

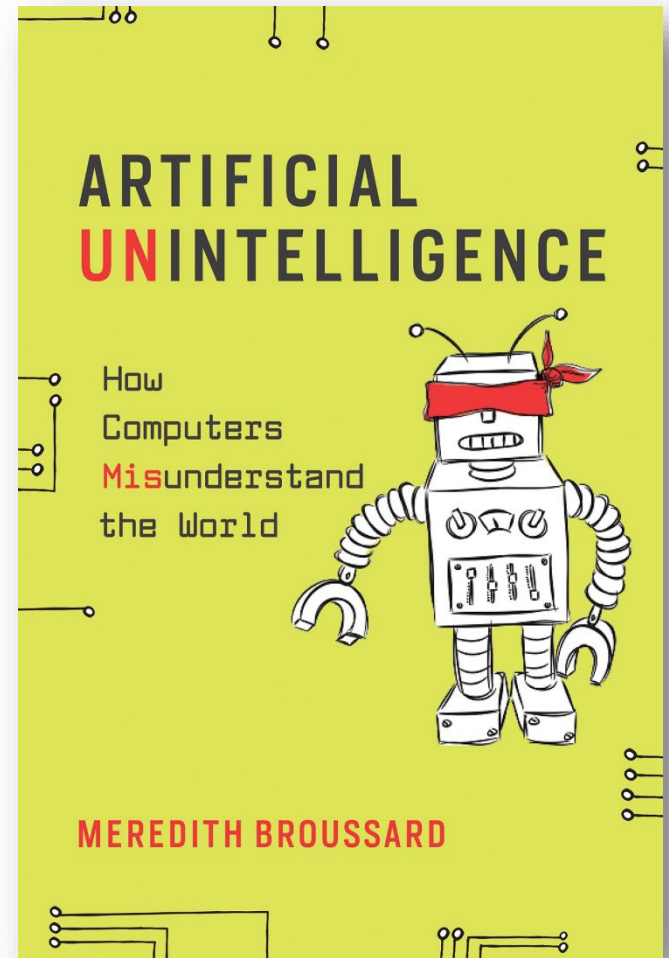
Common Curriculum Reading

Artificial Unintelligence
Fall 2024

Introduction to Artificial Intelligence

Michael Hahsler

Associate Professor
Computer Science
Bobby B. Lyle School of Engineering
SMU



Agenda

What is
AI?

AI
Today

?

AI Ethics & Safety

What is AI?



ASIMO (Advanced Step in Innovative Mobility) is a humanoid robot created by Honda in 2000

What is Intelligence?

Definition from Merriam-Webster

intelligence **noun**

in·tel·li·gence (in-ˈte-lə-jən(t)s)

[Synonyms of intelligence >](#)

1 a (1) : the ability to learn or understand or to deal with new or trying situations :

REASON

also : the skilled use of reason

(2) : the ability to apply knowledge to manipulate one's environment or to think abstractly as measured by objective criteria (such as tests)

b : mental acuteness : **SHREWDNESS**

c **Christian Science** : the basic eternal quality of divine Mind

2 a : **INFORMATION, NEWS**

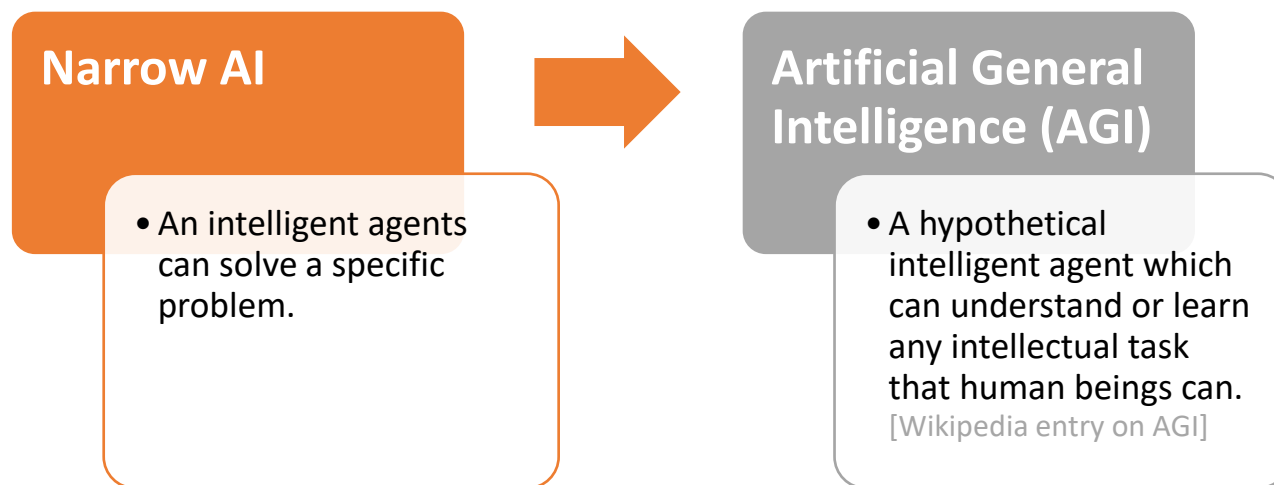
b : information concerning an enemy or possible enemy or an area

also : an agency engaged in obtaining such information

The Goal of AI

“Have machines solve problems that are challenging for humans.”

We call such a machine an **intelligent agent**.



How can we achieve this? Create an agent that can

Think like a human?

Act like a human?

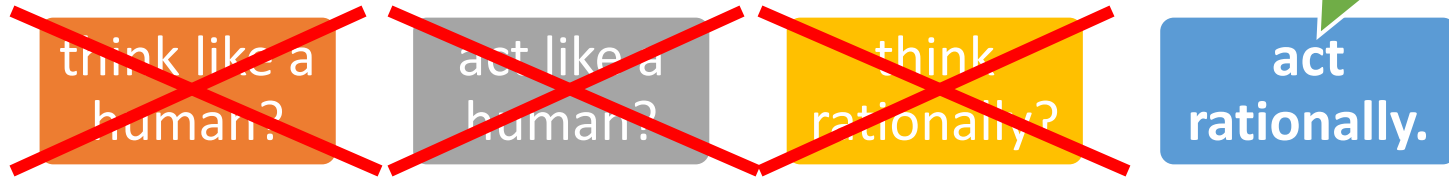
Think rationally?

Act rationally?

What type of AI do we see today?

Acting rational means to try to achieve the “best” outcome = solving an optimization problem.

Narrow AI agents that can



That is, machines that act in a way to solve a specific hard problem that traditionally would have been thought to require human intelligence.

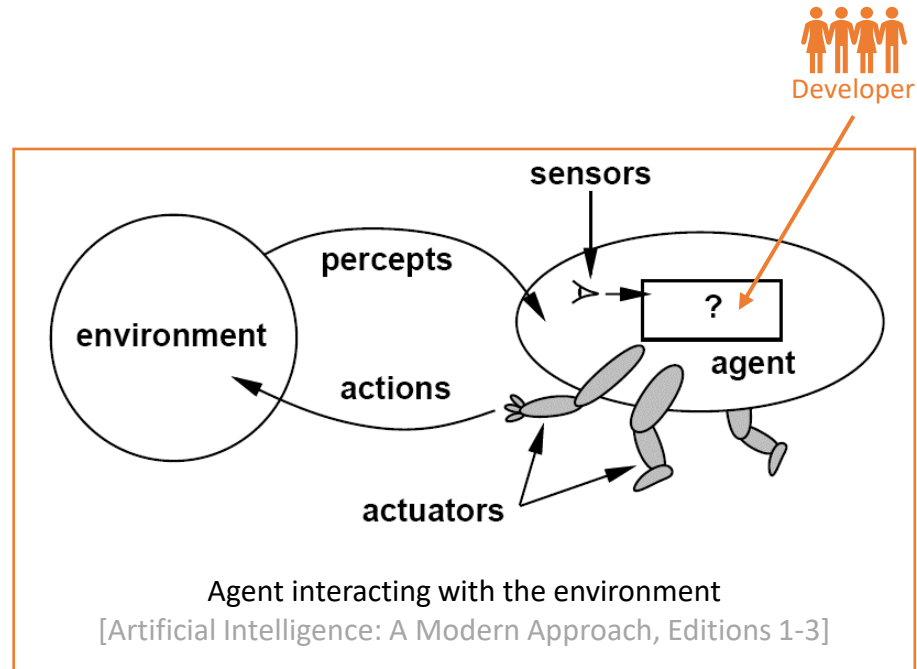
What are the Components of an Intelligent Agent?

Intelligent agents need to

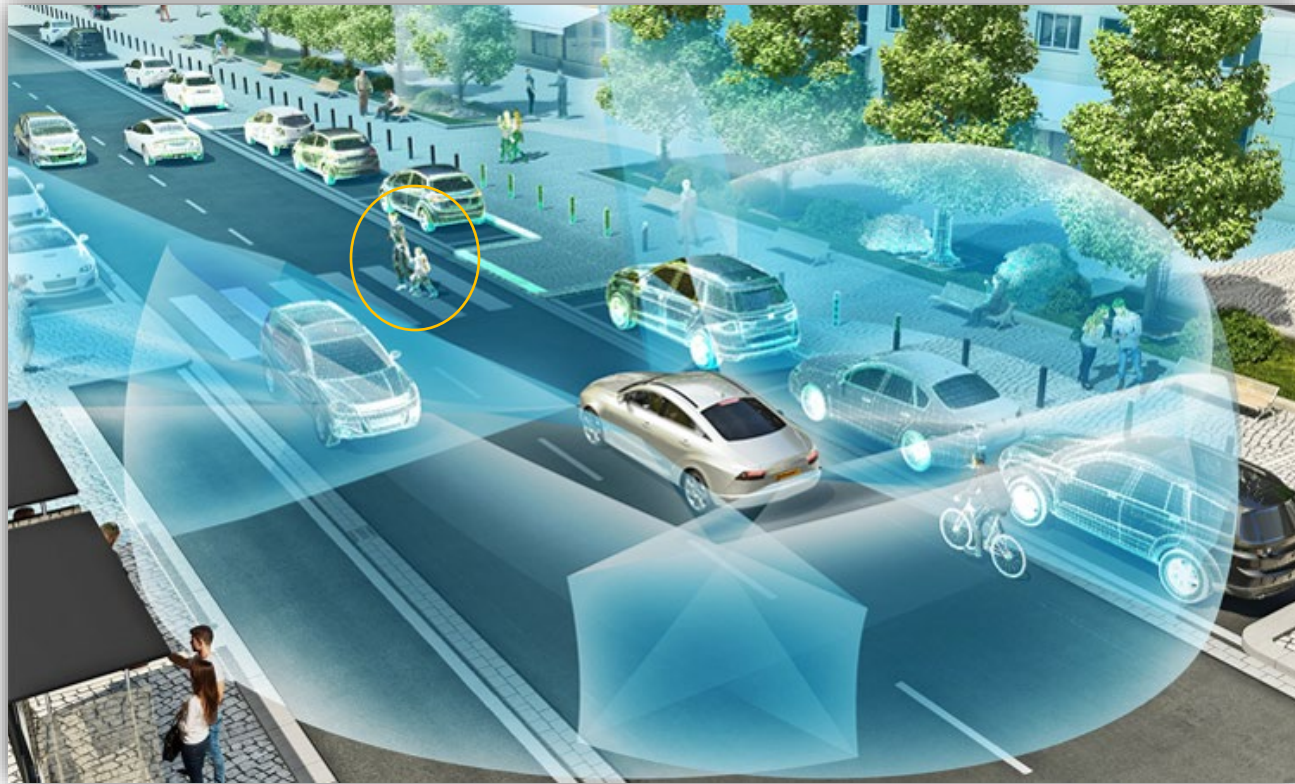
- **Communicate** with the environment.
- **Represent knowledge, reason** and **plan** to achieve a desired outcome.

Optional

- **Learn** to improve performance.



Example: Self-Driving Car



Percept: People crossing the street

Action: Stop the car

Example: LLM



Default (GPT-3.5)



the sun is shining. It is



a beautiful day. The warm rays of sunlight illuminate the surroundings, creating a cheerful and inviting atmosphere. People are out and about, enjoying outdoor activities, and the clear blue sky adds to the overall sense of positivity. Whether you're taking a leisurely stroll, having a picnic, or simply basking in the sun's glow, it's a perfect time to appreciate the natural beauty and soak in the uplifting energy that the sunshine brings.

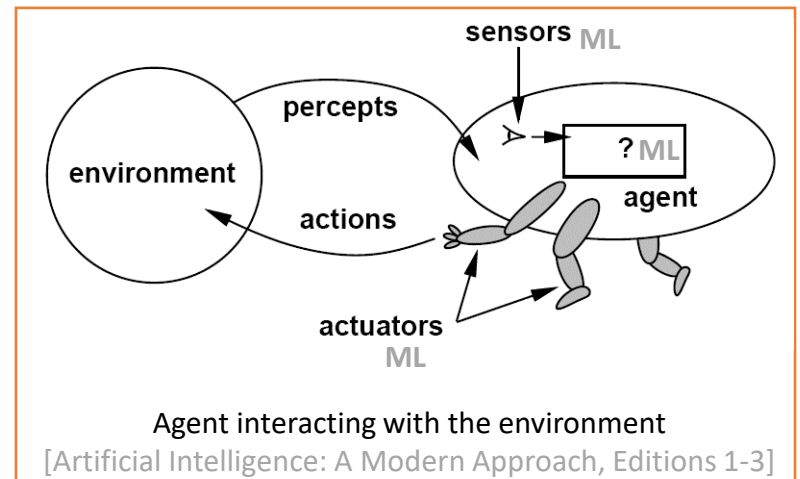
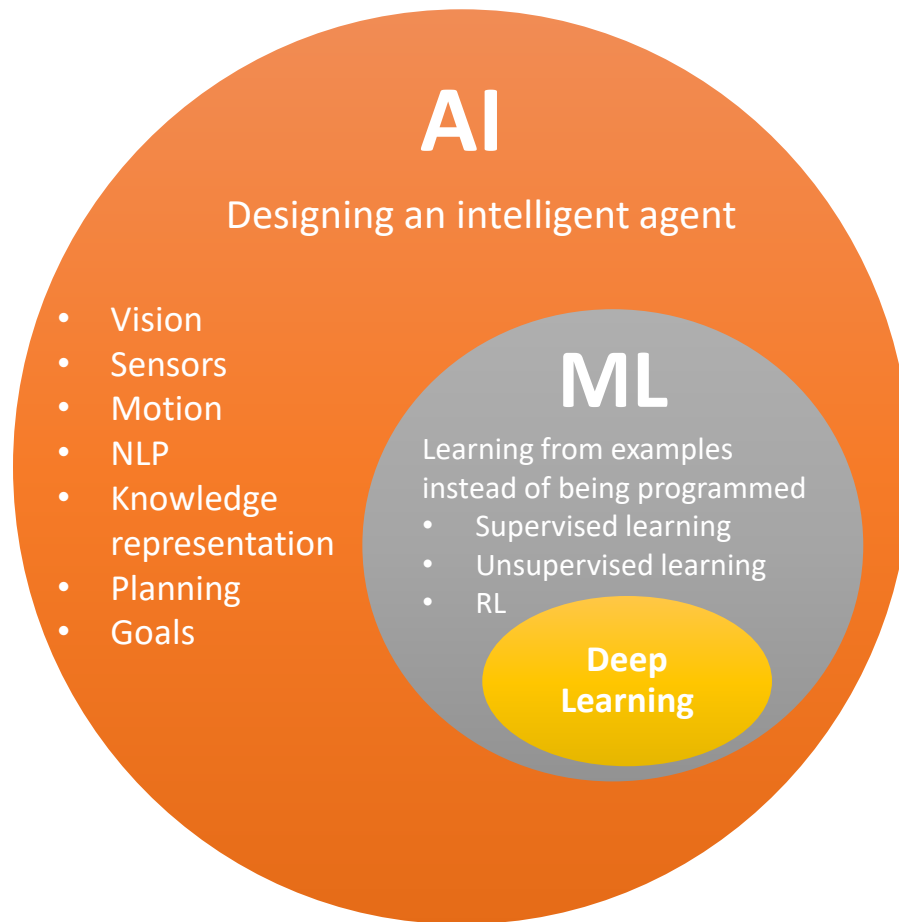


Percept: prompt

Action: next most likely word

More words are created word-by-word.

Machine Learning vs. Artificial Intelligence



ROOMBA



SELF-DRIVING CAR



AI ROBOT



EMAIL SPAM FILTER



SMART SPEAKERS

AI Today

Vision and Image Processing

- **OCR:** read license plates, handwriting recognition (e.g., mail sorting).
- **Face detection:** now standard for smart phone cameras.
- **Vehicle safety systems**
- **Visual search**
- **Image generation**

All these technologies operate now with superhuman performance.

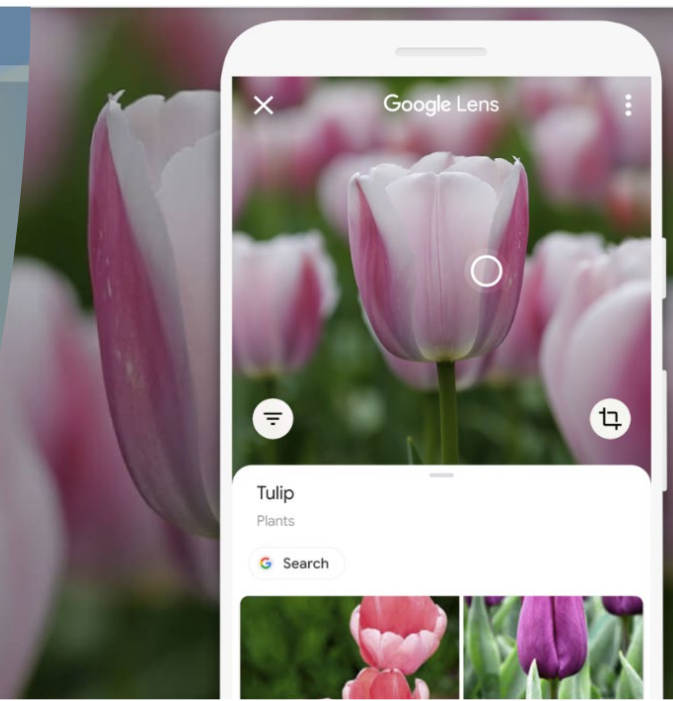


» Vision Applications

Road, Vehicle, Pedestrian Protection and more



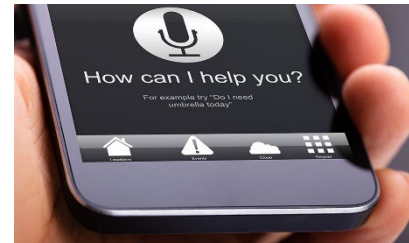
» AWS Advanced Warning



Natural Language Processing



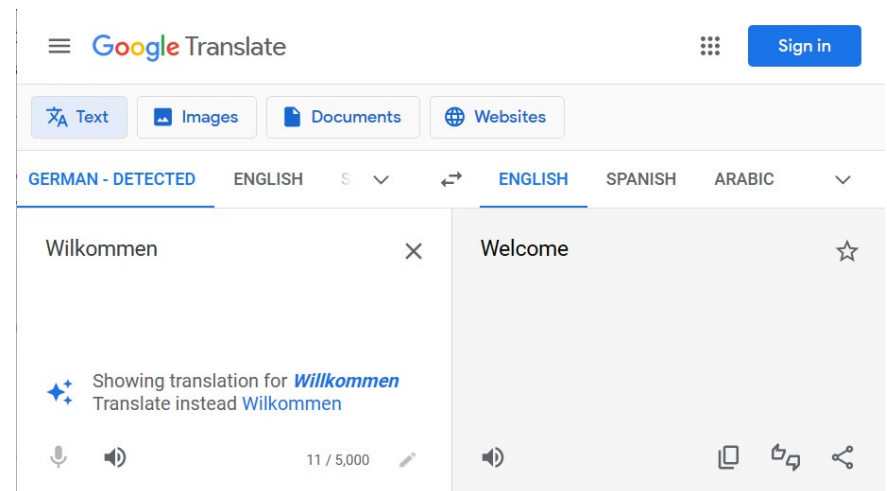
- Text-to-speech
- Speech-to-text to detect voice commands
- Machine translation
- Text generation (Q/A systems) using Large Language Models



These technologies operate now with close to or even superhuman performance.

Humans use language to reason. Does that mean AI that can create good language can reason?

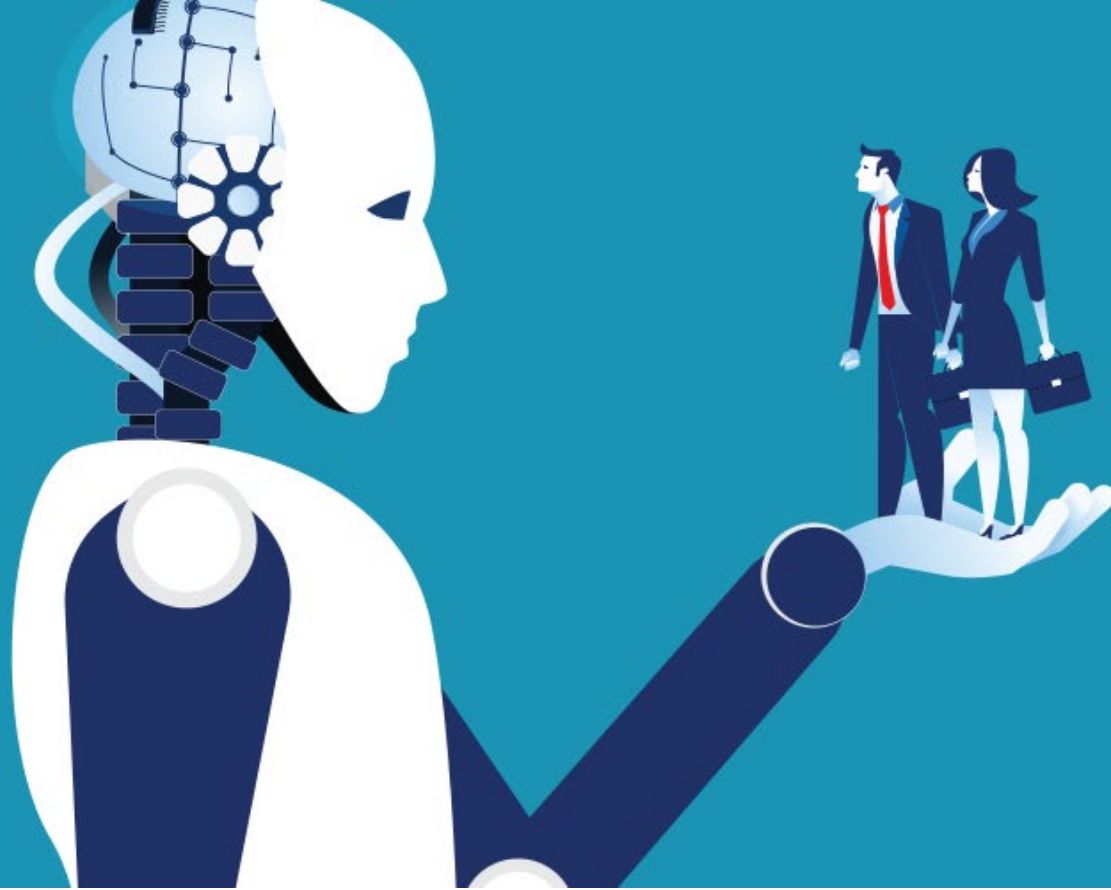
Language understanding is still elusive!



Robotics

- Mars rovers
- Autonomous vehicles
 - [DARPA Grand Challenge](#)
 - Google self-driving cars
- [Autonomous helicopters](#) and drones
- Robot soccer
 - [RoboCup](#)
- Personal robotics
 - Humanoid robots
 - [Robotic pets](#)
 - Personal assistants?





AI Ethics & Safety

A new Frontier for Fairness and Freedom

Commonly-Cited Safety and Ethics Principles

Use of AI by companies and organizations

- Ensure safety
- Limit harmful uses of AI
- Establish accountability: Liability?
- Avoid concentration of power: Winner-takes-All

Protect individuals

- Uphold human rights and values.
- Ensure fairness: Equal opportunity/equal impact. Reflect diversity/inclusion.
- Provide transparency: Right to explanations.
- Respect privacy: Surveillance?
- Contemplate implications for employment: Income and purpose.

Governance

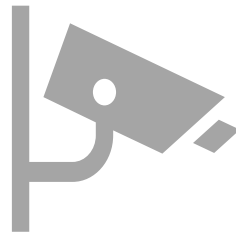
- Acknowledge legal/policy implications

AI Safety

“Prevent accidents, misuse, or other harmful consequences of AI.”



AI Testing



Monitoring AI



Adversarial
robustness

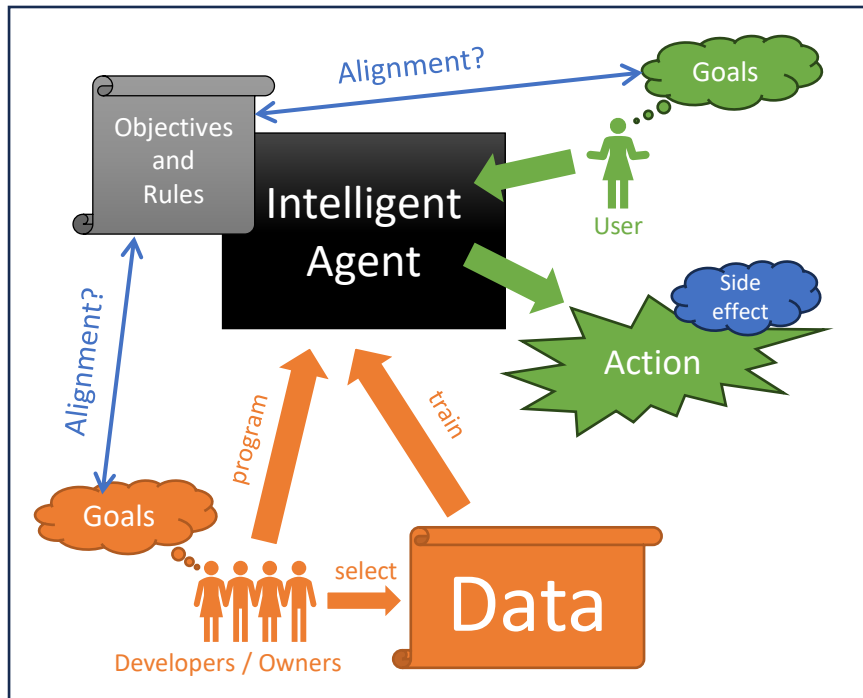
How should this be ensured?

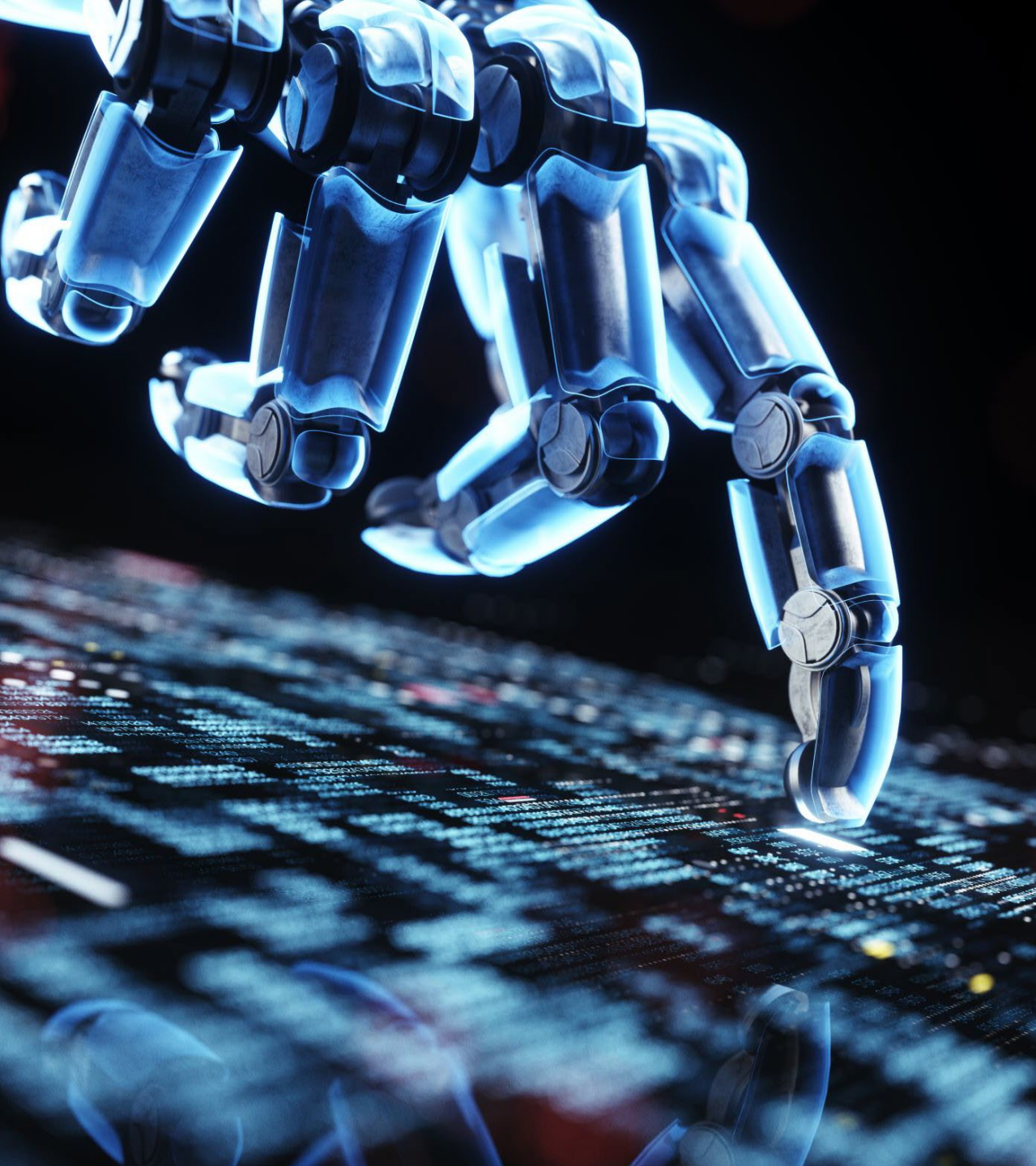
- Corporate self-regulation
- Government action

AI Safety

Intelligent Agents are “optimizers!”

- **Goal/reward alignment:** How do we specify a robust objective function?
- **Reward hacking** creates unintended side effects. AI needs to follow social norms.
- **Instrumental convergence:** All intelligent agents will pursue common subgoals like the need for more power.





Outlook

AI is a technology that is on the verge of significant leaps...

- New technologies always had a **profound impacted** on the way we live and work (e.g., electricity, the internet, mobile communication).
- We can expect unprecedented gains in productivity from better **narrow AI**.
- New technologies always also present **dangers** and need to be regulated.

Questions & Answers

Contact info

Michael Hahsler

Associate Professor of Computer Science
Bobby B. Lyle School of Engineering
SMU

Web: <http://michael.hahsler.net>

Email: mhahsler@lyle.smu.edu

Online material about AI: <https://github.com/mhahsler/CS7320-AI>

