

Gut oder Böse?

—

Technische Autonomie im Diskurs

Begriffsklärung und Stand der Dinge

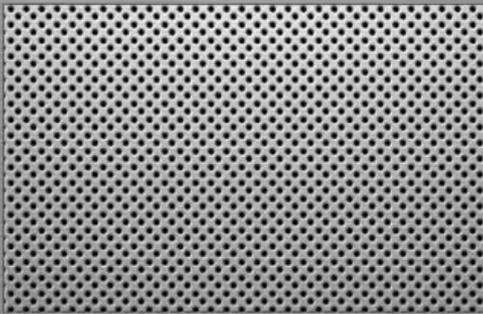
Ein Beitrag im Rahmen der TREE-Forschungskonferenz am 09.06.2017
an der Hochschule Bonn-Rhein-Sieg

Michael Lauster

Fakultät für Maschinenwesen – FB 4
Lehrstuhl für Technologieanalyse und –
vorausschau auf dem Gebiet der
Sicherheitsforschung

Univ.-Prof. Dr. Dr. Michael Lauster
Fraunhofer INT Euskirchen
(02251) 18 117
michael.lauster@int.fraunhofer.de

HAL 9000



Gut oder Böse? – Gut und Böse!

„Deontische“ Grundbegriffe¹⁾:

sie zielen auf die **1. Grundfrage der Ethik:**

Wie soll ich handeln?

¹⁾Sie beziehen sich auf **ethische Verpflichtungen**

Gut oder Böse? – Gut und Böse!

Dies zieht sofort die **2. Grundfrage der Ethik** nach sich:

Was ist gutes oder böses Handeln?

oder radikalisiert:

Warum sollte ich ethisch handeln?

Gut oder Böse? – Gut und Böse!

Kant:

„Gut“ und „Böse“ sind Kategorien *a priori*

Gut oder Böse? – Gut und Böse!

Kants Kategorischer Imperativ:

„handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, daß sie ein allgemeines Gesetz werde.“

(Was Du nicht willst, das man Dir tu‘, das füg‘ auch keinem anderen zu!)

Gut oder Böse? – Gut und Böse!

Folgerung 1:

„Gut“ und „Böse“ sind Kategorien, die **autonomen** Lebewesen (u.a. Menschen) zuzuschreiben sind.

Zum Autonomiebegriff

Autonomie

(altgriechisch αὐτονομία *autonomía*
,Eigengesetzlichkeit', ,Selbstständigkeit', aus
αὐτός *autós* ,selbst' und νόμος *nómos* ,Gesetz')

Zum Autonomiebegriff

Ursprünglich:

politischer Begriff (innere und äußere Freiheit von Tyrannis und Gewaltherrschaft)

(idealistische) Philosophie:

Fähigkeit, sich als Wesen der Freiheit zu begreifen und aus dieser Freiheit heraus zu handeln; unabdingbare Grundbedingung ethischen Handelns

Zum Autonomiebegriff

Freiheit:

Fähigkeit, sich Ziele (Zwecke) zu wählen und sie mit geeigneten Mitteln in die Tat umzusetzen.

Zum Autonomiebegriff

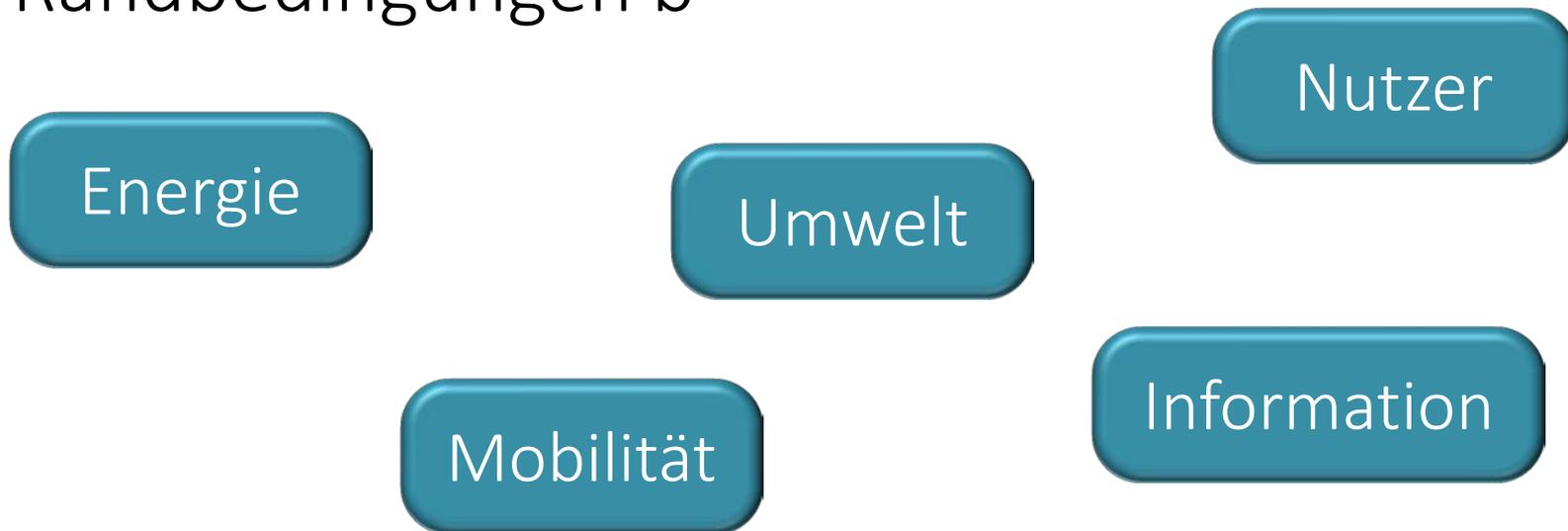
Folgerung 2:

„Gut“ und „Böse“ charakterisieren die Wahl der Zwecke sowie der angewandten Mittel zu deren Umsetzung

Zum Autonomiebegriff

Technische Autonomie:

System ist autonom in Hinsicht α unter
(normalerweise nicht ausgesprochenen)
Randbedingungen b



Zum Autonomiebegriff

Folgerung 3:

Autonomie ist moralisch relevante Intentionalität
von Personen

Technische Autonomie ist Relationalität in Bezug
auf Personen und Umwelt

Zum Autonomiebegriff

Folgerung 4:

Technische Autonomie ist weder gut noch böse.
In diesen Kategorien lassen sich nur die Zwecke ihrer Verwendung und die Art und Weise ihres Einsatzes beurteilen.

Autonomie und Intelligenz

Eine grundsätzliche Frage:

Können technisch autonome Systeme **intelligent** sein und damit die Kantsche Forderung nach **Moralität** erfüllen?

oder radikalisiert:

Sind technisch autonome Systeme **lebendige Wesen**?

Künstliche Intelligenz

Eine grundsätzliche Unterscheidung:

Schwache Künstliche Intelligenz: das technische System, beschränkt auf spezielle Aufgaben, simuliert intelligentes Verhalten, ist aber selbst nicht intelligent.

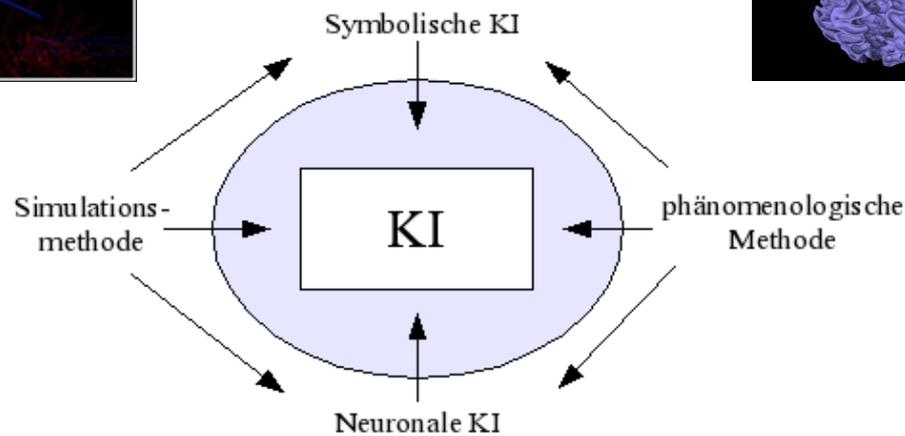
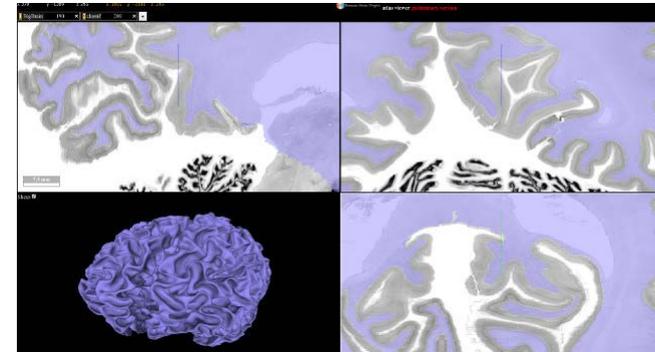
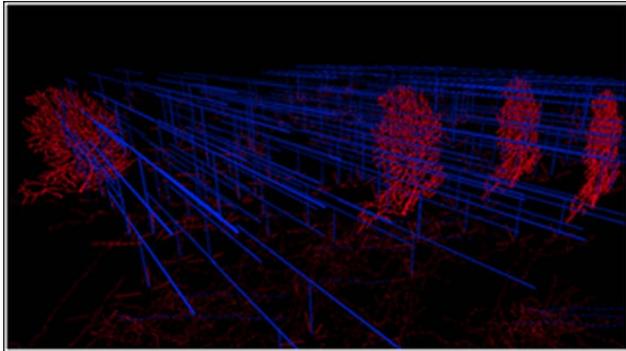
Starke Künstliche Intelligenz:

Das technische System ist „wirklich intelligent“
(kreativ, planend, lernend, empathisch?, selbstbewusst?)¹⁾

¹⁾ vgl. http://www.informatik.uni-oldenburg.de/~iug08/ki/Grundlagen_Starke_KI_vs._Schwache_KI.html abgerufen am 30.05.2017 16

Der Stand der Dinge

Starke künstliche Intelligenz:

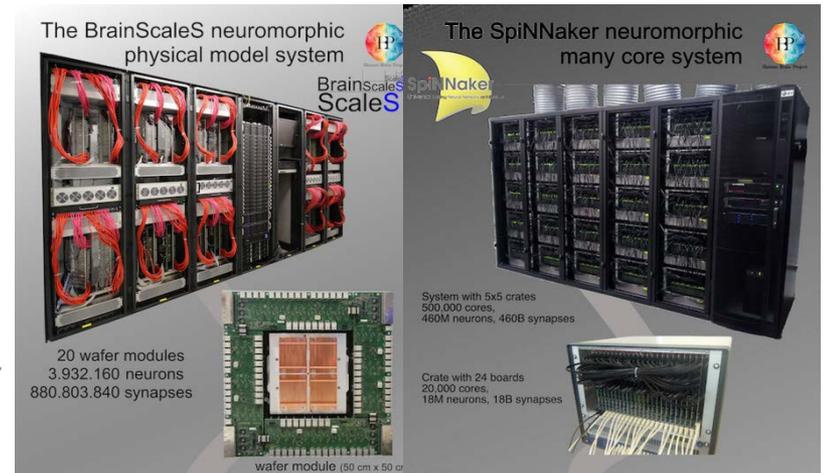


Der Stand der Dinge

Starke künstliche Intelligenz:

- Prinzipielle Bedenken bezüglich der Machbarkeit
- Expertenschätzungen zur „Singularität“ mit starker Streuung
- Viele Ansätze
- Ungelöstes Problem: Bewusstsein

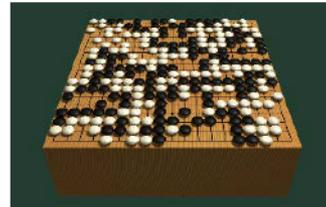
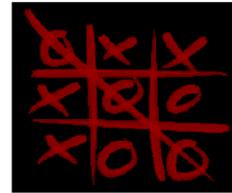
Zum Vergleich: menschliches
Gehirn ca. 100 Mrd. Neuronen!



Der Stand der Dinge

Kurze Historie der „maschinellen Intelligenz“

- 1950 Definition des Turing Tests
- 1952 erster Sieg eines Programmes bei Tic Tac Toe
- 1956 erster Sieg eines Schachprogrammes gegen einen Menschen
- 1996 Deep Blue verliert gegen Kasparow
- 1997 Deep Blue gewinnt Revanche mit 3,5:2,5
- 2011 IBMs Watson gewinnt Jeopardy
- 2016 AlphaGo gewinnt Go gegen Großmeister



Der Stand der Dinge

Einsatzgebiet	Fähigkeit ¹⁾
Schach	übermenschlich
Poker	unterschiedlich
Med. Diagnose	menschenähnlich
Kreuzworträtsel	Expertenniveau
Übersetzungen	Brauchbar
Go	meisterlich

¹⁾ vgl. Bostrom, N., Superintelligenz, Berlin 2014, S.27 ff.

Der Stand der Dinge

Starke künstliche Intelligenz:

Turing-Test:

Kann man in einer Unterhaltung nicht mehr unterscheiden, ob man einer Maschine oder einem Menschen gegenüber sitzt, dann ist die Maschine als intelligent zu bezeichnen.

Der Stand der Dinge

**Alle aktuellen technisch autonomen
Systeme sind schwache künstliche
Intelligenzen!**

Der Stand der Dinge

Die kognitiven Fähigkeiten des Menschen:¹⁾



¹⁾ Nach Bloom, B. S. (1956). Cognitive Domain. Longman.

Quelle: <https://www.bitkom.org/noindex/Publikationen/2017/Leitfaden/Bitkom-Leitfaden-KI-verstehen-als-Automation-des-Entscheidens-2-Mai-2017.pdf> am 31.05.2017

Der Stand der Dinge

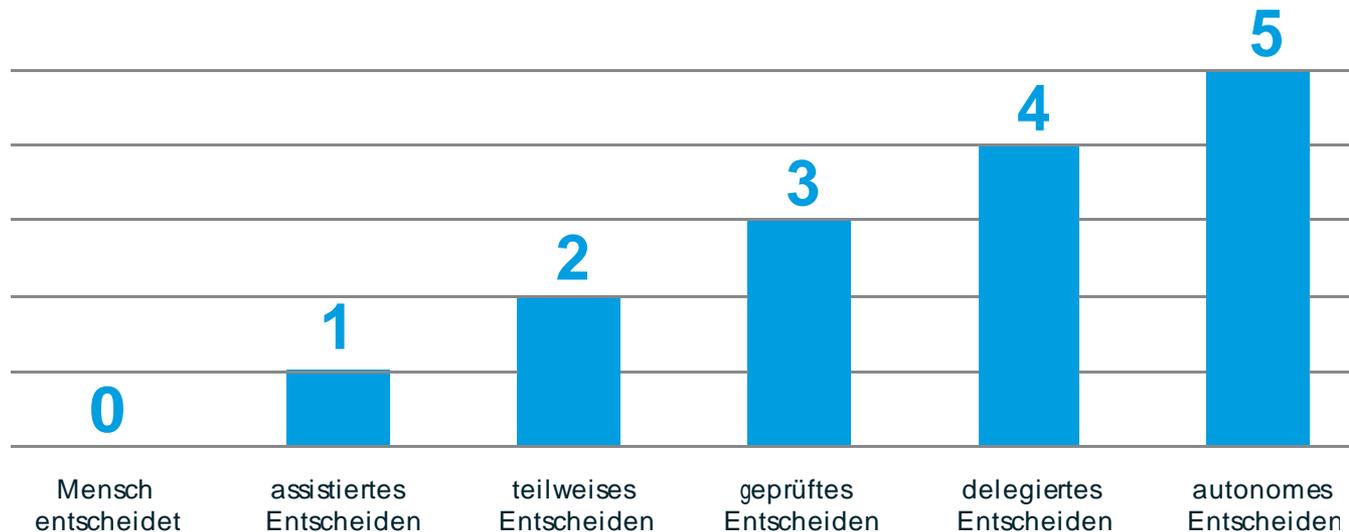
Alternative Sichtweise:

Es geht weniger um die bloße Simulation menschlicher Fähigkeiten, sondern um die Frage, auf welcher Stufe (und bis zu welchem Ausmaß) autonome Systeme Entscheidungen unterstützen können.



Der Stand der Dinge

Unterstützung durch „Cognitive Computing“



Der Stand der Dinge

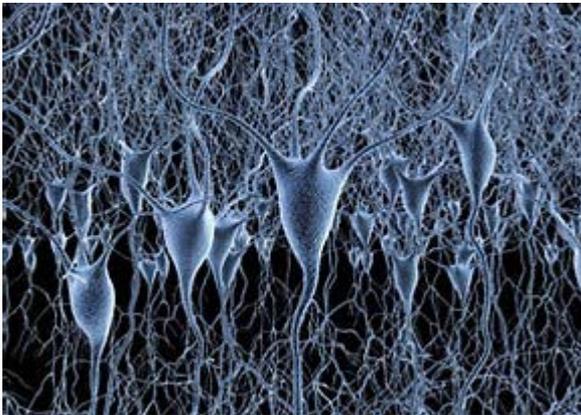
Beispiele:

- KI bei RWE: Kundendialoge führen
- H bei Hitachi: Lagerarbeiter delegieren
- Watson bei Fukoku: Schadensfälle beurteilen
- NeuroBayes bei Otto Versand: Retouren vermeiden
- Amelia bei der Bezirksverwaltung North London Borough: Formulare ausfüllen

Der Stand der Dinge

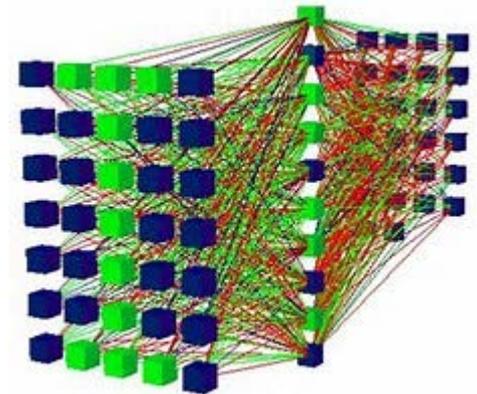
Methoden des „Cognitive Computing“:

- Neuronale Netze



Vorteil: neuronale Netze dienen der Mustererkennung. Sie können lernen und benötigen keine vorab abstrahierten Regeln über diese Muster.

Nachteil: sie müssen trainiert werden

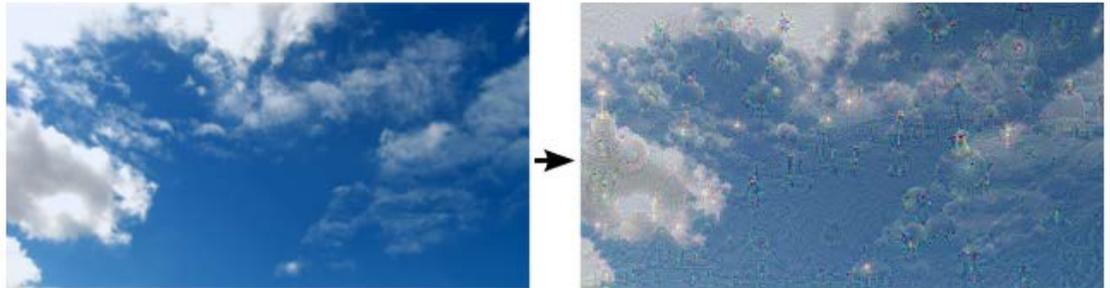


Der Stand der Dinge

Methoden des „Cognitive Computing“:

- Neuronale Netze

Manchmal klappt es –
und manchmal nicht ...



"Admiral Dog!"



"The Pig-Snail"



"The Camel-Bird"



"The Dog-Fish"

Der Stand der Dinge

Methoden des „Cognitive Computing“:

- Deep Learning
 - Bereits relativ alt, scheiterte aber bisher an Hardware und Content-Voraussetzungen
 - Basiert auf mehr(viel-)schichtigen neuronalen Netzen
 - Sehr erfolgreich z.B. bei der Bild- und Spracherkennung

Der Versuch einer Zusammenfassung

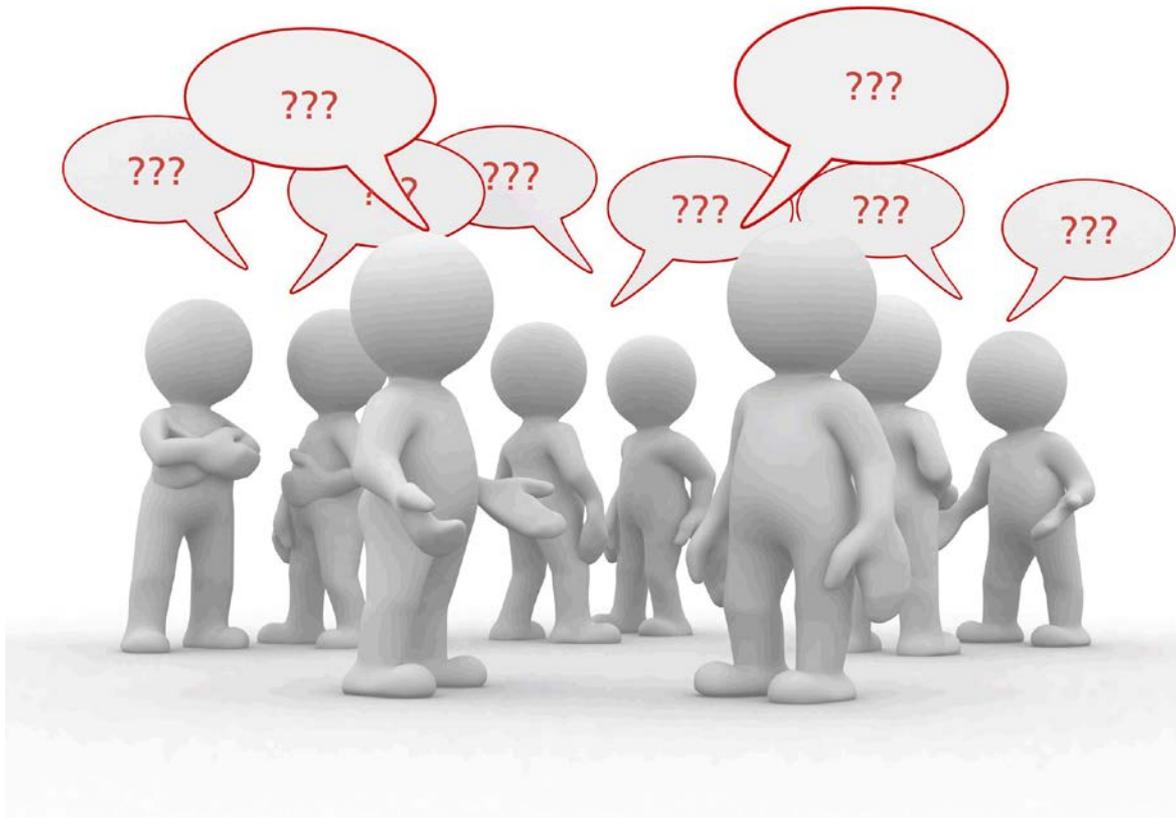
- Autonome technische Systeme sind (noch) keine Lebewesen
- Sie sind Werkzeuge - und damit weder gut noch böse
- Wir (Menschen) entkommen unserer Verantwortung nicht
- Die Situation ist vergleichbar mit derjenigen bei der Einführung von Großrechnern in den 60er und 70er Jahren

Ausblick

„Ich denke, dass es einen Weltmarkt für vielleicht 5 Computer gibt.“

(Watson, Chef von IBM, 1945)

Vielen Dank für Ihre Aufmerksamkeit!



©amasterpics123