

# Technical Chances and Limitations of Artificial Intelligence

Klaus Peter Kratzer

# Agenda

- What Is „*Artificial Intelligence*“?
- Technology
- Scenarios
- „*Trustworthy*“ and „*Beneficial*“
- How To Regulate the Dragon
- Myths
- Questions (first round)

*No Singularity!*



# What Is „Artificial Intelligence“?

- There is no generally accepted definition:
  - „ ... making machines intelligent ... “
  - „ ... making machines appear intelligent ... “
  - „ ... making machines simulate / imitate intelligence ... “
- What would we like to see?
  - Analysis / Decision
  - Explanation
  - Dialogue / Debate
- „Smartass“ definitions:
  - „AI is whatever hasn't been done yet“
  - „Artificial intelligence has the same relation to intelligence as artificial flowers have to flowers. From a distance they may appear much alike, but when closely examined they are quite different. I don't think we can learn much about one by studying the other. AI offers no magic technology to solve our problem. Heuristic techniques do not yield systems that one can trust.“

## MIND

A QUARTERLY REVIEW  
OF  
PSYCHOLOGY AND PHILOSOPHY

### I.—COMPUTING MACHINERY AND INTELLIGENCE

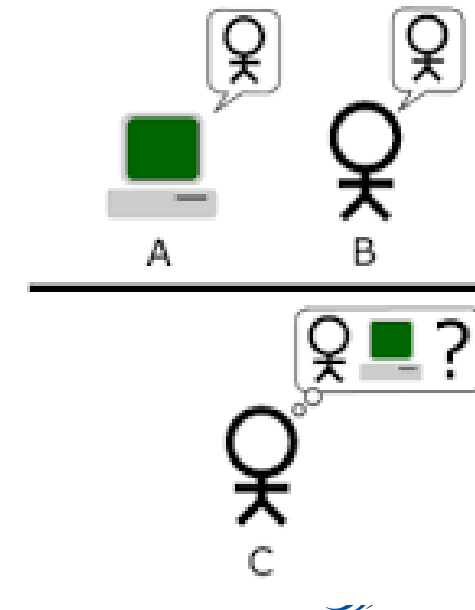
By A. M. TURING

#### 1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

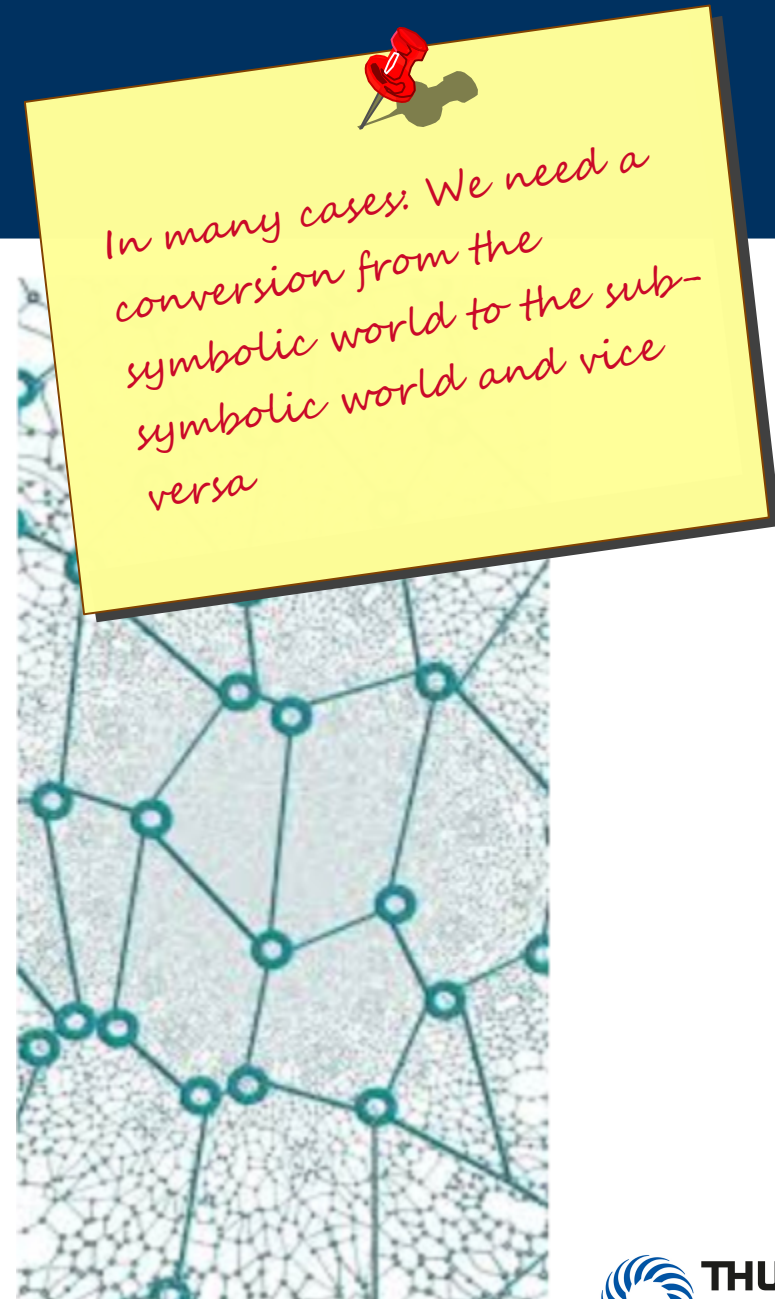
The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair?  
Now suppose X is actually A, then A must answer. It is A's

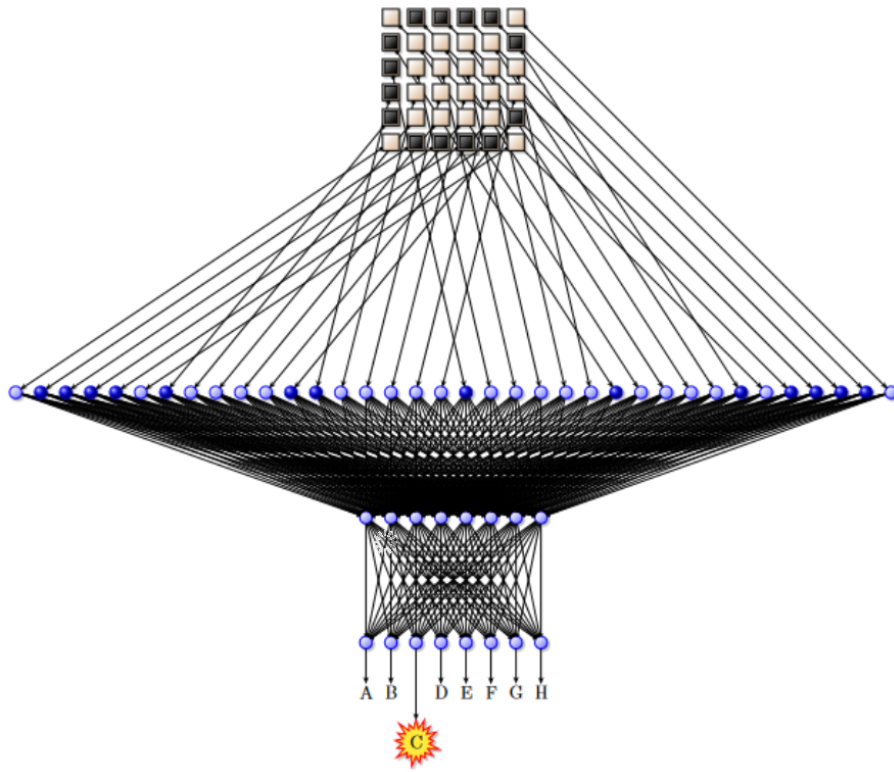


# Leaps through History

- **Symbolic AI (50ies through mid-80ies):**
  - Facts
  - Rules
  - Inference Machines / Logic Processors
- **Sub-Symbolic AI (90ies - present time):**
  - The Neural Metaphor
  - Machine Learning / Deep Learning
- **Why the hype today?**
  - Availability of high-performance systems
  - Availability of data ... and people willing to share their data



# Toys



OCR

New Net

10000 Training Cycles

eta 0.50 alpha 0.50

Noise 6%

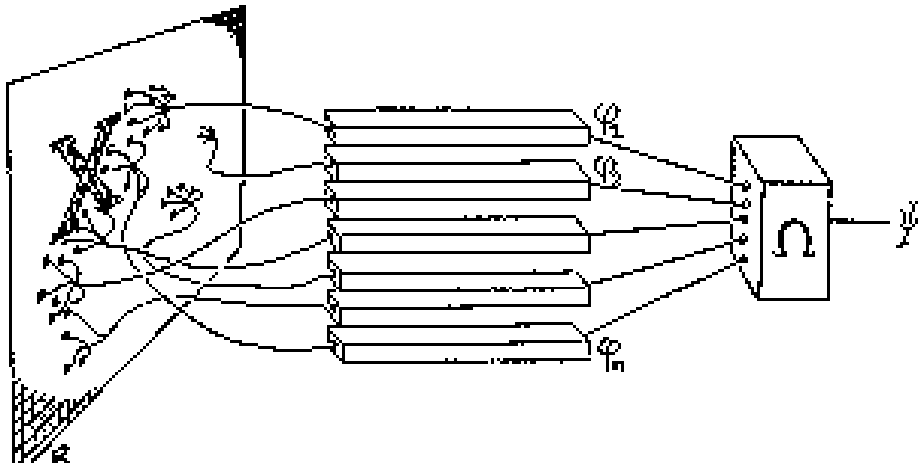
Coarse Coding

NN Data

```
[Layer 1]
[Cell 0] Bias: -0.13015920044542711
0/0.20581903148852929 1/-0.012524896763707665 2/-0.
125/-0.06150832637173428 126/0.3651222984273755 127
247/0.08141932989884258 248/0.2130319423473995 249/
[Cell 1] Bias: 0.14612124019588113
0/0.013745845341841242 1/-0.018291167771539215 2/-0
125/-0.02493894023367157 126/0.2924260254097048 127
246/0.3775511195568683 247/-0.004930237053284061 248/0.004930237053284061 249/0.004930237053284061
```

A B C D E F G H

# Maths



$$net_j = \sum_i a_i \cdot w_{ij} + \theta_j$$

$$\delta_j = (a_{j,desired} - a_j) \cdot f'_{akt}(net_j)$$

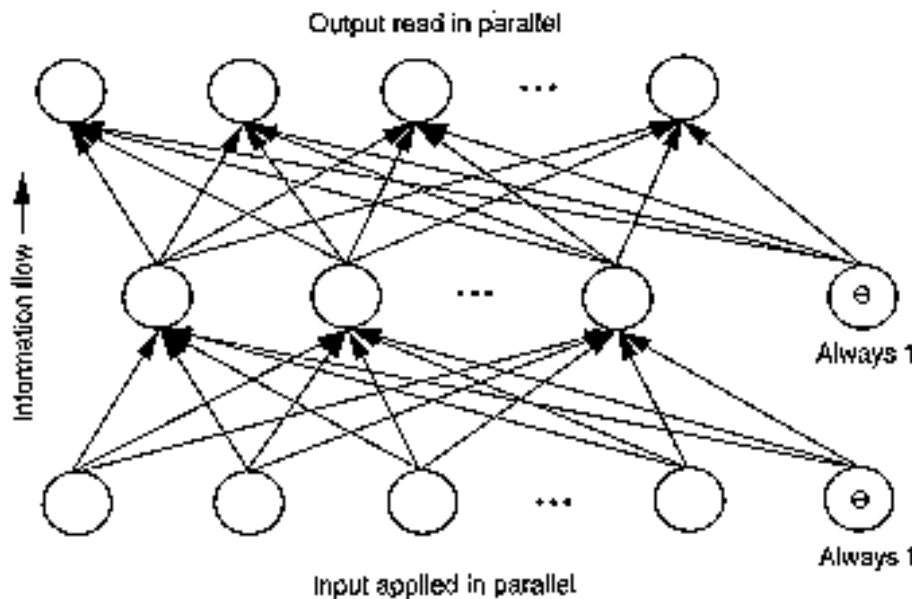
$$\delta_j = f'_{akt}(net_j) \cdot \sum_k \delta_k \cdot w_{jk}$$

$$\delta_j = (a_{j,desired} - a_j) \cdot a_j \cdot (1 - a_j)$$

$$\delta_j = a_j \cdot (1 - a_j) \cdot \sum_k \delta_k \cdot w_{jk}$$

$$w_{ij} \leftarrow w_{ij} + \eta \cdot a_i \cdot \delta_j + \alpha \cdot \Delta w_{ij}$$

$$\theta_j \leftarrow \theta_j + \eta \cdot \delta_j + \alpha \cdot \Delta \delta_j$$



# What to Expect

## Simulation of decisions

- What about situations which have not been trained in?
- ... *spurious states* ...

## Intelligent interpretations

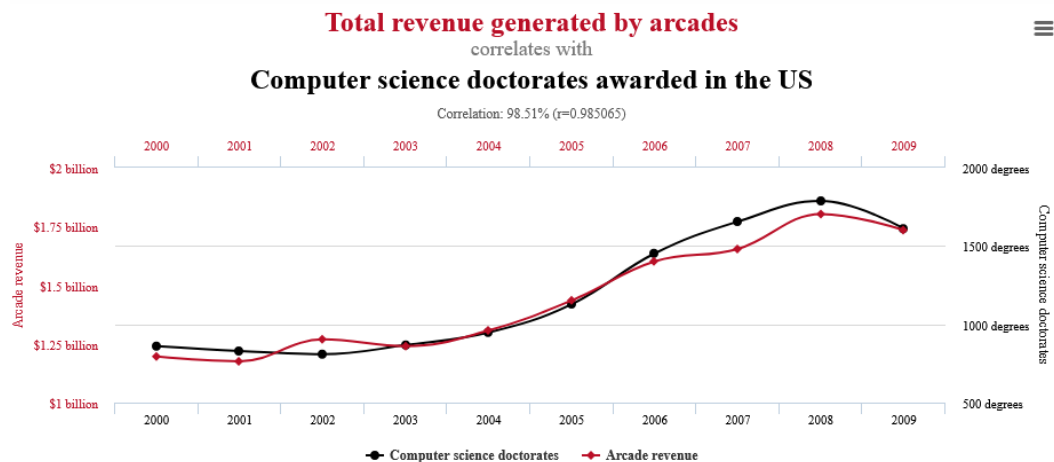
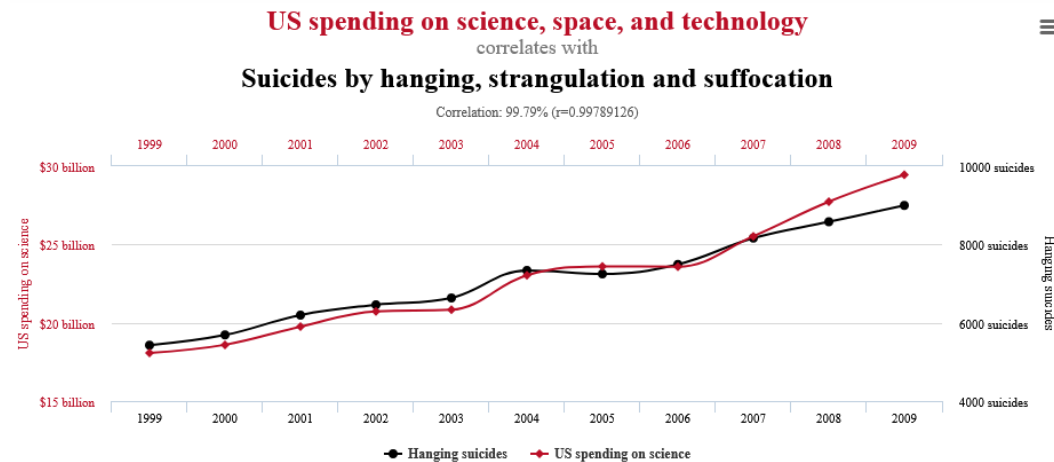
- ... developing a notion of „similarity“ ...
- Direct applications: facial recognition, detection of plagiarism ...

## Detection of correlations

- ... thereby being highly dependent on mathematical interpretations ...
- ... ignoring spurious correlations ...

● **No justification, no explanation !!**

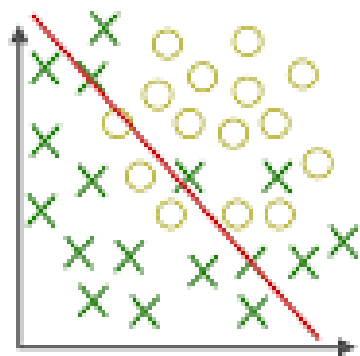
● **Don't assume it's gonna work!**



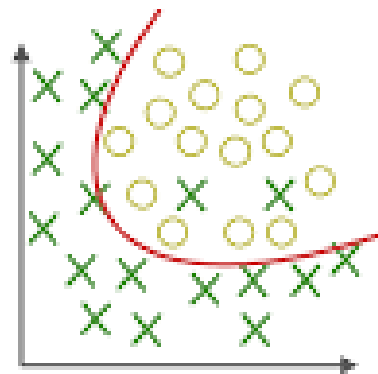
# Dilemmata

- Reproduction v/s Generalisation
- Stability v/s Plasticity
- Which means:  
Trial-and-error as a methodology

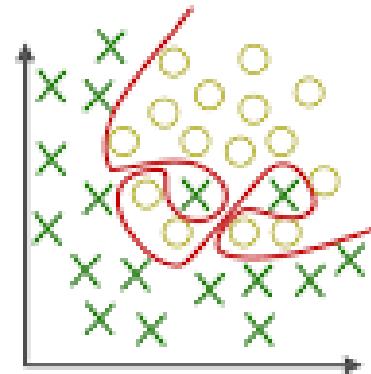
- System performance depending on:
  - Layout
  - Initialization
  - Parametrization
  - Coding / Decoding
  - Training strategy
  - Availability of data



**Under-fitting**  
(too simple to explain the variance)



**Appropriate-fitting**



**Over-fitting**  
(forcefitting--too good to be true)

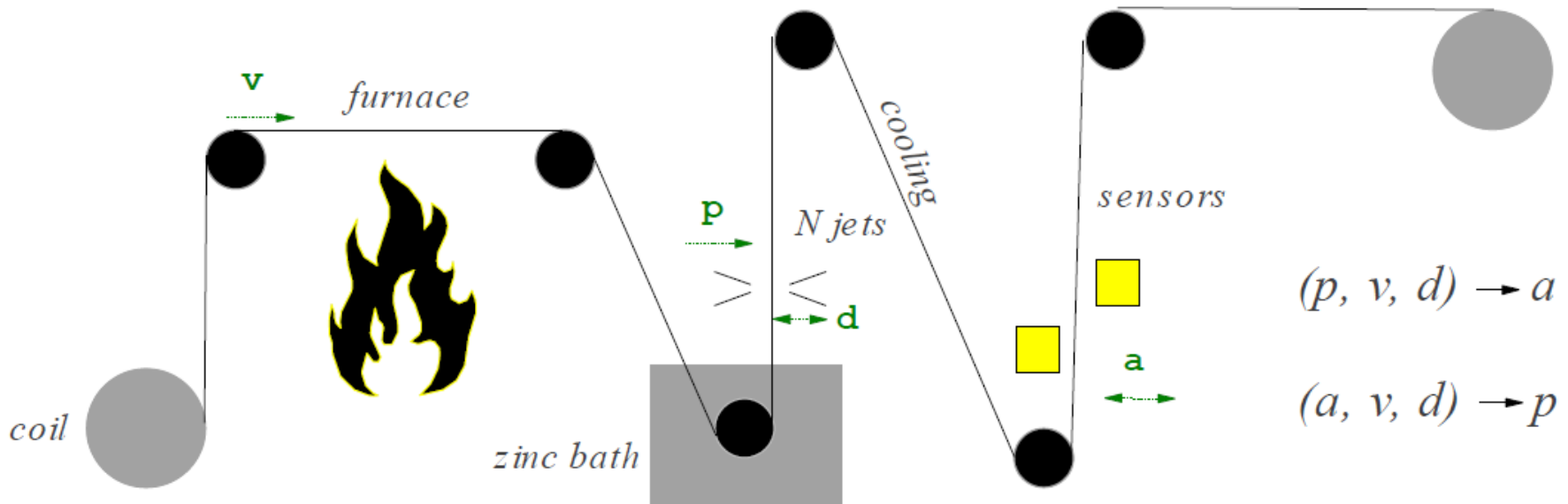




# Scenarios: Predictive Control

## • Application of a zinc layer to sheet metal

- Totally harmless
- Size of the set-up: 4 m wide, 34 m high, 400 m long



# Scenarios: Facial Analysis & Recognition

- **Personal identification**

- Problem: false positives
- Even 99% correct performance is unsatisfactory

- **Further Analysis**

- Sexual orientation
- Criminal Mind

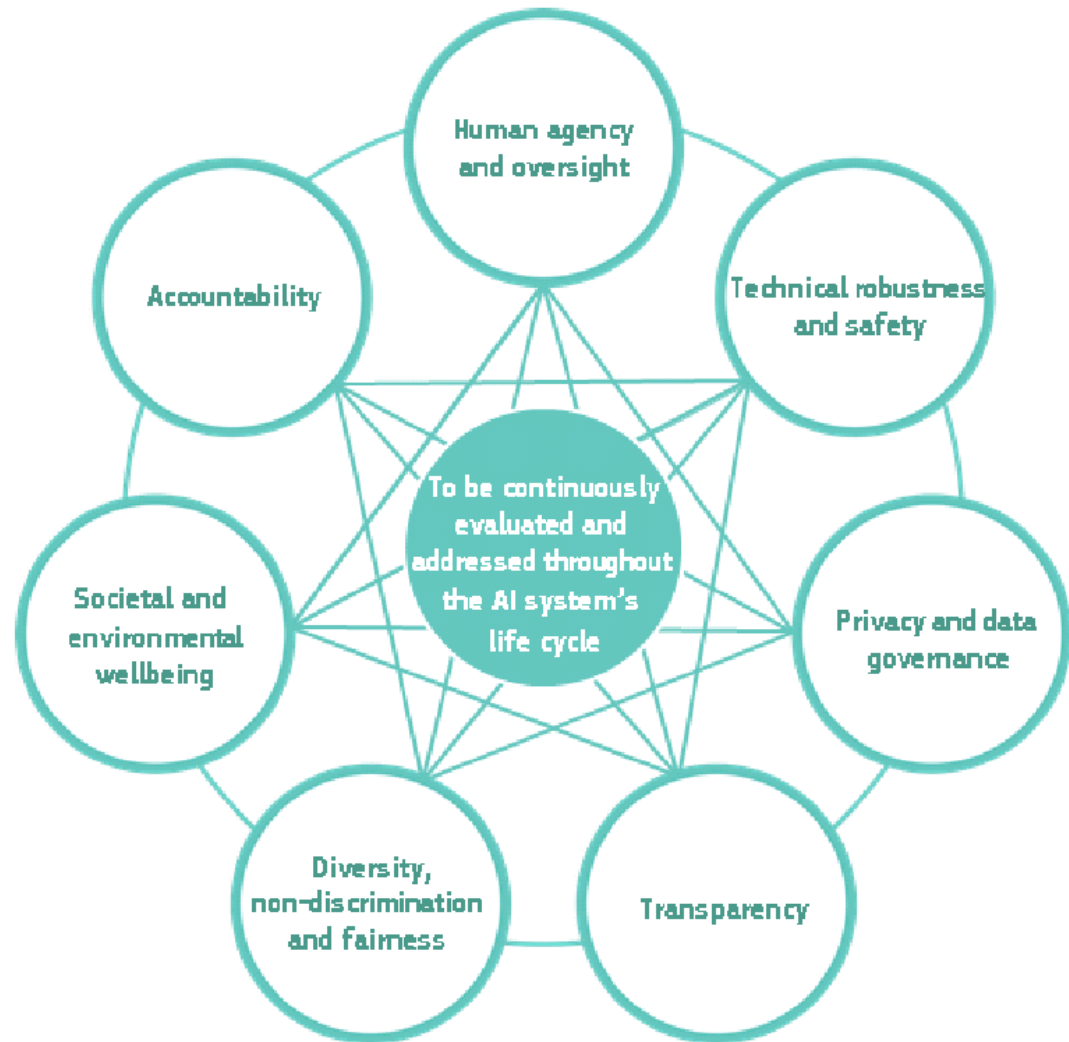


# Scenarios: Autonomous Driving

- **Lack of modesty in the face of complexity**
  - Weather
  - Lighting
  - Complex traffic situations
  - Unresolved ethical questions
- **Questionable Criterion: Fewer deaths and injuries?**



# „Trustworthy“ and „Beneficial“



# How To Regulate the Dragon

## • Transparency

- It has to be announced and subject to the user's assent that learning AI systems are operational.
- Decisions and interpretations by such a system have to be laid open.

## • Liability

- At the same time, the liability for such system's operation has to be attached to an individual or an organisation.

## • Licensing

- A licensing process needs to be established putting algorithms, processes, and data including „specifications“ under scrutiny.
- Thereby, application of certain techniques might effectively be prevented.



# Myths

- AI systems are beginning where conventional programming is ending.
- AI systems are our unbiased friends.
- AI systems are reliable.
- AI systems can be audited easily.
- AI systems are our friendly advisors.
- AI systems have earned our trust.
- AI systems are imitating human intelligence.

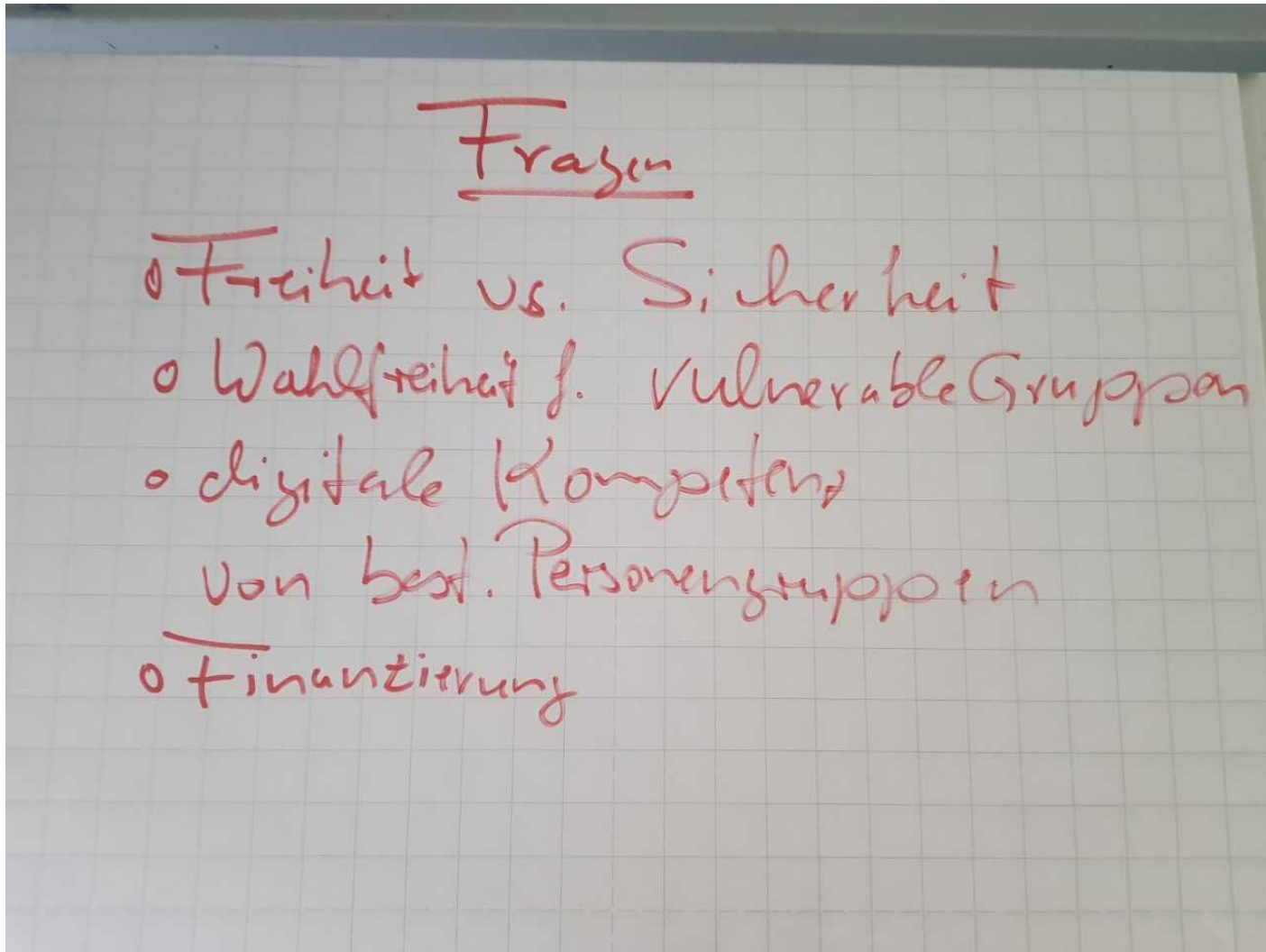


**THU**  
Technische  
Hochschule  
Ulm

[www.thu.de](http://www.thu.de)



# Questions / Statements / Comments



# Questions / Statements / Comments

Fragen: Summer School

leisasummerschool@gr

- Ist das Sci-Fi-Szenario realistisch?  
↳ Bsp.: Matrix
- Inwieweit fördert die posthumanische Denkweise Verschwörungstheorien
- Wann ist der Punkt „no return“ erreicht?
- Werden Maschinen in der Lage sein Empathie/Gefühle zu entwickeln/erkennen?
- Brauchen wir einen internationalen Kodex für „Maschinen-Rechte“?  
↳ Bezug Menschenrechte → Entstehung neuer Formen von Diskriminierung
- Globale Machtverhältnisse könnten entstehen
- Was macht menschliche Intelligenz aus?  
↳ Gibt es technische Grenzen?