

# Ethische Implikationen Künstlicher Intelligenz

Beitrag im Rahmen der Vorlesung  
“Artificial Intelligence” (AIN), 22. 6. 2021

Prof. Dr. Annette Kleinfeld



- **Künstliche Intelligenz:** Was ist das eigentlich
- **Ethik:** Worum geht es?
- **KI-Anwendungen** in der Arbeitswelt
- (Angewandte) KI **aus ethischer Sicht**
- **Orientierungshilfen** für einen ethischen Umgang mit KI



# Künstliche Intelligenz: Was ist das eigentlich?



# Kurze Diskussion – Was meinen Sie?

- Was bedeutet „**Künstliche Intelligenz**“?
- Müssen wir uns vor ihr fürchten?  
Warum – warum nicht?



# Künstliche Intelligenz(en) (KI): Versuch einer Definition

- Übersetzt aus dem Englischen „**Artificial Intelligence**“ (AI)
- **Künstliche Intelligenz:**
  - Versuch, **mithilfe von Algorithmen** bestimmte **Entscheidungsstrukturen** des Menschen **nachzubilden**;
  - **Simulation** intelligenten/ menschlichen Verhaltens
- **Problem:** Keine eindeutige Definition von menschlicher Intelligenz
- **Synonyme:** Auffassungsvermögen, Verstand, Klugheit
- **Varianten menschlicher Intelligenz**
  - Kognitive Intelligenz
  - Sensomotorische Intelligenz
  - Emotionale Intelligenz
  - Soziale Intelligenz
- **Medium von KI:** Maschinen/  
Computer/ Roboter, die selbstlernfähig sind





***„Die KI erforscht,  
ob und wie  
Computer Dinge  
tun können, die  
wir Menschen  
heute noch  
besser können“***

Elaine Rich



# KI-Verständnis der Bundes- regierung

## Grundlage: Unterscheidung in starke und schwache KI

- „Die **„starke“ KI** formuliert, dass KI-Systeme die gleichen intellektuellen Fertigkeiten wie der Mensch haben oder ihn darin sogar übertreffen können.
- Die **„schwache“ KI** fokussiert auf die **Lösung konkreter Anwendungsprobleme** auf Basis der Methoden aus der Mathematik und Informatik, wobei die entwickelten Systeme zur Selbstoptimierung fähig sind. Dazu werden auch **Aspekte menschlicher Intelligenz nachgebildet** und formal beschrieben bzw. Systeme zur Simulation und Unterstützung menschlichen Denkens konstruiert.“



Quelle: [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de)

# KI-Verständnis der Bundesregierung (2)

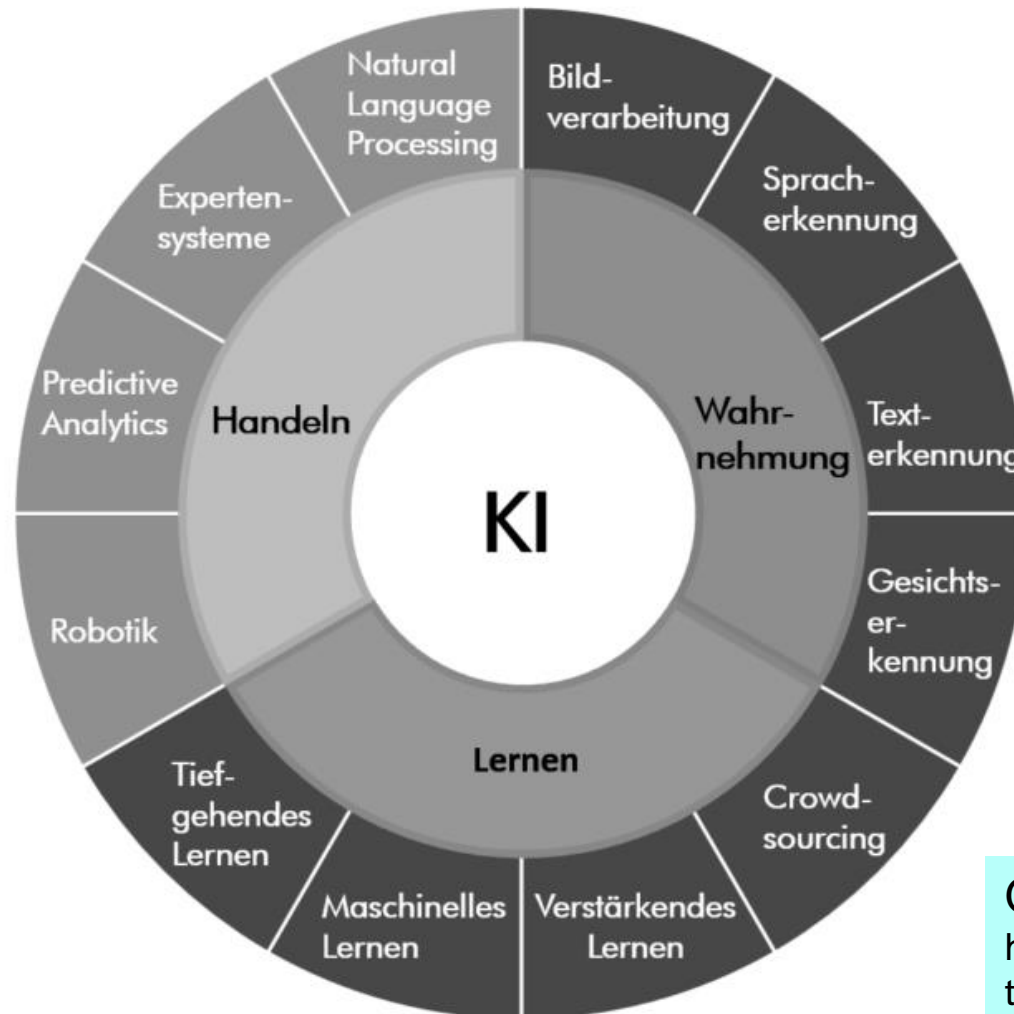
Bundesregierung orientiert sich bei ihrer Strategie an der **Nutzung der KI für die Lösung von Anwendungsproblemen** (schwache KI):

1. **Deduktionssysteme, maschinelles Beweisen**: Ableitung (Deduktion) formaler Aussagen aus logischen Ausdrücken, Systeme zum Beweis der Korrektheit von Hardware und Software;
2. **Wissensbasierte Systeme**: Methoden zur Modellierung und Erhebung von Wissen; Software zur Simulation menschlichen Expertenwissens (...)
3. **Musteranalyse und Mustererkennung**: induktive Analyseverfahren allgemein, insbesondere auch maschinelles Lernen;
4. **Robotik**: autonome Steuerung von Robotik-Systemen, d. h. autonome Systeme;
5. **Intelligente multimodale Mensch-Maschine-Interaktion**: Analyse und „Verstehen“ von Sprache (in Verbindung mit Linguistik), Bildern, Gestik und anderen Formen menschlicher Interaktion.

Quelle: [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de)



# Teilgebiete der Künstlichen Intelligenz: Überblick



Quelle:  
<https://www.marketinginstitut.biz/blog/kuenstliche-intelligenz/>

# Ethik: Worum geht es?



# Grundbegriffe: Moral, Ethos, Ethik

- **Moral** (von lat. *mos/mores* = Sitten)

- Antworten, die Menschen im Laufe der Geschichte in verschiedenen Kulturkreisen auf die Frage nach dem Guten und Bösen/sittlich Gesollten gefunden haben
- **Beispiele:** die 10 Gebote, die goldene Regel

- **Ethos** (griech. Begriff für *mores*)

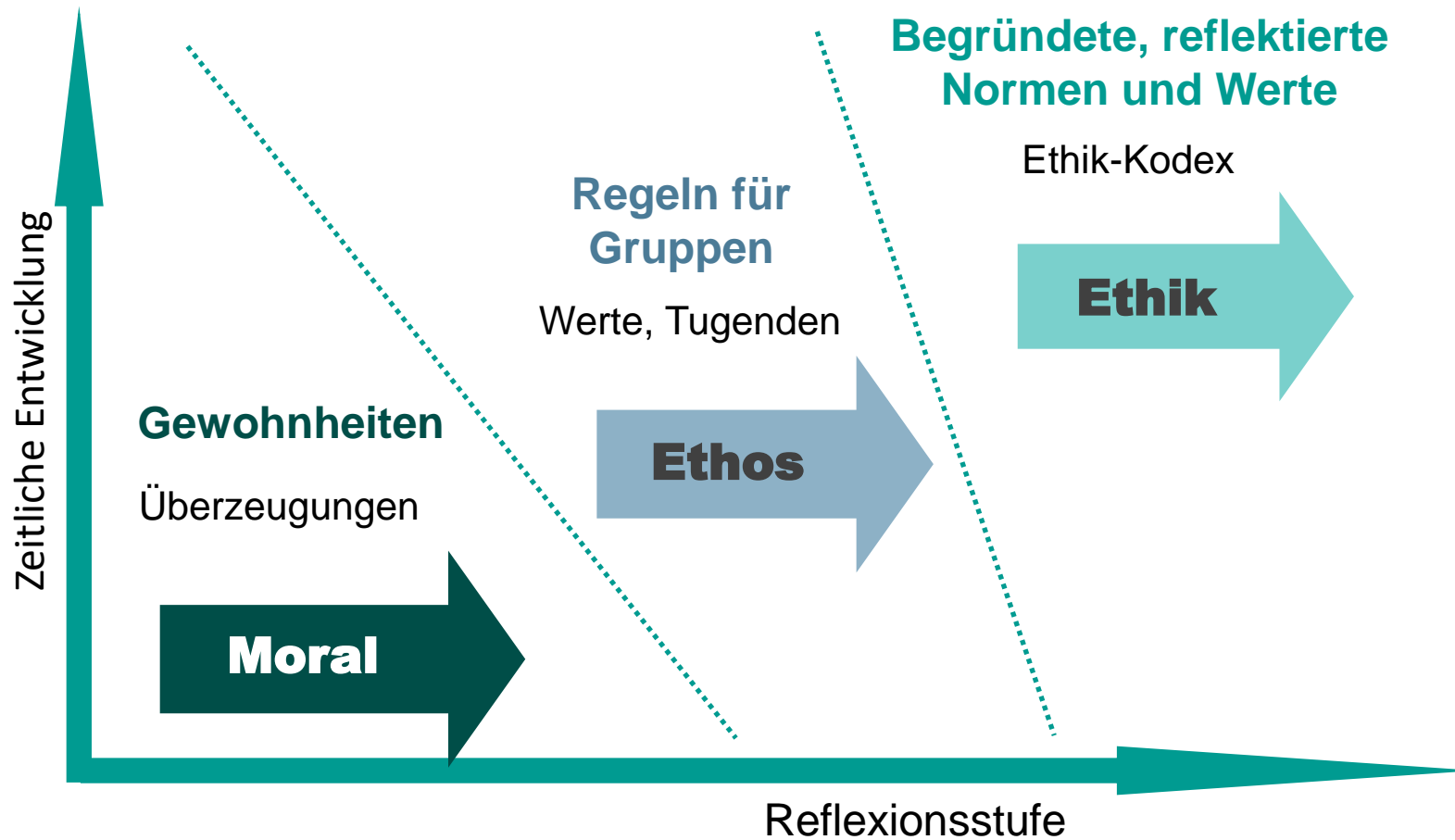
- Gesamtheit an Normen, Wertvorstellungen und Prinzipien Riten, Sitten und Gebräuche eines bestimmten sozialen Systems
- **Beispiele:** Wirtschaftsethos, Arbeitsethos

- **Ethik** (griech. „das sittliche Verständnis“)

- Aufstellung und Begründung von moralischen Orientierungen mit einem berechtigten Anspruch auf Allgemeingültigkeit

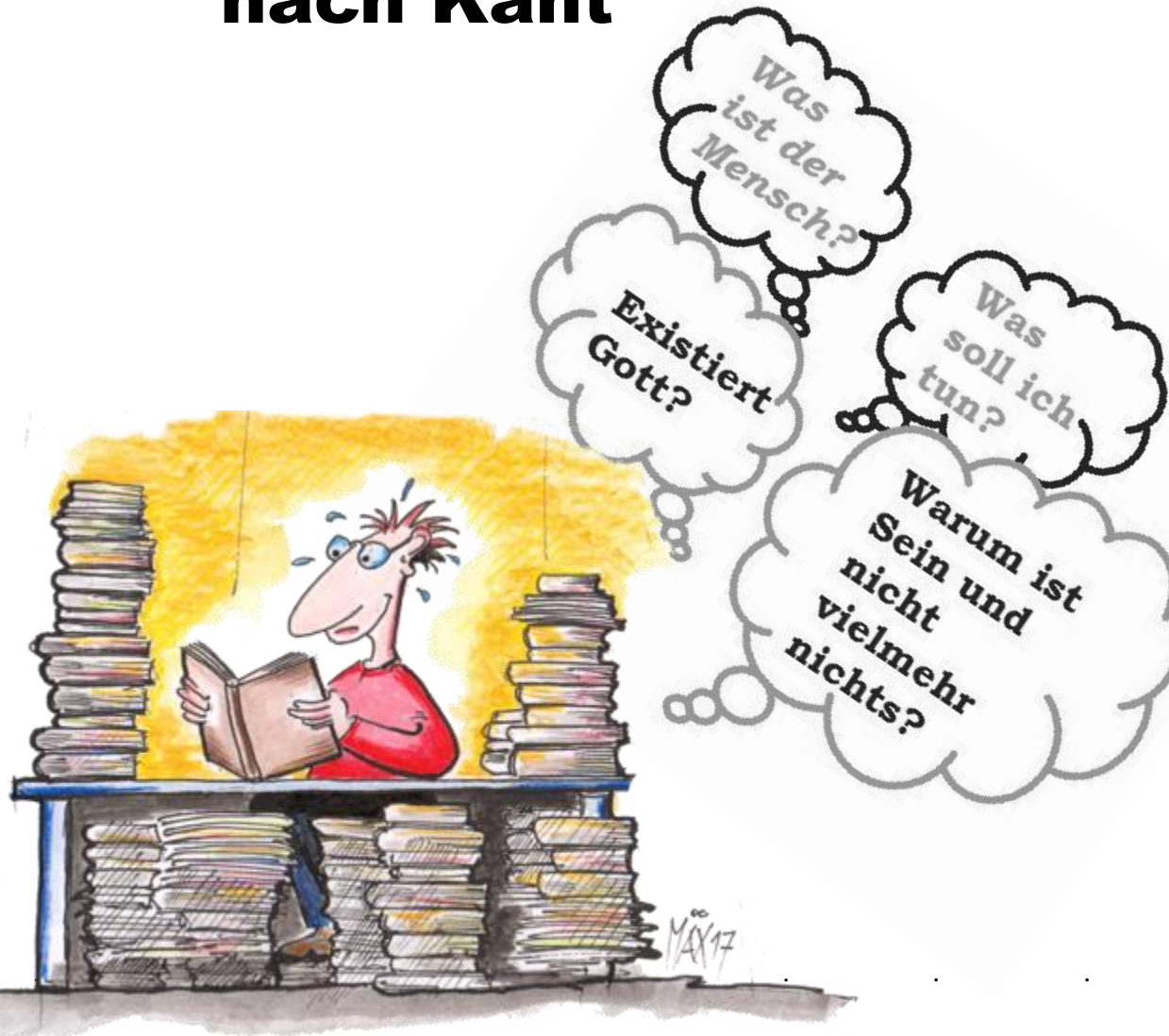


# Moral und Ethik: Unterschiede



(Dietzfelbinger, Daniel (2008): Praxisleitfaden Unternehmensethik, S. 65)

# Philosophie: Zentrale Fragen nach Kant



Alle münden  
in die Frage:  
**„Was ist der  
Mensch“?**



- „**Was soll ich tun**“? (Immanuel Kant)
- An welchen **Maßstäben** soll ich mein Handeln und Entscheiden orientieren?
- Nach welchen **Werten und Prinzipien** sollen wir unser Leben ausrichten?



## Übergeordnetes Ziel:

**Gutes, im Sinne eines glücklichen und gelingenden Lebens in einer Gemeinschaft ermöglichen**

# Aufgaben der Ethik als philosophischer Disziplin

- 1. Begründen:** Moralische Orientierungen mit einem berechtigten Anspruch auf Allgemeingültigkeit formulieren;
- 2. Maßstäbe definieren:** Höchste Werte und letzte Ziele des Menschen als Basis geltend machen
- 3. Reflexion und Überprüfung:** Entspricht die gelebte Moral/ Ethos diesen Maßstäben?

→ „ethisch“ = das, was

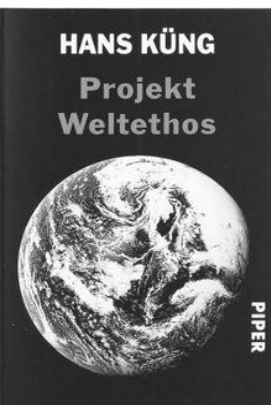
- moralisch gerechtfertigt ist,
- nicht nur legal, sondern auch legitim ist!

- „**ethisch relevant**“ = das, was Ziele, Aufgaben und Maßstäbe der Ethik betrifft

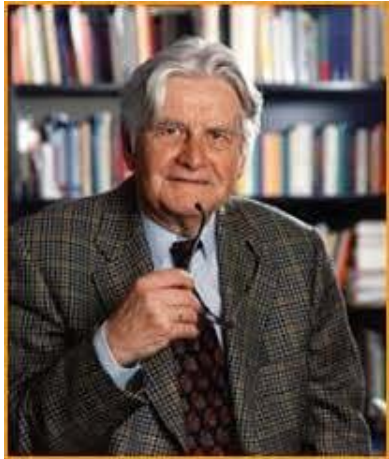


# Was sind die „richtigen“ Maßstäbe?

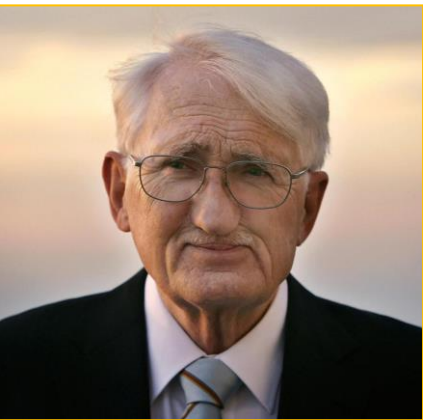
- Übergeordnete **Orientierungen**, die der **Mehrheit der Menschheit wichtig** sind. Insbesondere
  - **Frieden**
  - **Gerechtigkeit** im Sinne von Chancengleichheit und Fairness,
  - **Vernunft** im Sinne von „gesundem Menschenverstand“ und rationaler Argumentationsfähigkeit
  - **Freiheit** als Möglichkeit und Fähigkeit, sein Leben und Handeln selbst zu bestimmen (Autonomie)
  - **Menschenwürde** als Grund für die Pflicht zur Achtung der Menschenrechte (z.B. körperliche und seelische Unversehrtheit)
- Aus ihnen lassen sich **verallgemeinerbare Grundsätze** und **Spielregeln** des (zwischen)menschlichen Handelns und Entscheidens ableiten, wie z.B.
  - die „**goldene Regel**“ / der **Kategorische Imperativ (Kant)**
  - das „**Konsensprinzip**“ (**Apel/ Habermas**)
  - das „**Prinzip Verantwortung**“ (**Hans Jonas**)



# Beispiel: Die formale Diskurs-Ethik von Apel und Habermas



**Karl Otto Apel**



**Jürgen Habermas**

**Begründer:** **Karl-Otto Apel** (geb. 1922) und **Jürgen Habermas** (geb. 1929)

- **Ziel:** Das moralisch Gute bzw. Richtige im Rahmen eines demokratischen Diskurses ermitteln, der auf die Zustimmung aller Beteiligten und potentiell Betroffenen abzielt (= **Konsensprinzip**)
- **Weg:** Führen von idealen (= qua Gedankenexperiment) oder realen Diskursen, die bestimmten Voraussetzungen unterliegen, u.a.:
  - Teilnehmer sind zum Austausch von rationalen Argumenten **fähig und willig**
  - Teilnehmer **respektieren sich** als Personen „auf Augenhöhe“ (= herrschaftsfreier Diskurs)
- **Mittel:** Austausch von Argumenten nach definierten Regeln (allem voran Verständlichkeit, Wahrheit, Korrektheit, Wahrhaftigkeit)
- **Instrumente:** **Diskursprinzip + Universalisierungsprinzip**
- **Aktualität:** Anwendung z.B im Rahmen internationaler Normungsprojekte, im Bereich der angewandten Ethik, im Kontext gesellschaftspolitischer Prozesse, als Stakeholder-Dialoge von Unternehmen, etc.

# Beispiel: Verantwortungsethik von Hans Jonas (1903 – 1993) (Quelle: Das Prinzip Verantwortung, 1979)

**Kontext:** Kritik an der einseitigen Konzentrierung bisheriger Ethik auf den Menschen und dessen unmittelbaren Handlungsradius



**Hans Jonas**

- **Ziel:** Entwicklung einer erweiterten, **planetarischen Ethik**, die die Technologisierung ebenso wie die Rechte von Tieren und der Natur mit berücksichtigt und als Gegenstand menschlicher Verantwortung anerkennt
- **Weg: Verantwortung** als obersten Wert festlegen, definieren und als Maßstab geltend machen
- **Mittel:** Ausrichtung am **Ökologischen Imperativ** in den folgenden beiden Varianten
  - „Handle so, dass die Wirkungen deiner Handlungen verträglich sind mit der Permanenz echten menschlichen Lebens!“
  - „Handle so, dass die Wirkungen deiner Handlungen nicht zerstörerisch sind für die künftige Möglichkeit solchen Lebens!“
- **Aktualität: Grundlage einer Ethik im Sinne** der Leitidee der nachhaltigen Entwicklung und **des Prinzips der Nachhaltigkeit**

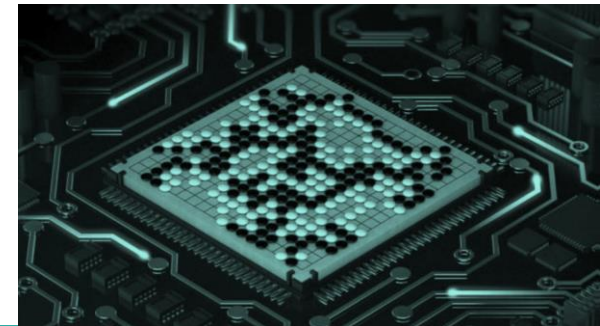


# KI-Anwendungen in der Arbeitswelt



# Anwendungen Künstlicher Intelligenz: Beispiele

- **Suchmaschinen**
- **Maschinelle Übersetzung**
- **Spracherkennung / Spracherfassung**, z.B. bei Navigationssystemen, Mobiltelefonen
- **Texterkennung / Textgenerierung**, z.B. von Eilmeldungen, Werbung oder für besonders strukturierte Daten
- **Data Mining**: Methoden zur Extraktion von Kerninformationen aus nicht- oder nur schwach strukturierten Texten, wie es etwa zur Erstellung von Inhaltsanalysen benötigt wird.
- **Argumentation Mining**: Analyse von Argumentationsstrukturen in Texten
- .....



# Professionelle Anwendungsgebiete von KI

## Zentrale Bereiche und Branchen

- **Medien**
- **Medizin und Pflege**
- **Juristik**
- **Hochschulbildung**
- **Militär**
- .....

## Branchenunabhängig

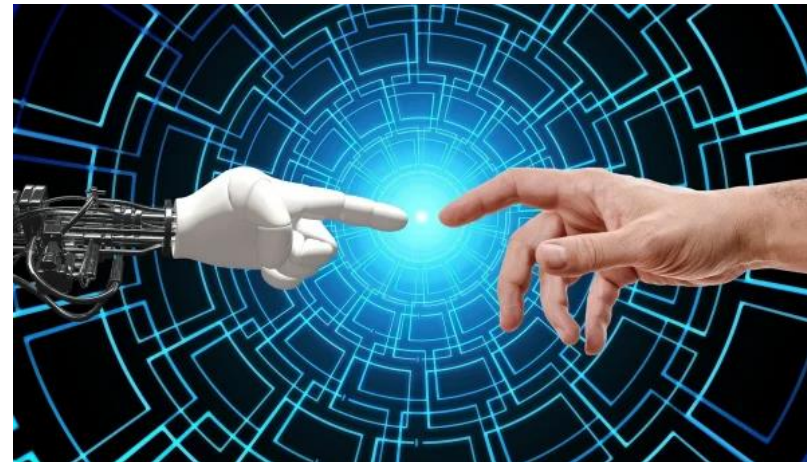
- **Verwaltung**
- **HR-Management**
- **Produktion**
- **Marketing**
- **Weiterbildung**
- **Logistik .....**



# (Angewandte) KI aus ethischer Sicht



- Damit angewandte KI ihr **positives Potenzial für Mensch, Gesellschaft und Arbeitswelt** entfalten kann, müssen sowohl ihre Entwicklung als auch ihre Nutzung **von ethischer Reflexion begleitet** sein!
- Nur der Mensch ist dazu in der Lage!
- **Künstliche Intelligenz** – schwache ebenso wie starke KI – wird dazu **nicht fähig** sein!
- Denkbar und machbar erscheint es lediglich, **KI „Moral zu lehren“**
- Geforscht wird daher an einer entsprechenden Programmierung und **Ausgestaltung von Algorithmen**





# Lassen sich Algorithmen „moralisch erziehen“?

*„Es gibt kein „**Moralzentrum**“ **im Gehirn**, sondern ein Zusammenspiel verschiedener neuronaler Netzwerke.*

*Die Entwicklung der Fähigkeit zu moralischem Handeln ist nicht nur an die Netzwerke des Gehirns gebunden, sondern eingebettet in die Prozesse des gesamten Körpers in seiner Lebenswelt mit sozialen Traditionen.“*

Prof. Dr. med. MARKUS FRINGS,  
Neurologe

Quelle: <https://www.ghst.de/ki-ethik/>



# Zentrale Fragen aus ethischer Perspektive



Bildquelle: <https://www.boell.de/de/podcast-kuenstliche-intelligenz>

- Wo und wie können KI-Anwendungen ein **gutes, menschenwürdiges Leben fördern**?
- **Unter welchen Voraussetzungen/** Rahmenbedingungen unterstützen sie Mensch und Gesellschaft, anstatt ihnen zu schaden?

# W W Verantwortung von Organisationen: G S Eruierung von Chancen durch KI

- Kann durch den Einsatz von KI das eigene **Geschäftsmodell im Interesse von Mensch und Gesellschaft (um)gestaltet** werden?
- Kann mithilfe von KI ein (noch) **größerer Beitrag zur nachhaltigen Entwicklung**, insbesondere zur **AGENDA 2030** geleistet werden?
- Wo und wie kann KI **zur Lösung globaler gesellschaftlicher Herausforderungen** eingesetzt werden?



## Beispiele:

- **Klima-resiliente Landwirtschaft** mittels Wettersensoren,
- „**Smarte Mobilität**“ mittels autonomer Fahrzeuge
- **Verbesserung der Energie-Effizienz** durch digitale Steuerung
- Einsatz von KI in **Medizin** und Pflege

• ..... • • • • •

# Herausforderung: Identifizierung von/ Umgang mit Risiken und Ängsten





# Ethische Fragen für KI nutzende Organisationen

- **Transparenz:** Wie können wir die Nachvollziehbarkeit algorithmisch gesteuerter Entscheidungsprozesse ermöglichen?
- **Rechenschaftspflicht:** Wer steht für die (negativen) Folgen gerade (z.B. rassistischer Bias bei Auswahlprozessen)?
- **Verantwortbarkeit:** Wissen wir, worauf wir uns beim Einsatz bestimmter Systeme einlassen (Folgenabschätzung, Dual Use-Problem)?
- **Nachhaltiges Datenwirtschaften:** Wie können wir Nutzung und Verarbeitung von Daten umwelt- UND sozialverträglich gestalten?
- **Privacy „made in Europe“:** Vertrauenswürdigkeit sicherstellen und sogar als Wettbewerbsvorteil nutzen?
- **Selbstverpflichtung zu wertebasiertem Design** (Ethics by Design)



Quelle: U.a. Petra Grimm, Hochschule der Medien, Stuttgart



# Orientierungshilfen für einen **ethischen** **Umgang mit KI**



# Kodizes, Richtlinien und Selbstverpflichtungen als Rahmen

- **2017: EU Charta der Digitalen Grundrechte** der Europäischen Union
- **2017: 10 Gebote der Digitalen Ethik** (Institut für Digitale Ethik – IDE, Stuttgart)
- **2018: 10 Leitlinien für die Digitalisierung von Unternehmen** (Institut für Digitale Ethik – IDE, Stuttgart)
- **2018: Algorithmen-Ethik** (Bertelsmann Stiftung)
- **2018: AI Principles** von Google
- **2018: 7 KI-Grundsätze** von SAP
- **2019: AI-Guidelines der EU** („Guidelines for Trustworthy Artificial Intelligence“)
- (.....)
- **2020: Einrichtung „Observatorium Künstliche Intelligenz in Arbeit und Gesellschaft“** durch das BMAS



Quelle: <https://www.ki-observatorium.de/>

# Das **KI-Observatorium** – ein Projekt der Denkfabrik des BMAS

**Arbeit des KI-Observatoriums ist in fünf Handlungsfelder unterteilt:**

1. Foresight und Technikfolgenabschätzung
2. KI in der Arbeits- und Sozialverwaltung
3. **Ordnungsrahmen für KI und Soziale Technikgestaltung**
  - Mitarbeit des KI-Observatoriums am KI-Weißbuchprozess zur Gestaltung eines zukünftigen Ordnungsrahmens auf EU-Ebene
  - Beschäftigung mit der Frage, wie Mensch und Maschine zukünftig zusammenarbeiten und wie sozio-technische Systeme gestaltet werden sollen;
4. Aufbau internationaler und europäischer Strukturen
5. **Gesellschaftlicher Dialog und Vernetzung**



Observatorium Künstliche Intelligenz  
in Arbeit und Gesellschaft

# Beispiel: Robert Bosch GmbH – das Vermächtnis des Gründers (1)



***„Die Fortschritte in der Entwicklung der Technik dienen in vollem Umfange dazu, der Menschheit die größten Dienste zu leisten.***

***Die Technik ist dazu bestimmt, der gesamten Menschheit ein Höchstmaß an **Lebensmöglichkeit** und **Lebensglück** zu verschaffen.“***

***Robert Bosch***

Quelle: <https://www.bosch.com/de/unternehmen/unsere-geschichte/>

# Beispiel: Robert Bosch GmbH – das Vermächtnis des Gründers (2)

„Sei Mensch und ehre Menschenwürde.“

Robert Bosch, 1920

„Immer habe ich nach dem  
Grundsatz gehandelt: **Lieber  
Geld verlieren als Vertrauen.**

Die **Unantastbarkeit meiner  
Versprechungen**, der Glaube  
an den Wert meiner Ware **und**  
an **mein Wort** standen mir  
stets höher als ein  
vorübergehender Gewinn.“



Robert Bosch

Quelle: <https://www.bosch.com/de/unternehmen/unsere-geschichte/>



# Beispiel Google (1): Die 7 Ethik-Regeln für KI-Anwendungen

- **Projekt Jedi:** Google verzichtet auf Milliarden-Deal mit Pentagon und verweist als Begründung auf **selbst auferlegte KI-Regeln**.

## „Objectives for AI applications



1. **Be socially beneficial.**
2. **Avoid creating or reinforcing unfair bias.**
3. **Be built and tested for safety.**
4. **Be accountable to people.**
5. **Incorporate privacy design principles.**
6. **Uphold high standards of scientific excellence.**
7. **Be made available for uses that accord with these principles.**

Sundar Pichai

# Beispiel Google (2): Die 7 Ethik-Regeln für KI-Anwendungen

## “AI applications we will not pursue



*In addition to the above objectives, we will not design or deploy AI in the following application areas:*

- ***Technologies that cause or are likely to cause overall harm. (...).***
- ***Weapons or other technologies whose principal purpose or implementation is to cause or directly facilitate injury to people.***
- ***Technologies that gather or use information for surveillance violating internationally accepted norms.***
- ***Technologies whose purpose contravenes widely accepted principles of international law and human rights”***

# Ergebnisse einer Meta-Studie von 2019

## Gemeinsamer Nenner aller (bis dahin) veröffentlichter KI-Richtlinien:

- **Schadensverhütung** / Nicht-Missbräuchlichkeit
- **Gerechtigkeit und Unparteilichkeit** (Fairness)
- **Freiheit und Autonomie**
- **Benefizienz** (Wohltätigkeit)
- **Verantwortung i.S.v. Rechenschaftspflicht**
- **Transparenz**
- **Kontrolle**
- **Erklärbarkeit**

Instrumentelle  
Grundsätze

Quelle: Jobin/Ienca/Vayena, 2019)

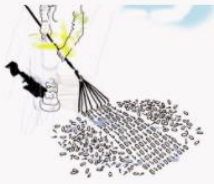


Quelle: SCOPAR-Umfrage 2013

## Algo.Rules

## Regeln für die Gestaltung algorithmischer Systeme

De | En



1. Kompetenz  
aufbauen



2. Verantwortung  
definieren



3. Ziele und erwartete Wirkung  
dokumentieren



4. Sicherheit  
gewährleisten



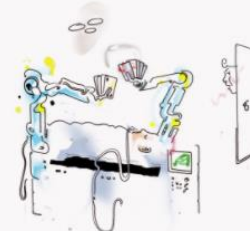
5. Kennzeichnung  
durchführen



6. Nachvollziehbarkeit  
sicherstellen



7. Beherrschbarkeit  
absichern



8. Wirkung  
überprüfen



9. Beschwerden  
ermöglichen

Quelle: <https://algorules.org/de/startseite>

# **Vielen Dank für Ihre Aufmerksamkeit!**

## **Und... Tschüss!**



**Dr. Kleinfeld CEC GmbH & Co. KG**  
**Corporate Excellence Consultancy**

Wollgrasweg 10  
38518 Gifhorn

**Prof. Dr. Annette Kleinfeld**

**Business & Society**  
**Fakultät Wirtschafts-, Kultur- und**  
**Rechtswissenschaften**  
**HTWG Konstanz**

Alfred-Wachtel-Straße 8  
78462 Konstanz

Tel.: +49 7531 206-404  
Fax: +49 7531 206-427

✉ [annette.kleinfeld@htwg-konstanz.de](mailto:annette.kleinfeld@htwg-konstanz.de)  
🖱 [www.htwg-konstanz.de](http://www.htwg-konstanz.de)

Tel.: +49 5371 941 67 67  
Fax: +49 5371 941 67 68

✉ [annette.kleinfeld@kleinfeld-cec.com](mailto:annette.kleinfeld@kleinfeld-cec.com)  
🖱 [www.kleinfeld-cec.com](http://www.kleinfeld-cec.com)

