

Transkript zum Podcast

## **Wir müssen auf Weitsicht fahren und fragen: Was wäre wenn?**

Frauke Rostalski im Gespräch mit  
Karsten Wendland



*Zur Podcast-Folge*

Podcast-Reihe „Selbstbewusste KI“  
Folge 6

Erstveröffentlichung: 20.10.2020

Forschungsprojekt:

Abklärung des Verdachts aufsteigenden Bewusstseins in der  
Künstlichen Intelligenz – KI-Bewusstsein

[www.ki-bewusstsein.de](http://www.ki-bewusstsein.de)

Projektleitung:

Prof. Dr. Karsten Wendland  
Institut für Technikfolgenabschätzung und  
Systemanalyse (ITAS)

Förderkennzeichen: 2016ITA202

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung

**Herausgeber:**

Karsten Wendland, Nadine Lahn, Pascal Vetter

**Empfohlene Zitationsweise/Suggested citation:**

Wendland, K., Lahn, N. & Vetter, P. (Hg.) (2021). Wir müssen auf Weitsicht fahren und fragen: Was wäre wenn?. Frauke Rostalski im Gespräch mit Karsten Wendland. Karlsruhe: KITopen.

<https://doi.org/10.5445/IR/1000139825>

**Hinweis zum Copyright:**

Lizenz: Namensnennung 4.0 International (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0/deed.de>

**Copyright notice:**

License: Attribution 4.0 International (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0/deed.en>

## Inhalt

1	Zum Projekt „KI-Bewusstsein“ .....	4
2	Podcast-Serie „Selbstbewusste KI“ .....	5
3	Bibliometrische Angaben zur Folge .....	6
4	Transkription des Gesprächsverlaufs .....	7
5	Erwähnte Quellen .....	28
6	Kontakt .....	29

# 1 Zum Projekt „KI-Bewusstsein“



Zum Projekt

Im Projekt „Abklärung des Verdachts aufsteigenden Bewusstseins in der Künstlichen Intelligenz (KI-Bewusstsein)“ am Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) des Karlsruher Instituts für Technologie (KIT) untersuchen und kartieren wir, welche Gruppen wissenschaftlich, wirtschaftlich und weltanschaulich zu „aufsteigendem Bewusstsein“ in der KI arbeiten. Wir fragen danach, welche Motive, Intentionen und Verankerungen jeweils dahinterstecken und welche Zukunftsszenarien angedacht sind oder auch in Zweifel gezogen werden. Dabei klären wir technische Entwürfe ab und hinterfragen Mythen und Narrative, die in die Welt gesetzt werden und bestimmte Zuschreibungen auslösen.

Die Idee einer „erwachenden“, sich selbst bewusst werdenden Künstlichen Intelligenz hat in den vergangenen Jahren zunehmend Popularität erfahren, u.a. durch verbrauchernahe KI-gestützte Systeme wie *Siri* von Apple und den auf *Alexa* hörenden Smart Speaker, den eingebürgerten Roboter *Sophia* und auch IBMs dialogfähigen *Watson*. Renommiertere KI-Akteure melden sich begeistert, mahnend oder warnend zu Wort und stellen die Entwicklung einer „Superintelligenz“ in Aussicht, die alles planetarisch Dagewesene in den Schatten stellen und den Menschen in seinen Fähigkeiten überholen werde.

In der KI-Community ist die Fragestellung zum sogenannten „maschinellen Bewusstsein“ zwar bekannt, aber kaum systematisch erforscht – das mystisch aufgeladene Nischenthema geht mit dem Risiko einher, sich einen wissenschaftlichen Reputationsschaden einzuhandeln. Gleichzeitig nähren KI-Forschung und -Marketing genau diese Mystik, indem sie vermenschlichende Sprachbilder verwenden, die ein aufkeimendes Bewusstsein verheißen, wenn etwa der Roboter „denkt“ oder „fühlt“, das autonome Fahrzeug mit einer „environment perception“ ausgestattet ist oder das Smart Home „weiß“, wie es seinen Bewohnern „helfen“ kann. Hierdurch werden Redeweisen und Narrative aufgebaut, die in der (medialen) Öffentlichkeit Vorstellungen zu einer „bewussten KI“ verbreiten, ohne dass hierzu wissenschaftlich belastbare Aussagen geliefert würden. Auch der transdisziplinäre Dialog zur Frage, was am sogenannten maschinellen Bewusstsein „dran“ sein könnte, ist bislang kaum vorhanden.

An diesem Defizit setzt das Projekt an mit dem Ziel, eine Abklärung zum Verdacht aufsteigenden Bewusstseins in der KI herbeizuführen, indem bestehende Diskurse analysiert, empirisch untersucht, einschlägige Akteure fächerübergreifend zusammengeführt, offene Fragen identifiziert und bearbeitet werden, ein gemeinsames, transdisziplinär tragfähiges Verständnis erarbeitet wird und die Ergebnisse in den öffentlichen Diskurs eingebracht werden.

„KI-Bewusstsein“ im Netz:  
Web: [www.ki-bewusstsein.de](http://www.ki-bewusstsein.de)  
Twitter: [@KIBewusstsein](https://twitter.com/KIBewusstsein)

Institut für Technikfolgenabschätzung  
und Systemanalyse (ITAS):  
<http://www.itas.kit.edu/>

## 2 Podcast-Serie „Selbstbewusste KI“



Zur Podcast-  
Serie

Kann Künstliche Intelligenz ein Bewusstsein entwickeln? Wie könnte das überhaupt funktionieren, und was würde das für uns bedeuten? 12 Folgen, 12 Gespräche mit Expertinnen und Experten und eine gemeinsame Abschlussrunde.

Folge	Titel	Gast
01	Ohne Leben kein Bewusstsein (01:10:29)	Thomas Fuchs
02	Roboter bekommen eine menschliche Aura (59:37)	Andreas Bischof
03	In der heutigen KI ist kein Geist (57:45)	Ralf Otte
04	Die Gründerväter der KI machten sich über Bewusstsein keine Gedanken (01:03:36)	Christian Vater
05	Von Kühlschränklichtern, KI-Pubertät und Turnschuhen (01:08:58)	Thomas Metzinger
06	Wir müssen auf Weitsicht fahren und fragen: Was wäre wenn? (41:31)	Frauke Rostalski
07	Bewusstsein ist eine kausale Kraft und kein cleverer Programmier-Hack (55:24)	Christof Koch
08	Wir müssen Maschinen bauen, die Gefühle haben (39:04)	Antonio Chella
09	Dass Roboter uns Emotionen vorgaukeln, kann sehr wichtig sein (45:06)	Janina Loh
10	Die größte Hoffnung wäre, die Dystopien zu verhindern (52:46)	Joachim Weinhardt
11	Die meisten SF-Romane sind als Warnung gedacht, nicht als Gebrauchsanleitung (55:14)	Andreas Eschbach
12	Roboter werden bald ein Bewusstsein besitzen (34:55)	Junichi Takeno
Bonus	Das große Staffelfinale – Diskussionsrunde zu bewusster KI (01:41:08)	Gesprächsrunde

### Verfügbarkeit der Audiodaten



Zu KITopen

KITopen: <https://publikationen.bibliothek.kit.edu/serie/649>  
Anchor.fm: <https://anchor.fm/kibewusstsein>  
Spotify: <https://open.spotify.com/show/4BzUdFgR6o74H5sS2ait9z>  
Apple Podcasts: <https://podcasts.apple.com/de/podcast/selbstbewusste-ki/id1530833724>

### 3 Bibliometrische Angaben zur Folge

#### Teasertext

Für die Juristin Prof. Dr. Dr. Frauke Rostalski, Universität Köln und Mitglied des Deutschen Ethikrats, kann man KI-Robotern weder Schuld noch Verantwortung zuschreiben – zumindest nicht nach den heutigen Maßstäben und technischen Realisierungen. KI-Systeme können nur Teile unseres Denkens abbilden, und umgekehrt funktioniert der Mensch nicht nach Algorithmen. Allerdings wissen wir nicht, was noch kommt, und sollten verschiedene Zukünfte rechtzeitig vorausdenken. Zukünftige Gesellschaften etwa mit e-Personen sind schon heute debattierte Szenarien.

#### Metadaten

Titel: Wir müssen auf Weitsicht fahren und fragen: Was wäre wenn?

Dauer: 41:31

Erstveröffentlichung: 20.10.2020

Autor: Karsten Wendland

Gesprächsgast: Frauke Rostalski

Fragensteller: Gerard Blommestijn

Redaktion und

Produktion: Kayla Zoller

Aufnahmeleitung: Karsten Wendland

DOI Audiofile: <https://doi.org/10.5445/IR/1000124841>

DOI Transkription: <https://doi.org/10.5445/IR/1000139825>

#### Folgenlogo





Zur Podcast-Folge

## 4 Transkription des Gesprächsverlaufs

**Karsten Wendland:** Hallo und herzlich willkommen bei selbstbewusste KI, dem Forschungspodcast an der Grenze zwischen Mensch und Maschine. [00:00:09]

**Frauke Rostalski:** Und wenn wir jetzt zum Beispiel den autonomen Roboter nehmen und der verletzt jemanden, dann ist doch der natürliche Impuls, verantwortlich dafür ist doch nicht dieser Roboter, sondern der Mensch, der dahintersteht. [00:00:19]

**Karsten Wendland:** Mein Name ist Karsten Wendland. [00:00:22] Ich bin Forscher am Karlsruher Institut für Technologie und gehe Fragen nach, wie Technik, die gerade erst noch erforscht wird, morgen vielleicht schon unseren Alltag prägen könnte. [00:00:32] Kann Künstliche Intelligenz ein Bewusstsein entwickeln? [00:00:37] Wie könnte das überhaupt funktionieren und was würde das für uns bedeuten? [00:00:43] Mein heutiger Gast zu dieser spannenden Frage ist eine Rechtswissenschaftlerin und Philosophin, die sich im Grenzbereich zwischen Strafrecht, Medizin und Ethik einen Namen gemacht hat und mit Biotechnologie und Künstlicher Intelligenz vertraut ist. [00:00:58] Herzlich willkommen Frau Prof. Dr. Dr. Frauke Rostalski. [00:01:01]

**Frauke Rostalski:** Vielen Dank. [00:01:02]

**Karsten Wendland:** Hallo, schön, dass wir heute miteinander sprechen. [00:01:04] Sie sind heute mein Gast, aber genau genommen bin ich Gast bei Ihnen hier im schönen Köln, in einer Universität, die fast leer ist momentan. [00:01:12] Das liegt an den Rahmenbedingungen unserer Corona-Zeit. [00:01:16] Bevor wir starten, möchte ich Ihnen aber erstmal gratulieren, Sie sind frisch berufenes Mitglied im Deutschen Ethikrat [[Quellenverweis 1](#)]. [00:01:26]

**Frauke Rostalski:** Vielen Dank! [00:01:27]

**Karsten Wendland:** Frau Rostalski, wie konnte denn das passieren, haben Sie nicht ohnehin genug zu tun? [00:01:31]

**Frauke Rostalski:** Ach ja, also, für so etwas hat man immer noch Zeit. (**Karsten Wendland:** Okay!) [00:01:34] Das ist ja eine ganz ehrenvolle Berufung, ich freue mich sehr darüber. [00:01:39] Ich gehe davon aus, dass es dazu

gekommen ist aufgrund der besonderen Kombination meiner Forschungsschwerpunkte. [00:01:46] Also, wie Sie es schon gesagt haben, sowohl Medizinrecht, Medizinethik als auch eben Rechtsphilosophie, Strafrecht und dann eben natürlich dieser gesamte Bereich der Digitalisierung, der mich jetzt schon länger beschäftigt. [00:01:59] Und diese Kombination ist, glaube ich, dann doch noch relativ selten. [00:02:02]

**Karsten Wendland:** Ja, digitale Ethik in der Informationsgesellschaft habe ich gelesen als ein Themenstrang und der vertrauenswürdige Einsatz von Künstlicher Intelligenz und anderen digitalen Innovationen. (**Frauke Rostalski:** Genau.) [00:02:14] Wie machen Sie das mit der KI, sprechen Sie manchmal mit Ihrem Computer? [00:02:18]

**Frauke Rostalski:** (Lacht) Nein, ich fluche auch nicht mit meinem Computer. [00:02:32]

**Karsten Wendland:** Sie fluchen auch nicht, okay, geben ihm auch keinen Namen oder so? (**Frauke Rostalski:** Nein) [00:02:25] Okay. [00:02:26] Ich habe Ihnen etwas mitgebracht, das ist ein Foto. [00:02:29] Hier sehen wir einen chinesischen Mann, das ist ein Ingenieur, der seine Braut über die Schwelle trägt. [00:02:37] Da sieht man, die hat also eine rote Kopfbedeckung, ist traditionell gekleidet und jetzt ist die Besonderheit, diese Braut ist aber kein Mensch, sondern ein Roboter, ein KI-System, das dieser Ingenieur selbst gebaut hat. [00:02:52] Also es wurde getitelt in der Presse, Ingenieur heiratet, in Anführungszeichen, seine selbstgebaute Robo-Freundin [[Quellenverweis 2](#)]. [00:03:00] Frau Rostalski, wo verläuft die Grenze zwischen Mensch und Maschine? [00:03:04]

**Frauke Rostalski:** Jetzt im Kontext der Ehe vielleicht zunächst einmal, würde ich das, was wir hier sehen, doch in gewissem Umfang kritisch sehen, weil, wir müssen uns ja immer fragen, was hat denn die Ehe so für eine Funktion? [00:03:18] Sicherlich mindestens zwei, also zum einen als Symbol, als Symbol von besonderer Zuneigung, von Liebe zwischen zwei Personen, sage ich einmal. [00:03:28] Und auf der anderen Seite hat es ja eben auch schon immer einen gewissen Fürsorgeaspekt, ich habe besondere Pflichten gegenüber einer Person, die ich heirate. [00:03:38] Also, ich muss sie zum Beispiel besonders beschützen, aber ich muss auch finanziell in Krisensituationen für



den Betreffenden aufkommen. [00:03:47] Und wenn wir jetzt mal diesen zweiten Aspekt nehmen, Fürsorge, sowohl im Hinblick auf die Person als auch im Hinblick auf deren finanzielle Situation, sehe ich das hier schon sehr kritisch, denn der Roboter kann diese Funktion ja selbst nicht erfüllen, ihn kann man nicht verpflichten zu irgendetwas, er ist also insofern kein Adressat von unseren Pflichten, würden wir im Recht sagen. [00:04:11] Insofern ist schon mal einer von beiden hier ein untauglicher Partner in der Ehe. [00:04:15]

**Karsten Wendland:** Es ist nicht auf Augenhöhe? [00:04:17]

**Frauke Rostalski:** Es ist nicht auf Augenhöhe. [00:04:18]

**Karsten Wendland:** Das würden Sie auch sagen als Rechtswissenschaftlerin? [00:04:19]

**Frauke Rostalski:** Das würde ich so sagen, genau. [00:04:20]

**Karsten Wendland:** Okay, ja. [00:04:20]

**Frauke Rostalski:** Also sie kann nicht, sie in Anführungsstrichen, kann hier nicht dieselben Pflichten erfüllen wie er. [00:04:25]

**Karsten Wendland:** Ja, okay. [00:04:25]

**Frauke Rostalski:** Und was das Symbol der Liebe angeht, stelle ich auch mal ein großes Fragezeichen dahinter. [00:04:30] Ich kenne die Debatte und weiß, dass offenbar auch festgestellt wurde, dass ähnlich starke Emotionen eines Menschen zu einer sonstigen Entität aufgebaut werden können. [00:04:39] Ich frage mich aber, ob wir das als Gesellschaft mit dem Etikett der Ehe tatsächlich versehen wollen. [00:04:45] Ich wäre da kritisch. [00:04:46]

**Karsten Wendland:** Was wir jetzt gerade besprechen, ist jetzt auch eine europäische Haltung, die wir hier noch haben, auch mit traditionellen Rollenmodellen, die sich jetzt auch weiterentwickeln. [00:04:55] In Asien, wir haben jetzt hier das Beispiel von China, wird diese Art der Ehe ja auch noch nicht vom Staat anerkannt, gesellschaftlich aber schon, ja. [00:05:06] Es gibt auch Beispiele aus Japan, da hat jemand eine Manga-Puppe geheiratet und auch Hochzeiten, bei denen dann die Trauung von einem Roboter durchgeführt wurde, wo man sich auch fragen könnte, ist das denn dann eigentlich gültig oder nicht, ja. [00:05:23] Also die Gesellschaft geht da ein bisschen anders

damit um, ja, und wir verlangen doch noch schon den Menschen und machen diesen kleinen Unterschied, ja. [00:05:31] Und wir wollen herausfinden, wo sitzt dieser Unterschied eigentlich. [00:05:34] Ein Kollege von uns, der viele Jahre in Oxford war, Evolutionsbiologe Richard Dawkins, ist bekannt für ein Zitat, in dem er diesen kleinen Unterschied eigentlich auflöst. [00:05:47] Er sagt, Menschen sind nur Roboter, blind und programmiert zur Erhaltung der Gene [[Quellenverweis 3](#)]. [00:05:54] Und Daniel Dennett aus den USA, der ist schon in den 80ern angekommen, der hat jetzt natürlich auch damit zu tun, dass der Körper mit 80 nicht mehr so ist wie mit 25, ja, hat jetzt zwei neue Knie bekommen, es geht ihm gut und er sagt, I am a Bio-Robot. [00:06:15] Also er stellt sich selbst als Roboter dar und fängt in der Argumentation an mit den vielen Milliarden Zellen, die irgendwie im Zusammenspiel schließlich ihn als Person ausmachen. [00:06:27] Aber alles, was es dann noch zusätzlich geben könnte, was den Unterschied ausmacht, wo wir teilweise nach suchen, was ist denn das, haben wir so etwas wie eine Seele beispielsweise, ja, wird an dieser Stelle verneint. [00:06:39] Das ist so eine Art neuer Atheismus und wie ist denn das, dieser Zusammenhang in unserem Rechtssystem verankert? [00:06:47] Wenn wir aus Europa oder speziell aus Deutschland drauf schauen, sind wir Bioroboter oder sind wir nur zur Weitergabe der Gene da, was sagt der Jurist oder die Juristin dazu? [00:06:58]

**Frauke Rostalski:** Also bei uns Juristen gibt es auch schon ganz lange die Debatte, auf die ja auch die beiden zitierten Personen hinweisen, nach der Möglichkeit von Willensfreiheit, von Entscheidungsfreiheit. [00:07:10] Und ich glaube ganz genau läuft-, hier läuft auch wieder die Diskussion, sind wir wirklich nur Algorithmen, funktioniert unser Gehirn nach einem Algorithmus. [00:07:19] Ich bestreite das, wir haben lange diskutiert, beispielsweise über Hirnforschungsexperimente, das Libet-Experiment ist Ihnen sicherlich auch ein Begriff, in denen vermeintlich solche- [[Quellenverweis 4](#)]. [00:07:30]

**Karsten Wendland:** Können Sie das kurz erläutern? [00:07:31]

**Frauke Rostalski:** (Lachen) Genau, das kann ich gerne tun. [00:07:33] Also in aller Kürze, das Ergebnis, das festgestellt hat, dass das Gehirn funkt, bevor der Mensch meint, die Entscheidung getroffen zu haben und daraus dann eben die Schlussfolgerung, die der ein oder andere, vor allen Dingen Neurowissenschaftler, gezogen hat, das Gehirn der eigentliche Herr im Hause ist, also das so als Schlagwort. [00:07:55] Und ich sehe das sehr kritisch nicht

nur im Hinblick auf das Experiment, über das man sicherlich auch in der Interpretation in naturwissenschaftlicher Hinsicht streiten könnte, sondern eben auch, ob wir das tatsächlich als Aussagegehalt anerkennen wollen. [00:08:06] Also so ein Signal, dass der Mensch sich einbildet, nicht schon vorher die Entscheidung getroffen zu haben, dass ein vorheriger Impuls im Gehirn da war, erschüttert das denn wirklich unser Konzept von Willensfreiheit. [00:08:19] Ich glaube nämlich, das Entscheidende ist doch, dass wir uns in gewisser Weise auf Vorstellungen voneinander einigen. [00:08:26] Also wir reden ja heute viel über Begriffe, über den Begriff des Roboters, über den Begriff des Menschen und wir müssen, und das zeigt auch das schöne Ehebeispiel, das Sie eben genannt haben, uns immer fragen, wie funktioniert unsere Gesellschaft. [00:08:38] Nun die funktioniert nach den Narrativen, die wir uns auferlegen. [00:08:41] Und wenn wir uns eben das Narrativ auferlegen, dass wir von unserem Willen selbst-, also unseren Willen selbst beherrschen, ich also frei entscheiden kann, ob ich Sie heute zu mir ins Büro hineingebeten habe oder nicht und ob das nun eine Illusion ist, die mir meine Biochemie aufzwingt oder nicht, letztlich entscheidet doch, wie wir uns als Gesellschaft definieren. [00:09:03] Und da, glaube ich, gibt es einen ganz großen Konsens, dass die Menschen eben sagen, wir sind selbstgesteuert und es ist nicht das Gehirn oder irgendein anderer Algorithmus, der mir irgendwas vorgibt. [00:09:13] Also nicht determiniert, sondern Entscheidungsfreiheit. [00:09:16] Und deswegen muss ich den Kollegen, so namhaft sie sind, dann doch widersprechen und sagen, nein, wir sind keine Roboter, außer sie wollen einen Begriff des Roboters zugrunde legen, der letztlich eine sprachliche Verwirrung irgendwann aufwerfen würde. [00:09:28]

**Karsten Wendland:** Jetzt kann man ja beobachten, dass manche Menschen auf Roboter oder auf Systeme, die irgendwie funktionieren, weil sie algorithmisch organisiert sind, auch teilweise mit Aggressivität reagieren. [00:09:40] Ja, es werden Roboter geschlagen, es werden Maschinen zertrümmert. [00:09:44] Wäre das dann aus Ihrer Sicht noch Sachbeschädigung, oder wäre das möglicherweise etwas mehr? [00:09:51]

**Frauke Rostalski:** Also für den Augenblick ist das Sachbeschädigung, wobei ich natürlich als Juristin immer dazu gezwungen bin zu sagen, es kommt darauf an, denn ich kann mir durchaus Situationen vorstellen, in denen das eine

Körperverletzung ist. [00:10:03] Denn, stellen Sie sich vor, das ist mein Klassiker aus der Vorlesung, Sie sitzen in Ihrer Mietwohnung, Ihr Vermieter möchte zum Beispiel irgendwie erreichen, dass Sie eine höhere Miete bezahlen, Sie sagen nein. [00:10:13] Und dann sagt er, alles klar, ich schalte die Heizung mal aus oder noch besser, ich zertrümmere sie. [00:10:19] Dann kann das eine Körperverletzung sein. [00:10:20] Die Gewalt, die gegen die Sache ausgeübt worden ist, die sich dann eben auf Ihre Körperintegrität auswirkt. [00:10:25] Aber das ist natürlich nicht das, wonach Sie gefragt haben. [00:10:27]

**Karsten Wendland:** Ja, aber das wäre dann tatsächlich die Körperverletzung bei der-? [00:10:30]

**Frauke Rostalski:** Beim Menschen. [00:10:31]

**Karsten Wendland:** Beim Menschen. [00:10:31]

**Frauke Rostalski:** Genau, richtig. [00:10:32]

**Karsten Wendland:** Ist es so, vielleicht können Sie dazu auch eine Einschätzung geben, wenn es um psychische Verletzungen geht, dass dem auch immer eine physische Verletzung vorausgegangen sein muss? [00:10:45]

**Frauke Rostalski:** Das ist nicht zwingend. [00:10:46]

**Karsten Wendland:** Nicht der Fall? [00:10:46]

**Frauke Rostalski:** Das ist nicht zwingend. [00:10:48] Dazu haben wir glücklicherweise mal was ganz handfestes, nämlich Rechtsprechung des Bundesgerichtshofes, das ist ja immer schön. [00:10:54] Und zwar ist das nicht zwingend, aber wir brauchen einen bestimmten Grad an psychischer Beeinträchtigung. [00:10:59] Also nicht das bloße Unwohlsein genügt, um erheblich genug zu sein für eine Körperverletzung, aber zum Beispiel, wenn Sie Depressionen haben, auch schon Schlaflosigkeit genügt da schon für den Bundesgerichtshof zu sagen, das ist eine Körperverletzung. [00:11:12] Das halte ich auch für richtig. [00:11:14]

**Karsten Wendland:** Okay. [00:11:15] Also sind wir da über die reine physische Verletzung als Auslöser für Schmerzensgeldzahlungen beispielsweise hinweg? [00:11:22]

**Frauke Rostalski:** Hinweg. [00:11:22] Es genügt die psychische Beeinträchtigung, wenn sie eine gewisse Erheblichkeitsschwelle überschritten hat. [00:11:26]

**Karsten Wendland:** Okay. [00:11:27] Was ist denn dran an der Diskussion, Robotern einen eigenen Rechtsstatus zuzuweisen, worum geht es denn da? [00:11:34] Da ist mein Hintergedanke, vielleicht geht es ja da tatsächlich zum Schluss nur ums Geld, dass man Robotern, die in Industriebetrieben arbeiten, einen Rechtsstatus zuweist, damit man sie auch mit einer Steuernummer versehen kann? [00:11:47] Geht es in diese Richtung? [00:11:48]

**Frauke Rostalski:** Ja, Sie müssen da auf jeden Fall verschiedene Rechtsgebiete differenzieren. [00:11:52] Sie haben ganz recht, wenn es um das Zivilrecht geht. [00:11:54] Da wird ja zum Beispiel darüber diskutiert, so eine Art E-Person einzuführen und die dann etwa mit einem Fond auszustatten, also beispielsweise bei autonom fahrenden Kraftfahrzeugen, dass man schon das autonom fahrende Kraftfahrzeug als so eine E-Person etwa einordnet, wenn dahinter halt finanzielle Mittel stehen. [00:12:13] Das machen wir ja schon ewig im Recht. [00:12:14] Also Unternehmen sind nichts anderes als Zusammenschlüsse von natürlichen Personen, die aber eben für sich genommen nur eine juristische Person sind, und das ergibt guten Sinn, weil ja Vermögenswerte mit diesen Unternehmen verbunden sind. [00:12:27] Anders im Strafrecht, das muss ich natürlich als Strafrechtlerin dann betonen, da ist das nicht so einfach. [00:12:32] Im Strafrecht fragen wir uns ja immer, ist jemand verantwortlich? [00:12:35] Also, zum Beispiel, der eine schlägt den anderen, ist er dafür verantwortlich? [00:12:38] Zum Beispiel ist er nicht verantwortlich, wenn er in einem Zustand der geistigen Umnachtung gehandelt hat, schuldunfähig war. [00:12:45] Und jetzt die Frage, wer ist verantwortlich, wenn etwa der autonome Roboter einen Menschen schädigt. (**Karsten Wendland:** Genau.) [00:12:53] Und darum diskutieren wir gerade ganz heiß im Recht. [00:12:56] Und das ist eben nochmal was Anderes, da geht es nicht nur um Finanzielles, sondern da geht es tatsächlich um das, was wir als Verantwortung oder auch Schuld bezeichnen, genau. [00:13:05] Und ich vertrete da derzeit die Position, ich sage derzeit, weil ich auf der Basis des jetzigen wissenschaftlichen Fortschritts argumentiere, dass wir Verantwortung und Schuld noch keinen künstlichen Entitäten wie Robotern zuschreiben können. [00:13:20] Und das ist letztlich eine rechtsphilosophische Begründung, die da

dahinter steht. [00:13:24] Verantwortung ist ja etwas, was wir einander zuschreiben und wenn wir jetzt zum Beispiel den autonomen Roboter nehmen und der verletzt jemanden, dann ist doch der natürliche Impuls der großen Gruppe von Menschen zu sagen, verantwortlich dafür ist doch nicht dieser Roboter, sondern der Mensch, der dahintersteht. [00:13:41] Also wieder dieser Zuschreibungsprozess, wie haben wir uns als Gesellschaft aufgestellt. [00:13:46] Und solange das so ist, so lange ist auch keine Verantwortung zu konstruieren. [00:13:51] Und Sie sehen, so wie ich das beschreibe, dass das ein offener Begriff ist und alle Begriffe sind immer offen, denn sie dienen uns, es sind unsere Instrumente. [00:13:58] Aber solange wir als Gesellschaft eben sagen, ich schreibe die Verantwortung immer dem dahinter stehenden Menschen zu, solange gibt es keine sinnvolle Funktion eines Verantwortungsbegriffs, der anders wäre [00:14:08]. Also funktionale Betrachtung letztlich des Rechts, die ich hier zugrunde lege und deswegen weder Schuld noch Verantwortung beim Roboter, sondern bei dem dahinter stehenden Menschen. [00:14:17] Aber das wird natürlich dann kompliziert. [00:14:18] Wenn jetzt, zum Beispiel, KI dazu kommt, Transparenzprobleme, dann kann es Schwierigkeiten geben mit der Verantwortung. [00:14:25]

**Karsten Wendland:** Wenn diese E-Person jetzt käme und ein Roboter hätte einen eigenen Status, man würde vielleicht sogar darüber sprechen, ihn in die Verantwortung nehmen zu können. [00:14:35] Was macht man denn, wenn dann mal was passiert ist. [00:14:37] Also den Roboter wird man ja jetzt nicht sinnvoll ins Gefängnis sperren können, weil er vielleicht jemanden verletzt hat, ja. [00:14:45] Es gibt die Diskussion, dass man streng genommen, dem Roboter aber auch nicht einfach den Stecker ziehen dürfte, wenn man ihn vorher mit einem rechtlichen Status ausgestattet hat, weil er Anspruch hätte darauf, dass seine Rechte verteidigt werden. [00:15:01] Wie ist das aus Ihrer Sicht als Rechtswissenschaftlerin, ist das eine aufgeblasene Diskussion oder ist da tatsächlich etwas dahinter? [00:15:11]

**Frauke Rostalski:** Ja, also da ist schon was dahinter. [00:15:12] Also ich hatte ja eben gesagt, für mich ist es momentan kein Zustand, in dem wir sagen können, es gibt eine Verantwortung. [00:15:17] Die Kehrseite von Verantwortung sind Rechte. [00:15:20] Ja, also wenn ich Adressat von Pflichten bin, dann kann ich auch in Anspruch nehmen, dass andere mir gegenüber

Pflichten haben. [00:15:26] Das heißt, wenn sich etwas ändern würde an meiner eben genannten Beurteilung und wir sagen, da ist ein Roboter verantwortlich, dann hat er auch Rechte. [00:15:33] Und dann müssen wir seine Interessen auch in gewisser Weise schützen. [00:15:35] Wir kennen das als analoge Debatte aus dem Tierschutz, wie ich finde, mit einem traurigen Ergebnis bis heute, wo wir eben auch immer wieder diskutieren, haben denn vielleicht Tiere selbst Rechte und haben da ja auch ganz unterschiedliche Begründungsstränge. [00:15:51] Wenn wir jetzt mal weggehen von der Verantwortung, die wir ja auch bei Tieren nicht aufstellen, anders als im Mittelalter, wo ja Tiere noch Strafen bekommen haben, ja, aber auf dem Stand sind wir ja nicht mehr. [00:16:03] Und dennoch sagen wir zum Beispiel, die Leidensfähigkeit könnte ein Grund sein, den ich für sehr überzeugend halte persönlich, den Tieren doch Rechte zuzugestehen. [00:16:11] Das heißt, wir brauchen vielleicht gar nicht die Verantwortungsebene, müssen erst gar nicht dem Roboter Verantwortung zuschreiben, sondern vielleicht müsste man auch einfach feststellen, dass da Leidensfähigkeit ist. [00:16:21] Finde ich zum jetzigen Stand auch kompliziert, noch hat mein Staubsauger nicht geweint, wenn ich ihn zum dritten Mal am Tag verdonnert habe, den Boden sauber zu machen, aber wenn sich das irgendwann ändern würde, dann müssten wir darüber diskutieren. [00:16:37] Aber, und das noch als kleine Randnotiz der Tierschutzrechtlerin, vorher müssten wir auf jeden Fall bei den Tieren darüber nachdenken, denn da haben wir eine ganz klare Antwort auf die Frage, ob die leidensfähig sind. [00:16:47] Ja, sind sie. [00:16:48] Und das, was da eben läuft, wie wir mit Tieren umgehen, global, aber auch eben einfach nur in Deutschland, also ist meiner Ansicht nach nicht haltbar, Massentierhaltung. [00:16:57]

**Karsten Wendland:** Ja. [00:16:58] Nehmen wir mal als Beispiel einen Roboter, der den Rasen mäht. [00:17:04] Dem bringt man ja aus technischer Perspektive auch so etwas wie ein Schmerzempfinden bei, was man dann als Ausweichverhalten umsetzt. [00:17:13] Also wenn der zweimal gegen den Baum gefahren ist, dann speichert der sich, da war ein Baum, er versucht auszuweichen, er versucht auch, andere Objekte zu erkennen und Kollisionen zu verhindern und das kann man-, im technischen Sprachgebrauch kann man diese Funktionalität als Schmerzempfinden beschreiben. [00:17:30] Das machen wir Informatiker auch, ja, wir nennen das so. [00:17:33] Schwierig ist dann, wenn das die Marketingabteilung der Firma erfährt und das dann nach

draußen gibt, unser Roboter hat Schmerzempfinden. [00:17:39] Der hat ja vermutlich keins, aber die technische Implementierung arbeitet durchaus mit solchen Analogien, weil die Begriffe auch so gut verständlich sind. [00:17:48] So, und jetzt gibt es ja Haltungen, die sagen, na ja, also wenn der so ein Schmerzempfinden simuliert und wir das so beobachten, dass der ausweicht, schreiben wir ihm diesen Schmerz durch Beobachtung, diese Reaktion, dieses Schmerzempfinden, dieses Ausweichverhalten zu. [00:18:08] Aber genau genommen wissen wir ja auch bei unserem Gegenüber nicht, ob er wirklich Schmerzen empfindet oder das nur simuliert. [00:18:17] Und wenn wir das jetzt mal umdrehen, möglicherweise hat der Roboter tatsächlich ein Schmerzempfinden, aber wir wissen es nicht. [00:18:24] Also bewegen wir uns jetzt in so eine erkenntnistheoretische Richtung hinein, in der wir beide, Sie hatten es vorhin erwähnt, ja auch nicht wirklich wissen, ob wir gegenseitig existieren, ob der Andere ein Bewusstsein hat oder nur so tut als ob. [00:18:39] Und wenn wir das jetzt mal auf die Maschinen übertragen und diese Gedanken weiterspinnen, inklusive der Zuschreibung eigener Rechtspositionen, eines eigenen Rechtsstatus für die Maschine, ist die Frage, wie bringen wir denn das in unsere Gesellschaft unter? [00:18:55] Und da ist ja in den letzten Jahren viel passiert. [00:18:58] Wenn wir Stellenausschreibungen anschauen, wird ja jetzt nicht mehr nur nach männlich, weiblich gesucht oder ausgeschrieben, sondern männlich, weiblich und d für divers. [00:18:07] Haben wir da irgendwann noch eine vierte Kategorie? [00:19:09]

**Frauke Rostalski:** Ja, also ausgeschlossen ist nichts. [00:19:12] Wir befinden uns auf einem ständigen Weg des Erkenntniszugewinns, hoffe ich zumindest – nicht, dass unsere Erkenntnisse weiter abbauen. [00:19:20] Aber, um die Frage nochmal vom Anfang aus zu beantworten: Es ist immer letztlich eine risikoethische Abwägung, die wir vornehmen müssen, wenn wir Konsequenzen, ob sie nun ethischer oder rechtlicher Natur sind, an etwas binden. [00:19:35] Also, um jetzt nochmal zu dem Beispiel von uns zurückzukommen, ich kann tatsächlich niemals wissen, wie es ist, Karsten Wendland zu sein und Sie werden niemals wissen, wie es ist, Frauke Rostalski zu sein. [00:19:46] Aber es gibt gewisse Evidenzen, die dafür sprechen, dass ich aus meinem Selbsterleben, aus meinem Bewusstsein meiner eigenen Person heraus, Rückschlüsse ziehen kann aufgrund von Ähnlichkeiten, dass es bei anderen Personen anders ist. [00:20:00] Dafür gibt es einfach viele Evidenzen, die wir über die Zeit gesammelt haben. [00:20:04] Gehen wir rüber zu



den Robotern. [00:20:06] Wenn wir irgendwann einen Punkt erreicht hätten, wo diese Evidenzen so stark werden, dass wir sagen, da können wir die Augen nicht mehr vor verschließen, dann müssen wir auch Konsequenzen daran knüpfen wie zum Beispiel Rechte erteilen, nicht mehr gegen den Baum fahren lassen und so weiter. [00:20:19] Ich glaube, dass wir diesen Zustand noch lange nicht erreicht haben. [00:20:22] Und auch das Beispiel, was Sie eben genannt haben, deutet genau in die Richtung, denn da scheint es mir eher eine Begriffsproblematik zu sein, die dem zugrunde liegt, nämlich eine Zuschreibung, wie Sie gesagt haben, na ja, das lässt sich gut beschreiben durch einen menschlichen Terminus. [00:20:37] Es ist ja auch ganz klar, es ist unsere Sprache, gibt eben das her, was wir selbst sind und dann neigen wir natürlich auch als Menschen immer dazu, unsere Begriffe zu übertragen auf unsere Entitäten, vor allem, wenn wir sie geschaffen haben. (Lachen) [00:20:49]

**Karsten Wendland:** Natürlich, klar. [00:20:49] Und gerade der KI-Bereich, der wimmelt ja von Zuschreibungen, dieser Art, von Metaphern, von Narrativen, die sich auch soweit verselbständigt haben, dass viele Menschen den KI-Systemen ja viel mehr zutrauen, als sie tatsächlich können. [00:21:00] Wenn wir jetzt den Bewusstseinsbegriff mal fokussieren mit der Frage, können denn diese Geräte irgendwann mal, ich sage jetzt mal flapsig, aufwachen oder nicht, da gibt es eine ganz einfache Definition für den Hausgebrauch von einem französischen Kollegen, der sagt, na ja, also Bewusstsein, das ist das, was morgens da ist, wenn wir aufwachen und was abends verschwindet, wenn wir einschlafen. [00:21:22] Das ist erstmal so für ein erstes Verständnis ganz gut. [00:21:26] Jetzt kann man weiterfragen, was ist denn, wenn wir träumen, da sind wir nicht ganz unbewusst und so weiter, wie ist das in der Narkose. [00:21:32] Das, worum es uns jetzt hier in dieser Diskussion um KI und Bewusstsein geht, geht ja noch einen Schritt weiter in die Frage: Kann eine solche Maschine, mit KI ausgestattet, kann die sich wirklich selbst als existierend erleben? [00:21:48] Das wäre dieses phänomenale Bewusstsein, (**Frauke Rostalski:** Genau.) ja. [00:21:52] Und das ist ja die große Schwierigkeit, das können wir ja gegenseitig uns auch nur durch Beobachtung zuschreiben, (**Frauke Rostalski:** Genau.) [00:22:00] So, und jetzt gibt es dazu ja ganz unterschiedliche Positionen. [00:22:04] Da gibt es Leute, „das wird niemals kommen, das weiß die Philosophie schon seit hundert Jahren, Bewusstsein in der Maschine ist ausgeschlossen“, ja. [00:22:13] Und andere

sagen, das kann irgendwann mal viel schneller kommen als man denkt, wenn man diesen sprichwörtlichen einen Draht zu viel angelötet hat, dann ist, und so einen Draht anzulöten, das geht schnell, das kann ich Ihnen versichern, ja, dieser eine Draht zu viel und dann ist das Bewusstsein da. [00:22:29] Und dann könnte man natürlich auch überlegen, ob man sowas nicht vielleicht sogar nicht versucht zu verhindern, sondern absichtlich zu machen, um jetzt mal global, politisch gesehen, einen Vorteil zu haben, ja. [00:22:41] Und da ist die Frage, die ethische Frage, sollten wir das tun, sollten wir Maschinen bewusst versuchen zu erschaffen, in denen ein Bewusstsein dann entsteht, oder sollten wir genau das vermeiden? [00:22:55] Und da gibt es ja sehr harte Positionen. [00:22:58] Ich habe mit einem Kollegen gesprochen, den wir auch noch hier im Podcast hören werden, der sagt, „Lasst die Finger davon für die nächsten 30 Jahre, es könnte sein, dass ihr leidensfähige Computer baut, ja, und das sollten wir nicht tun.“ [00:23:10] Andere sagen, wir dürfen auf keinen Fall diese Technologie dem Klassenfeind überlassen. [00:23:15] Also müssen wir die ersten sein, weil der Gewinner in dieser Disziplin wird derjenige sein, der das gesellschaftliche System vorgibt und wir wollen hier nicht leben wie die Chinesen. [00:23:25] Was sagen Sie dazu? [00:23:25]

**Frauke Rostalski:** Oh je, da sind jetzt natürlich einige Fragen schon vorweg gegangen, die Sie als Prämisse gesetzt haben. [00:23:31] Vielleicht darf ich zu einer davon wenigstens noch was sagen und zwar zu der Frage, ob ich glaube, dass sowas wie Bewusstsein tatsächlich mal in die Maschinen hinein gelangen kann. [00:23:42] Das halte ich für schwierig, solange wir auf der Basis von Algorithmen sprechen. [00:23:47] Denn wie, glaube ich, vorher auch schon gesagt, der Mensch funktioniert nicht nach Algorithmen, wir lassen uns von Gründen affizieren, wir haben sehr komplexe Denkprozesse, die werden sich niemals in dieser Komplexität, wie die Lebenswirklichkeit sie enthält, in Maschinen übersetzen lassen. [00:24:04] Und dann gibt es eben auch immer noch von-, da gibt es eben auch noch immer Denkvorgänge, die sich nicht als Algorithmen erfassen lassen. [00:24:12] Also das heißt, ich glaube, dass das nie funktionieren wird, nicht zuletzt deswegen, weil wir Menschen mit Intentionalität ausgestattet sind. [00:24:23] Der auf Algorithmen basierende Computer oder der Roboter sagt A plus B, also C. [00:24:30] Aber wir sagen zum Beispiel, ein Glas ist deswegen ein Glas, weil es diese Eigenschaften hat und verständigen uns darauf, damit wir uns unterhalten können, dass das jetzt so ist. [00:24:41] Das ist etwas, was Computer nicht tun.

[00:24:43] Sie arbeiten mit Zeichen, die wir Ihnen schon vorgegeben haben, verstehen aber nicht die Bedeutung, die dahinter steht. [00:24:47] So, zu der letzten Frage. Es ist immer die Frage, wie gehen wir mit Risikoforschung um. [00:24:54] Ich bin prinzipiell relativ forschungsfreundlich und kenne diese Diskussion ja auch aus vielen bioethischen Kontexten. [00:25:04] Das Argument, für sich genommen zu sagen, dann erschafft ihr ja leidensfähige Computer, das finde ich kein überzeugendes Argument. [00:25:12] Mit dem gleichen Argument kann man nämlich auch, was es in der Bioethik, Biomedizin immer wieder als Argument gibt, sagen, wir möchten nicht, dass bestimmte Menschen auf die Welt kommen, weil sie eine bestimmte Krankheit haben, ja. [00:25:24] Und deswegen wollen wir eine entsprechende, ja, Kombination dieser Gene nicht und so weiter. [00:25:30] Und das halte ich nicht für überzeugend, denn wie schon gesagt, keiner von uns kann sich in den anderen hereinvertreten, keiner weiß, was eigentlich Lebensqualität für den anderen bedeutet. [00:25:39] Und nur, weil etwas leiden kann, heißt es ja nicht auch, dass es auch halt eine positive Seite haben kann. [00:25:44] Also nur, wenn wir etwas Leidensfähiges erschaffen, dann müssten wir ja nur verantwortungsvoll damit umgehen, damit es eben auch eine schöne Seite des Lebens erkennen kann. [00:25:51]. Also das ist nicht das Argument. [00:25:53] Aber es gibt natürlich andere Argumente, die wir hier wägen müssten, also wie zum Beispiel, wie gefährlich ist das für uns Menschen. [00:26:01] Wenn jetzt da tatsächlich eben das Bewusstsein kommt, kennen wir alle, diese schlimmen Science Fiction Filme, in denen dann auf einmal die Maschinen sich bemächtigen des Menschen und vielleicht nicht die Ethik zugrunde legen, die wir uns wünschen, die wir unserem eigenen Miteinander zugrunde legen, das sind eher Risiken, die wir da diskutieren sollten. [00:26:19]

**Karsten Wendland:** Ja, aber ist es da nicht ein Problem, dass wir nicht überall auf der Welt diese eine Ethik haben, sondern dass wir ja in sehr unterschiedlichen Kulturkreisen auch unterschiedliche Vorstellungen davon haben, was gut ist, was ein gutes Leben ausmacht, was man tun und was man nicht tun sollte. [00:26:37] Wenn es dann etwas komplizierter wird, dann reichen diese eingeübten moralischen Verhaltensweisen nicht mehr aus, dann braucht man die Ethik, die einem auch Kriterien offenlegt, aber die sind ja immer in kulturellen Kontexten. (**Frauke Rostalski:** Ja.) [00:26:49] Was bedeutet das für KI-Systeme? [00:26:51] Also wir arbeiten jetzt ja auch viel mit Systemen, die stammen von Firmen aus Kalifornien. [00:26:57] So, das heißt,

wie wir uns hier in unserem Alltag organisieren, ist letztlich auch geprägt durch Geschäftsmodelle, die im kalifornischen Raum entstanden sind. [00:27:07] Würden wir jetzt alle mit chinesischen Systemen arbeiten, hätten wir vielleicht jetzt auch ganz andere Arbeitsprozesse an manchen Stellen. (**Frauke Rostalski:** Ja.) [00:27:13] Ja, und auch die Frage, was ist denn richtig und was ist falsch, was ist ein Menschenleben wert, wie trifft man Entscheidungen. [00:27:20] Da haben wir die große Utilitarismus-Diskussion, ja, das größte Glück der größten Zahl zu erreichen und unsere sehr europäische Haltung in der kantschen Ethik, wo es um den Einzelnen geht. [00:27:31] Und das passt ja nicht so richtig zusammen. [00:27:33] So, jetzt soll man KI-Systeme bauen, die möglichst auch noch eine ethische Komponente haben. [00:27:41] Welche Ethik soll es denn sein? [00:27:43]

**Frauke Rostalski:** Jetzt ist die Frage natürlich ganz leicht gestellt, welche soll es sein. [00:27:46] Da kann ich jetzt natürlich meine persönliche Überzeugung mitteilen, die eine an Kant orientierte deontologische Ethik wäre, also den Menschen in den Mittelpunkt rückt, so wie Sie das eben auch beschrieben haben. [00:27:58] Ich muss bestreiten, dass man nicht doch auch sagen kann, dass es gute und schlechte Ethik gibt. [00:28:05] Also wenn wir alle uns hinter dem Schleier des Nichtwissens befinden würden, ich würde Sie fragen, wollen Sie lieber in Nordkorea leben oder lieber in Deutschland, dann bin ich ganz sicher, was Sie antworten würden. [00:28:14] Und zwar nicht Nordkorea. [00:28:16]

**Karsten Wendland:** Nein, ich würde natürlich lieber hier leben. [00:28:18].

**Frauke Rostalski:** Ja, richtig. [00:28:18] Und das, denke ich, ist schon eine gewisse Nagelprobe, zu sagen, dass es durchaus ethische Grundlinien gibt, die besser und schlechter sind. [00:28:26] Jetzt kommen wir natürlich dann irgendwann in die Untiefen von ganz komplexen Einzelproblemen und sicherlich da wird es dann auch unterschiedliche Positionen geben, wo es vielleicht auch nicht mehr so eindeutige Antworten gibt. [00:28:37] Aber ist das mit der KI jetzt nicht gerade ein guter Punkt, mal zu sagen, jetzt soll es mal zum Schwur kommen. [00:28:44] Lange Zeit haben wir ja viele Dinge einfach nur irgendwie hier in unserem Elfenbeinturm diskutiert. [00:28:48]

**Karsten Wendland:** Ja. Ja, genau. [00:28:49]

**Frauke Rostalski:** Jetzt könnte man doch mal sagen, so, jetzt lasst es uns doch mal bauen, lasst doch mal in Europa zum Beispiel sagen, wir machen mal einen Aufschlag und mal gucken, wie die anderen das finden. [00:28:57] Und da bin ich nicht sicher, ob sich da nicht doch Europa-. [00:29:01].

**Karsten Wendland:** Vielleicht schließen die Nordkoreaner sich uns an, vielleicht gelingt es uns auch, sie umzuerziehen. [00:29:06] Ich bin mir nicht ganz sicher, aber Sie weisen auf einen ganz wichtigen Punkt hin, nämlich dass man aufpassen muss. [00:29:11] Und dazu haben wir eine Zuhörerfrage, drei sogar, von einem Kollegen aus den Niederlanden, die ich Ihnen jetzt gerne vorspielen möchte. [00:29:22]

**Frauke Rostalski:** Ja. [00:29:23]

**Gerard Blommestijn:** Mein Name ist Gerald Blommestijn aus Amsterdam, der ursprünglich in der Kernphysik und später in einem biomedizinischen Forschungslabor gearbeitet hat. [00:29:33] Ich war auch immer an der quantenmechanischen Kommunikation zwischen Materie und Geist interessiert. [00:29:40] Liebe Frau Rostalski, ich habe folgende Fragen an Sie. [00:29:45] Frage 1: Hielten Sie es für klug, wenn für die Verantwortlichen der KI und Roboterentwicklung eine Art Hippokrates-Eid entworfen würde und was oder wo läge die Grenze der Komplexität und Macht der Systeme, ab der man einen solchen Eid verlangen könnte? [00:30:06]

**Karsten Wendland:** Ein Eid des Hippokrates für Verantwortliche im Bereich der künstlichen Intelligenz. [00:30:11] Wenn sie sowas hören, Frau Rostalski, was löst das in Ihnen aus? [00:30:15]

**Frauke Rostalski:** Also grundsätzlich frage ich mich dann erstmal, was ist denn eigentlich der Sinn und Zweck von so einem hippokratischen Eid, ja. [00:30:20] Bei den Ärzten so eine Art Selbstverpflichtung auf ethischer, vielleicht in Teilen auch berufsrechtlicher Ebene [[Quellenverweis 5](#)]. [00:30:27] Jetzt liest man den hippokratischen Eid als Juristin und stellt fest, dass da durchaus Sachen drinstehen, die sich mit dem geltenden Recht gar nicht in Einklang bringen lassen. [00:30:35]

**Karsten Wendland:** Okay, was zum Beispiel? [00:30:36]

**Frauke Rostalski:** Zum Beispiel, ich werde niemanden, auch auf eine Bitte nicht, ein tödlich wirkendes Gift geben und auch keinen Rat dazu erteilen. [00:30:43] Glücklicherweise werde ich keiner Frau ein fruchtatreibendes Zäpfchen geben. [00:30:47] Also sowohl der Schwangerschaftsabbruch ist in engen Grenzen zulässig als auch die Suizidassistentz. [00:30:53] Das heißt, hier ist etwas verboten im hippokratischen Eid, wo wir unbedingt Ärzte für brauchen, und was ja auch die Praxis der Ärzte ausmacht, glücklicherweise. (**Karsten Wendland:** Ja.) [00:31:03] Das heißt, wenn, dann müsste es ein Eid sein, der auch mit dem geltenden Recht in Einklang steht. [00:30:09]

**Karsten Wendland:** Zeitgemäß. [00:31:10].

**Frauke Rostalski:** Zeitgemäß, genau. [00:31:11] Und dann muss man sich immer noch fragen, wie viel bringt das. [00:31:14] Also bei den Ärzten hat es natürlich so ein gewisses Selbstverständnis hervorgerufen, dass es sicherlich gut wäre [00:31:22]. Allerdings unterscheidet die Tätigkeit von, sagen wir mal, Informatikern, Programmierern sich durchaus nochmal von den von Ärzten, weshalb es, glaube ich, da weniger Sinn ergibt, einen solchen Eid abzulegen. [00:31:29] Und zwar, der Arzt arbeitet mit seinen Händen, er ist als Person unmittelbar gefragt. [00:31:35] Bei den KI-Programmen geht es uns um die Anwendung selbst und weniger um die Tätigkeit des Menschen, der dahintersteht. [00:31:42] Das heißt, eigentlich bräuchten wir einen hippokratischen Eid für das KI-System selbst, also wie das eben programmiert sein soll. [00:31:49] Und jetzt hatten wir es am Anfang schon gesagt, ich bin eine große Anhängerin der Zertifizierung. [00:31:53] Das befindet sich auf einer teilweise erst ethischen, vorregulativen Ebene. [00:31:58] Und das halte ich für eine gute Idee, wenn wir sagen wollen, das ist eine Form von hippokratischem Eid übertragen auf KI-Systeme, dann gerne. [00:32:05] Also Zertifizierung. [00:32:06]

**Karsten Wendland:** Dazu passt auch die zweite Frage von Herrn Blommestijn. [00:32:09]

**Gerard Blommestijn:** Frage zwei: Wäre es auch klug, ausreichend komplexe und mächtige KI-Systeme selbst einen Eid schwören zu lassen? [00:32:19]

**Frauke Rostalski:** Eid schwören, ja, also jedenfalls müssen wir sie gewissen Regeln unterwerfen und eine Zertifizierung wäre, finde ich, eine Form von

solchen Regeln. [00:32:28] Zertifizierung reicht nicht immer, das hängt natürlich von der jeweiligen Anwendung ab. [00:32:33] Je eingriffsintensiver, wie man als Jurist sagen würde, also das heißt, umso mehr Sie in relevante, menschliche Interessen eingreift, umso mehr ist dann eben auch staatliche Regulierung durch Gesetze erforderlich. [00:32:44]

**Karsten Wendland:** Zertifizieren würde bedeuten, dass wir dem KI-System etwas zuschreiben. [00:32:52] Das würde aber jetzt nicht bedeuten, dass das KI-System selbst einen Eid schwören würde. (**Frauke Rostalski:** Genau.) [00:32:58] Dann müsste ja etwas mehr drin sein, so etwas wie eine, sagen wir mal, Persönlichkeit, (**Frauke Rostalski:** Richtig.) irgendeine Art von etwas, von Bewusstsein vielleicht, was in der Lage wäre, auch in die Verantwortung zu treten. [00:33:11]

**Frauke Rostalski:** Genau. [00:33:11] Und damit sind wir ja wieder bei der Ausgangsfrage. [00:33:13] Also nur, wenn ich der Maschine selbst Verantwortung zuschreiben würde, könnte ich ja sagen, dass die einen eigenen Eid schwören kann, weil sonst ist das letztlich eine leere Phrase ohne Sinngehalt. (**Karsten Wendland:** Ja.) [00:33:23] Und deswegen muss der Mensch das natürlich abgeben. [00:33:26]

**Karsten Wendland:** Und an welchem Punkt wären denn Juristen alarmiert? [00:33:30] Jetzt passiert hier gerade was im Bereich der KI-Systeme, das können wir uns nicht mehr so richtig erklären. [00:33:35] Und hier wird ein Bereich verlassen, in dem wir das System zertifizieren. [00:33:41] Jetzt müssten wir das System so langsam selber fragen. [00:33:43]

**Frauke Rostalski:** Ja, also das ist ja ganz schwierig, so etwas eine Juristin zu fragen. [00:33:48] Also wir haben den Turing-Test, der ja nach wie vor offenbar nie wirklich erfolgreich ist, wenn man ihn mit den verschiedensten vorhandenen Programmen durchspielt. [00:33:58] Das ist sicherlich ein ganz gutes Testmittel. [00:34:01] Ich-. [00:34:01]

**Karsten Wendland:** Ja aber, hier der Chinese, wir haben das Bild noch zwischen uns liegen, der hat diese Frau geheiratet, ich glaube, der Turing-Test ist dem egal. [00:34:07] Und der behandelt ja diesen Roboter als seine Frau und stellt von sich aus vielleicht so was wie Ebenbürtigkeit dar, zumindest haben sie gemeinsam dieses Hochzeitsritual vollzogen. [00:34:19] Also der Turing-Test ist ja leicht auszuhebeln, weil man dieses so tun als ob, gar nicht

so und auch zwischen uns Menschen nicht so genau entscheiden kann.  
[00:34:29]

**Frauke Rostalski:** Ja gut, aber dann, wenn Sie so fragen, muss ich natürlich so antworten wie eben schon. [00:34:33] Dann würde der Bruch nicht dadurch entstehen, dass sich an der Technologie was verbessert hat, sondern dadurch, dass wir als Gesellschaft zu einem anderen Ergebnis miteinander kommen, das auch tatsächlich mehrheitsfähig sein muss. [00:34:43] Ich halte das momentan für eine Randerscheinung. [00:34:46]

**Karsten Wendland:** Also könnte es sein, dass tatsächlich, rein hypothetisch, ein Bewusstsein in einer Maschine entstünde und auch da wäre, wir es ihr aber als Gesellschaft nicht zuschreiben und damit auch nicht zugestehen? [00:34:57] Dann hätten wir den Fall, den Herr Metzinger beschreibt, man hätte eine bewusste Maschine, KI-System, das in seinem Bewusstsein aber gefangen wäre und wir würden es einfach nicht gelten lassen? [00:35:14]

**Frauke Rostalski:** Das, denke ich, ist grundsätzlich in der Theorie möglich. [00:35:18] Die Frage ist nicht, ob der Mensch sensibel genug wäre, um ab einem bestimmten Punkt zu sagen, jetzt ist es nicht mehr bloß eine besonders gute Täuschung, sondern es spricht einiges dafür, dass wir es mit einer Entität zu tun haben, der wir mehr zugestehen müssen, als wir das bislang getan haben. [00:35:33]

**Karsten Wendland:** Okay. [00:35:36] Herr Blommestijn hat noch eine dritte Frage für uns. [00:35:38]

**Gerard Blommestijn:** Nach dem Aufkommen der Atomwaffen gründete die Internationale Gemeinschaft die Internationale Atomenergie Organisation. [00:35:46] Wäre es gut, eine vergleichbare Agentur zu gründen, um potenziell gefährliche KI und Roboter-Entwicklungen im Auge zu behalten? [00:35:57]

**Frauke Rostalski:** Das sehe ich ganz genauso. [00:35:59] Also wenn man eben sieht, ein paar von den Risiken haben wir ja schon angesprochen, welche Risiken im Raum stehen, welche weitreichenden gesellschaftlichen Konsequenzen die auch können, halte ich das für unbedingt erforderlich. [00:36:11] Ich glaube aber, dass wir insoweit hier schon auf einem ganz guten Weg sind. [00:36:15] Wir sind ja inzwischen auch wirklich auf europäischer



Ebene im Diskurs miteinander, sodass ich glaube, dass die Entwicklung durchaus auch dahin gehen wird. [00:36:22]

**Karsten Wendland:** Was für Leute sollten in einem solchen Gremium drinsitzen? [00:36:24]

**Frauke Rostalski:** Nicht nur Virologen. [00:36:26]

**Karsten Wendland:** Juristen. (Lachen) [00:36:28]

**Frauke Rostalski:** Nein, nicht nur Virologen. (**Karsten Wendland:** Okay.) [00:36:30] Also auf jeden Fall-. [00:36:32]

**Karsten Wendland:** Ja, es geht um die Ausgewogenheit, ja. [00:36:33]

**Frauke Rostalski:** Genau, es geht um die Ausgewogenheit in dem Fall- [00:36:35].

**Karsten Wendland:** Es geht um die Ausgewogenheit und möglichst darum, dass man nicht diejenigen ausgrenzt, die vielleicht aus so einer Außenseiterposition heraus aber trotzdem das entscheidende Wichtige beitragen können. [00:36:43] Ich glaube, das ist die Kunst, (**Frauke Rostalski:** Ja.) auch diejenigen, auch die Störenfriede an dieser Stelle anzuhören (**Frauke Rostalski:** Ja.) und nicht gleich raus zu negieren, weil es gerade nicht passt. [00:36:52] Also wir hätten-, in einem solchen Fall hätten wir eine komplett neue Situation und die erfordert vielleicht auch mal Gesprächsverfahren. [00:36:59]

**Frauke Rostalski:** Ja. [00:37:00] Zunächst freut mich das, dass Sie das als Informatiker so sehen, weil ja oftmals der Informatiker als Stereotyp doch eher sagt, ach, die Normwissenschaftler sollen mal lieber nicht so viel, weil am Ende verbieten die mir noch irgendwas, was ich gerne machen möchte. [00:37:14]

**Karsten Wendland:** Das ist richtig. [00:37:15]

**Frauke Rostalski:** Insofern ist das ja doch ein hohes Maß an Kooperation, was mir gerade entgegenschlägt, über das ich mich sehr freue. [00:37:19] Also diese Zusammensetzung müsste vielleicht nicht genauso ausgestattet sein wie im Deutschen Ethikrat, wo wir ja nicht so viele Informatiker haben. [00:37:27] Also wenn es um etwas wie KI geht, würde ich sagen, müsste der

prozentuale Anteil der Technikwissenschaftler doch noch deutlich höher sein. [00:37:34] Aber unbedingt müssten die Normwissenschaftler dabei sein und damit meine ich eben Ethiker, Juristen natürlich auch, Vertreter von verschiedenen besonderen ethischen Vorstellungen, da meine ich, Religionsauffassungen gehören da sicherlich auch mit dazu. [00:37:47] Ob das ein Thema für den Deutschen Ethikrat ist? [00:37:49] Ich glaube, dass es ein Zukunftsthema ist. [00:37:52] Ich halte es für sehr sinnvoll, wenn wir uns damit näher befassen würden. [00:37:55]

**Karsten Wendland:** Die Gefahr besteht ja dann darin, dass die Technikwissenschaftler bestimmte Probleme gar nicht wahrnehmen, weil sie die Begrifflichkeiten und die Instrumentarien nicht haben oder aber in Analogien arbeiten oder Metaphern verwenden, die begrifflich am Gleichen ansetzen, aber letztlich doch was Anderes meinen. (**Frauke Rostalski:** Ja.) [00:38:15] Also versprechen Sie sich an der Stelle nicht zu viel von den Technikern, da macht es tatsächlich der gesamte Mix. [00:38:22]

**Frauke Rostalski:** Wobei ich sagen muss, ich arbeite ja selbst auch in interdisziplinären Konsortien und es ist am Anfang natürlich sehr schwierig, eine gemeinsame Sprache zu finden. [00:38:30] Aber, also ich habe es auch sehr positiv erlebt, dass man sich da durchaus auch beieinander zubewegen kann. [00:38:35]

**Karsten Wendland:** Ja, Frau Rostalski, wir sprachen jetzt in einem großen Bogen über KI, selbstbewusste Künstliche Intelligenz, die, wenn es sie denn dann gäbe, tatsächlich mehr wäre, als eine bloße Maschine. [00:38:50] Frage zum Abschluss an Sie: Wie lange wird es noch dauern? [00:38:53]

**Frauke Rostalski:** Oh mein Gott, lange, wenn je. [00:38:59] Recht pessimistisch, aber mal sehen, (**Karsten Wendland:** Heißt?) ich bin ja für alles offen. [00:39:07] Ja, also lange, wenn überhaupt jemals. [00:39:09] Mit einem großen Fragezeichen dahinter. [00:39:13]

**Karsten Wendland:** Was sollten wir tun, wenn wir nicht wissen, ob etwas kommt oder ob nicht. [00:39:19] Sollten wir uns trotzdem vorbereiten? [00:39:21]

**Frauke Rostalski:** Ja, unbedingt. [00:39:21]

**Karsten Wendland:** Und das Szenario des schwarzen Schwans zu versuchen in den Griff zu halten und auf Vorrat uns vorzubereiten, auch wenn wir es nicht brauchen [[Quellenverweis 6](#)]? [00:39:30]

**Frauke Rostalski:** Ja, unbedingt. [00:39:30] Also wir haben das in der Biomedizin viel zu häufig erlebt, dass plötzlich irgendwelche Entwicklungen da waren und keine Regeln dafür vorhanden waren und wir als Gesellschaft dastanden und gesagt haben, wie sollen wir denn jetzt damit umgehen. [00:39:39] Wollen wir das, wollen wir das nicht und die Reaktion ist dann häufig totale Überregulierung. [00:39:43] Also ich glaube, wir müssen unbedingt heute schon auf Sicht-, nicht nur auf Sicht fahren, sondern wirklich auf Weitsicht fahren und uns eben fragen, was wäre wenn [00:39:52]. Nicht nur, weil es schöne Gedankenexperimente sind, sondern weil wir es vielleicht doch mal irgendwann als Gesellschaft gebrauchen können. [00:39:58]

**Karsten Wendland:** Vielen Dank für das tolle Gespräch. [00:40:00]

**Frauke Rostalski:** Ja, vielen Dank, war sehr schön. [00:40:03]

**Karsten Wendland:** Das war Frauke Rostalski, Rechtswissenschaftlerin in Köln, Experte für Biotechnologie und Künstliche Intelligenz und Mitglied des Deutschen Ethikrats, in unserer Podcast-Serie selbstbewusste KI, Ihrem Forschungspodcast an der Grenze zwischen Mensch und Maschine. [00:40:21] Sind Ihnen beim Zuhören weitere Fragen eingefallen oder geniale Ideen gekommen? Wir freuen uns über ihre Gedanken. [00:40:28] Lassen sie uns daran teilhaben und eine Nachricht über unsere Projektwebseite zukommen, die sie im Internet unter [www.ki-bewusstsein.de](http://www.ki-bewusstsein.de) finden. [00:40:37] Oder schreiben und folgen sie uns auf Twitter, dort finden sie unser Projekt unter dem gleichen Namen, [@KIBewusstsein](https://twitter.com/KIBewusstsein). [00:40:44] In der nächsten Folge sprechen wir mit Christof Koch, der als Neurowissenschaftler seit vielen Jahrzehnten der Künstlichen Intelligenz und dem Bewusstsein auf der Spur ist und der angeblich auf einer kleinen Insel lebt [00:40:59]. Redaktion und Produktion dieser Folge lagen in den guten Händen von Kayla Zoller. Die Aufnahmeleitung vor Ort im schönen Köln lag diesmal bei mir selbst. [00:41:09] Ich freue mich, wenn es Ihnen gefallen hat und diese Folge auch für sie ein Beitrag dazu war, KI-Bewusstsein etwas mehr zu entmystifizieren. [00:41:17] Bleiben sie gesund, hoffnungsvoll und gestaltungstark. [00:41:21] Das war Ihr und euer Karsten Wendland, bis bald. [Ende 00:41:25]

## 5 Erwähnte Quellen

Folgende weiterführende Quellen wurden in der Podcast-Folge genannt:

- [1] Frauke Rostalski ist Mitglied im Deutschen Ethikrat.  
<https://www.ethikrat.org/>
- [2] Artikel auf Welt.de: Ingenieur „heiratet“ seine selbstgebaute Robo-Freundin (vom 07.04.2017).  
<https://www.welt.de/kmpkt/article163506203/Ingenieur-heiratet-seine-selbstgebaute-Robo-Freundin.html>
- [3] Zitat des britischen Evolutionsbiologen Clinton Richard Dawkins: “We are survival machines – robot vehicles blindly programmed to preserve the selfish molecules known as genes. This is a truth which still fills me with astonishment.” (Aus seinem Buch „The Selfish Gene“, Oxford University Press: 1976).  
[https://global.oup.com/academic/product/the-selfish-gene-9780198788607?q=The selfish gene&lang=en&cc=us](https://global.oup.com/academic/product/the-selfish-gene-9780198788607?q=The%20selfish%20gene&lang=en&cc=us)
- [4] Artikel über das Libet-Experiment bei Planet Wissen.  
<https://www.planet-wissen.de/natur/forschung/hirnforschung/pwiedaslibetexperiment100.html>
- [5] Der hippokratische Eid.  
<https://www.aerztekammer-bw.de/10aerzte/40merckblaetter/20recht/10gesetze/hippoeid.pdf>
- [6] Das Szenario des schwarzen Schwans nach dem Publizisten und Börsenhändler Nassim Nicholas Taleb. (Artikel bei Risknet.de vom 20.07.2020).  
<https://www.risknet.de/themen/risknews/zum-umgang-mit-seltenen-ereignissen/>

## 6 Kontakt



*Zur Website des  
ITAS*

Prof. Dr. Karsten Wendland

[karsten.wendland@kit.edu](mailto:karsten.wendland@kit.edu)

Karlsruher Institut für Technologie (KIT)

Institut für Technikfolgenabschätzung und Systemanalyse (ITAS)

Karlstraße 11

76133 Karlsruhe

GERMANY