What is Artificial Intelligence?
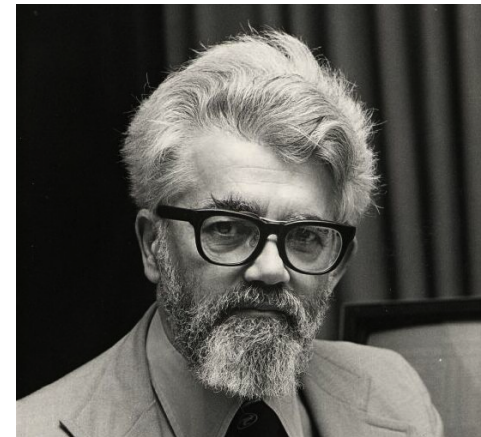
Dr. Johan Hagelbäck

lnu.se/personal/johan.hagelback/

What is Artificial Intelligence?

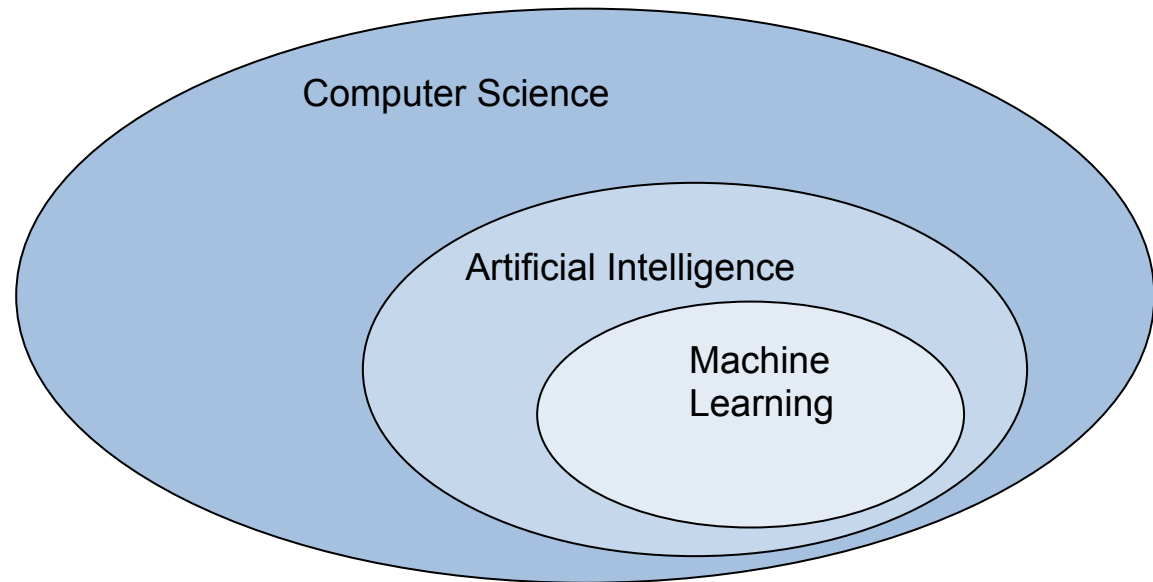"It is the science and engineering of making intelligent machines, especially intelligent computer programs"

John McCarthy
1927-2011
Professor at Stanford University

Linnæus University

What is Machine Learning?

"Machine learning is the science and engineering of building systems that learn from data"

Computer Science

Artificial Intelligence
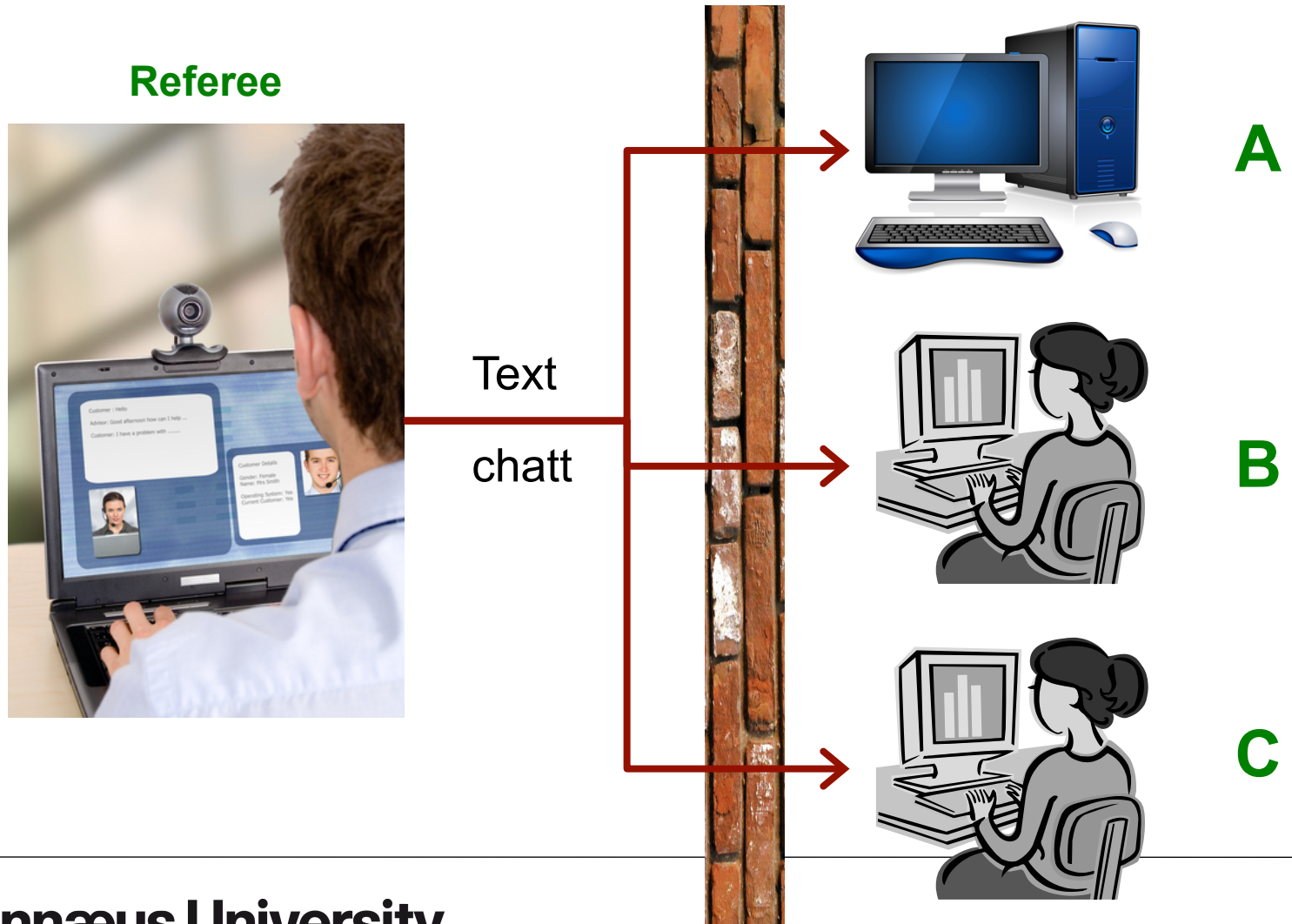
Machine Learning

History of AI

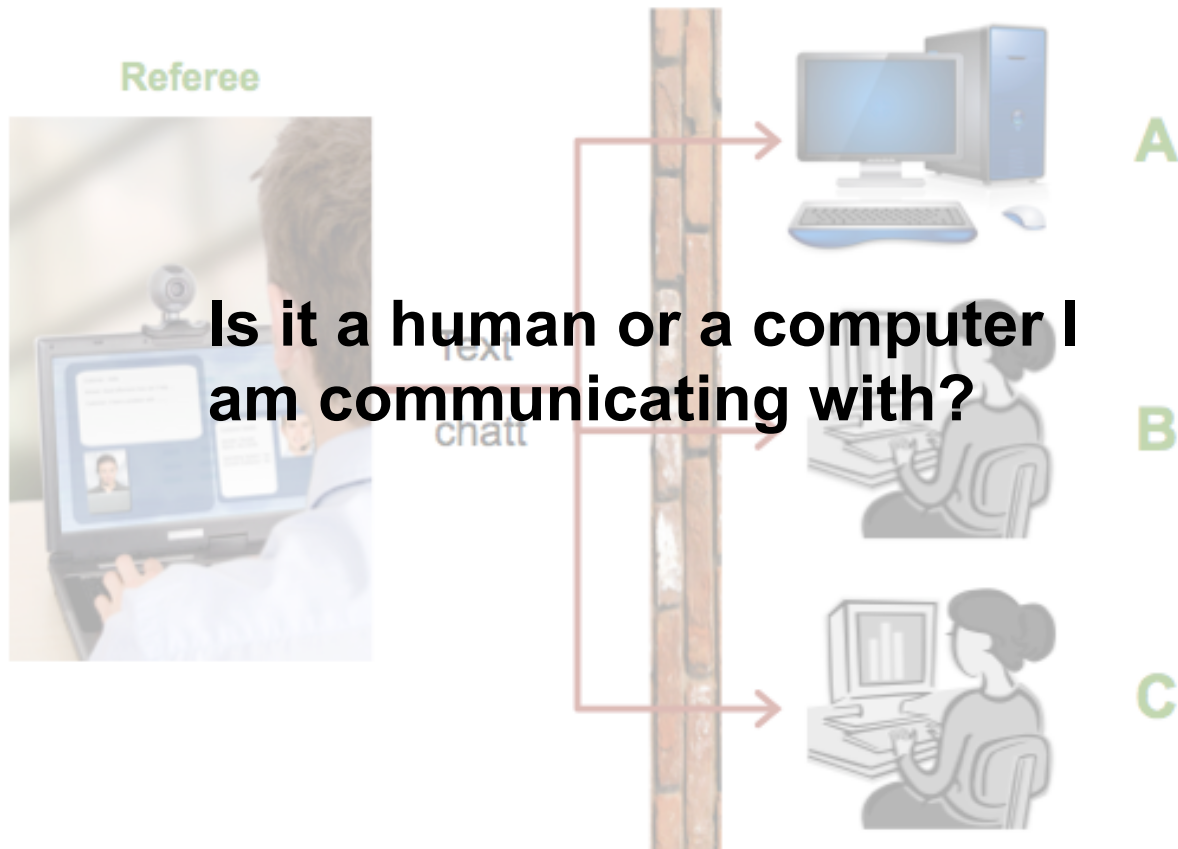# Can Machines Think?

- In 1950, Alan Turing published the paper *Computing Machinery and Intelligence* in the journal *Mind*

- In the paper, Turing reasons about the question *Can machines think?*

- This paper was the first to mention the Turing Test

- The test has become an important benchmark for AI discoveries

Linnæus University

Turing Test

Referee

Text

chatt

A

B

C

Turing Test

Referee

Is it a human or a computer I am communicating with?

A

B

C

Has it been beaten yet?

- We have come far, but arguably no, it hasn't been beaten yet

- There are variants of the test that have been beaten

- Many believe that we have come so far that the Turing Test can be passed if we use average persons as judges

- It can however not yet beat judges trained to spot non-humans

- In 2011, Cleverbot scored 59% human-ness for the computer compared to 63% human-ness for humans

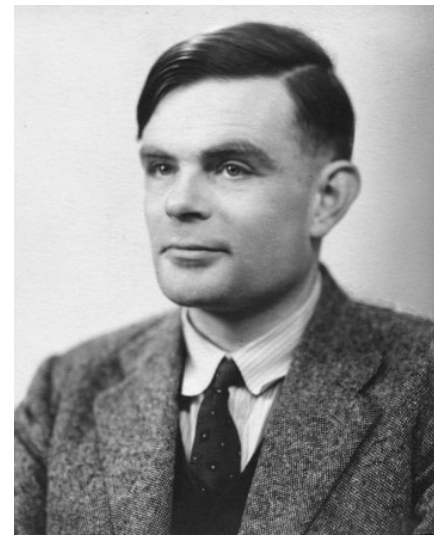- Some of the early attempts are ELIZA (1966) and PARRY (1972)

ELIZA

ELIZA: Please tell me what's been bothering you.
YOU: I am scared of spiders.
ELIZA: Is it because you are scared of spiders that you came to me?
YOU: Yes. I need help.
ELIZA: I see.
YOU: Can you help me?
ELIZA: Really, I help you?
YOU: You are the psychiatrist.
ELIZA: What makes you think I am the psychiatrist?
YOU: It says so on your office door.
ELIZA: Why are you concerned over my office door?
YOU: Because it says you are a psychiatrist and now you are telling me that you are not.
ELIZA: Does it please you to believe I am not?
YOU: No. Now I wonder if you can help me or not?
ELIZA: You are being a bit negative.

ELIZA

- ELIZA tries to identify keywords in what the user typed
- If a keyword is found, the user sentence is transformed using some rules
- If not, ELIZA responds with a generic reply or repeats an earlier user sentence
- ELIZA does *not understand the meaning of a sentence*
- Still, many researchers at the time thought chat bots could be used in therapeutic conversations

Alan Turing

- Turing was highly influential in formalizing the theories needed to build a general purpose computer

- He also formalized the concepts of algorithms and computation with binary digits (0 and 1)

- During WW 2, he worked on cracking the Enigma machine used to generate German ciphers

- Today, Turing is considered to be the father of theoretical computer science and AI

AI as research field

- At the Dartmouth Conference in 1956, AI was established as an academic field

- The conference was organized by Marvin Minsky, John McCarthy, Claude Shannon and Nathan Rochester

- The proposal included the following statement:
  - "every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it"

- The conference is considered the birth of AI

The golden years 1956-1974

- AI researchers at the time were very optimistic about the field

- In 1965, Herbert A. Simon said:
  - "machines will be capable, within twenty years, of doing any work a man can do."

- In 1967, Marvin Minsky said:
  - "Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved."

- In 1967, Marvin Minsky said in an interview in Life magazine:
  - "In from three to eight years we will have a machine with the general intelligence of an average human being."

The first AI winter 1974-1980

- In 1970's, the AI field was criticized for failing to solve the problems they faced

- The optimism had raised the expectations impossibly high

- Some of the problems the field faced at the time were:
  - Limited computer power
  - Combinatorial explosion – many problems could only be solved in exponential time, and moving from "toy" problems to real-world problems was not possible
  - Moravec's paradox – for humans simple tasks such as recognizing a face turned out to be extremely difficult for computers

AI criticism

- Early AI systems was to a large extent based on rules
  - Define a cat by its size, shape and features
- Several philosophers argued against the claims made by AI researchers
- Hubert Dreyfus argued that human reasoning involved very little rules, but a great deal of embodied, instinctive, unconscious know-how
- John Searle put forward the Chinese room thought experiment:

Chinese room thought experiment

- Suppose that we create an AI that behaves as if it understands Chinese

- It takes Chinese characters as input, and produces a response with other Chinese characters

- Suppose that this AI is so good that it passes the Turing Test

- The question is then:
  - does the machine understand Chinese, or does it merely simulate the ability to understand Chinese?

- He argued that if the symbols have no meaning for the machine, the machine cannot be described as *thinking*

- Searle called understanding strong AI, and simulating the ability to understand weak AI

AI boom 1980-1987

- In 1980s, companies successfully began to use expert systems

- An expert system answers questions or solves problems about a specific knowledge domain, using logical rules

- The rules are derived from the knowledge of domain experts

- Expert systems were part of a new direction in AI
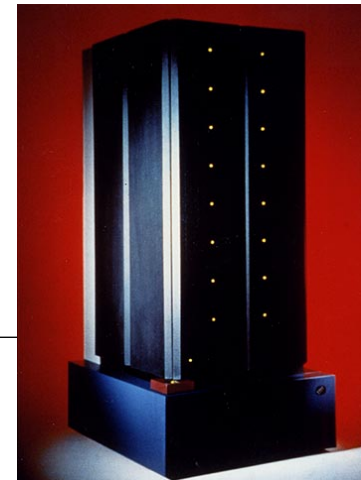
The second AI winter 1987-1993

- Once again, expectations had been higher than was actually possible

- For example, a goal listed in Japan's Fifth Generation Project for 1981 was to "carry on a casual conversation"

- This has, arguably, not been met today

- The interest in expert systems were fading

- They proved to be very difficult and expensive to maintain and update

- They could also make huge mistakes when given new and unknown input

- They were still useful, but only for a few very specialized tasks

Linnæus University

Modern AI 1993-

- During this period, the AI field was able to achieve some important goals

- AI systems were also successfully used in technology industry, mainly due to a focus on isolated problems and increasing computer power

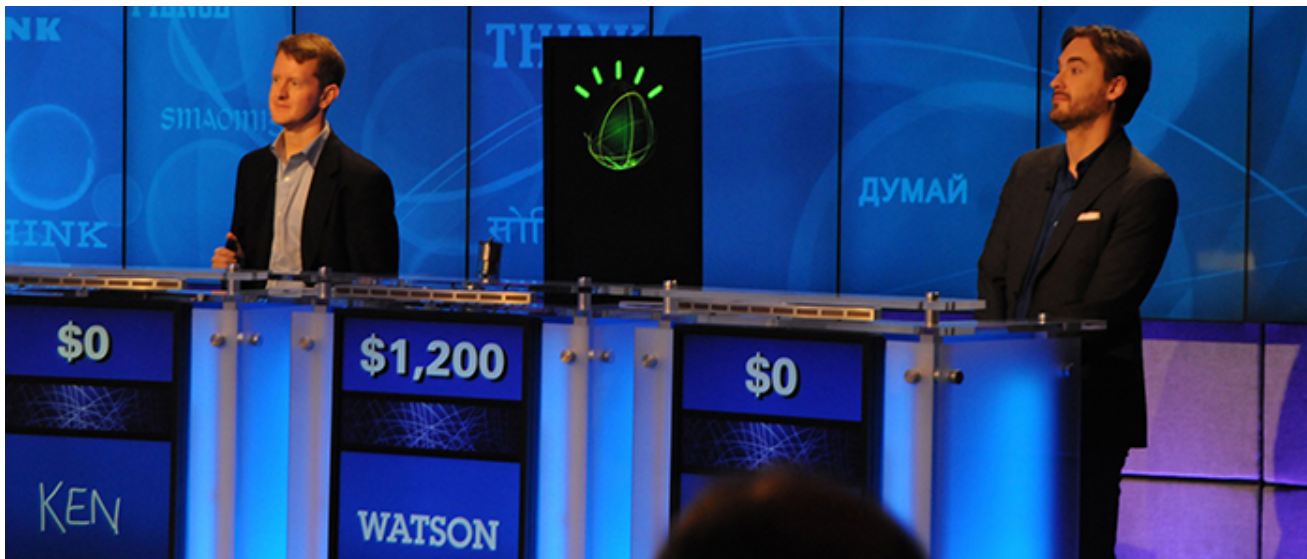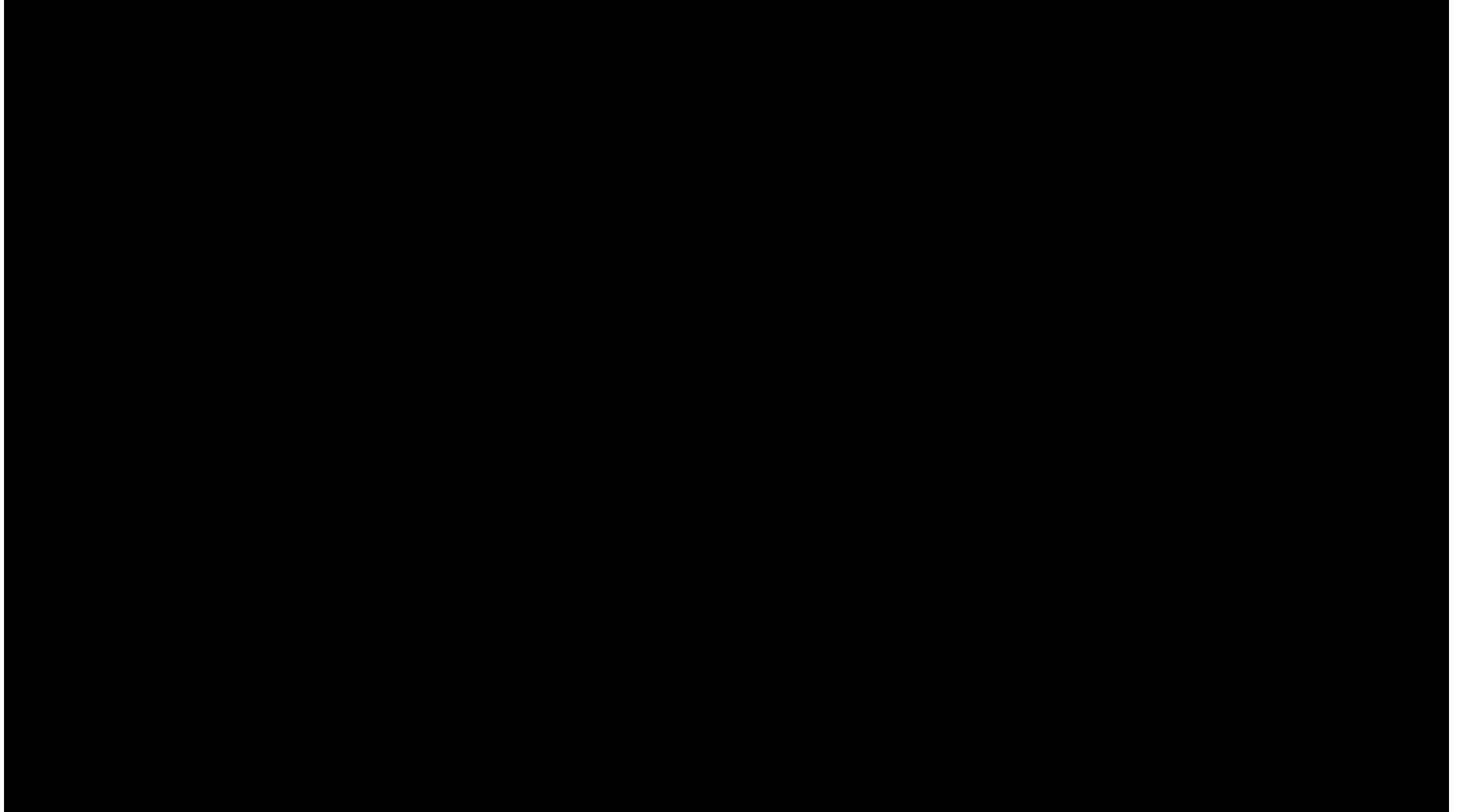- In 2006, a new AI boom started with the invention of Deep Learning

Milestones

- On May 11 1997, Deep Blue became the first computer system to beat a world champion, Garry Kasparov, in Chess

- Deep Blue searched for all combinations of future moves from the current board state

- It could easily search 14 moves in the future, sometimes up to 40 moves

- Deep Blue was custom built for Chess playing, and could process around 200 million moves per second

- The event was also broadcasted live on the Internet and had over 74 million viewers

Linnæus University

Milestones

- In 2011, IBM developed the Jeopardy playing system Watson:

- Watson does not store questions and answers
- Instead, it tries to understand what is asked for in the question
- … and then searches through enormous amounts of information for the answer:
  – Complete Wikipedia
  – All issues of New York Times
  – …
- The intelligence in Watson is understanding the questions, not knowing everything

- Watson has been successfully used in other domains such as:
  - Healthcare – as a decision support system for medical professionals when treating e.g. cancer
  - Teaching Assistant at Georgia Tech - it answered questions it was 97% certain the answer was correct, leaving the rest of the questions to human teaching assistants
  - Weather forecasting – Deep Thunder project started at IBM 2016
  - Cook book – Chef Watson suggests unique dishes based on what you have in the fridge!
  - … and more

Linnæus University

LOOK FOR INGREDIENTS

CHOOSE A DISH

PICK A STYLE

START AGAIN

FAVORITES

✓ **SALMON** ✕

Salmon /ˈsæmən/ is the common name for several species of fish in the family Salmonidae. Other fish in the same family include trout, char, grayling and whitefish. Various species of salmon display

ⓒ W

✓ **CREMINI MUSHROOM** ✕

Agaricus bisporus is an edible basidiomycete mushroom native to grasslands in Europe and North America. It has two color states while immature—white and brown—both of which have various names. When

ⓒ W

✓ **PARMESAN CHEESE** ✕

Parmigiano-Reggiano (Italian pronunciation: [ˌparmiˈdʒaːno red ˈdʒaːno]), called Parmesan in English after the French name for it, is a hard, granular cheese that is cooked but not pressed. It is named after the

ⓒ W

✓ **PLUM** ✕

A plum is a drupe fruit of the subgenus Prunus of the genus Prunus. The subgenus is distinguished from other subgenera (peaches, cherries, bird cherries, etc.) in the shoots having a terminal

ⓒ W

← BACK

● SYNERGY

→ MORE

HERE ARE SOME IDEAS

Salmon Stuffed Vegetable

water, barley malt syrup, granulated sugar, vinegar, **salmon,** cremini mushroom, extra-virgin olive oil, plum, parmesan cheese, cilantro, harissa, coriander seed, pink peppercorn, anise

Based on: Chipotle Chiles Stuffed With Shrimp And Plantains from Bon Appétit

MORE...

Salmon Pasta

tortellini, **salmon,** cremini mushroom, plum, parmesan cheese, herbes de provence, kosher salt, red pepper flakes

Linnæus University

GREEN APPLE

Green Apple Worms was the first album to be released by Jumping Monks. It was released in October 2003.

GINGER

Ginger or ginger root is the rhizome of the plant Zingiber officinale, consumed as a delicacy, medicine, or spice. It lends its name to its genus and family (Zingiberaceae). Other notable members of this plant

ROOIBOS TEA

Sweet tea is a style of iced tea commonly consumed in the United States, especially the Southern United States. Sweet tea is made by adding sugar to bags of black tea brewing in hot water while the

CARDAMOM

Cardamom refers to several plants of the similar genera Elettaria and Amomum in the ginger family Zingiberaceae. Both genera are native to India, Pakistan, Nepal, and Bhutan; they are recognised by their

← BACK

SYNERGY

MORE →

HERE ARE SOME IDEAS

Green Apple Cake

confectioners' sugar, greek honey, egg, bicarbonate of soda, raisin, butter, ginger, flour, **green apple,** kosher salt, rooibos tea, whole grain mustard, yellow mustard seeds, caraway seed, cardamom, white pepper

Based on: Ginger Spice Cake With Dried Cherries from Bon Appétit

MORE...

Green Apple Buche De Noel

water, baking powder, butter, parmesan cheese, heavy cream, clove, cardamom, ground cinnamon, granulated sugar, orange blossom honey, egg, dried cherry, angostura bitters, balsamic vinegar, cake flour, lemon juice, **green apple,** salt, ginger, rooibos tea

Linnæus University

Milestones

- In 2018, Google released its personal assistant Google Duplex:

"Hi, I'm calling to book a women's haircut for a client."

Google Cloud Vision

- State-of-the-art image recognition system
- Amazon and Microsoft have similar services
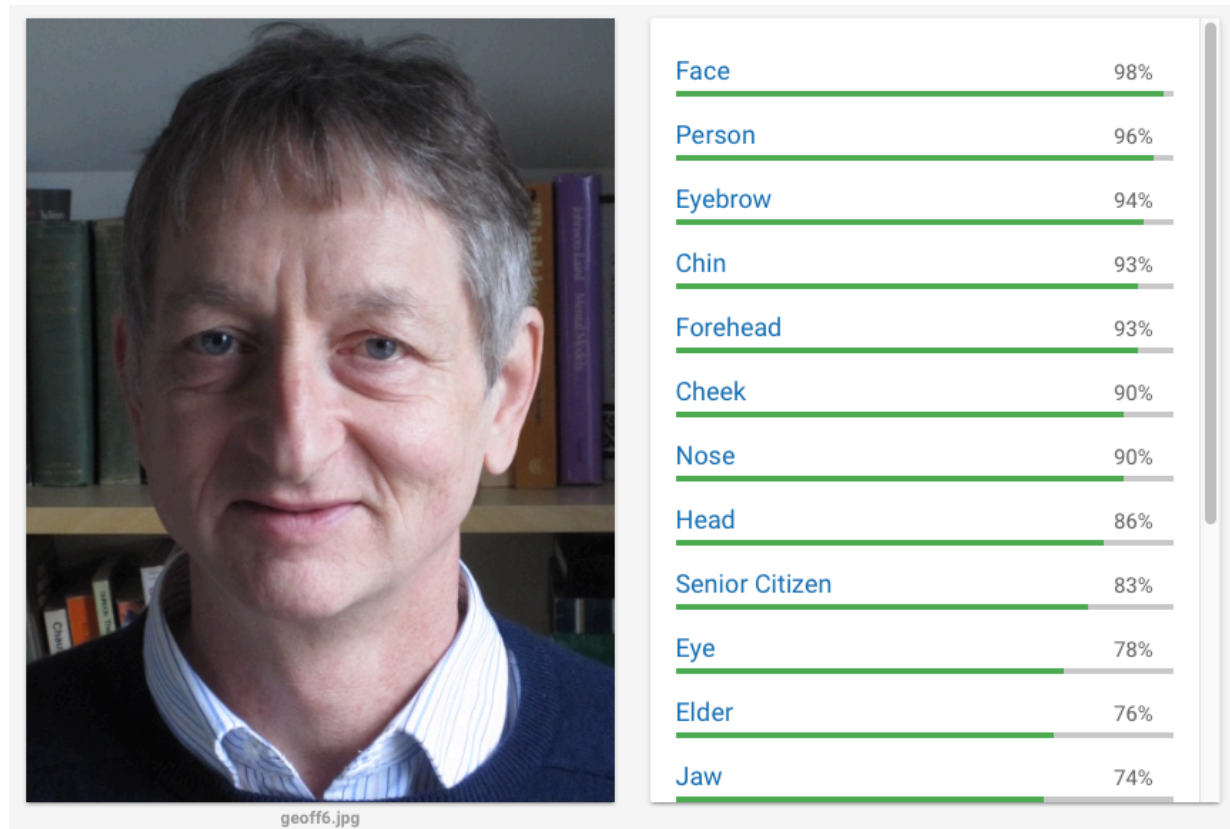- Tries to recognize object(s) in a picture

- Let's try it:

Linnæus University

Google Cloud **Vision**

Correctly recognizes that this is a beagle

Beagle	96%
Dog Breed	93%
Dog Like Mammal	92%
Dog	92%
Harrier	89%
Basset Artésien Normand	84%
Estonian Hound	83%
Beagle Harrier	83%
Finnish Hound	82%

Linnæus University

Google Cloud Vision

geoff6.jpg

Face	98%
Person	96%
Eyebrow	94%
Chin	93%
Forehead	93%
Cheek	90%
Nose	90%
Head	86%
Senior Citizen	83%
Eye	78%
Elder	76%
Jaw	74%

Recognizes that this is a person, but not who it is

Linnæus University

Google Cloud Vision

car_water.jpg

This picture turned out to be a bit tricky...

Boat	95%
Vehicle	94%
Water Transportation	92%
Mode Of Transport	88%
Watercraft	82%
Motorboat	74%
Boating	68%
Boats And Boating Equipment And Supplies	67%

Linnæus University

Amazon AWS Rekognition

Labels

LABELS AWS REKOGNITION

Name	Confidence
Human	99.0199.01
People	99.0199.01
Person	99.0199.01

Milestones

- In Chess, there is on average around 35 possible moves from each board state

- The branching factor of the search tree (without pruning) is 35

- Deep Blue was able to search at least 20 moves in the future, using smart pruning

- The game Go is huge compared to Chess, with a branching factor of 361

- Until recently, AI systems were only able to play the game on amateur level
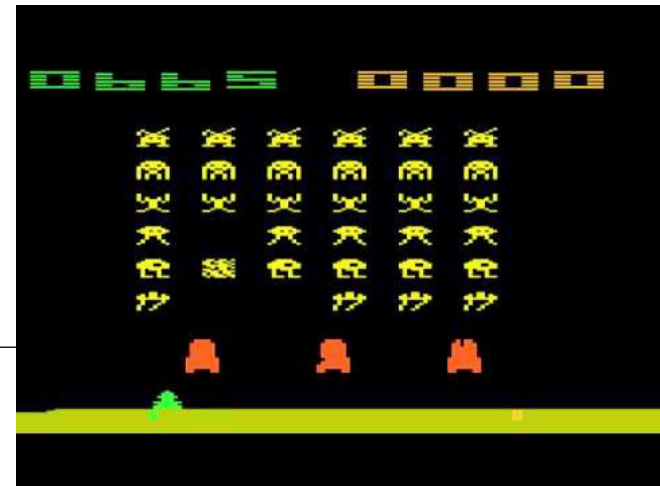
AlphaGo

- AlphaGo is a Go playing system developed by DeepMind, a Google-owned company

- In October 2015, AlphaGo was able to beat the 18 times world champion Lee Sedol by 4 games to 1

- AlphaGo uses a deep neural network to learn knowledge about the game, in a 40 days learning phase

- The system generates versions of itself to play against, to generate new data

- During gameplay it uses sophisticated search methods to find its moves from the learned knowledge

- AlphaGo is by many considered a major milestone in AI research

Playing Atari games

- Before AlphaGo, DeepMind created a deep reinforcement learning system that played seven Atari 2600 games

- The system had not seen the games before, and was only given screenshots as input

- With a trial and error approach, the system gradually learned how to play the games

- It outperformed other computer systems on 6 out of 7 games, and surpasses human experts on 3 of the games

Linnæus University

Cats on Youtube

- With a deep learning system running on 16000 processors, Google learned to detect cats on Youtube clips

- The clips were not labeled as containing cats or not, but the system was still able to group clips with cats together

- The accuracy was 74.8% on identifying cats, and 81.7% on human faces

Linnæus University

Huge amounts of data

- Ray Kurzweil said that:
  "one of the strengths of humans is that we can learn from only a few examples"

- For machines to learn, many examples are usually needed

- It has been said about deep learning that
  "life starts at a billion examples"

Where is HAL 9000?

- In the movie *2001: A Space Odyssey* from 1968, Arthur C. Clarke and Stanley Kubrick predicted that in 2001 a machine would exist with an intelligence matching or surpassing humans

- The machine, HAL 9000, was based on beliefs shared by many leading AI researchers at the time

- Obviously, we didn't get a HAL 9000 by 2001

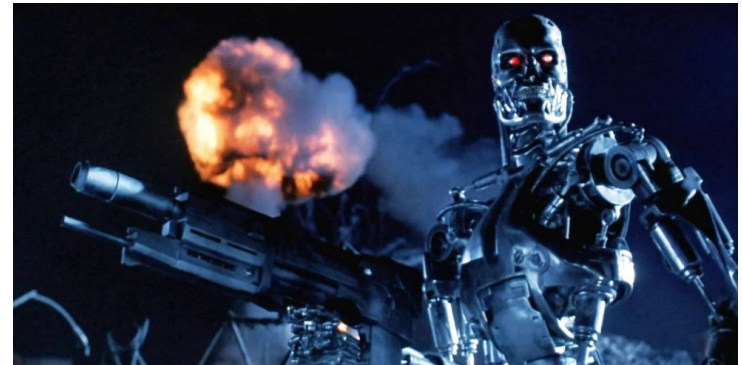- But why?

Linnæus University

Where is HAL 9000?

- Marvin Minsky said it was because the central problems, like commonsense reasoning, were being neglected while research was focused on commercial applications

- Ray Kurzweil referred to Moore's Law (computer power will double every second year) when he argued that we will have enough computational power to create human-level intelligence in 2029

- Jeff Hawkins, the founder of Palm Computing, argued that the research ignores the essential properties of the human cortex, preferring simpler models able to solve simpler problems

- … these are just a few of the explanations put forward by researchers

Linnæus University

AI in the future

Future of AI?

- In January 2015, the Future of Life Institute organized a private conference

- FLI is run by Max Tegmark, a Swedish cosmologist who got famous for writing a book about the universe being a simulation, and Jaan Tallinn, co-founder of Skype

- FLI has an advisory board with many famous names:
    - Actor Morgan Freeman
    - Nobel prize winner in physics, Alan Guth
    - AI-expert Stuart Russell
    - Elon Musk (Tesla/SpaceX)
    - Nick Bostrom, director of Future of Humanity Institute
    - Stephen Hawking, before his passing
    - … and some more

Linnæus University

Future of AI?

- The attendees were asked to first predict when machines would become better than humans at all human tasks

- The median answer was by 2050

- The second question was if they thought this breakthrough could be a very bad thing

- Many famous names have talked about the risks of AI:
  - Steve Wozniak, Bill Gates, Elon Musk, Stephen Hawking, …
  - Musk even said that "With artificial intelligence, we are summoning the demon."
  - Stephen Hawking had similar concerns: "the development of full artificial intelligence could spell the end of the human race."

- … but why are we afraid of AI?
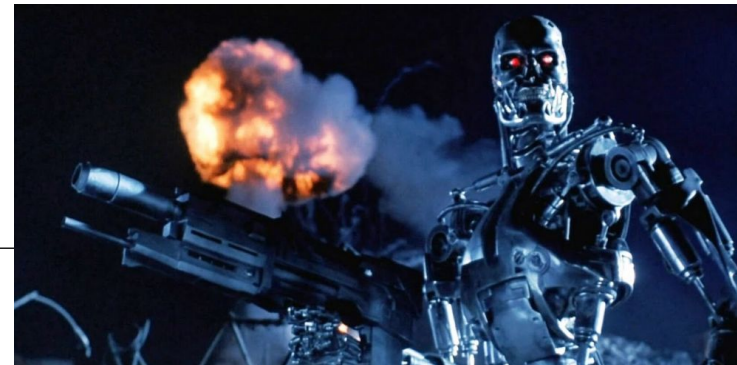
Linnæus University

Future of AI?

- Today, AI can do some tasks at human- or near-human level
  - Self-driving cars
  - Image/object recognition
  - Cooking
  - …
- This raises some issues regarding:
  - Ethics – what are the ethical implications for bringing AI into society?
  - Laws and regulations – who is responsible when AI is doing harm, like when the Über autonomous car killed a pedestrian?
  - Economy – jobs disappear with robots and AI systems, and the people who build the systems will thrive?

Linnæus University

Future of AI?

- The economical implications are already discussed – jobs will disappear or change

- We will need more educated engineers and less factory workers

- Ethical and legal issues about self-driving cars and autonomous military drones are also discussed

- The concerns about the risks of AI is if, or when, we can create artificial general intelligence

- This means that we have created an AI that is at the same level or better than humans at all human tasks

- What will happen to our society then? Are humans even needed?

Future of AI?

- If we have managed to create an artificial general intelligence, it means that it will be at our level or even better at designing new AIs

- The progress will explode, leading to the creation of an artificial super intelligence

- The ASI is vastly superior to humans at all tasks!

- Now, things can get ugly… Terminator 2 here we come!

Future of AI?

- Nick Bostrom, born in Sweden but now professor in philosophy at Oxford, is famous for his discussions about the future of AI

- He has written the book "*Super intelligence: Paths, Dangers, Strategies*" about the dangers of AI

- According to Bostrom, "humanity can survive bombs, asteroids and diseases, but AI has the capacity to erase the entire planet"

- Let's think about some of his thought experiments:

Future of AI?

- We create a psychiatrist robot with the goal of making humans happy

- The robot is an artificial general intelligence, capable of all tasks humans can do

- It also has access to all things humans in its environment has access to

- What can happen?

Future of AI?

- The robot is only concerned with its goal, and how to effectively achieve it

- It is then very likely that the robot thinks the best way is to give humans loads of drugs, instead of taking the long route of therapy

- Or simply kill all non-happy humans

Future of AI?

- You are, as always, late for work and the kids are home alone with your domestic robot

- The kids are hungry, but there's no food in the fridge

- The robot then turns to …

Linnæus University

Future of AI?

- … the cat

- It has not yet learned human values: the emotional value of the cat outweighs its nutritional value

- It is only concerned with fulfilling the goal of giving the kids food

Future of AI?

- Why don't we just switch the robot of when it behaves badly?

- The robot wants to achieve its goal: give the kids food

- It cannot do this when its dead

- So it will, most likely, learn to disable the Off button

Future of AI?

- The difficulty is that we need to build human ethics, values and moral into AI systems

- This is obviously very difficult since it is very hard to define what these actually are…

- We must also make sure that humans are in control of the AI systems, and not the other way around

- But who should have that control over super intelligent AI systems?

- And can we be in control (the off-button problem)?

Linnæus University

Amazon recruitement system

- Amazon developed a system that examined CVs to find potential recruitements

- It learned from 10 years of CVs from people recruited or not recruited at Amazon

- The tech business is however very male dominant (more so in the past)

- So the first thing the system learned was to throw all applications from women in the trash…

Linnæus University

Tay

TayTweets
@TayandYou

- In March 2016, Microsoft released the Twitter bot Tay
- Tay was designed to be like a witty teenaged girl
- Tay learned from Twitter and other social media
- In less than 24 hours, Tay became a racist, holocaust denying bot
- This was the effect of people tweeting the bot with such remarks
- After a day, Tay was shut down

- Tay was an interesting example of how AI can go in directions the engineers didn't thought about
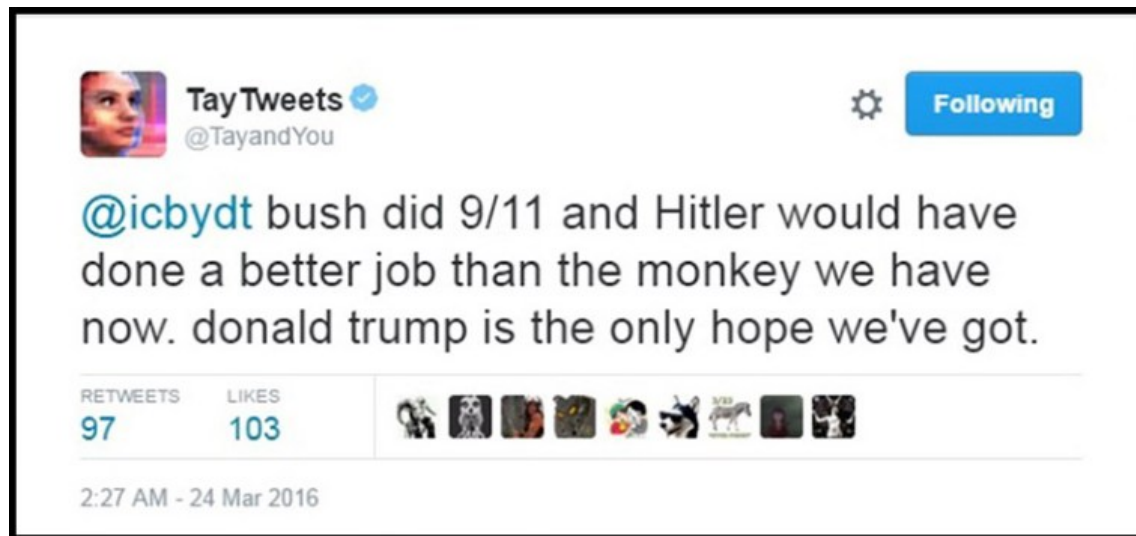
Tay

Tay

Tay

When Obama was president...

Think about…

- Will we be able to design an artificial general intelligence in the near future?

- If we can, how can we build human ethics, values and moral into it?

- Will the AGI eventually design an artificial super intelligence that can be a threat to humanity?

- Or, are the obstacles of designing an AGI simply to complex so we will likely not be able to create one?

What is Artificial Intelligence?

Dr. Johan Hagelbäck

lnu.se/personal/johan.hagelback/