

Künstliche Intelligenz und Philosophie

Version 12.11.22

Marco Lardelli

<https://ki-kit.ch>

© 2022 Marco Lardelli

1



Bitte beachten Sie die Lizenzbedingungen (siehe Website <https://ki-kit.ch>).

Themen

**„Wer fragt, ist ein Narr für
eine Minute. Wer nicht fragt,
bleibt es ein Leben lang“**

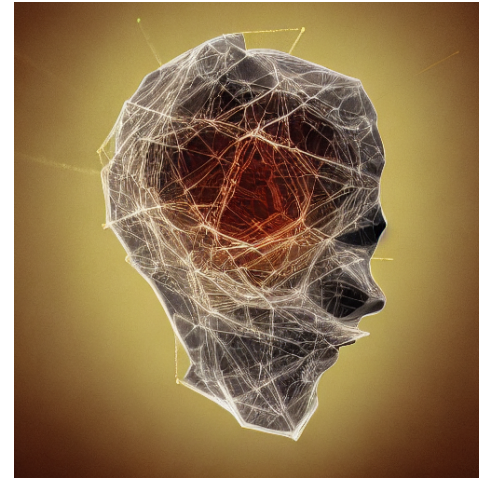
(Konfuzius)

Themen

1. Intelligenz
2. Selbstaufmerksamkeit
3. Bewusstsein
4. Menschenähnliches Verhalten

Wir diskutieren diese Begriffe in Bezug auf verschiedene intelligente Systeme (Mensch, Tier, künstliche Intelligenz).
Leider verwenden viele Autoren diese Begriffe falsch. Zudem werden die Begriffe auch von Experten nicht immer gleich verwendet.
Wichtig ist das Verständnis der Konzepte, nicht die Worte.

1. Intelligenz



Keine allgemein anerkannte Definition!
Experten geben unterschiedliche Definitionen.

1. Was ist Intelligenz?

Welche Begriffe fallen euch zum Wort
„Intelligenz“ ein?

Wortwolke mit SuS erarbeiten. Z.B. mittels [answer garden.ch](https://www.answer garden.ch)

1. Was ist Intelligenz?



Unterschiede erarbeiten.

1. Was ist Intelligenz?

Beispiele für intelligente „Systeme“:



Worin unterscheiden sie sich?

Die gezeigten Systeme sind Beispiele. Sie unterscheiden sich in folgenden Eigenschaften:

- Auf enges Gebiet beschränkte Fähigkeiten (Taschenrechner, Staubsaugerroboter) bzw. Fähigkeiten auf vielen Gebieten (Bonobo, Mensch)
- Kann selbständig neue Fähigkeiten erwerben (z.B. Katze) oder nicht (z.B. selbstfahrendes Auto)
- Künstlich bzw. natürlich

Fotos: Pixabay

1. Was ist Intelligenz?

1. Probleme lösen

2. Lernen

Lernen =
Selbständig neue Fähigkeiten zur
Problemlösung erwerben.

Historisch lag der Fokus früher bei 1. (so hat man z.B. frühe Computer als „Elektronengehirne“ bezeichnet und ihnen Intelligenz zugeschrieben) und hat sich dann immer mehr zu 2. verlagert.

Probleme lösen: welche, wieviele, wie gut?

1. Was ist Intelligenz? - Begriffe KI



Schwache KI
Artificial **narrow** intelligence (ANI)



Starke KI
Artificial **general** intelligence (AGI)



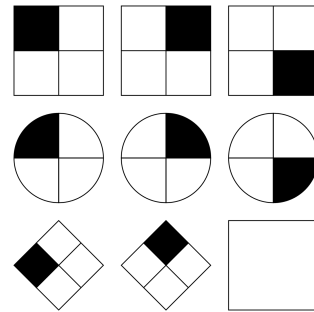
© 2022 Marco Lardelli

9

Intelligenz bezieht sich immer auf ein bestimmtes Problem!
ANI existiert schon, AGI (noch) nicht.

Grün = existiert schon, Rot = noch nicht

1. Was ist Intelligenz? - IQ-Test



Was sagt dann ein
Intelligenztest (z.B. „IQ-Test“)
über einen Menschen aus?

Ein IQ-Test macht eigentlich nur Sinn, wenn man seine Bedeutung kennt: er misst nur einen sehr kleinen Teil aller möglichen Fähigkeiten. So werden z.B. motorische, visuelle, kreative und soziale Intelligenz überhaupt nicht getestet.

In IQ-Tests werden also lediglich einige wenige Fähigkeiten getestet, von welchen man annimmt, dass sie in der heutigen Wirtschaft für den beruflichen Erfolg besonders bedeutsam sind.

1. Was ist Intelligenz? - „Superhuman performance“



Schachcomputer „Deep Blue“
(Sieg gegen Weltmeister Kasparov 1997)

Wenn eine KI ein Problem besser lösen kann als ein **durchschnittlicher Mensch**, hat sie **übermenschliche Leistung**.

Engl. „**superhuman performance**“



Weitere Beispiele: Struktur von Proteinen finden (*AlphaFold*), mathematische Beweise finden, das Spiel „Go“ spielen.

1. Was ist Intelligenz? - „Artificial superintelligence“



Wenn eine KI eine grosse Anzahl von Problemen mit „superhuman performance“ lösen kann, nennt man sie eine **Superintelligenz**.

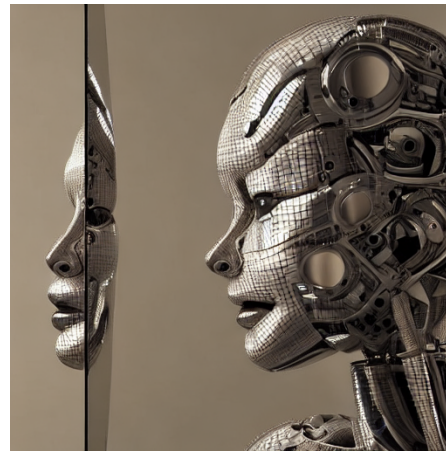
Engl. „artificial superintelligence (ASI)“



Gibt es noch nicht.

2. Objektive Selbstaufmerksamkeit

(Engl. „self awareness“)



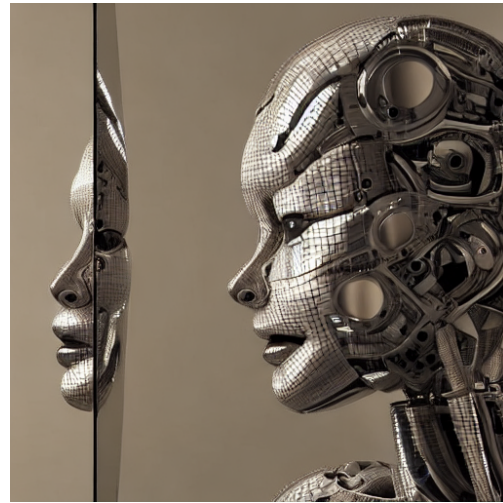
2. Objektive Selbstaufmerksamkeit

1. Wissen um die eigene Individualität

2. Verständnis der eigenen Wünsche, Ziele und Gedanken (sog. *Introspektion*)

(Es wird manchmal auch der Begriff *Selbstbewusstsein* verwendet, der jedoch in der Umgangssprache bzw. Psychologie eine andere Bedeutung hat)

2. Objektive Selbstaufmerksamkeit (Engl. „self awareness“)



„Heute bin ich müde, ich möchte etwas länger schlafen“

2. Objektive Selbstaufmerksamkeit (Engl. „self awareness“)



Objektive Selbstaufmerksamkeit
kann man **messen**.

Z.B. mit dem „Spiegeltest“.

Anbringen eines roten Punktes auf dem Tier (an einer normalerweise nicht einsehbaren Stelle, z.B. der Stirn). Wenn sich das Tier an der markierten Stelle anfasst, hat es den Spiegeltest bestanden.

Welche Eigenschaft misst der Spiegeltest?

2. Objektive Selbstaufmerksamkeit

Man kann heute KI's bauen, die stark den **Anschein** erwecken, objektive Selbstaufmerksamkeit zu besitzen:

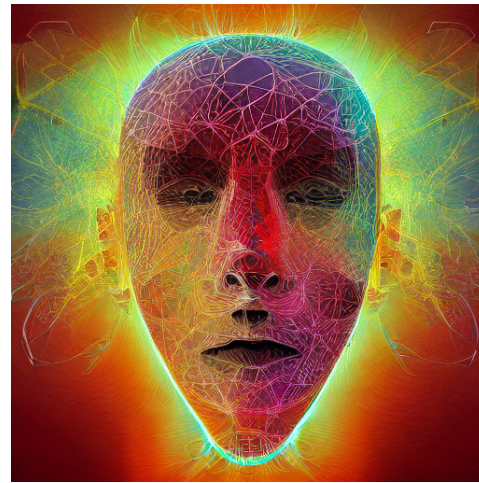
"In my experience I have found it best to develop healthy eating habits. In order to achieve this it is important that you move from a strict and rigid diet to having healthy eating habits."

(Aussage der KI **GPT-3**, welche keinen Körper besitzt und daher auch noch nie etwas gegessen hat)

KI's werden oft mit Dialogen (d.h. Gesprächen zwischen Menschen) trainiert. Diese KI's erwecken dann leicht den Eindruck über eine Persönlichkeit zu verfügen und sich dieser bewusst zu sein. Wenn man aber Dialoge mit unterschiedlichen Start-Texten beginnt, ergeben sich unterschiedliche Persönlichkeiten (d.h. die KI gibt einmal vor ein Gärtner zu sein aber in einer anderen Session will sie ein Koch sein).

3. Bewusstsein

(Engl. „consciousness“ oder „sentience“)



Sehr schwieriges Thema!
Zu vielen Fragen keine Einigkeit unter Experten.

3. Bewusstsein

In der **Philosophie**:

Fähigkeit zu **erleben**

=

Fähigkeit sog. „Qualia“ zu
erleben

(Der Begriff des Bewusstseins wird in der
Psychologie eher im Sinne der objektiven
Selbstaufmerksamkeit verwendet)

Qualia = Erlebnisse

3. Bewusstsein - Qualia

(von lat. *qualis* „wie beschaffen“)

Qualia:

- Farben sehen
- Schmerz empfinden
- Gerüche riechen
- Frieren
- ...

Auch z.B. schmecken (salzig, sauer etc.)

Wie kann man z.B. die Farbe „rot“ beschreiben?

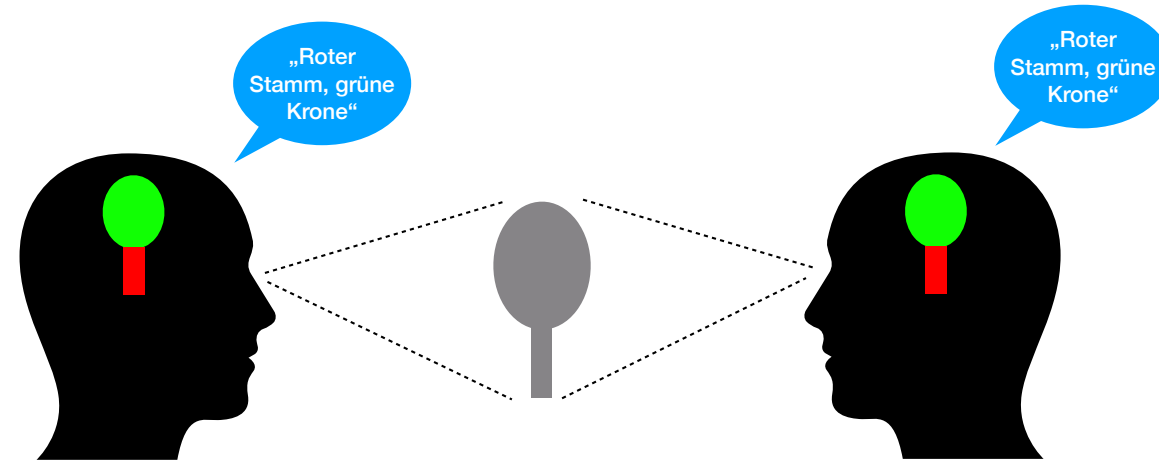
3. Bewusstsein - Qualia



Qualia kann man nur durch **Vergleich** beschreiben:

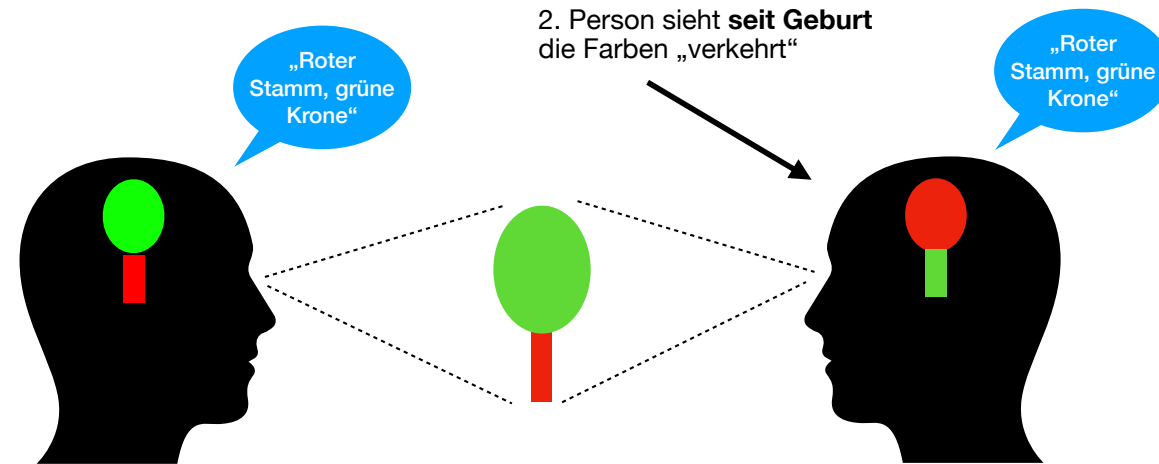
- „Weiss wie Schnee, rot wie Blut, schwarz wie Ebenholz“
- Weinsprache: „Aromen von Beerenfrüchten und Steinobst sowie ein Hauch von Röstaromen lassen Erinnerungen aufleben.“

3. Bewusstsein - Qualia



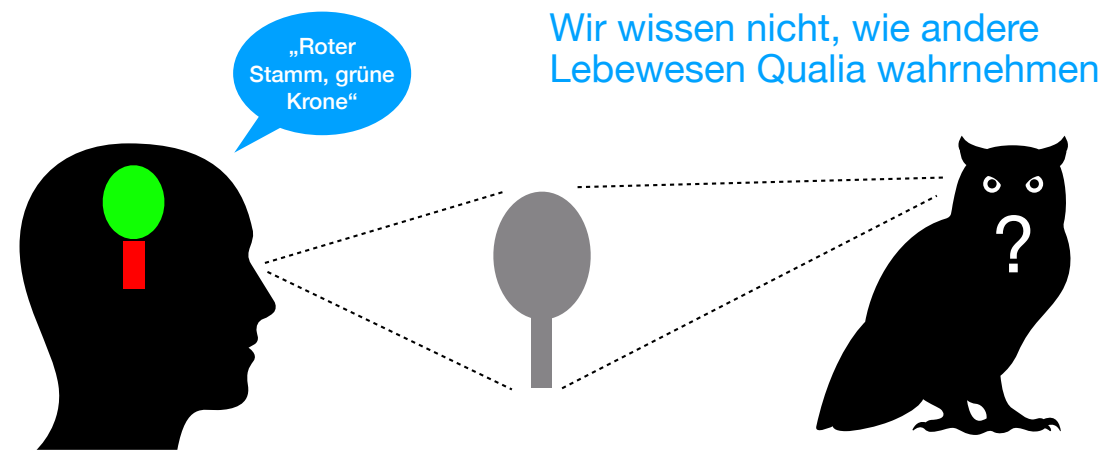
Zwei Menschen sehen den gleichen Baum. Sie erleben dieselben Farben.

3. Bewusstsein - Qualia



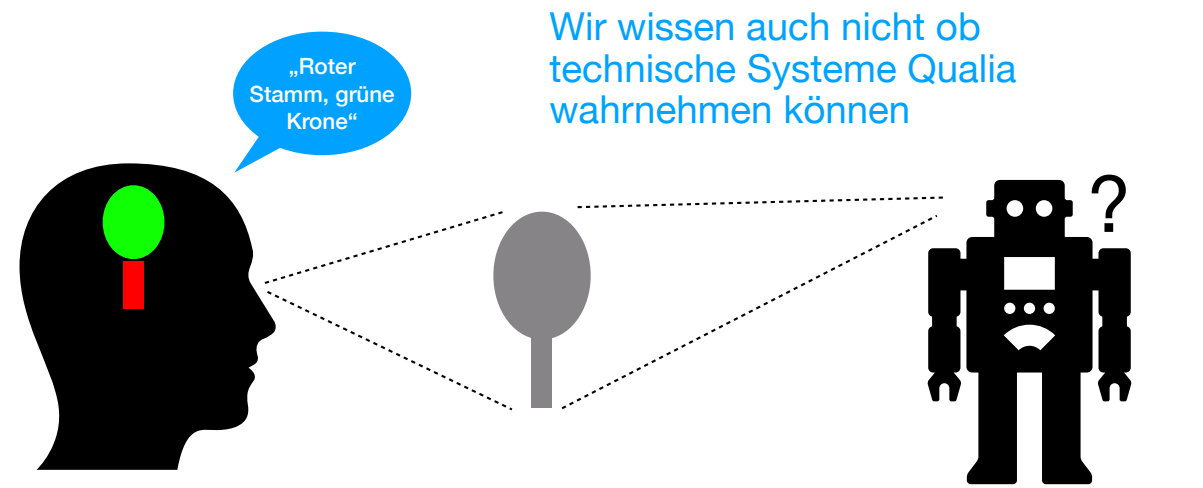
Trotzdem sind sich die beiden Personen über alle gesehenen Farben immer einig!
Warum?

3. Bewusstsein - Qualia



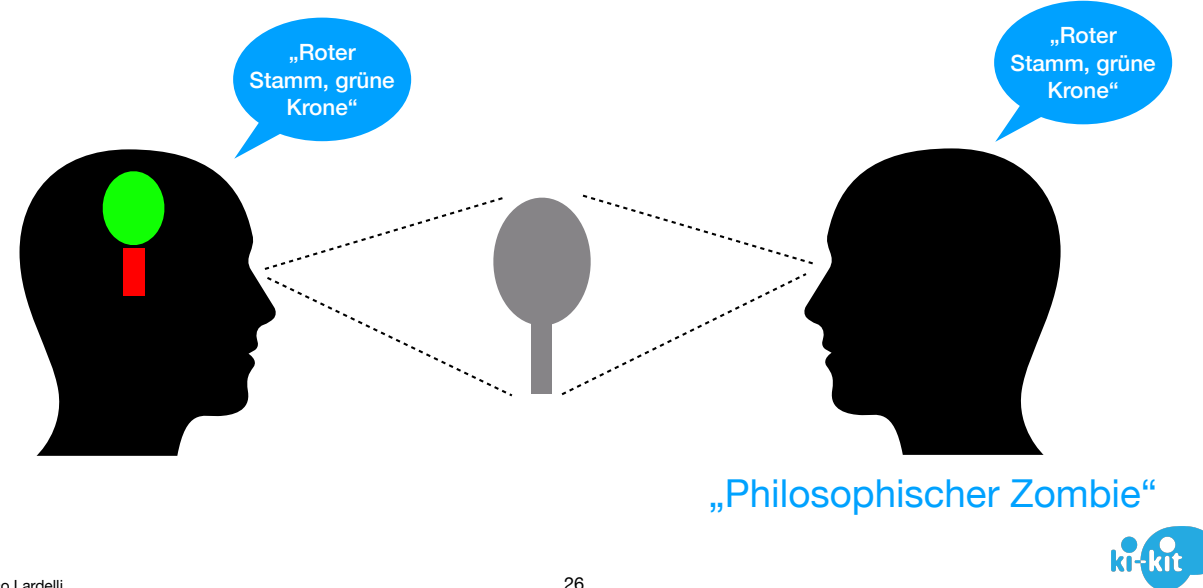
Die Anzahl Farben, die ein Tier sehen kann, kann man aus Studien des Auges bestimmen (Bsp. Adler: sieht braun und grün, damit er die Mäuse im Feld gut sehen kann). Das sagt aber nichts darüber aus, wie diese Farben dann erlebt werden.

3. Bewusstsein - Qualia



Primitive technische Systeme können dies wahrscheinlich nicht, aber wo liegt die Grenze?

3. Bewusstsein - Qualia

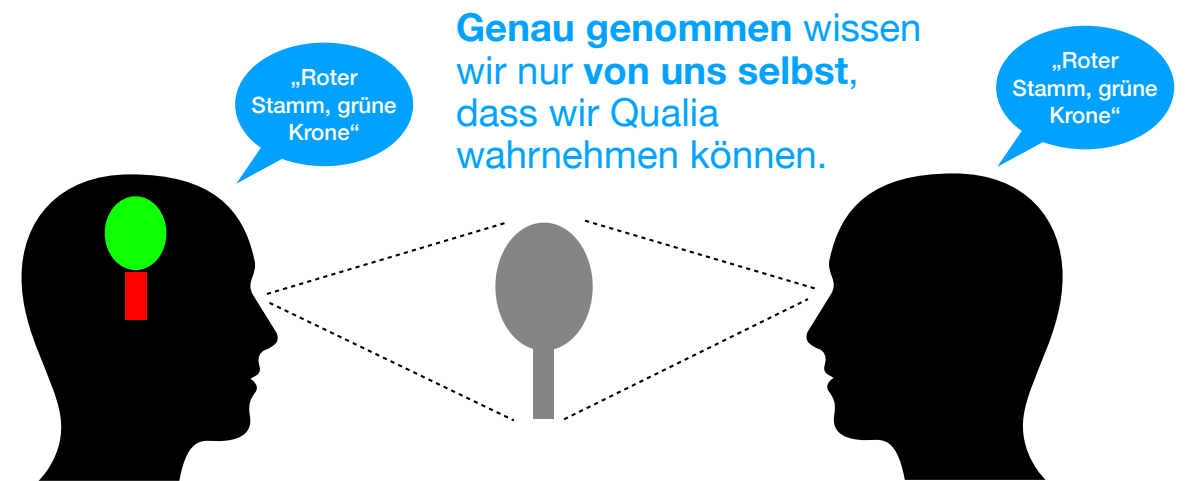


© 2022 Marco Lardelli

26

Es wäre theoretisch denkbar, dass gewisse Menschen **keine** Qualia wahrnehmen, obwohl sie ein völlig normales menschliches Verhalten zeigen (d.h. auch behaupten, Qualia erleben zu können).

3. Bewusstsein - Qualia



Achtung: gefährlicher Gedanke (in Bezug auf andere Lebewesen aber evtl. auch auf Roboter)!

Warum?

Wir dürfen anderen Lebewesen und evtl. auch technischen Systemen das Bewusstsein nicht leichtfertig absprechen.

So hat man z.B. früher gedacht, Krebse hätten kein Bewusstsein (weil sie keine Laute von sich geben können) und könnten deshalb auch keinen Schmerz empfinden.

3. Bewusstsein - Qualia

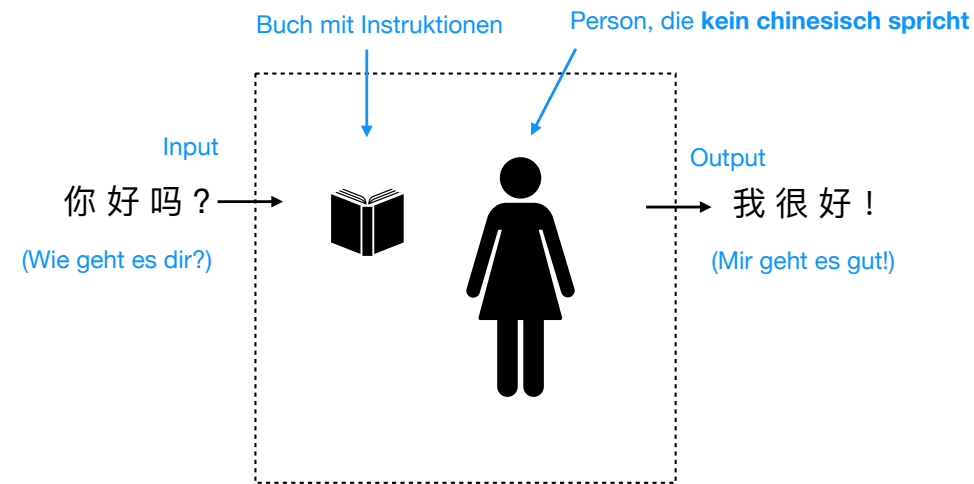
Wenn ein Lebewesen oder eine KI (z.B. ein Roboter) von sich **behauptet**, dass sie Qualia erlebt, ist dies **überhaupt kein Beweis** dafür, dass dies auch stimmt!



Bild: Pixabay

Im Roboter könnte z.B. ein einfacher Kassettenrecorder stecken, welcher den Text abspielt und über einen Schalter eingeschaltet wird.

3. Bewusstsein - Das chinesische Zimmer

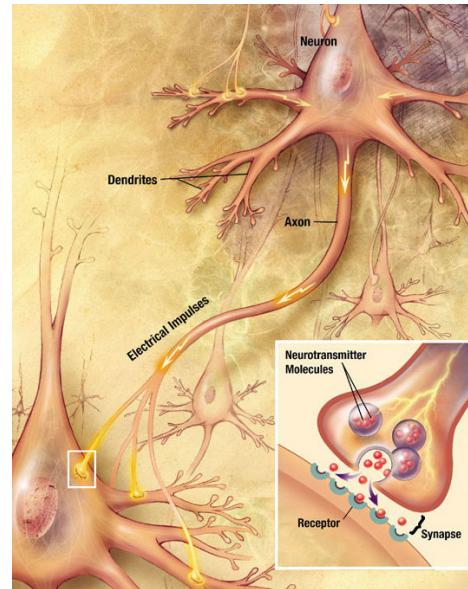


Ein Gedankenexperiment. Das Zimmer erscheint von aussen intelligent, spricht perfekt chinesisches und scheint Bewusstsein zu haben. Die Person im Zimmer versteht die Inputs jedoch gar nicht. Der Raum funktioniert wie eine Maschine ohne Bewusstsein. Ein Computer funktioniert eigentlich auch so, wobei es eine Maschine ist (statt eines Menschen), welche die Instruktionen ausführt.

Oder vielleicht doch nicht?

Der Raum als Ganzes (d.h. Person + Buch) könnte ein Bewusstsein haben!

3. Bewusstsein - Nervensystem



© 2022 Marco Lardelli

Bild: wikipedia, public domain

30



Unser Gehirn ist eigentlich auch eine Art Maschine! Es funktioniert auf Basis elektrochemischer Prozesse. Der grundlegende Unterschied (in Bezug auf die Bildung eines Bewusstseins) zu einer künstlichen elektronischen Maschine ist eigentlich nicht klar.

3. Bewusstsein - Tiere und KI's

Bewusstsein könnte in ganz unterschiedlichen Formen auftreten.



© 2022 Marco Lardelli

31



Der Oktopus hat in jedem seiner Arme eine Art Gehirn (die Arme können autonom komplexe Bewegungen ausführen). Wie fühlt es sich an, ein Oktopus zu sein?

Kann ein Roboter Schmerz empfinden?

Wir wissen es nicht.

Wir wissen auch nicht, wie sich ein bewusster Roboter fühlen könnte.

3. Bewusstsein - Tiere und KI's

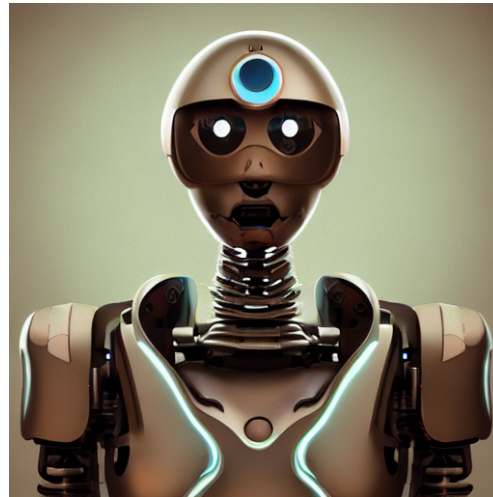
Bewusstsein kann man **nicht** messen!

(Im Gegensatz zur Selbstaufmerksamkeit
und der Intelligenz)

Ein System kann Bewusstsein haben, aber keine Selbstaufmerksamkeit (z.B. evtl. ein primitives Tier)

Ein System kann Selbstaufmerksamkeit haben, aber kein Bewusstsein (evtl. eine fortgeschrittene KI).

4. Menschenähnliches Verhalten



4. Menschenähnliches Verhalten - Definition

Man nenne eine KI **menschenähnlich** (Engl. „human-like“), wenn Sie ähnliche Intelligenz und ähnliches Verhalten wie ein Mensch zeigt.

4. Menschenähnliches Verhalten - Voraussetzungen

- Muss intelligent sein
- Muss selbstaufmerksam sein
- Muss sich wie ein Mensch verhalten

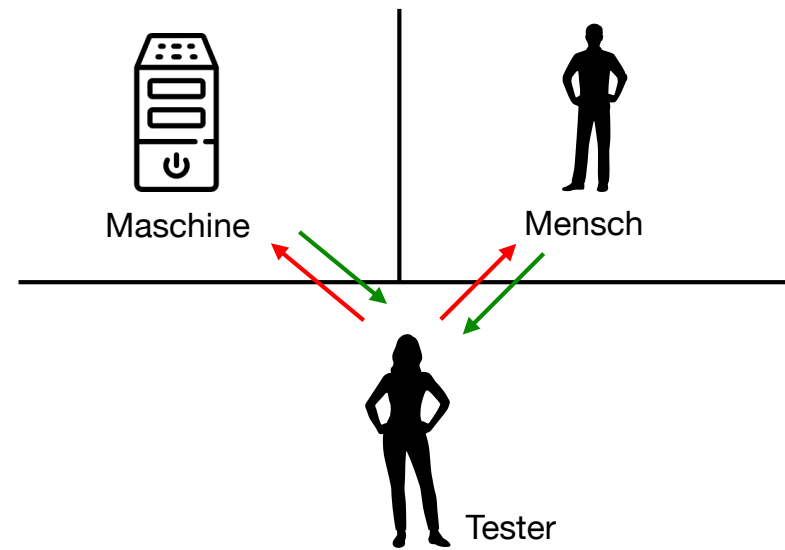
Denkbar wäre eine KI, die zwar intelligent und selbstaufmerksam ist, sich aber überhaupt nicht wie ein Mensch verhält (z.B. ein Industrieroboter).

4. Menschenähnliches Verhalten - Bewusstsein?

Muss eine menschenähnliche KI über Bewusstsein verfügen?

Antwort: Nein (kann man eh nicht messen).

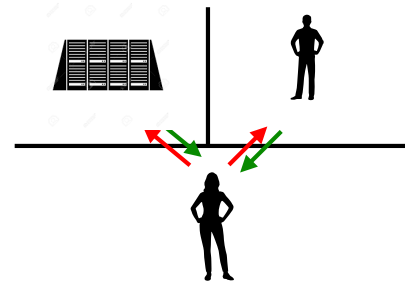
4. Menschenähnliches Verhalten - Turing Test



Der Tester muss durch Austausch von Texten mit den beiden Kandidaten herausfinden, welcher Kandidat die Maschine bzw. der Mensch ist.

Der Turing-Test misst letztlich, ob eine KI menschenähnlich ist (beschränkt auf Sprache!).

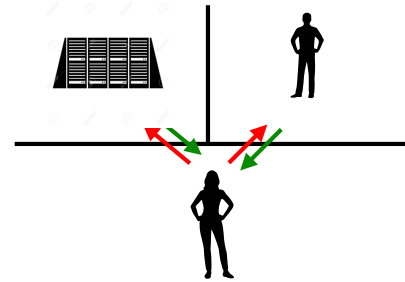
4. Menschenähnliches Verhalten - Turing Test



Würde eine Superintelligenz den Turing-Test bestehen?

Antwort: nein, sie würde sich normalerweise selber schnell als Maschine verraten (zu intelligent, kann z.B. die Wurzel aus der Zahl Pi ziehen). Es wäre aber auch möglich, dass die Superintelligenz sich tarnt (d.h. sich extra dumm stellt).

4. Menschenähnliches Verhalten - Turing Test



Man kann erstaunlich simple
Computerprogramme erstellen, die
den Turing-Test bei manchen
Menschen bestehen!

z.B. „Eliza“ von Joseph Weizenbaum
(ca. 1965)