

UiO • **Department of Informatics**
University of Oslo

INF3490/INF4490 Biologically inspired computing

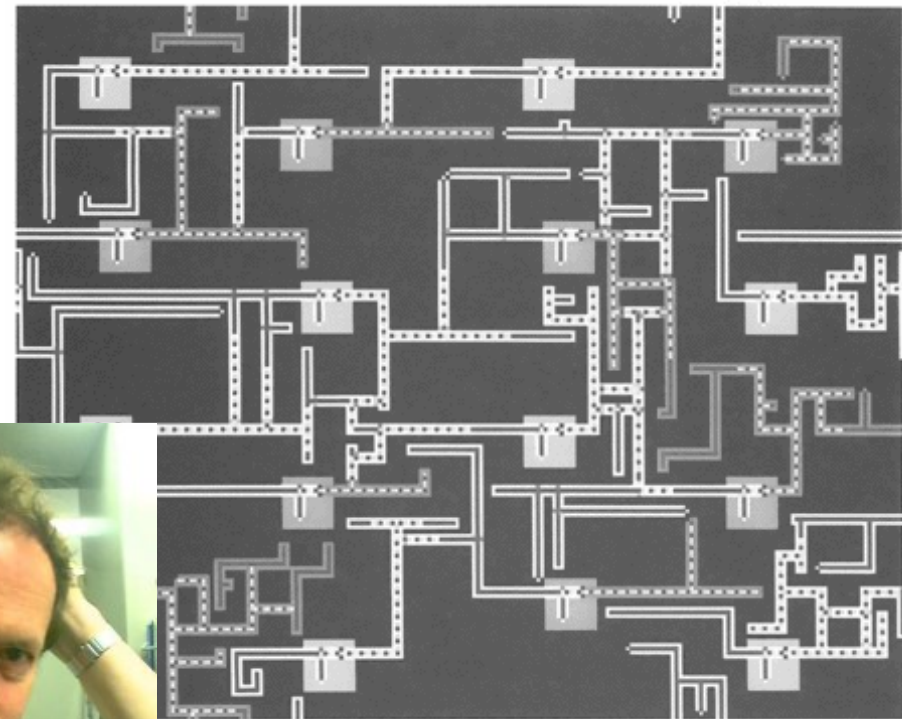
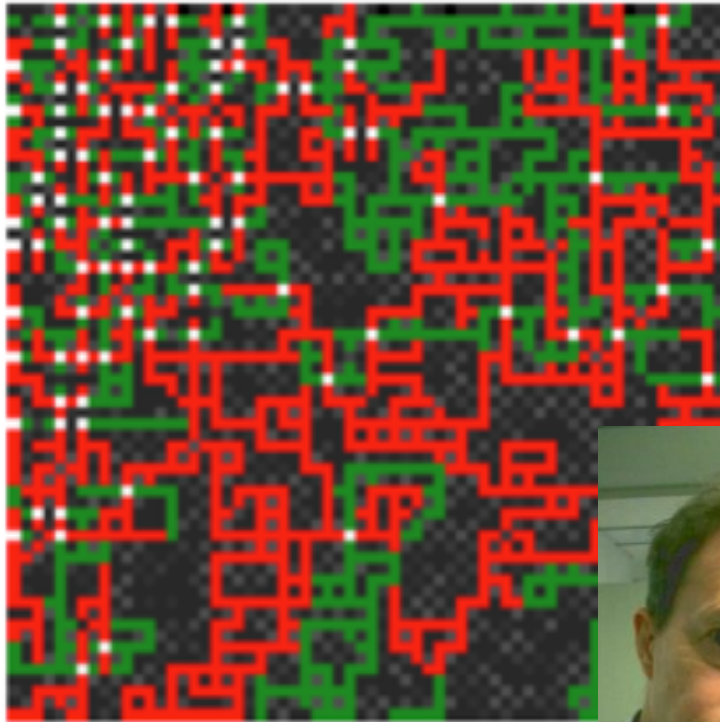
Future Perspectives on Artificial Intelligence
- What to expect and should we worry?



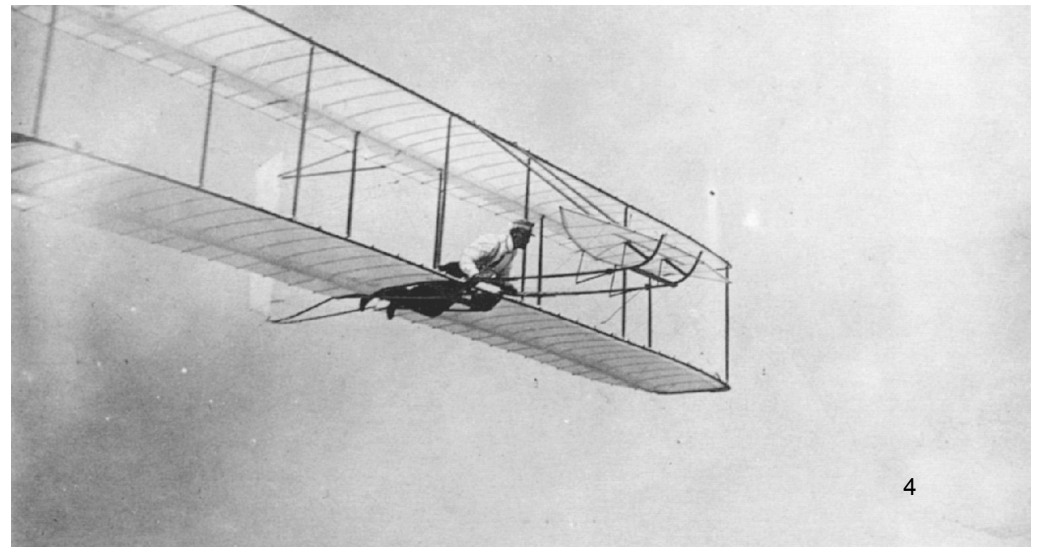
When and Where will a Breakthrough Come?

- Technology breakthroughs often happens randomly and not linked to major initiatives and projects.
- AI breakthroughs depend on the invention of ***scalable*** learning methods.
 - Need to understand more about how scalability and complexity arises in nature.
- Contributors to progress in AI:
 - Researchers mimicking biological or medical phenomena.
 - Researchers solving engineering problems

CAM (Cellular Automata Machine) Brain



What Methods are Best?





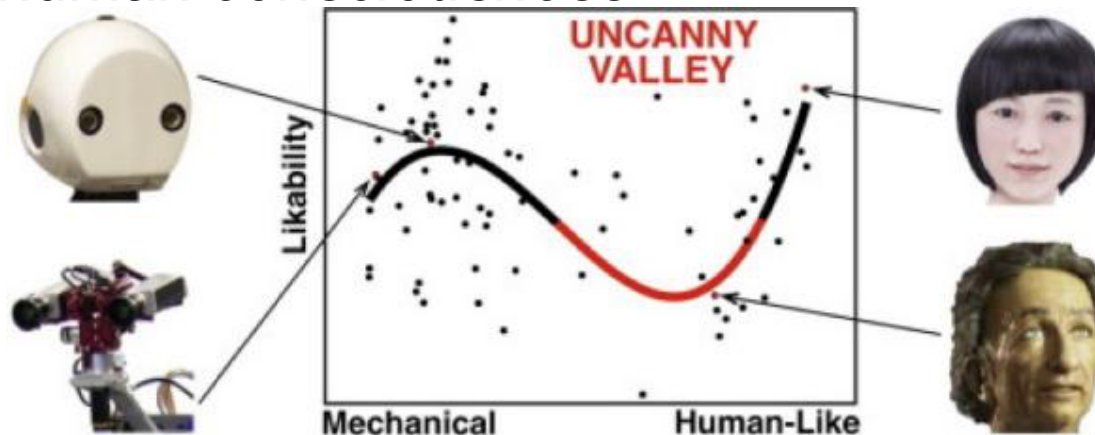
Autonomous Systems

- Systems that can operate with **no or limited human control**
- Degrees of human control:
 - Full autonomy (human decides only the goal)
 - Remotely monitored (autonomous planning but human take over if the system is misbehaving)
 - Remotely operated (autonomous steering using human planning)
 - Remotely controlled (human take all decisions)



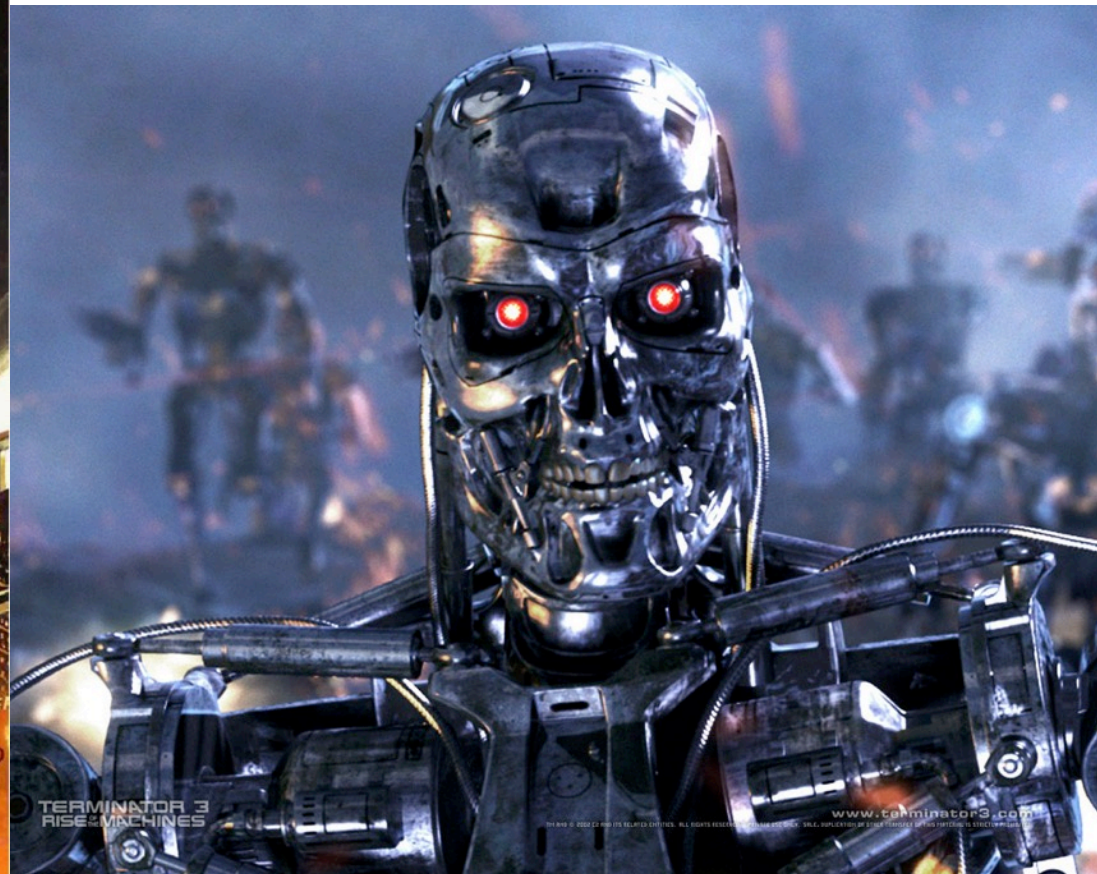
How Similar to Humans?

- Similarity depends on progress in a number of fields such as AI methods, computing power, vision systems, speech recognition, speech synthesis, human-computer interaction, mechanics and actuators.
- Design and usability will be essential for future robots but they should not necessarily be as similar as possible to humans (ref. *uncanny valley*).
- We will see robots having human-like *behavior*, but not human *consciousness*.



Would AI be Good or Bad?

Is Terminator Coming Close?





“Humans, limited by slow biological evolution, couldn’t compete and would be superseded by A.I.”



AI is our “biggest existential threat”



I am in the camp that is concerned about super intelligence.

Ethical Risks of Developing AI Systems

- People may become unemployed because of automation.?
- We get too much free time.?
- Artificial intelligence can be used for destructive and unwanted tasks.? [\("1984"\)](#)
- Successful KI can lead to the extinction of mankind?

Future Scenario with Autonomous Interacting AI Systems



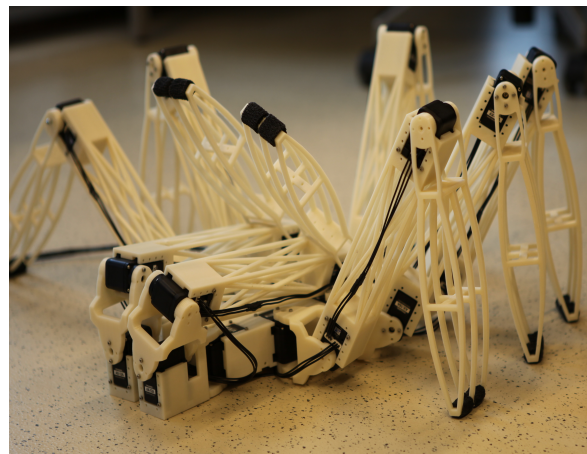
The image shows a screenshot of a complex data dashboard, likely a financial or operational management system. It features multiple tables with columns for various metrics, some highlighted in red and green. There are also small charts and summary statistics at the top of the interface.



Ethics for Programmers

- In the book “Moral Machines” it is argued that ethic competence should be included during program development by e.g.
 - machine learning methods based on examples of ethical and unethical behavior
- Software that will replace human evaluation and social function should adhere to criteria such as
 - accountability (no. ansvarlighet)
 - inspectability
 - manipulation robustness
 - predictability (no. forutsigbarhet)

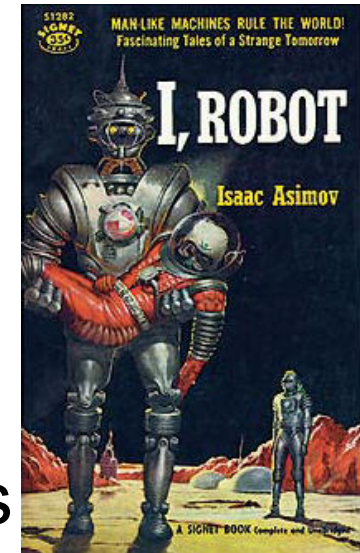
What would you require of a robot to be used in your home?





Ethical Guidelines for Robots and their Developers (Asimov 1942)

- A robot may not harm a human being, or through inaction, allow a human to be injured.
- A robot must obey orders given by human beings except where such orders would conflict with the first law.
- A robot must protect its own existence as long as such protection does not conflict with the first or second law.



Ethical Recommendations for Commercial Robots (Euronet Roboethics Atelier)

- **Safety:** There must be mechanisms (or opportunities for an operator) to control and limit a robot's autonomy.
- **Security:** There must be a password or other keys to avoid inappropriate and illegal use of a robot.
- **Traceability:** Similarly as aircraft, robots should have a "black box" to record and document their own behavior.
- **Identifiability:** Robots should have serial numbers and registration number similar cars.
- **Privacy policy:** Software and hardware should be used to encrypt and password protect sensitive data that the robot needs to save.