

Webinar Unvorhersehbare (Fehl-) Entscheidungen von KI-Systemen

Sprecherin:

Claudia Otto | Counsel

Dentons Europe LLP

+49 69 450 012 392

claudia.otto@dentons.com

Unser Gast:

Prof. Dr. Anne Riechert

AI Frankfurt Rhein-Main e.V., Vorstand

University of Applied Sciences

riechert@fb2.fra-uas.de

Agenda

- I. Definition „Künstlicher Intelligenz“
- II. Trifft KI eine „eigene Entscheidung“?
- III. Kann eine künstlich intelligente „Entscheidung“ falsch sein?
- IV. Eine eigene Rechtspersönlichkeit für die KI?
- V. Wer ist also verantwortlich?
- VI. Was bedeutet das für die (Rechts-)Praxis? (Otto & Prof. Dr. Riechert)
- VII. Ethics by Design (Prof. Dr. Riechert)
- VIII. Fazit und Key Takeaways

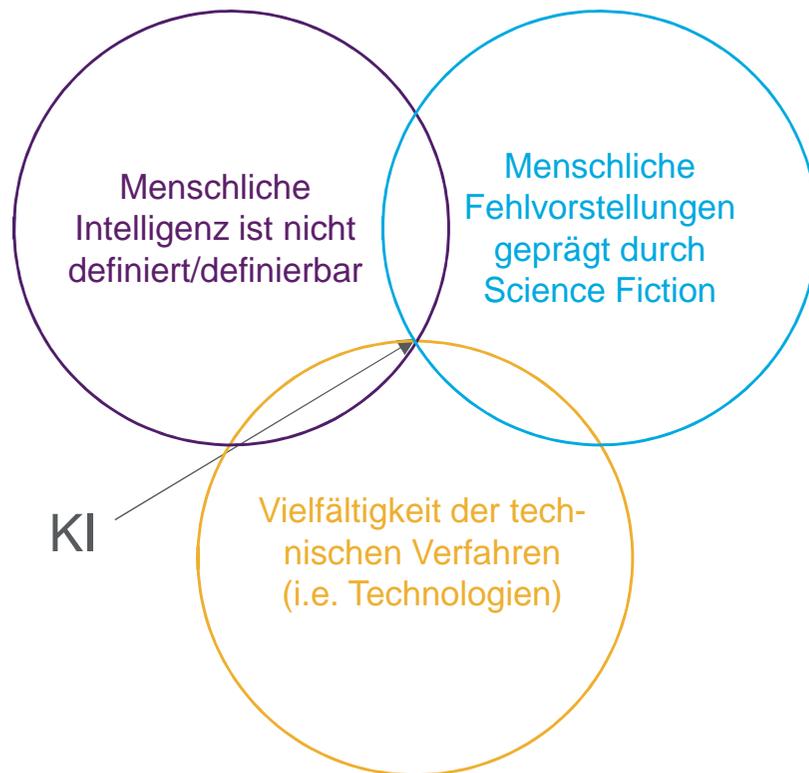
I. Definition „Künstlicher Intelligenz“

Was ist die Definition von „Definition“?

„Künstliche Intelligenz“ ist nicht definiert und, jedenfalls mit dem Ziel der sachgerechten Rechtsanwendung, nicht definierbar.

I. Definition „Künstlicher Intelligenz“

1. Herausforderungen und aktuelle Ansätze



2. Anforderungen

- korrekt
- leicht verständlich
- fortschrittsresistent

3. Definitionsversuch

KI ist ein Sammelbegriff für „automatisierte datenbasierte technische Verfahren zur Lösung von komplexen Problemen.“

- Lässt sich mit “Teilgebiet der Informatik” gut vereinbaren.
- “Höher entwickelt” ignoriert den technischen Fortschritt.
- Definitionen mit Konkretheitsanspruch sind schnell überholt.
- Schließt Verwechslung mit menschlicher Intelligenz aus.
- Diskriminiert nicht.
- Im Einklang mit wiss. Bewertung (“Fehlbenennung”, “unsinnig”).

II. Trifft KI eine „eigene Entscheidung“?

Auch der Begriff „(algorithmisches) Entscheidungssystem“ suggeriert menschenähnliche Entscheidungen. Besser ist der Begriff einer „(algorithmischen) Entscheidungshilfe“.

(Gesetz und) Rechtsanwendung dürfen nicht auf Sprachverwirrung aufbauen.

Die “Entscheidung” im Rechtssinne ist eine andere als die “Entscheidung” im Kontext der Informatik.

Die “Handlung” im Rechtssinne ist eine andere als die “Handlung” im Kontext der Informatik.

Beispiel: Der Algorithmus, die “Handlungsanweisung” an den Computer.

II. Trifft KI eine „eigene Entscheidung“?

1. Die eigentliche Frage ist doch...

„Kann ein technisches Verfahren einen eigenen (freien) Willen bilden?“

- Was ist eine „Entscheidung“?

Eine Entscheidung ist, wie auch die Wahl oder der Entschluss, vom allgemeinen Begriffsverständnis her ein Resultat des menschlichen (Nach-) Denkens.

- Was ist Denken?

Etwas Höchstpersönliches. Etwas, das erlernt werden muss. Das Eigeninitiative und intrinsischer Denkbereitschaft bedarf.

„In das was Denken heißt, gelangen wir, wenn wir selber denken. Damit ein solcher Versuch glückt, müssen wir bereit sein, das Denken zu lernen.“

Heidegger, Was heißt denken?, I, 1954

Denken ist hiernach nicht nur ein schlichter Informationsverarbeitungsvorgang, der auf der Kommunikation und Funktion von physischen und /oder Software-Komponenten nach bestimmten Vorgaben eines anderen beruht.

2. Ein technisches Verfahren denkt nicht.

Höchstpersönlichkeit ergibt sich auch nicht aus mangelnder Nachvollziehbarkeit des Prozesses der maschinellen Informationsverarbeitung. Der Eindruck hiervon folgt aus mangelnder Übersicht des schaffenden Menschen.

3. Ein technisches Verfahren entscheidet nicht (frei).

- Freiheit ergibt sich aus der Abwesenheit bzw. dem Abweichenkönnen von Vorgaben.
- Eine Fehlermeldung ist keine Entscheidung.
- Eine Entscheidungshilfe trifft keine Entscheidung.

Der Mensch bildet den entscheidenden Willen.

Uns unterscheidet vom Algorithmus die grundsätzliche Fähigkeit zur Unterscheidung von Fakt und Fiktion. Die Frage ist nur, ob wir uns um diese Fähigkeit wirklich „entlasten“ wollen.

III. Kann eine künstlich intelligente „Entscheidung“ falsch sein?

Was bedeutet „richtig“ und was bedeutet „falsch“?

**What will
you fight
for?**

#actuallycurious

Die Antwort ergibt sich stets aus dem konkreten Kontext, z.B. Zeit und Kulturkreis.
Gesellschaftliche Bewertungen unterliegen zudem dem gesellschaftlichen Wandel.

III. Kann eine künstlich intelligente „Entscheidung“ falsch sein?

1. Die eigentliche Frage ist doch...

„Kann richtig zugleich falsch sein?“

Inhaltlich fehlerhafte Ergebnisse beruhen auf unzureichenden (falschen oder unvollständigen) Informationen oder Vorgaben. Das Ergebnis ist dennoch idR rechnerisch richtig. Das Ergebnis kann also nur unerwünscht sein infolge mangelhafter menschlicher Übersicht oder Bewertung.

People with no idea about AI, telling me my AI will destroy the world

Me wondering why my neural network is classifying a cat as a dog..



2. „In meinem Kopf hat es gerade noch Sinn gemacht...“

Was macht eine menschliche Entscheidung aus?

Dass wir nicht wissen, wie genau und wieso sie zustande kam. Weil eine menschliche Entscheidung individuelle Facetten hat, sich aus unterschiedlichen individuellen Erfahrungen und Emotionen speist, die menschliche Entscheidungen nicht rechnerisch nachvollziehbar machen. Oft sind Entscheidungen selbst für den sie treffenden Menschen nicht (mehr) nachvollziehbar.

Diese Komplexität einer menschlichen Entscheidung wohnt einem technischen Verfahren nicht inne.

3. „Rationale“ Vorhersagen aufgrund „irrationaler“ Entscheidungen

Wenn von Vorhersagen im Zusammenhang mit KI gesprochen wird, geht es um Schlussfolgerungen aus erkannten menschlichen Entscheidungen der Vergangenheit (Mustererkennung)

Problem: „gute“ Entscheidungen verrauschen in einer Vielzahl von umstands- und nicht eigeninteressebedingten Entscheidungen. Die Qualität der Vorhersagen wird dadurch schlecht, selbige aber nicht „falsch“.

Die (Vorgabe der) Interpretation der Daten ist falsch.

Ein in der Folge sinnloses Ergebnis macht das technische Verfahren sinnlos.

IV. Eine eigene Rechtspersönlichkeit für die KI?

Überzeugte David Copperfield 1983 die Freiheitsstatue davon, sich von ihrem Sockel herunterzubeben und zu verstecken?



Die menschliche Täuschung ist keine taugliche Grundlage für Schlussfolgerungen dieser Art.

Die elektronische Person ist v.a. ein Resultat von fehlenden Informationen, darauf basierenden Fehlannahmen und damit Fehlern in der menschlichen Kommunikation.

IV. Eine eigene Rechtspersönlichkeit für die KI?

1. Die eigentliche Frage ist doch...

„Warum sollte ein technisches Verfahren Träger von Rechten und Pflichten sein?“

Es gibt keinen Grund. Die Schaffung einer elektronischen Person passt auch nicht ins Rechtsgefüge:

- Es fehlt an der Nähe zur natürlichen Person mit Willensbildungs(- und Geschäfts- sowie Schuld-)fähigkeit.
- Die juristische Person braucht Menschen: sie ist ohne Menschen (Organe, § 31 BGB) nicht handlungsfähig. Die KI soll hingegen gerade menschenunabhängig handlungsfähig sein.
- Tiere sind gemäß § 90a S. 1 BGB keine Sachen. Auf sie sind gemäß § 90a S. 3 BGB die für Sachen geltenden Vorschriften entsprechend anzuwenden, soweit nichts anderes bestimmt ist.

Der Mensch bleibt immer involviert. Ihm dienen die hier diskutierten technischen Verfahren lediglich als Entscheidungshilfe.

2. Überschätzung der technischen Möglichkeiten...

„From a technical perspective, [the idea] is based on an overvaluation of the actual capabilities of even the most advanced robots, a superficial understanding of unpredictability and self-learning capacities and, a robot perception distorted by Science-Fiction and a few recent sensational press announcements.“

Robotics Open Letter to the European Commission, robotics-openletter.eu, eine Stellungnahme zum Bericht 2015/2103/(INL)

„The original question, `Can machines think?` I believe to be too meaningless to deserve discussion.“

Alan M. Turing, Mechanical Intelligence

3. ...infolge von vermeidbaren Missverständnissen

Wir sollten unsere Kommunikation verbessern, um Fehlannahmen und hierauf gründende Rechtssätze zu verhindern. Aber dabei berücksichtigen:

Hier treffen mindestens zwei verschiedene Disziplinen aufeinander, die jeweils für sich dieselben Worte anders definieren und Probleme in anderer Herangehensweise lösen. Sie sind nicht austauschbar, ohne Widersprüche und Unsicherheiten innerhalb der jeweils eigenen Disziplin auszulösen.

V. Wer ist also verantwortlich?

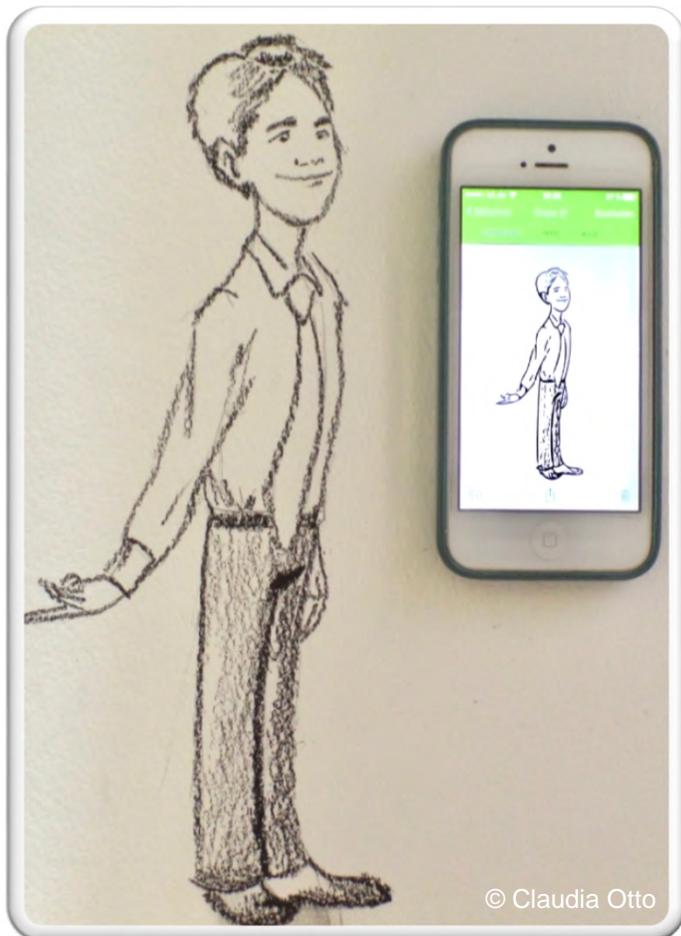


V. Wer ist also verantwortlich?

Eine Rechtsperson.

Eine Trägerin von Rechten und Pflichten.

Keine Fantasiegestalt.



1. Beispiel: Wechselwirkungs-App

„Wir müssen uns eingestehen, dass wir viele Krankheitsprozesse noch nicht vollständig verstehen. Eine Maschinenintelligenz wird nur dann ursächliche Zusammenhänge zwischen chemischen Strukturen und ihren Effekten lernen können, wenn wir sie mit den entsprechenden Daten füttern.“

Prof. Dr. Gisbert Schneider, Professor für Computer-Assisted Drug Design an der ETH Zürich (Zukunftsblog der ETHZ vom 28.03.2019)

2. Beispiel: Vordiagnose-App

„Es erfordert ein gehöriges Fachwissen, die Kenntnis englischer Fachbegriffe, haptische Arbeit, Software-Knowhow und viel Zeit, um diese Tools zum Wohle der Patienten einzusetzen. Fachzentren für seltene Erkrankungen greifen häufig auf solche Werkzeuge zurück. Im eng getakteten Alltag niedergelassener Haus- und Fachärzte hingegen ist dafür kaum Platz, zumal die ärztliche Gebührenordnung zeitaufwendige diagnostische Rechercharbeit derzeit nur unzulänglich entlohnt.“

Prof. Dr. Martin Hirsch, Leiter des Instituts für Künstliche Intelligenz in der Medizin an der Universität Marburg (FAZ, Verlagsspezial)

„KI“ ist das Ergebnis von Menschenwissen, das vom Menschen korrekt verknüpft werden muss.

VI. Was bedeutet das für die (Rechts-)Praxis?

Nicht an Magie glauben. Nicht auf Wunder hoffen.

Das tun auch die Gerichte nicht.

Beispiel: OLG Hamm, Urteil vom 22. März 2019 – 9 U 93/17 (Rn 49)

„Der Senat schließt sich der vom BGH vertretenen weiten Auslegung der Haftungsnorm des § 7 Abs. 1 StVG an. Für sie spricht entscheidend der vom BGH im Tiefgaragenfall angeführte weite Schutzzweck der Haftungsvorschrift des § 7 Abs. 1 StVG, nämlich Dritte von allen von einem Kraftfahrzeug ausgehenden Gefahren zu schützen. Für den Dritten macht es keinen Unterschied, ob ihm der Schaden durch ein Versagen von Fahrzeugkomponenten entstanden ist, die für die Fortbewegungs- und die Transportfunktion des Fahrzeugs zwingend erforderlich sind oder nur der Bequemlichkeit der Fahrzeugnutzer oder anderen Zwecken des Fahrzeuges wie etwa dessen Wohnfunktion dienen.“

VI. Was bedeutet das für die (Rechts-)Praxis?

1. Die Hauptaufgabe ist...

sich durch bestmögliches Beherrschen der Bedingungen zu schützen und zu befähigen.

Das bedeutet, geltende rechtliche Vorgaben zu kennen und umzusetzen, insbesondere durch Sicherstellung von

- Entwicklungs-/Produktionsverständnis (Frage: Sind Design-Fehler nach Möglichkeit ausgeschlossen?)
- Dokumentation für dauerhaftes Verständnis (Frage: Erlaubt die Dokumentation das nachträgliche Ausschließen von Fehlern?)
- Produktverständnis (Frage: Ist es eine Gefahrenquelle?)
- Rechtsverständnis (Frage: Welche besonderen Anforderungen ergeben sich aus der konkreten Gefahrenquelle?)
- Instruktion und Information (Frage: Kann der Nutzer das Risiko der Gefahrenquelle realistisch einschätzen?)
- Vertragsgestaltung (Frage: Wofür wird eingestanden? – AGB) und
- Schutzmaßnahmen (z.B. Produktbeobachtung)

„Responsibility finds a way. Irresponsibility makes excuses.“

Gene Bedley

2. Die antizipierte (Abgabe einer) Willenserklärung

Hier werden im Voraus die Bedingungen festgelegt, unter denen ein Vertragsschluss zustande kommen soll. Diese Bedingungen können auch einen Entscheidungsspielraum abstecken, innerhalb dessen der Festlegende bereit ist, sich rechtlich zu binden.

Beispiele:

- Warenautomat
- Service-Roboter der Bahn
- M2M-Gestaltungen (selbst-tankendes E-Auto)

3. Angemessene Risikoverteilung ermöglichen durch

- Erkenntnis: Perfektion gibt es nicht. Mit Fehlern ist immer zu rechnen. Vor allem mit menschengemachten Fehlern.
- Bestimmung des Fehlgebrauchs (vorhersehbare bestimmungswidrige Nutzung), v.a. durch Festlegung des Einsatzkontextes und Warnfunktionen.
- (Soweit sinnvoll) Bestimmung, dass nur Fachkräfte Zugang bzw. nutzenden Zugriff haben sollen.

A photograph of a Gothic cathedral tower, likely the spire of a church, silhouetted against a sunset sky. The tower features two large circular clock faces. To the right of the tower, a construction crane is visible, and further right, a smaller domed structure is silhouetted. The sky is filled with soft, scattered clouds, and the sun is low on the horizon, creating a warm, golden glow.

I-VI. Zusammenfassung und To-Dos

Stürme sind keine Zufälle.

Fehler auch nicht.

I-VI. Zusammenfassung und To-Dos

I. Was ist „KI“?

KI ist ein Sammelbegriff für „automatisierte datenbasierte technische Verfahren zur Lösung von komplexen Problemen.“

II. Trifft KI „eigene Entscheidungen“?

Nein. Der Mensch bildet den entscheidenden Willen.

III. Kann eine KI-Entscheidung falsch sein?

Das Ergebnis dürfte idR rechnerisch richtig sein. Inhaltliche Falschheit ist menschursächlich. Falsch meint eher unerwünscht.

IV. Braucht es eine elektronische Person?

Nein. Sie passt nicht in das Rechtsgefüge und basiert vor allem auf der Überschätzung der technischen Möglichkeiten.

V. Wer ist verantwortlich?

Ein Träger von Rechten und Pflichten. V.a. der Mensch, denn „KI“ ist das Ergebnis von Menschenwissen, das vom Menschen korrekt verknüpft werden muss.

VI. Was bedeutet das für die (Rechts-)Praxis?

Ziel muss das bestmögliche Beherrschen der Bedingungen sein.

To-Dos

- ✓ Kennen Sie Ihren rechtlichen Handlungsspielraum.
- ✓ Legen Sie klare Bedingungen für Ihr Produkt fest.
- ✓ Analysieren Sie Ihre Prozesse auf Nachvollziehbarkeit.
- ✓ Nehmen Sie eine Risiko-Analyse vor.
- ✓ Richten Sie Prozesse zur Bewältigung dieser Risiken ein.
- ✓ Schaffen Sie klare Rahmenbedingungen, z.B. AGB.
- ✓ Dokumentieren Sie umfassend.

Vielen Dank für Ihre Aufmerksamkeit!

Nun freuen wir uns über Ihre Fragen und Meinungen.



Claudia Otto

Dentons Europe LLP | Counsel

+49 69 450 012 392

claudia.otto@dentons.com



Prof. Dr. Anne Riechert

AI Frankfurt Rhein-Main e.V., Vorstand

University of Applied Sciences

rieichert@fb2.fra-uas.de