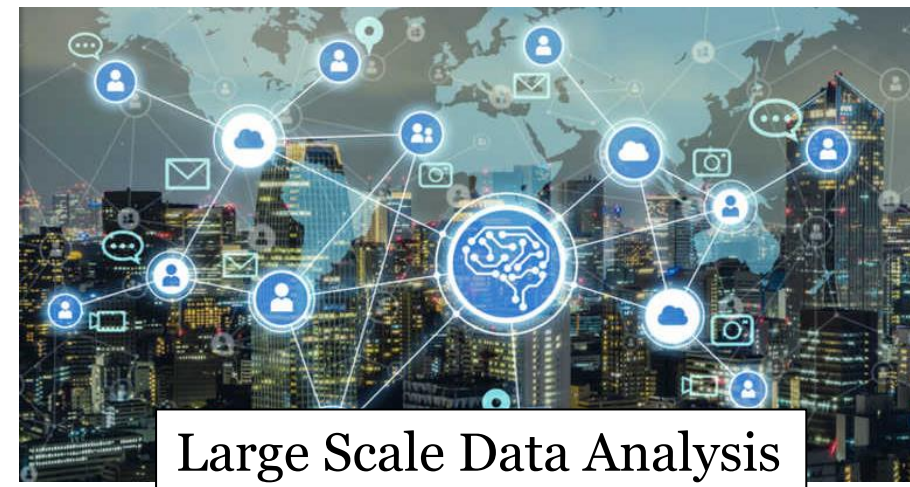
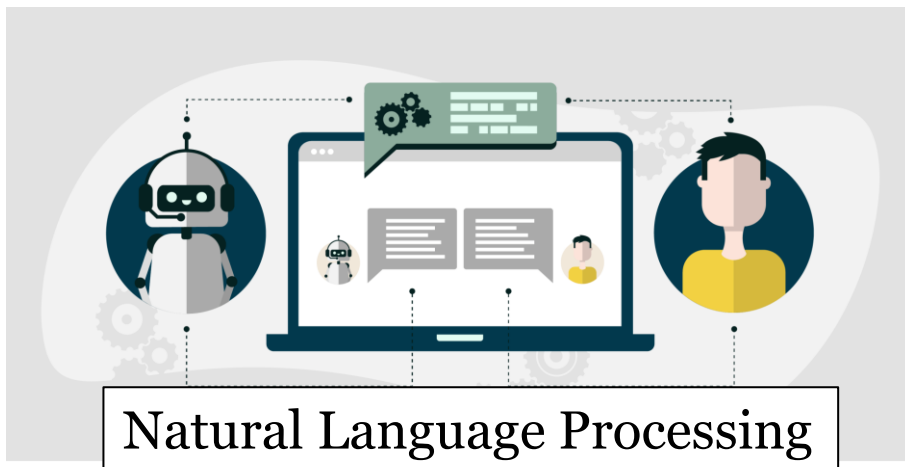
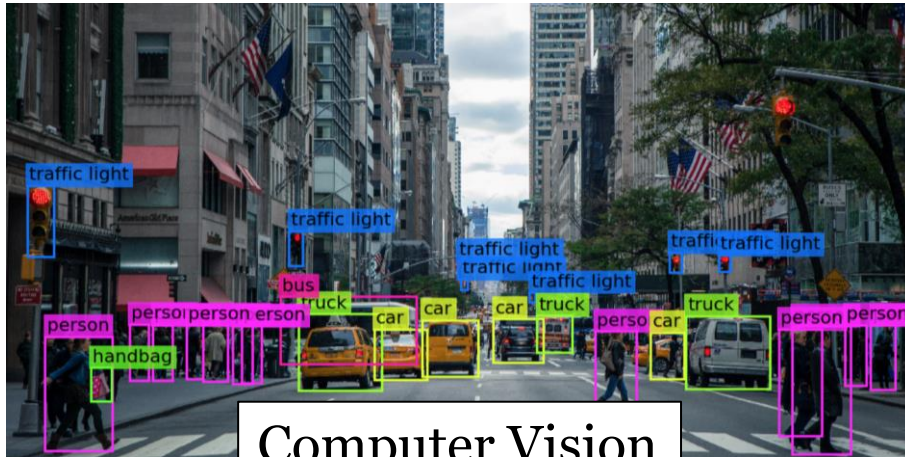


# AI



# DALL·E

# Applications of AI



# Artificial Intelligence – What is it? – Definitions

“Artificial Intelligence is the **science and engineering of making intelligent machines**, especially intelligent computer programs.”

- John McCarthy, Stanford

“Artificial intelligence (AI) refers to **systems that display intelligent behaviour** by analysing their environment and taking actions – with some degree of **autonomy** – to achieve specific **goals**.”

- EU Communication 25 April 2018

“the scientific understanding of the **mechanisms underlying thought and intelligent behavior** and their embodiment **in machines**.”

- AAAI

# What is intelligence?

- Legg and Hutter made a survey of 71 different definitions of intelligence in 2007.
- Commonly occurring features:
  - Is a property that an individual agent has as it interacts with its environment or environments.
  - Is related to the agent's ability to succeed or profit with respect to some goal or objective.
  - Depends on how able the agent is to adapt to different objectives and environments
- Based on this, they came up with: “Intelligence measures an agent's ability to achieve goals in a wide range of environments.”

# Artificial Intelligence – Four Views

Empirical Sciences

Fidelity to human performance

## Human-Centered

Mathematics/Engineering

Ideal concept of Intelligence

## Rationality-Centered

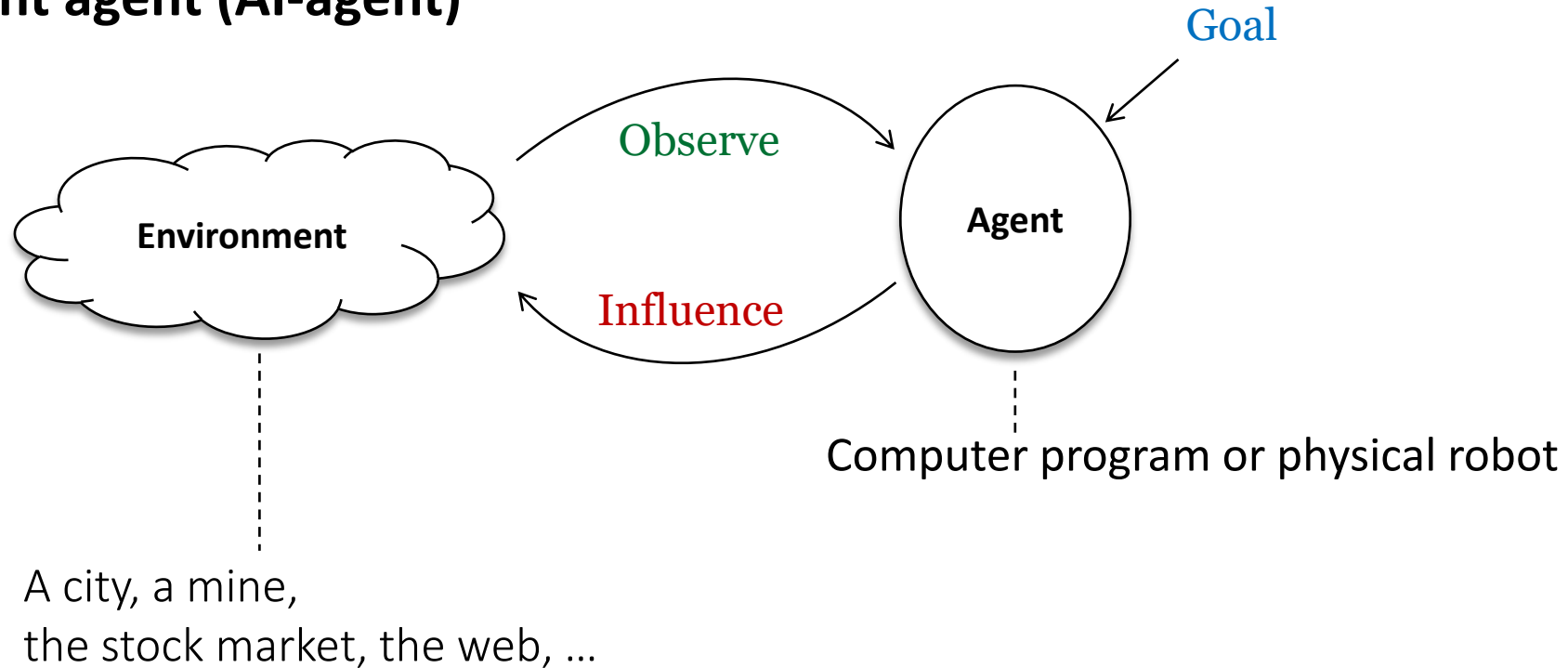
Thought Processes  
Reasoning

Systems that <u>think</u> like humans	Systems that <u>think</u> rationally
<p>”The exciting new effort to make computers think. . .machines with minds, in the full and literal sense.” (Haugeland, 1985)</p> <p>”[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning...”(Bellman, 1978)</p>	<p>”The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985)</p> <p>”The study of computations that make it possible to perceive, reason, and act.” (Winston, 1992)</p>
Systems that <u>act</u> like humans	Systems that <u>act</u> rationally
<p>”The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990)</p> <p>”The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)</p>	<p>”Computational Intelligence is the study of the design of intelligent agents.” (Poole et al., 1998)</p> <p>”AI . . . Is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)</p>

Behavior

# Artificial Intelligence – The Intelligent Agent Paradigm

- Intelligent agent (AI-agent)



- Intelligent capabilities through integration of AI techniques

- **Computer Science:**

Doing the right thing, efficiently, when you can define what that means in advance.

- **Artificial intelligence:**

Doing the right thing, efficiently, when you don't know what that means in advance.

- Shifting from design-time to run-time.

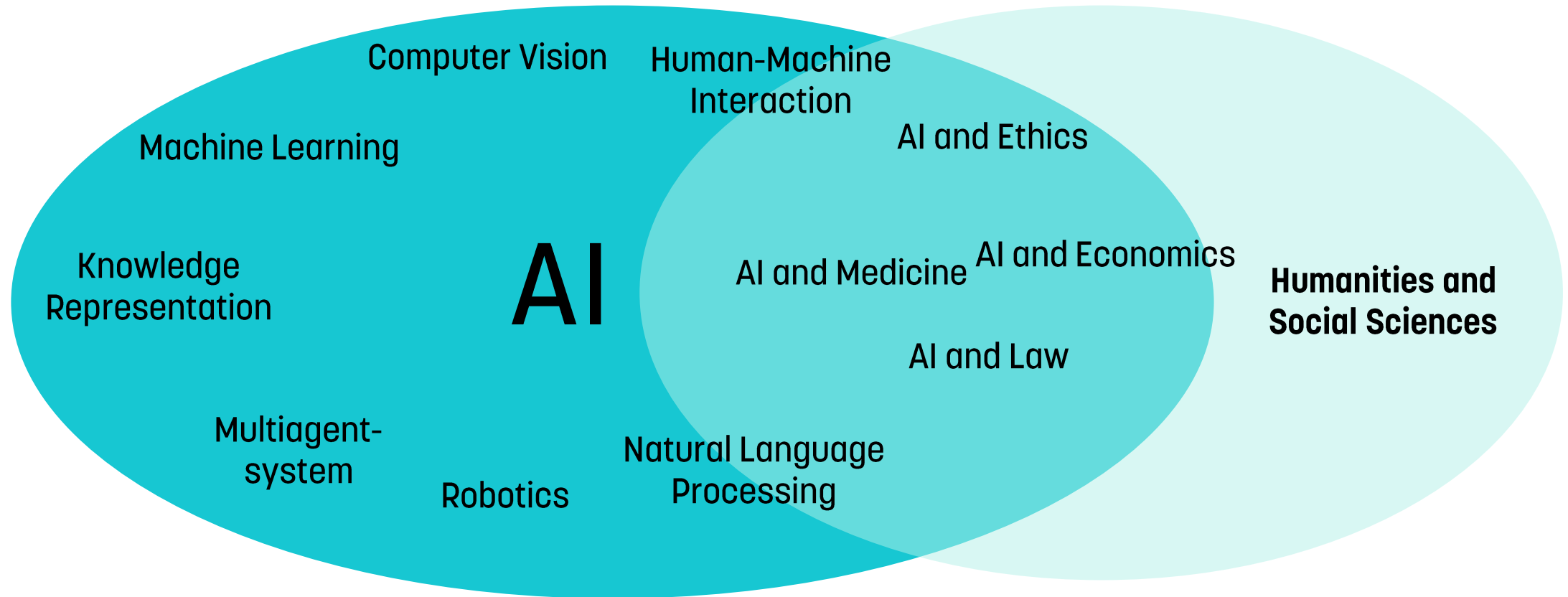


Artificial Intelligence  
(AI)

Machine Learning  
(ML)

Deep Learning (DL)

# Topics within Artificial Intelligence





Pure Logic

Pure Learning

- Slow thinking: deliberative, cognitive, model-based, extrapolation
- Amazing achievements until this day
- “*Pure logic is brittle*”  
noise, uncertainty, incomplete knowledge, ...





**Pure Logic**

**Pure Learning**

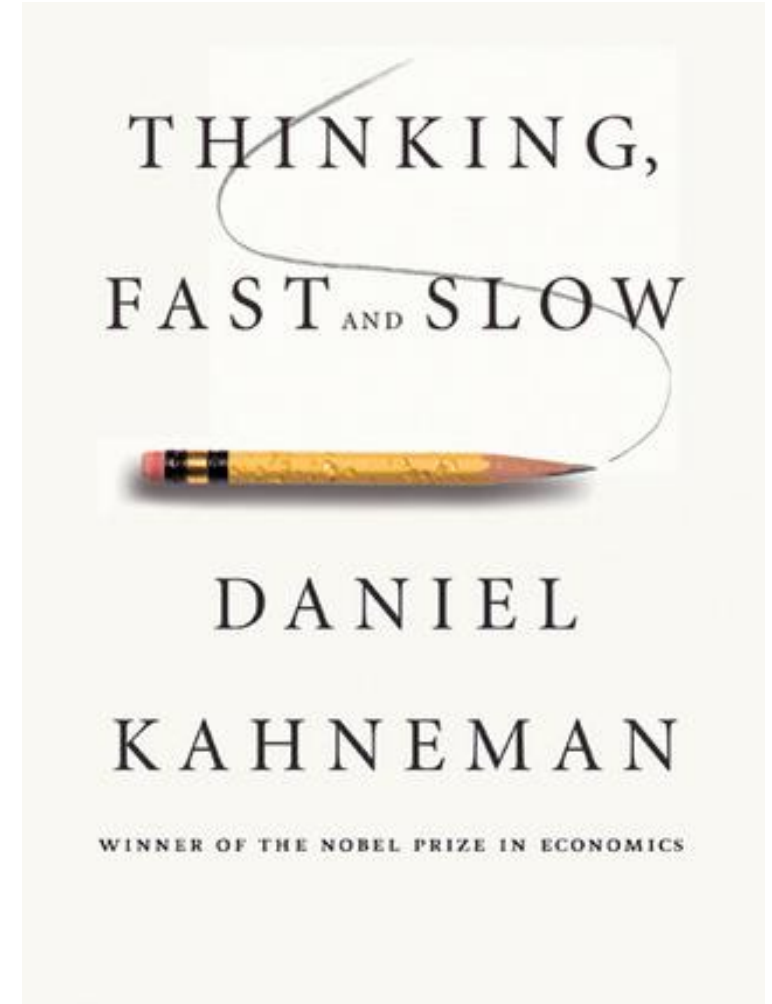
- Fast thinking: instinctive, perceptive, model-free, interpolation
- Amazing achievements recently
- “*Pure learning is brittle*”
  - bias, algorithmic fairness, interpretability, explainability, adversarial attacks, unknown unknowns, calibration, verification, missing features, missing labels, data efficiency, shift in distribution, general robustness and safety
  - fails to incorporate a sensible model of the world

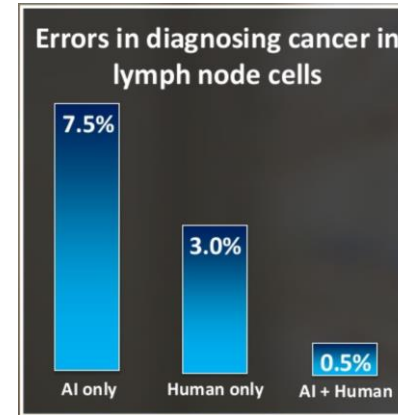


# Human and Computational Thinking

Figure 1: A Comparison of System 1 and System 2 Thinking

<b>System 1</b> "Fast"	<b>System 2</b> "Slow"
<b>DEFINING CHARACTERISTICS</b> Unconscious Effortless Automatic	<b>DEFINING CHARACTERISTICS</b> Deliberate and conscious Effortful Controlled mental process
<b>WITHOUT</b> self-awareness or control  "What you see is all there is."	<b>WITH</b> self-awareness or control  Logical and skeptical
<b>ROLE</b> Assesses the situation Delivers updates	<b>ROLE</b> Seeks new/missing information Makes decisions





“Weak human + machine + superior process was greater than a strong computer and, remarkably, greater than a strong human + machine with inferior process.”

Garry Kasparov