Technical Chances and Limitations of Artificial Intelligence

Klaus Peter Kratzer

THU
Technische
Hochschule
Ulm

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.*"

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
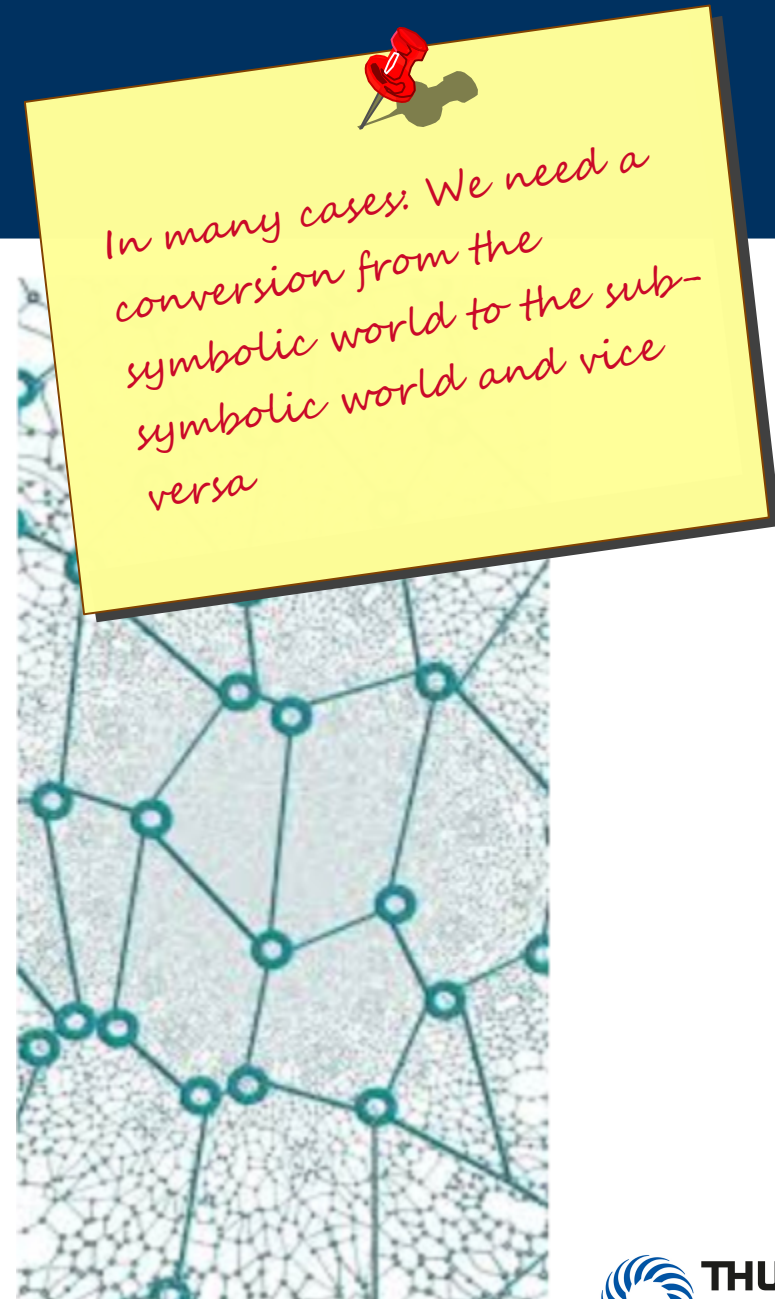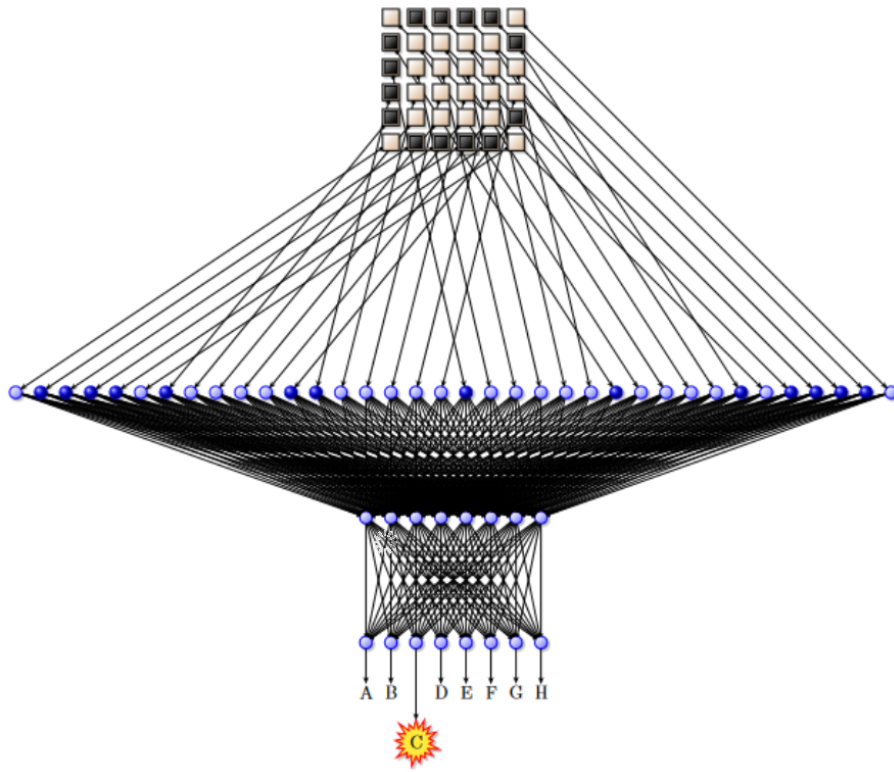
In many cases: We need a conversion from the symbolic world to the sub-symbolic world and vice versa

Toys

Maths

Output read in parallel

Information flow →

Input applied in parallel

Always 1

Always 1
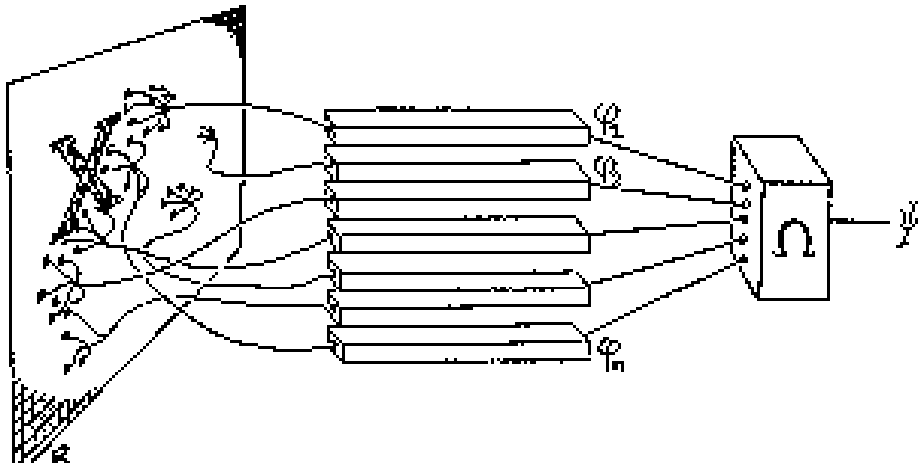
$$net_j = \sum_i a_i \cdot w_{ij} + \theta_j$$

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

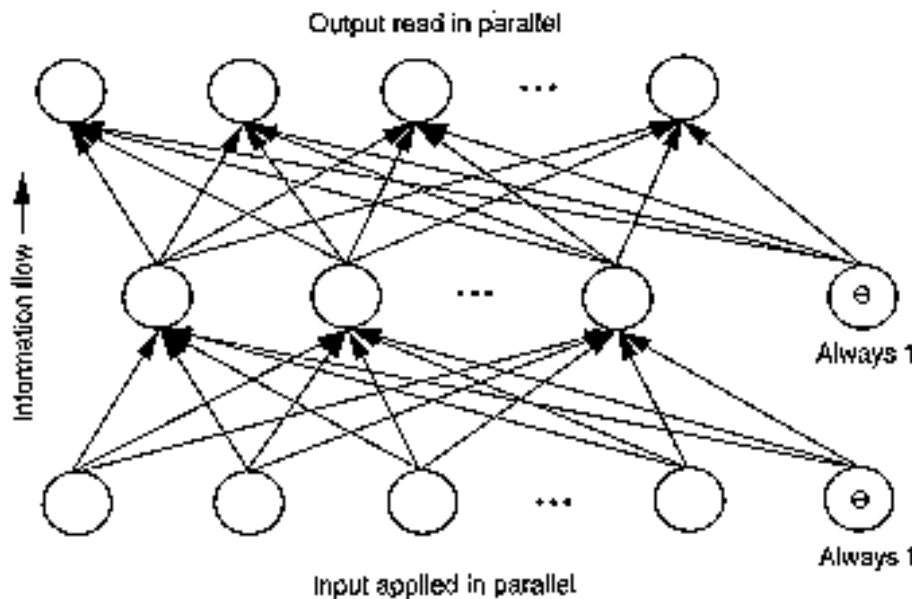$$\delta_j = f'_{akt}(net_j) \cdot \sum_k \delta_k \cdot w_{jk}$$

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

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

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

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

THU
Technische
Hochschule
Ulm

What to Expect

- ## Simulation of decisions
  - What about situations which have not been trained in?
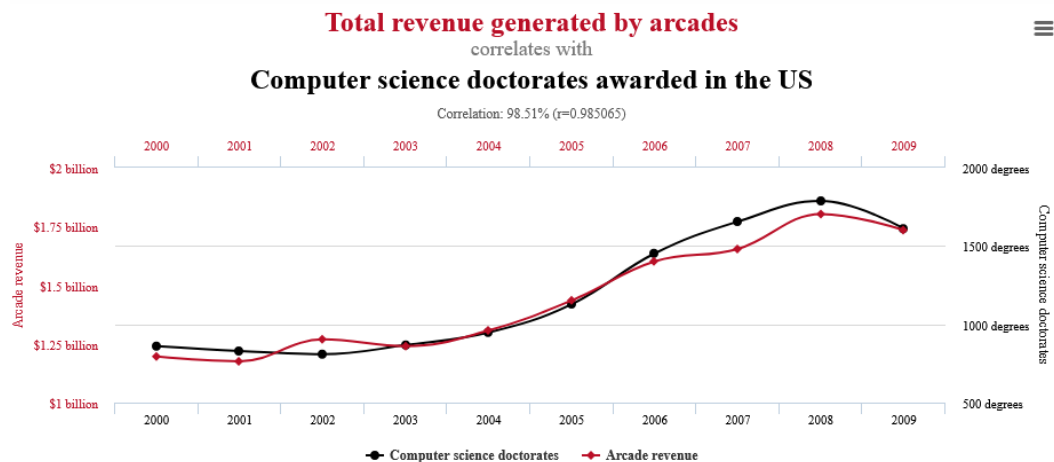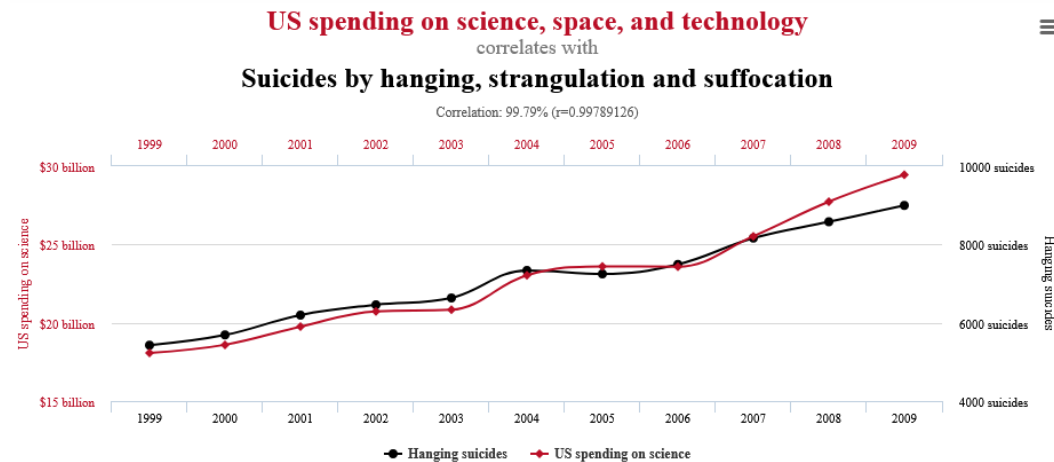 - … *spurious states* …

- ## 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!

US spending on science, space, and technology
correlates with
Suicides by hanging, strangulation and suffocation
Correlation: 99.79% (r=0.99789126)

Total revenue generated by arcades
correlates with
Computer science doctorates awarded in the US
Correlation: 98.51% (r=0.985065)

THU
Technische
Hochschule
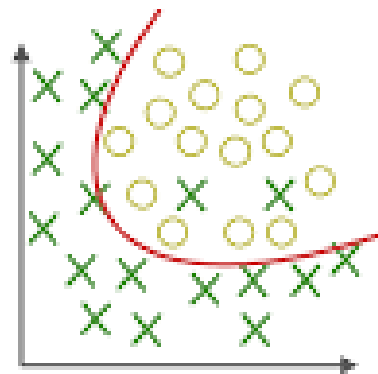Ulm

Dilemmata

- **Reproduction v/s Generalisation**
- **Stability v/s Plasticity**

- **Which means:
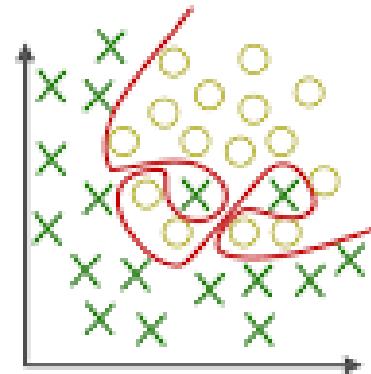 Trial-and-error as a
  methodology**

- **System performance
  depending on:**
  - Layout
  - Initialization
  - Parametrization
  - Coding / Decoding
  - Training strategy
  - Availability of data
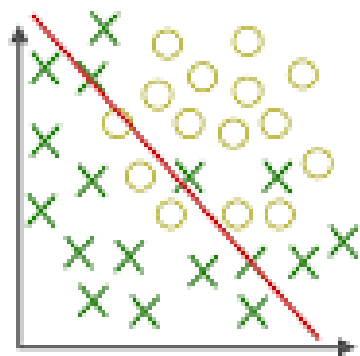
Under-fitting
(too simple to
explain the variance)

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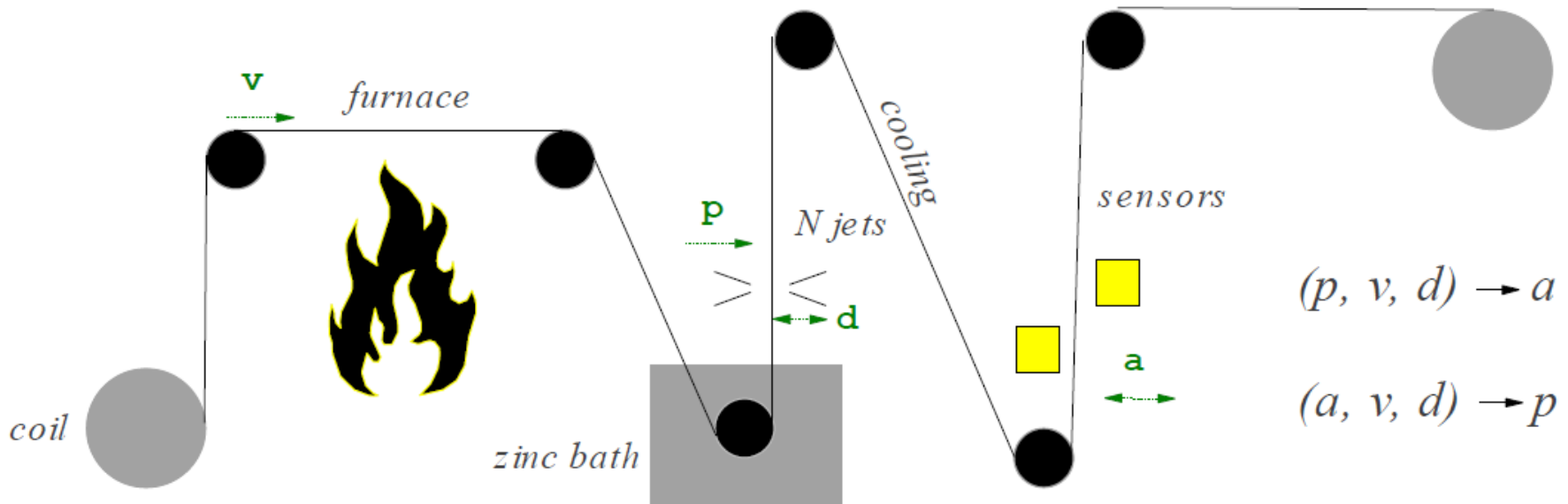
Over-fitting
(forcefitting--too
good to be true)

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**
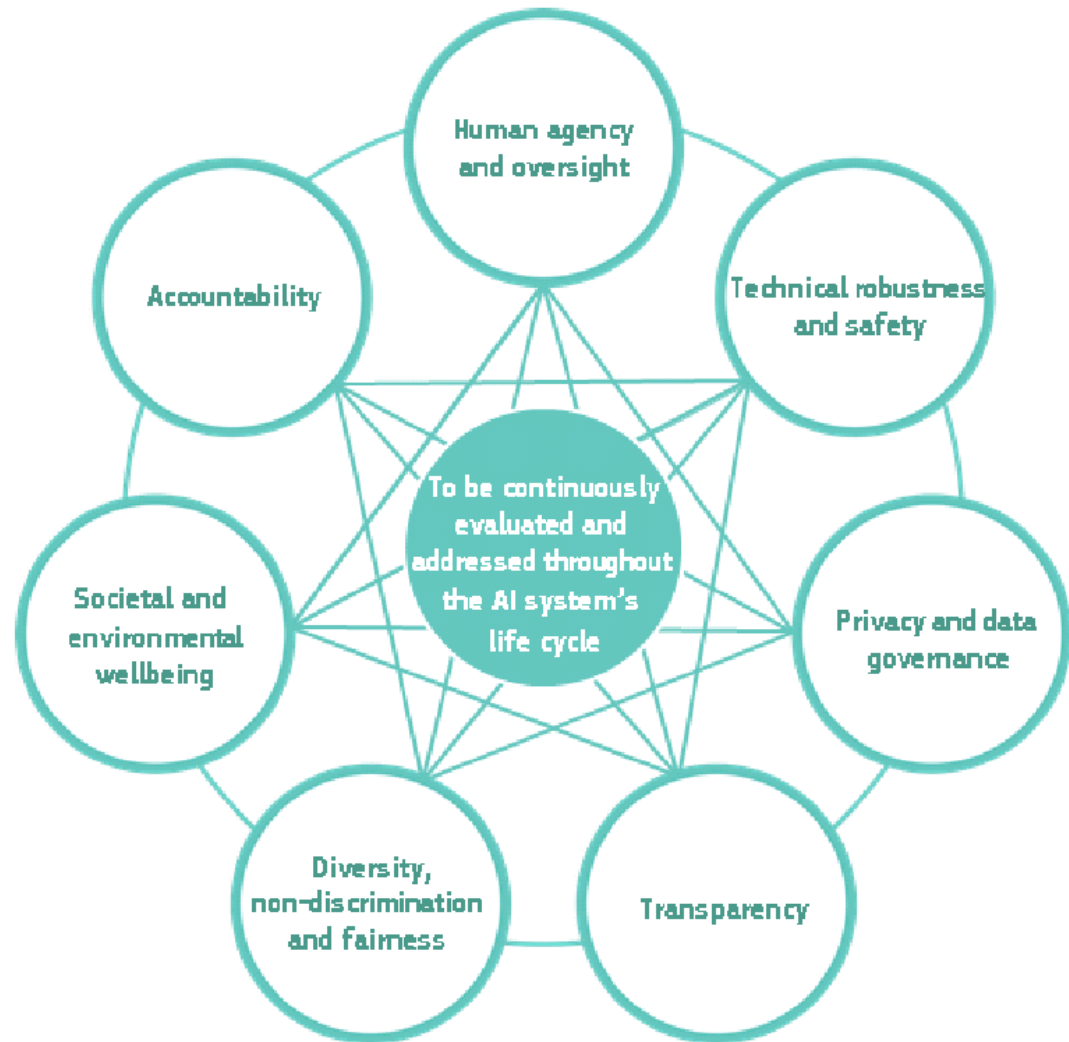 - Sexual orientation
  - Criminal Mind

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.

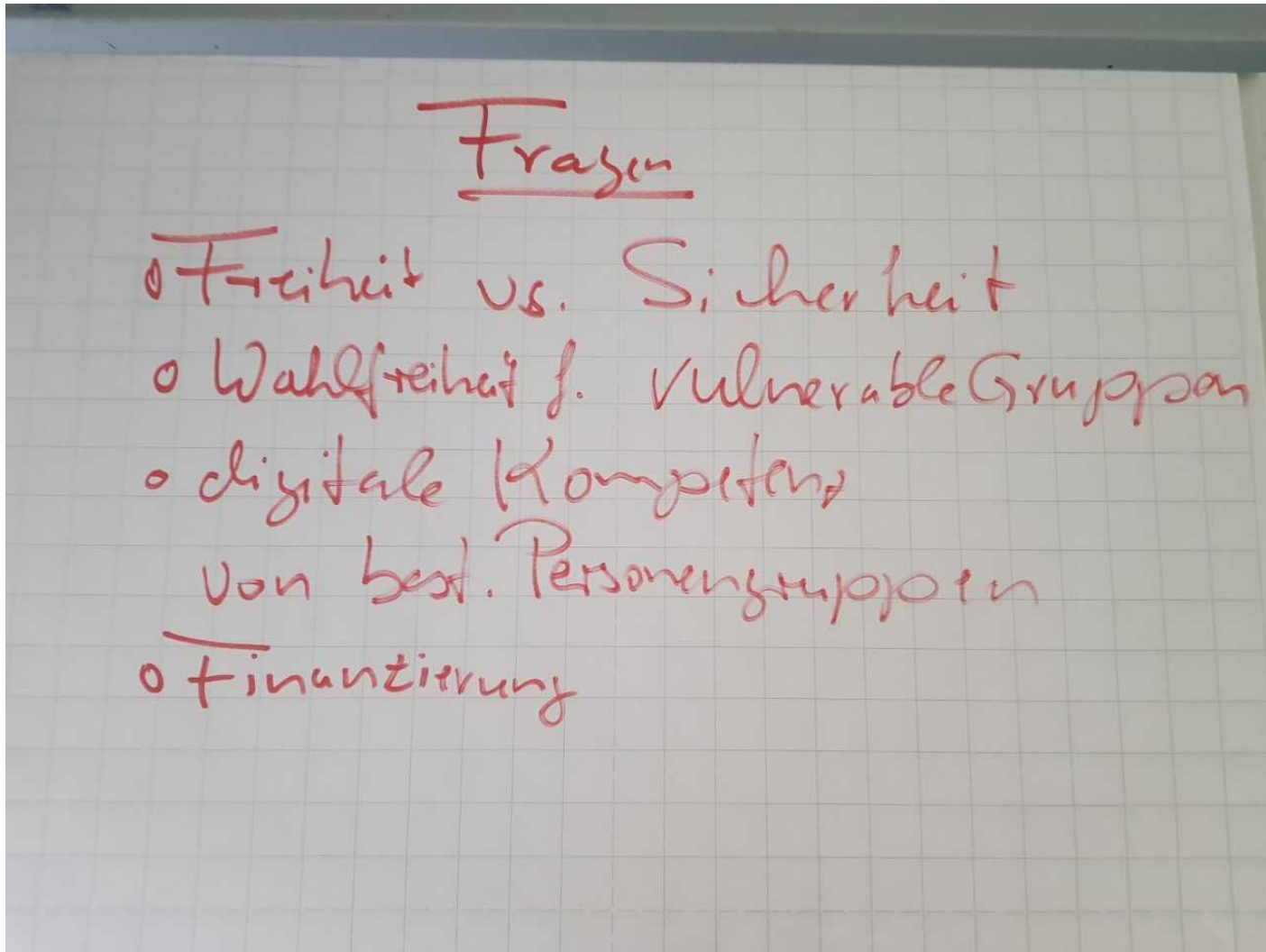- AI systems are imitating human intelligence.

THU
Technische
Hochschule
Ulm

www.thu.de

Questions / Statements / Comments

Questions / Statements / Comments